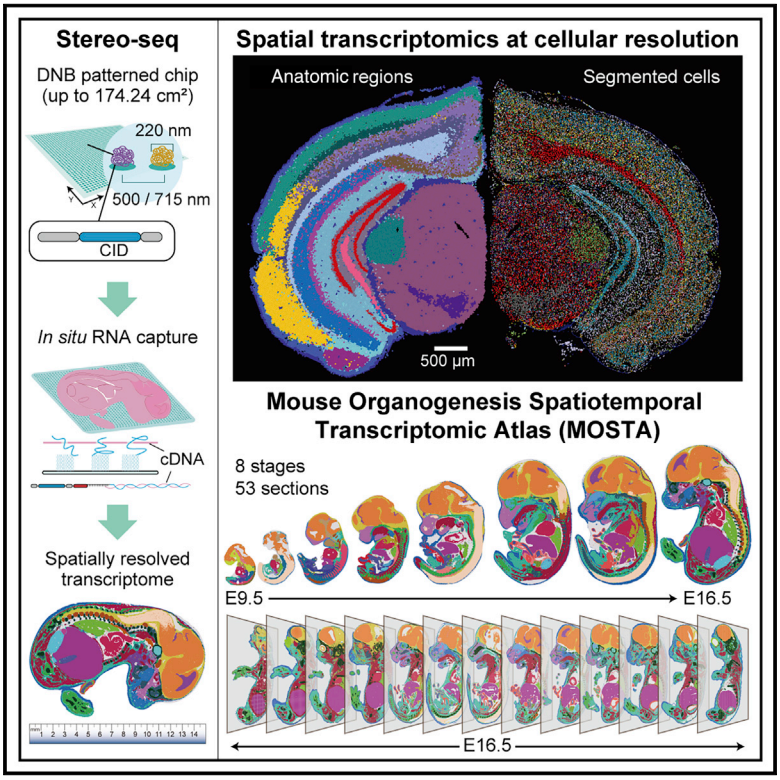
Cell

Resource

Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays

Graphical abstract



Highlights

- Stereo-seq enables large field-of-view spatial transcriptomics at cellular resolution
- Stereo-seq reveals the spatial cell-type heterogeneity of mouse embryonic tissues
- Stereo-seq maps the spatiotemporal transcriptomic dynamics during mouse organogenesis
- Stereo-seq defines the spatiotemporal window of developmental disease vulnerability

Authors

Ao Chen, Sha Liao, Mengnan Cheng, ..., Longqi Liu, Xun Xu, Jian Wang

Correspondence

miguelesteban@genomics.cn (M.A.E.), liyuxiang@genomics.cn (Y.L.), liulongqi@genomics.cn (L.L.), xuxun@genomics.cn (X.X.), wangjian@genomics.cn (J.W.)

In brief

Stereo-seq combines DNA nanoball-patterned arrays and tissue RNA capture to achieve large field-of-view spatial transcriptomics at cellular resolution, enabling the dissection of spatial cell-type heterogeneity of mouse embryonic tissues.

Chen et al., 2022, Cell *185*, 1777–1792
May 12, 2022 © 2022 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.cell.2022.04.003

CellPress

Cell

CellPress
OPEN ACCESS

Resource

Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays

Ao Chen,[1,2,26] Sha Liao,[1,26] Mengnan Cheng,[1,3,26] Kailong Ma,[1,26] Liang Wu,[1,3,4,26] Yiwei Lai,[1,5,26] Xiaojie Qiu,[6,7,26] Jin Yang,[8] Jiangshan Xu,[1,3] Shijie Hao,[1,3] Xin Wang,[1] Huifang Lu,[1] Xi Chen,[1] Xing Liu,[1] Xin Huang,[1,3] Zhao Li,[1] Yan Hong,[1] Yujia Jiang,[1,9] Jian Peng,[1] Shuai Liu,[1] Mengzhe Shen,[1] Chuanyu Liu,[1,10] Quanshui Li,[1] Yue Yuan,[1] Xiaoyu Wei,[1] Huiwen Zheng,[1,9] Weimin Feng,[1,3] Zhifeng Wang,[1,4] Yang Liu,[1] Zhaohui Wang,[1] Yunzhi Yang,[1,9] Haitao Xiang,[1,3] Lei Han,[1] Baoming Qin,[5] Pengcheng Guo,[5] Guangyao Lai,[5] Pura Muñoz-Cánoves,[11,12] Patrick H. Maxwell,[13] Jean Paul Thiery,[14] Qing-Feng Wu,[15] Fuxiang Zhao,[1] Bichao Chen,[1] Mei Li,[1] Xi Dai,[1,3] Shuai Wang,[1,3] Haoyan Kuang,[1] Junhou Hui,[1] Liqun Wang,[16] Ji-Feng Fei,[16] Ou Wang,[1] Xiaofeng Wei,[17] Haorong Lu,[17] Bo Wang,[17] Shiping Liu,[1,4] Ying Gu,[1,18] Ming Ni,[8] Wenwei Zhang,[1,19] Feng Mu,[8] Ye Yin,[1,20] Huanming Yang,[1,21] Michael Lisby,[2] Richard J. Cornall,[22] Jan Mulder,[23,24] Mathias Uhlén,[23,24] Miguel A. Esteban,[1,5,25,*] Yuxiang Li,[1,*] Longqi Liu,[1,9,10,*] Xun Xu,[1,18,27,*] and Jian Wang[1,21,*]

[1]BGI-Shenzhen, Shenzhen 518103, China
[2]Department of Biology, University of Copenhagen, Copenhagen 2200, Denmark
[3]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[4]Shenzhen Key Laboratory of Single-Cell Omics, BGI-Shenzhen, Shenzhen 518120, China
[5]Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
[6]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[7]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[8]MGI, BGI-Shenzhen, Shenzhen 518083, China
[9]BGI College & Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450000, China
[10]Shenzhen Bay Laboratory, Shenzhen 518000, China
[11]Department of Experimental and Health Sciences, Pompeu Fabra University (UPF), ICREA and CIBERNED, Barcelona 08003, Spain
[12]Spanish National Center on Cardiovascular Research (CNIC), Madrid 28029, Spain
[13]Cambridge Institute for Medical Research, Department of Medicine, University of Cambridge, Cambridge CB2 0XY, UK
[14]Guangzhou Laboratory, Guangzhou 510320, China
[15]State Key Laboratory of Molecular Development Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[16]Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China
[17]China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China
[18]Guangdong Provincial Key Laboratory of Genome Read and Write, Shenzhen 518120, China
[19]Shenzhen Key Laboratory of Neurogenomics, BGI-Shenzhen, Shenzhen 518103, China
[20]BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China
[21]James D. Watson Institute of Genome Sciences, Hangzhou 310058, China
[22]Medical Research Council Human Immunology Unit, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, UK
[23]Department of Protein Science, Science for Life Laboratory, KTH-Royal Institute of Technology, Stockholm 17121, Sweden
[24]Department of Neuroscience, Karolinska Institute, Stockholm 17177, Sweden
[25]Institute of Stem Cells and Regeneration, Chinese Academy of Sciences, Beijing 100101, China
[26]These authors contributed equally
[27]Lead contact
*Correspondence: miguelesteban@genomics.cn (M.A.E.), liyuxiang@genomics.cn (Y.L.), liulongqi@genomics.cn (L.L.), xuxun@genomics.cn (X.X.), wangjian@genomics.cn (J.W.)
https://doi.org/10.1016/j.cell.2022.04.003

SUMMARY

Spatially resolved transcriptomic technologies are promising tools to study complex biological processes such as mammalian embryogenesis. However, the imbalance between resolution, gene capture, and field of view of current methodologies precludes their systematic application to analyze relatively large and three-dimensional mid- and late-gestation embryos. Here, we combined DNA nanoball (DNB)-patterned arrays and *in situ* RNA capture to create spatial enhanced resolution omics-sequencing (Stereo-seq). We applied Stereo-seq to generate the mouse organogenesis spatiotemporal transcriptomic atlas (MOSTA), which maps with single-cell resolution and high sensitivity the kinetics and directionality of transcriptional variation during mouse organogenesis. We used this information to gain insight into the molecular basis of spatial cell heterogeneity and cell fate specification in developing tissues such as the dorsal midbrain. Our panoramic atlas will facilitate in-depth investigation of longstanding questions concerning normal and abnormal mammalian development.

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

INTRODUCTION

Since the time of Aristotle, understanding how a single totipotent cell, the zygote, develops into a complex organism such as a mammal in a precisely controlled manner, over time and space, is one of the most fascinating scientific challenges. The laboratory mouse is an excellent model animal to study development and establish parallels with larger mammalian species. Recent cell atlases using high-throughput single-cell RNA-sequencing (scRNA-seq) and single-cell assay for transposase accessible chromatin sequencing (scATAC-seq) technologies have provided an increasingly detailed view of mouse developmental gene expression and gene regulatory dynamics (Cao et al., 2019; He et al., 2020; Pijuan-Sala et al., 2020). However, the lack of topographical information greatly complicates the interpretations of the hierarchical mechanisms underlying the emergence of cell positioning and identities. High-throughput spatially resolved analysis of gene expression will be a critical enabler for further dissecting the intricacies of mammalian development.

Several remarkable methodologies have been developed recently that allow systematic spatially resolved transcriptomic profiling of tissue sections (Rao et al., 2021). These techniques all have specific advantages and disadvantages in resolution and applicability but share a major caveat in the limited field of view. Early mouse embryos (Lohoff et al., 2022; Peng et al., 2019), specific late-stage-embryonic tissues such as the cerebral cortex (Di Bella et al., 2021; Stickels et al., 2021), and whole midgestation embryos (Liu et al., 2020; Srivatsan et al., 2021; Yao et al., 2020) have been analyzed using these approaches. However, no available technology can profile whole mouse embryos in mid- or late-stage gestation in an unbiased manner and with high definition (single-cell resolution and high sensitivity). This poses a tremendous challenge for studying the differentiation continuum across separate areas of the same tissue and for comparing different tissues simultaneously. Here, we report the development of a DNB-based genome-wide technology, Stereo-seq, that combines single-cell resolution, high sensitivity, and a large field of view, and report its use to generate a spatially resolved transcriptomic atlas of mouse organogenesis.

RESULTS

DNB-patterned arrays enable large field-of-view spatially resolved transcriptomics with high definition

DNB sequencing is based on lithographically etched chips (patterned arrays) for *in situ* sequencing (Drmanac et al., 2010). We used these features as the foundation for a spatially resolved transcriptomic technology with high resolution and large field of view as follows. Standard DNB chips have spots with approximately 220 nm diameter and a center-to-center distance of 500 or 715 nm (Figure 1A, step 1), providing up to 400 spots for tissue RNA capture per 100 μm^2. DNB templates containing random barcodes are deposited on the patterned array, incubated with primers, and sequenced to obtain the data matrix containing the coordinate identity (CID) of every DNB (Figure 1A, step 2). The use of random barcode-labeled DNB achieves a large spatial barcode pool size (4^{25} distinct spots). Next, unique molecular identifiers (UMI) and polyT sequence-containing oligonucleotides are ligated onto each spot through hybridization with an oligonucleotide sequence containing the CID (Figure 1A, step 3). Frozen tissue sections (10 μm thickness) are loaded onto the chip surface, followed by fixation, permeabilization to capture the tissue polyA-tailed RNA, and finally reverse transcription plus amplification (Figure 1A, step 4). Amplified-barcoded cDNA is collected, used as template for library preparation, and sequenced together with the CID (Figure 1A, step 5). Computational analysis of the sequencing data allows high-resolution spatially resolved transcriptomics (Figure 1A, step 6). We named this approach spatial enhanced resolution omics-sequencing (Stereo-seq). Importantly, Stereo-seq has a larger number of spots per 100 μm^2, with smaller spot size and center-to-center distance, than any other published method (Figure 1B). So far, we have used Stereo-seq chips of up to an effective area of 13.2 cm × 13.2 cm for different tissue sizes, subdivided into 50, 100, and 200 mm^2 chips for profiling sections from the mouse olfactory bulb (~10.5 mm^2), a mouse hemibrain (~24.2 mm^2) and whole mouse embryos (~7.1 to ~76.1 mm^2), respectively (Figure S1A). These capture areas are substantially larger than those achieved by other reported technologies (Figure 1B), and we anticipate the application of unsliced chips to much larger tissues (Figure S1A).

To benchmark our technology, we profiled the mouse olfactory bulb (Lebrigand et al., 2020; Ståhl et al., 2016; Stickels et al., 2021; Vickovic et al., 2019). Stereo-seq captured UMI counts ranging on average from 69 per 2 μm (diameter) bin (bin 3, 3 × 3 DNB), 1,450 per 10 μm (diameter) bin (bin 14, 14 × 14 DNB, equivalent to ~1 medium size cell), and 133,776 per 100 μm (diameter) bin (bin 140, 140 × 140 DNB) (Figure 1C). This is superior to other reported technologies including Seq-Scope (848 UMI on average per 10 μm diameter bin) when compared at the same resolution. The distribution of genes and UMI per bin was highly consistent between adjacent sections (R^2 = 0.963) (Figures S1B and S1C). Of note, conventional cell taxonomy analysis based on scRNA-seq uses algorithms such as Leiden clustering to group cells based on transcriptome similarity (Traag et al., 2019). However, applied to spatially resolved data, this is problematic due to the lack of consideration of spatial coordinates. To fully utilize spatially resolved transcriptomic information, several algorithms have been applied to the analysis of datasets derived from technologies such as Visium and Slide-seqV2 (Dries et al., 2021; Hu et al., 2021; Zhao et al., 2021). Considering the specific Stereo-seq features, we developed a spatially constrained-clustering (SCC) algorithm optimized for analysis of larger, more complex tissues (see STAR Methods). This approach clustered bins over a continuous area from which tissue domains are annotated (Figure S1D). The distribution of specific markers (*Pcp4* and *Slc17a7*) captured by Stereo-seq exhibited the expected similarity to *in situ* hybridization (ISH) data taken from Allen Brain Atlas (ABA) (Lein et al., 2007) and clearer patterns than HDST (Vickovic et al., 2019) or Slide-seqV2-resolved (Stickels et al., 2021) expression data (Figure S1E).

To further demonstrate the robustness of Stereo-seq and prove its ability to dissect tissues with cellular resolution, we profiled an adult mouse coronal hemibrain section. We applied a

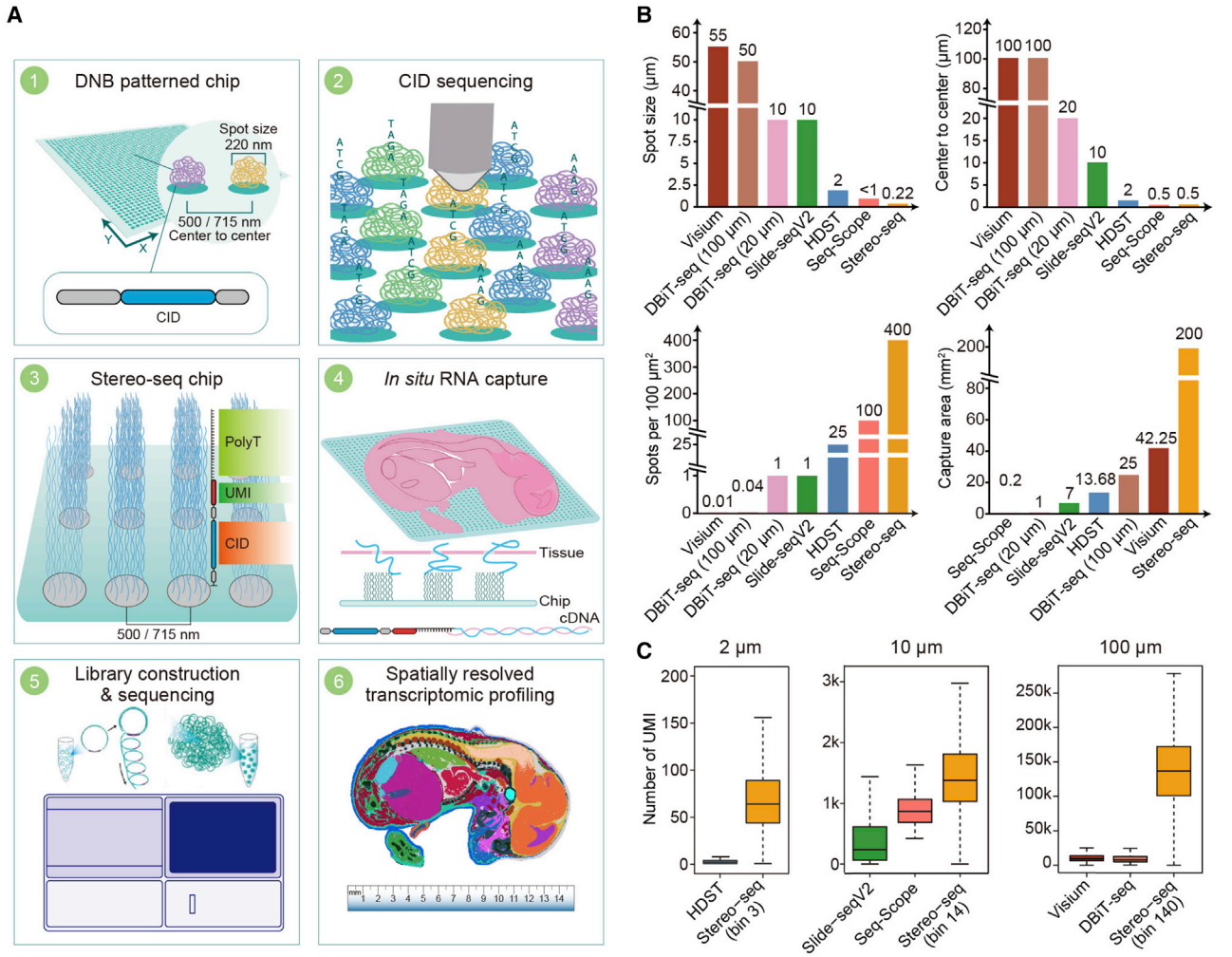


Figure 1. Stereo-seq enables high-definition spatially resolved transcriptomics with large field of view

(A) Stereo-seq pipeline. Step 1, design of the DNB patterned array chip. Step 2, *in situ* sequencing to determine the spatial coordinates of uniquely barcoded oligonucleotides. Step 3, preparation of capture probes by ligating the UMI-polyT containing oligonucleotides to each spot. Step 4, *in situ* RNA capture from tissue. Step 5, cDNA amplification, library construction, and sequencing. Step 6, data analysis.

(B) Stereo-seq achieves a smaller spot size (upper left), higher resolution (upper right), higher number of spots per 100 μm^2 (bottom left), and larger capture area (bottom right) than other reported methods. Samples used for the comparison included mouse olfactory bulb (Stereo-seq, Visium [Lebrigand et al., 2020], Slide-seqV2 [Stickels et al., 2021], and HDST [Vickovic et al., 2019]), E10 mouse embryo (DBiT-seq) (Liu et al., 2020), and mouse liver (Seq-Scope) (Cho et al., 2021). Note that since Seq-Scope uses a random array, the size of each pixel was estimated according to the published dataset.

(C) Box plots showing the number of transcripts captured by Stereo-seq at the indicated resolution in comparison with reported HDST, Slide-seqV2, Visium, DBiT-seq, and Seq-Scope datasets. Samples in those datasets used for comparison are as in (B).

See also Figure S1.

nucleic acid dye before *in situ* permeabilization to assess the correlation between the obtained signals and the Stereo-seq data after sequencing. There was substantial colocalization of the dye and the aggregated transcripts detected by Stereo-seq (Figures 2A and 2B), illustrating that cells can be segmented in an image-dependent manner. To analyze the brain data, we performed unsupervised SCC of binned (bin 50, 50 × 50 DNB bins, 25 μm diameter) Stereo-seq data to identify the different anatomic regions. These included the cortex and subcortical regions such as hippocampus, thalamus, and striatum, among others (Figure 2C). Then, we applied an image-based cell segmentation (see STAR Methods) that enables accurate integration of nucleic acid and transcript images into an image-guided gene-by-cell matrix. After filtering out the low-capture cells, we obtained 50,140 segmented cells with an average of 1,910 UMI and 792 genes per cell (Figure S1F). The segmented cells displayed differential spatial distribution between nucleus-localized transcripts (*Malat1* and *Neat1*) and cytoplasmic-enriched mitochondrial transcripts within each cell, consistent with the subcellular resolution of Stereo-seq (Figure S1G). We also examined the lateral RNA diffusion by measuring the spatial distribution of the neuropeptide *Vip* in *Vip*$^+$ interneurons, observing that the

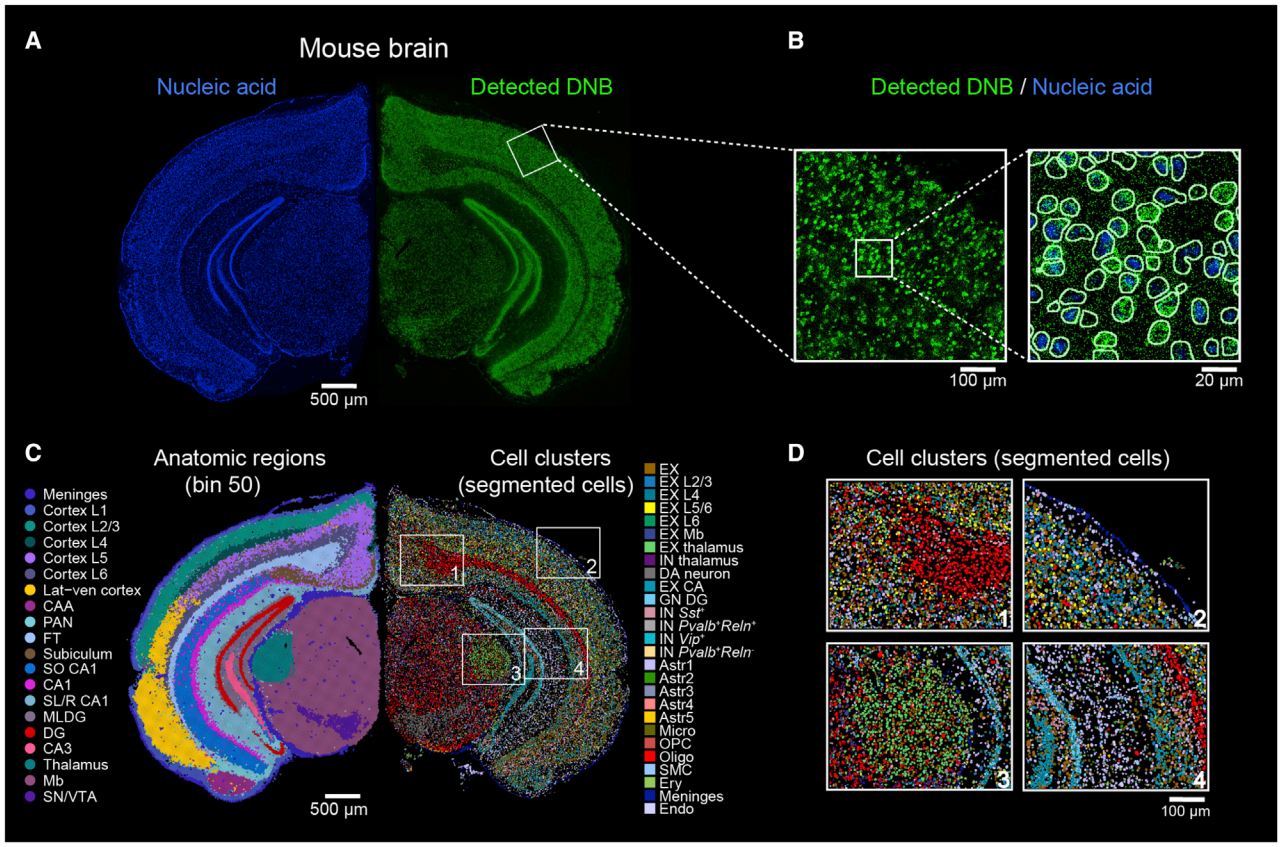


Figure 2. Stereo-seq dissects the adult mouse brain with cellular resolution

(A) Left: nucleic acid staining image of an adult coronal mouse hemibrain section. Right: spatial visualization of the detected DNB signals in the same section. Scale bars, 500 μm.

(B) Left: magnification showing the detected DNB signals from the region squared in (A). Right: superimposed nucleic acid staining and captured DNB signals from the region squared in the left. Scale bars, 100 μm (left) and 20 μm (right). Outlines are segmented cell boundaries.

(C) Left: unsupervised SCC of the same mouse hemibrain section analyzed by Stereo-seq at bin 50 resolution. Right: spatial visualization of segmented cell clusters by unsupervised clustering. Bins and cells are colored by their annotation. Lat-ven, lateral-ventral cortex; CAA, cortical amygdalar area; PAN, posterior amygdalar nucleus; FT, fiber tract; SO CA1, stratum oriens area 1; CA1, cornu ammonis area 1; SL/R CA1, stratum lacunosum/raditum cornu ammonis area 1; MLDG, molecular layer of dentate gyrus; DG, dentate gyrus; CA3, cornu ammonis area 3; Mb, midbrain; SN/VTA, substantia nigra/ventral tegmental area. EX, excitatory glutamatergic neuron; IN, GABAergic interneuron; DA, dopaminergic neuron; GN DG, granule cell of dentate gyrus; Astr, astrocyte; Micro, microglia; OPC, oligodendrocyte precursor cell; Oligo, oligodendrocyte; SMC, smooth muscle cell; Ery, erythrocyte; and Endo, endothelial cell. Scale bars, 500 μm.

(D) Magnification images showing cell-type localization in the regions squared in (C). Scale bars, 100 μm.

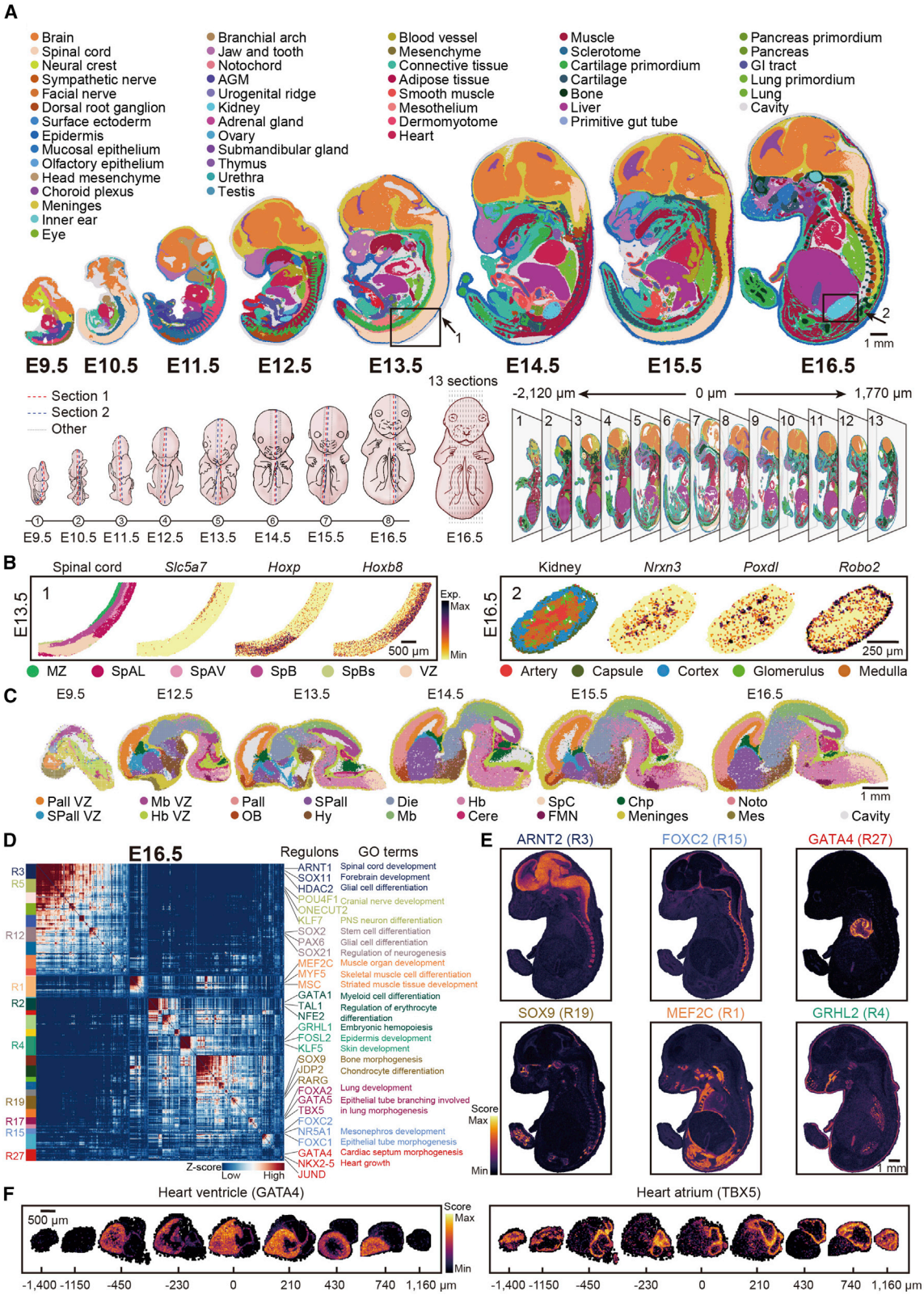
See also Figure S1.

diffusion distance is moderately greater with Stereo-seq (on average, 6.84 μm) than smFISH (on average, 5.32 μm) (Codeluppi et al., 2018) (Figure S1H). Unsupervised clustering of the gene-by-cell matrix and subsequent detailed annotation based on the expression of marker genes identified different types and subtypes of cells. These included neurons (i.e., eight types of excitatory glutamatergic neurons from different layers of cortex or regions and four types of GABAergic interneurons expressing diverse neurochemical markers including *Sst*, *Pvalb*, *Vip*, and *Reln*) as well as non-neuronal cells (i.e., five types of astrocytes, oligodendrocytes, oligodendrocytic precursors, microglia, and vascular cells) (Figures 2C, 2D, S1I, and S1J). The validity of these annotations was confirmed by comparison with a reported scRNA-seq dataset for the concordant spatial domains (Zeisel et al., 2018) (Figures S1K and S1L). Moreover, the location of the different cell types was consistent with existing knowledge of the corresponding anatomic regions annotated using bins. Therefore, Stereo-seq data can be used to both identify functional anatomic regions in large tissues through bin-based clustering, and cell types based on image-guided cell segmentation.

Hence, Stereo-seq spatially characterizes the transcriptomic organization and individual cell-type composition of complex tissues in an unbiased manner with high sensitivity and large field of view.

Spatially resolved transcriptomic atlas of mouse organogenesis

As proof of principle of the technological strength of Stereo-seq, we initiated the mouse organogenesis spatiotemporal transcriptomic atlas (MOSTA) to map the spatiotemporal transcriptomic



(legend on next page)

dynamics of the developing mouse embryo at full scale. We profiled 53 sagittal sections from C57BL/6 mouse embryos spanning E9.5–E16.5 with one-day intervals. For E9.5–E15.5 stages, four to six sections were included from different embryos. For E16.5, 18 sections were profiled, with a series of 13 sections from one single embryo and five from another to assess reproducibility, allowing coverage of all major tissues and organs (Table S1). These stages include most of the key events in mouse organogenesis (Cao et al., 2019).

To gain an unbiassed global view of the spatial transcriptomic patterns, we aggregated our embryo datasets for individual sections into bins (bin 50). In total, we retrieved transcriptomic information for 3,511,403 bins. The average number of captured genes per bin ranged from 1,770 at E15.5 to 3,900 at E10.5 and of UMI from 4,357 at E15.5 to 13,789 at E9.5 (Figure S2A; Table S1). The distribution of genes and UMI per bin varied between sections from different time points (Figure S2B). Unsupervised SCC of these bins showed transcriptomic configurations matching the localization of major tissues and organs (e.g., skin, bone, muscle nervous system, thymus, heart, lung, liver, pancreas, kidney, adrenal gland, gastrointestinal tract, genitourinary tract, and ovary) at each time point. Remarkably, the external boundaries of these clusters closely resembled the anatomic regions (Figures 3A and S2C and MOSTA website). We also reconstructed the developmental trajectories for the annotated regions in whole embryos using TOME (Qiu et al., 2022a). This showed spatiotemporal connections largely consistent with our contemporary understanding of mouse development (Figure S2D). Tissue-specific identities were confirmed by visualizing specific marker genes (e.g., *Acta2* in the smooth muscle, *Myog* in the skeletal muscle, *Krt5* in the epidermis, *Sftpc* in the lung, *Afp* in the liver, and *Col2a1* in the cartilage primordium) on the spatial maps across all embryonic time points (Figure S3A). Notably, the extended coverage of the E16.5 embryo allowed us to obtain serial images in developing tissues and organs, such as the meninges (*Atp1a2*$^+$), the spinal cord (*Hoxb8*$^+$), and the lung (*Sftpc*$^+$) (Figure S3B). In addition, since Stereo-seq captures high density of signals, we performed spatial reclustering of bins from areas corresponding to selected tissues, which revealed detailed subregions. For example, in the E13.5 spinal cord, we identified the ventricular zone (*Hopx*$^+$), marginal zone (*Slc5a7*$^+$), basal plate (*Vsnl1*$^+$), ventral (*Fut9*$^+$) and lateral parts of the spinal alar plate (*Hoxb8*$^+$), and superficial stratum of spinal basal plate (*Pdyn*$^+$); in the E16.5 kidney, we localized artery (*Myl9*$^+$), cortex (*Lhx1*$^+$), medulla (*Nrxn3*$^+$), capsule (*Robo2*$^+$), and glomeruli (*Poxdl*$^+$) (Figure 3B). We also reclustered the E9.5, E12.5, E13.5, E14.5, E15.5, and E16.5 brains (Figure 3C). The subsequent clusters matched well with anatomically defined brain regions including the ventricular and mantle zones of the pallium, subpallium, midbrain, hindbrain, diencephalon, cerebellum, hypothalamus, olfactory bulb, and choroid plexus. The expression of many known marker genes in these brain clusters is consistent with known cellular distributions across the corresponding anatomic regions (Figures S3C and S3D). Projection of publicly available scRNA-seq data (La Manno et al., 2021) of the developing brain onto the reclustered E9.5 brain using Tangram (Biancalani et al., 2021) identified the localization of progenitor cell types including secondary organizers (e.g., midbrain basal plate, midbrain floor plate, ventricle roof plate, and hindbrain floor plate) (Figure S3E), which are specialized substructures responsible for regional patterning.

Next, we studied functional enrichments within specific areas contained in the individual tissue clusters generated with Stereo-seq by applying Hotspot (DeTomaso and Yosef, 2021), an algorithm that measures nonrandom variation to recognize informative gene programs. This is important because amalgamation of different signals in complex tissues can complicate the interpretations. The resulting gene programs showed well-defined spatial patterns that, to a large extent, match anatomic regions and contained sets of genes that could be validated by ISH (e.g., *Kitl* colocalizing with *Shh* in the E9.5 notochord and *Pantr1* with *Sox2* in the E11.5 brain) (Figures S4A–S4E). *Kitl* is a pleiotropic factor that acts in the ontogeny of different cell types (Broudy, 1997). The colocalization of *Pantr1* with *Sox2* is consistent with the regulatory function of this long noncoding RNA in early neuronal differentiation (Goff et al., 2015). Likewise, gene ontology (GO) enrichment analysis demonstrated that most of these spatially restricted gene programs correspond to features related to organ- or region-specific biological processes (Figures S4A and S4B).

We also examined whether Stereo-seq can be used to dissect the gene regulatory networks controlling cell identity and cell

Figure 3. Spatiotemporal transcriptomic atlas of mouse organogenesis

(A) Top: unsupervised SCC of mouse embryo sections across E9.5–E16.5. Embryo sections including E9.5 E1S1, E10.5 E1S1, E11.5 E1S1, E12.5 E1S1, E13.5 E1S1, E14.5 E1S1, E15.1 E1S1, and E16.5 E1S1 are shown. Bins are colored by their annotation. The squares indicate the regions for analysis in (B). Bottom left: overview of the sampled embryonic time points and sections shown in (A) and Figure S2C. Bottom right: overview of the sampled sections at E16.5 (E2S1–E2S13). The distance of each section from the midline is indicated. Scale bars, 1 mm.

(B) Reclustering of the squared areas in (A) identified detailed anatomic regions for the spinal cord (E13.5) and kidney (E16.5). The expression of marker genes for selected anatomic regions is shown on the right side of each panel. MZ, marginal zone of spinal cord; SpAL, lateral part of spinal alar plate; SpAV, ventral part of spinal alar plate; SpB, spinal cord basal plate; SpBs, superficial stratum of spinal basal plate; VZ, ventricular zone of spinal cord. Scale bars, 500 μm (left) and 250 μm (right).

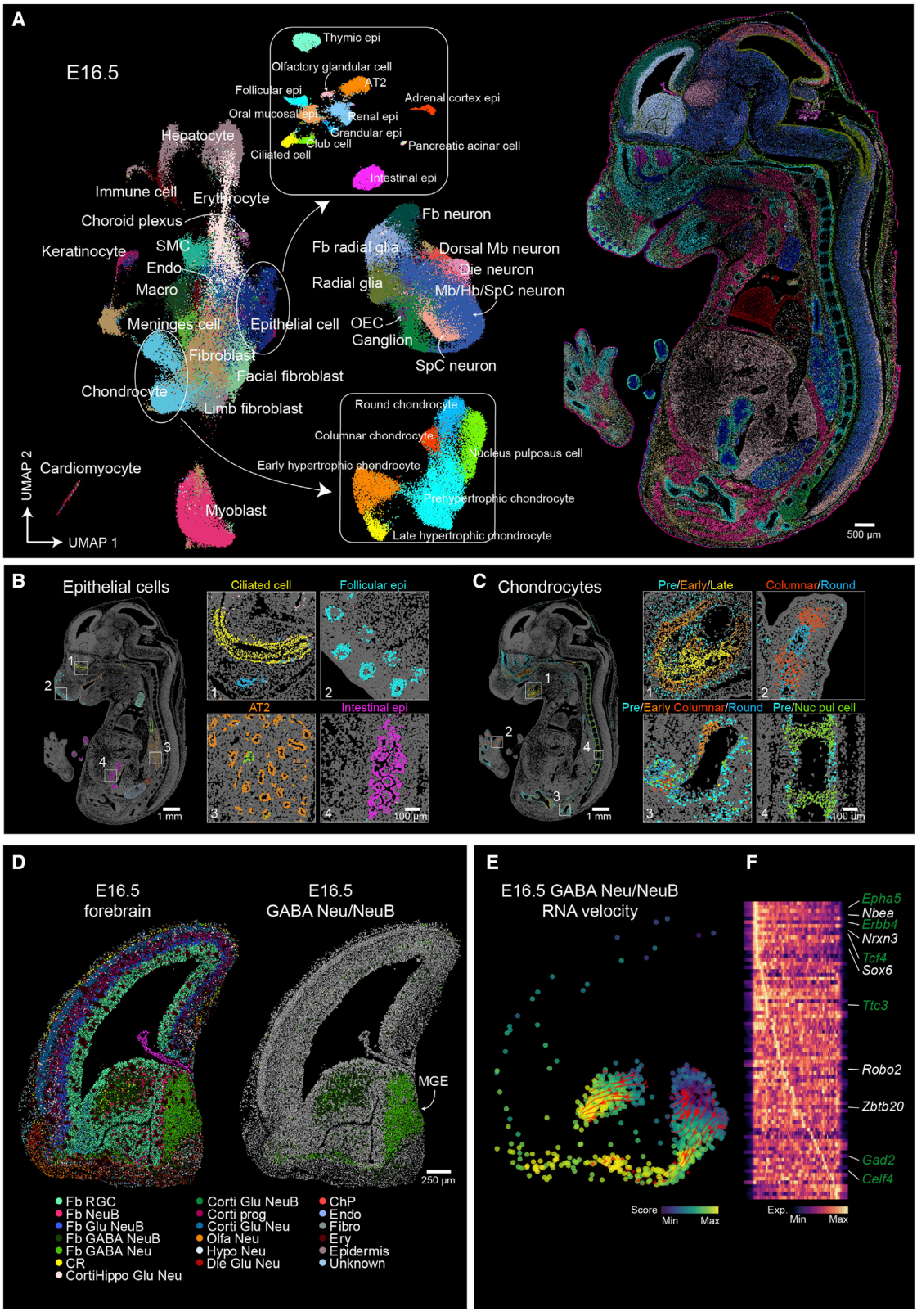
(C) Unsupervised clustering of mouse embryonic brain from E9.5 and E12.5–E16.5 identified anatomic regions of the developing brain. Brain regions were retrieved from E9.5 E1S1, E12.5 E1S2, E13.5 E1S2, E14.5 E1S1, E15.5 E1S1, and E16.5 E1S1. MZ, mantle zone; VZ, ventricular zone; Pall, pallium; SPall, subpallium; Die, diencephalon, OB, olfactory bulb; Hy, hypothalamus; Mb, dorsal midbrain; SpC, spinal cord; Hb, hindbrain; Cere, cerebellum; ChP, choroid plexus; Noto, notochord; FMN, facial motor nucleus; and Mes, mesenchyme. Scale bars, 1 mm.

(D) Heatmap showing the regulons with significant spatial autocorrelation grouped into different modules based on pairwise spatial correlations of an E16.5 embryo section (E1S1). Selected regulons and their corresponding GO terms related to representative regulon modules are highlighted on the right side.

(E) Spatial visualization of representative regulons from the indicated modules in (D). Scale bars, 1 mm.

(F) Spatial visualization of the regulons GATA4 (left) and TBX5 (right) from the heart regions across sections from E16.5 shown in (A) (E2S3–E2S11). Scale bars, 500 μm.

See also Figures S3 and S4.



(legend on next page)

state in the developing embryos. These networks are often studied in the form of modules of coexpressed genes that are coregulated by common sets of transcription factors or regulons. These data can be extracted from scRNA-seq datasets using algorithms such as single-cell regulatory network inference and clustering (SCENIC) (Aibar et al., 2017), but the lack of spatial information is likely to confound important associations. We applied SCENIC to the binned Stereo-seq maps of each of the individual time points in MOSTA to identify coexpressed genes and calculate the frequencies of predicted transcription factor-binding events. Then, we used again Hotspot (DeTomaso and Yosef, 2021) to find more precisely the spatial domains controlled by these regulons and to define modules of interconnected regulons. This identified multiple regulon modules associated with regional specificity at each stage. For example, at E9.5, we identified 31 modules containing a total of 498 regulons, and at E16.5, we identified 35 modules containing a total of 469 regulons that correspond to different tissues (Figures 3D, S4F, and S4G). Among the E16.5 regulons, ARNT2 was enriched in the brain, FOXC2 in the meninges, GATA4 in the heart, SOX9 in the bone, and MEF2C in the muscle (Figure 3E). Some regulons such as GRHL2, which belongs to a family of master epithelial transcription factors (Hinze et al., 2018), were highly enriched in multiple regions including the hair follicle, aorta-gonad-mesonephros, submandibular gland, lung, primitive gut tube, kidney, and skin. In the E16.5 embryo, we could define the localization of regulons across serial sections of the same organ such as GATA4 and TBX5 in the ventriculus and atrium of the developing heart, respectively (Figure 3F).

MOSTA generated by Stereo-seq constitutes a vast and unbiased topographic transcriptomic resource to investigate the molecular basis of tissue patterning during mouse organogenesis and can be searched using our interactive data portal at https://db.cngb.org/stomics/mosta/.

Spatial heterogeneity of cell types in mouse embryonic tissues

To understand the developmental dynamics of cell fate decisions leading to tissue patterning, it is important to topographically dissect the embryo transcriptome at single-cell level. We performed image-based cell segmentation for a section of the E16.5 embryo to showcase the strength of Stereo-seq in presenting whole embryo transcriptomic analysis at single-cell resolution. We selected this time point because of the near complete level of patterning in embryonic tissues at this stage. After filtering out the low-capture cells, we obtained 281,377 segmented cells with an average of 1107 UMI and 529 genes per cell (Figures 4A and S5A). Unsupervised clustering of the gene-by-cell matrix and uniform manifold approximation and projection (UMAP) visualization revealed 25 major cell types based on known markers (Figures 4A and S5B and MOSTA website). These cell types were either located in specific anatomic regions (e.g., cardiomyocytes, hepatocytes, and different types of neurons) or represented common cell types spread across multiple sites (e.g., chondrocytes, epithelial cells, myoblasts, and endothelial cells). Reclustering of specific cell populations could further identify cell subtypes or cell states. For example, the epithelial cell cluster (*Epcam*$^+$, *Krt8*$^+$) could be grouped into specialized epithelial cell subtypes including hair follicle cells (*Krt5*$^+$, *Krt17*$^+$), alveolar type II cells (*Foxp2*$^+$, *Nkx2-1*$^+$), club cells (*Scgb3a2*$^+$, *S100a6*$^+$), and ciliated cells (*Dynlrb2*$^+$, *Foxj1*$^+$) from the lung, zona glomerulosa epithelial cells (*Akr1b7*$^+$, *Star*$^+$) from the adrenal cortex, glandular epithelium (*Gstm1*$^+$, *Wfdc18*$^+$) from the salivary gland, stratified squamous epithelium (*Krt5*$^+$, *Krt15*$^+$) from the oral mucosa, renal tubular cells (*Dach1*$^+$, *Calb1*$^+$), and thymic epithelial cells (*Tcrg-C1*$^+$, *Tbata*$^+$) (Figures 4B, S5D, and S5E). Similarly, the developing bone showed gradual appearance and maturation of chondrocytes at different anatomic regions. Reclustering of the chondrocyte population identified distinct cell states including round (*Sox5*$^+$, *Sox9*$^+$), columnar (*Sulf1*$^+$, *Robo1*$^+$), prehypertrophic (*Prrx1*$^+$, *Fbxl7*$^+$), early hypertrophic (*Runx2*$^+$, *Ibsp*$^+$), and late hypertrophic (*Spp1*$^+$, *Mmp9*$^+$) chondrocytes, as well as the nucleus pulposus cells (*Pax1*$^+$, *Fmod*$^+$) (Figures 4C, S5F, and S5G).

We then studied the cellular heterogeneity of the E16.5 brain. The precise regulation of cell identity in the developing brain is critical for the establishment of brain architecture but is yet poorly understood. Clustering of the segmented cells in different

Figure 4. Spatial diversification of cell types at whole embryo scale

(A) Left: UMAP of the segmented cells from an E16.5 section (E1S3). The UMAP visualization of the reclustering results of epithelial cells (upper) and chondrocytes (bottom) are showed in the side panels. Right: spatial visualization of cell types shown in the left panel for the whole E16.5 embryo section. Cells are colored by their annotation. Fb radial glia, forebrain radial glia cell; OEC, olfactory epithelial cell; Fb neuron, forebrain neuron; Dorsal Mb neuron, dorsal midbrain neuron; Die neuron, diencephalon neuron; Mb/Hb/SpC neuron, mid-/hind-brain, and spinal cord neuron; SMC, smooth muscle cell; Endo, endothelial cell; and Macro, macrophage. Scale bars, 500 μm.

(B) Left: spatial visualization of different subtypes of the epithelial cell clusters shown in (A). Scale bars, 1 mm. Right: magnification images showing cell-type localization for the regions squared in the left. Scale bars, 100 μm.

(C) Left: spatial visualization of different cell states of chondrocytes shown in (A). Scale bars, 1 mm. Right: magnification images showing cell-type localization for the regions squared in the left. Scale bars, 100 μm.

(D) Left: spatial visualization of cell types in the E16.5 telencephalon. Cells are colored by the annotations. Right: spatial visualization of GABAergic neuroblast and GABAergic neuron clusters. The MGE region is indicated. Fb RGC, forebrain radial glia cell; Fb NeuB, forebrain neuroblast; Fb Glu NeuB, forebrain glutamatergic neuroblast; Fb GABA NeuB, forebrain GABAergic neuroblast; Fb GABA Neu, forebrain GABAergic neuron; CR, Cajal-Retzius cell; Corti Glu NeuB, cortical glutamatergic neuroblast; Corti prog, cortical intermediate progenitor; Corti Glu Neu, cortical glutamatergic neuron; Olfa Neu, olfactory neuron; Hypo Neu, hypothalamus neuron; Die Glu Neu, diencephalon glutamatergic neuron; ChP, choroid plexus; Endo, endothelial cell; Fibro, fibroblast; Ery, erythrocyte. Scale bars, 250 μm.

(E) RNA velocity streamline plot visualizes the migratory trajectory of the GABAergic neurons/neuroblasts. Cells are colored by the vector field-based pseudotime score and positioned in the same coordinates as in (D).

(F) Gene expression heatmap of all PCA genes in a pseudotemporal order of tangential migration. Example genes are labeled.

See also Figures S5 and S6.

anatomic regions identified different types of progenitor cells (e.g., radial glia cells, neuroblasts, and glioblasts), neurons (e.g., GABAergic interneurons, excitatory glutamatergic neurons, and dopaminergic neurons), and non-neuronal cell types (e.g., endothelial cells and fibroblasts) based on the expression of marker genes (Figures S6A and MOSTA website). To validate our results, we aligned with scRNA-seq data using Tangram (Biancalani et al., 2021), which confirmed the reported probabilistic spatial profile for specific cell types such as choroid plexus, hindbrain glioblasts, forebrain GABAergic neuroblasts, midbrain glutamatergic neuroblasts, cortical and hippocampal glutamatergic neurons, and forebrain GABAergic neurons in the whole brain (La Manno et al., 2021) (Figure S6B).

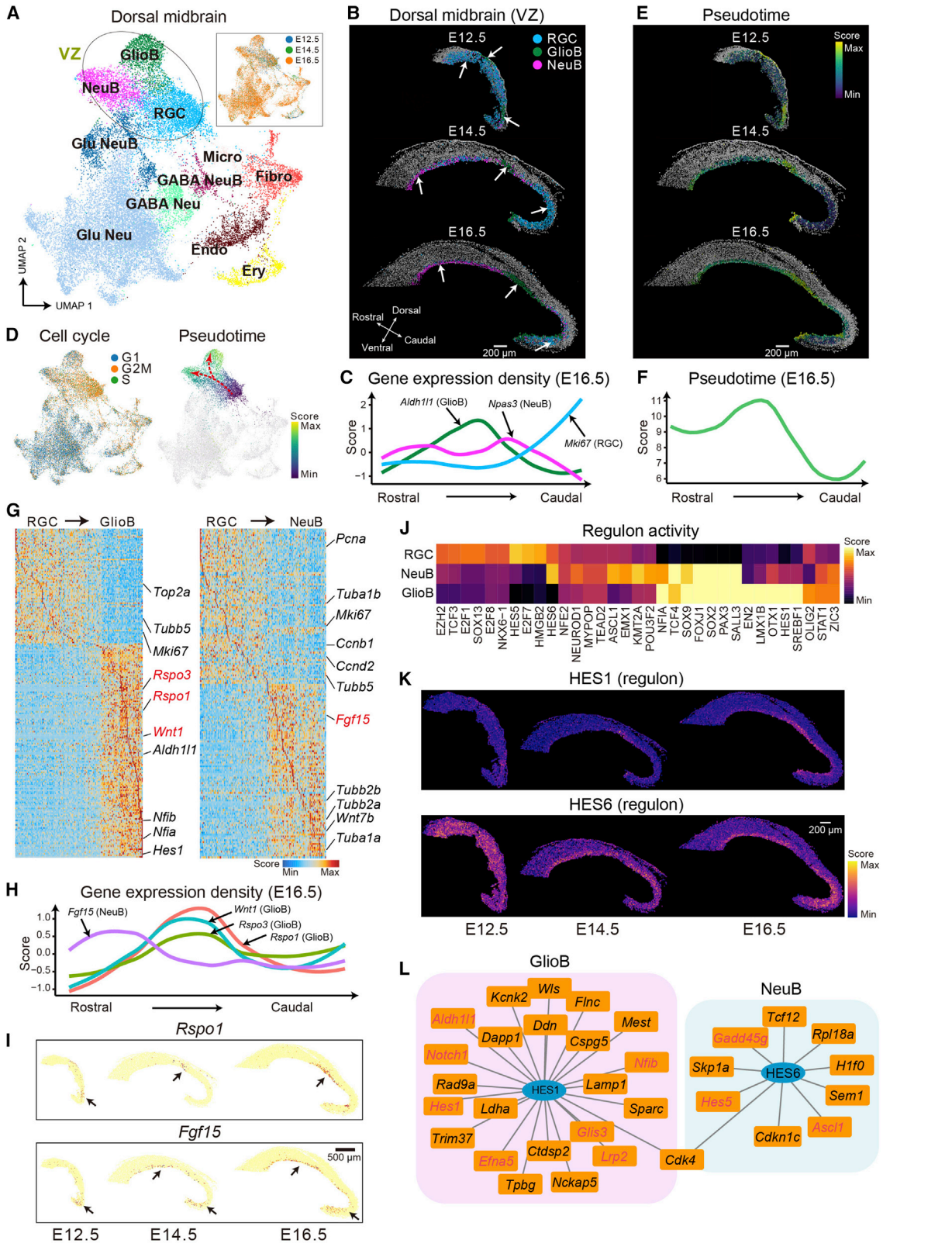
We zoomed into the telencephalon for more detailed analysis. Dorsal cortical ventricular zones give rise to glutamatergic neurons that populate the cerebral cortex and hippocampus. The ventral ventricular zones of the lateral, medial, and caudal ganglionic eminence give rise to GABAergic interneurons that populate the striatum and globus pallidus and migrate into the cerebral cortex. We also aligned the Stereo-seq data with a scRNA-seq study from the developing mouse cortex (Di Bella et al., 2021) using Tangram, confirming the reported spatial profiles of these cell types (Figure S6C). Our E16.5 Stereo-seq map clearly identified the expression of interneuron progenitor markers (*Dlx1*$^+$, *Dlx2*$^+$, *Lhx6*$^+$, *Nkx2-1*$^+$) (Lim et al., 2018) within the medial ganglionic eminence (MGE) and the migratory trajectory of interneurons into the cerebral cortex (Figures 4D and S6D). To determine the directionality of gene expression changes during interneuron migration and maturation along the rostro-caudal cortical axis, we studied spatial RNA velocity using dynamo (Qiu et al., 2022b) (see STAR Methods), an analytical framework that measures transcriptional dynamics to predict the emergence of cell fates by reconstructing a continuous vector field. In this regard, over 10% of the Stereo-seq captured transcripts contained intronic reads (Figure S6E), comparable with the proportion captured by scRNA-seq (La Manno et al., 2018; Qiu et al., 2022b). Spatial RNA velocity analysis revealed strong directional flows consistent with the tangential migration of interneurons from the MGE into the cerebral cortex (Figure 4E). To discern the regulators guiding this process, we performed vector field-based pseudotime analysis along the migration path (see STAR Methods) (Figure 4F and MOSTA website). This showed the trajectory toward mature interneurons as indicated by the shift from the early progenitor markers *Epha5*, *Tcf4*, and *Erbb4* to the late maturation marker *Gad2* (Lim et al., 2018). The expression patterns of *Erbb4* and *Gad2* are consistent with ISH images taken from ABA (Lein et al., 2007) (Figure S6F). We also observed unknown potential regulators (*Ttc3* and *Celf4*) of interneuron migration that could be interesting for further exploration (Figure 4F). In addition, we used SCENIC to analyze the distribution of telencephalon regulons corresponding to the segmented cell clusters (Figure S6G). This identified enrichment for transcription factors with specific activities associated with the maturation of projection neurons and interneurons (i.e., PAX6 in apical progenitors, MEF2C in cortical projection neurons, and DLX1 in the interneurons) (Di Bella et al., 2021; Lim et al., 2018) (Figure S6H). Reanalysis of a reported scATAC-seq dataset of the developing cortex (Di Bella et al., 2021) showed consistency between the top transcription factor hits identified with Stereo-seq and the DNA-binding motif deviation score in the corresponding cell types (Figure S6I).

MOSTA comprehensively characterizes cell-type and cell-state heterogeneity at whole embryo level, which is relevant for reconstructing cell differentiation processes involving regional and subregional specificities.

Spatial heterogeneity of progenitor cells in the developing dorsal midbrain

Despite its relatively small size, the midbrain (or mesencephalon) is associated with essential functions including vision, hearing, motor control, and temperature regulation. The dorsal midbrain contains complex multilayered structures like the cerebral cortex that form the superior and inferior colliculi (Arimura et al., 2019). However, its development has been less well characterized in terms of cellular taxonomy and organization compared with other brain regions. We thus chose the dorsal midbrain to investigate the dynamics of progenitor cell differentiation across multiple embryonic time points. We collected image-based cell segmentation of the developing dorsal midbrain regions at E12.5, E14.5, and E16.5. Reclustering of the gene-by-cell expression matrix revealed multiple cell types including terminally differentiated cells (e.g., GABAergic neurons, glutamatergic neurons, microglia, fibroblasts, and endothelial cells) and progenitors (radial glia cells, neuroblasts, and glioblasts) based on marker gene expression (Figures 5A, S7A, and S7B). Interestingly, we noticed variations in the pattern of progenitor cell appearance and maturation across the different embryonic stages (Figures 5B, 5C, S7C, and S7D). Radial glia cells (*Mki67*$^+$, *Dbx1*$^+$), which are the source of both neuroblasts (*Npas3*$^+$, *Adgrb3*$^+$) and glioblasts (*Aldh1l1*$^+$, *Efna5*$^+$), were the most abundant progenitors at E12.5, with widespread distribution. At this stage, there was only occasional appearance of sparsely distributed neuroblasts, whereas glioblasts localized mostly in the central and caudal parts. At E14.5, the radial glia cells disappeared from the rostral part and remained in the more caudal domains of the ventricular zone, revealing the spatiotemporal development of this brain region. Intriguingly, both neuroblasts and glioblasts were not evenly distributed along the ventricular zone but organized in glioblast- and neuroblast-rich patches. This indicates the programmed organization of units along the ventricular zone that give rise to different cell types that populate the neuron-rich and glia-rich regions of the superior colliculus. At E16.5, radial glia cells further decreased and concentrated in the caudal part, with neuroblasts showing a more restricted distribution and glioblasts expanding in number in the original locations. These findings suggest an asynchronous and spatially heterogeneous model of neurogenesis and gliogenesis, with glioblasts appearing early at specific anatomic locations where they remain until later differentiation. Spatial pseudotime analysis of progenitor cells in these three stages using monocle 3 (Cao et al., 2019) confirmed these results (Figures 5D–5F and S7E). Cell-cycle score analysis confirmed that radial glia cells show higher proliferation compared with neuroblasts and glioblasts.

To further probe potential molecular mechanisms underlying neurogenesis and gliogenesis in the dorsal midbrain, we applied monocle 3 to the global UMAP of E12.5, E14.5, and E16.5 cell



(legend on next page)

populations. The resulting developmental trajectories revealed two branches arising from radial glia cells, one to neuroblasts and the other to glioblasts (Figure 5D). We examined the differentially expressed genes (DEG) associated with the neuroblast and glioblast lineage bifurcation (Figure 5G). In addition, we identified highly specific expressions of extrinsic cues (ligands) in glioblasts (e.g., *Rspo1/2/3*, *Wnt1*, and *Wnt9a*) and neuroblasts (*Fgf15, Wnt7b, and Tnc*) (Figures 5G–5I and MOSTA website), providing potential mechanistic insight for the further specification of these progenitors. These results are in line with ISH images of the *Rspo1* and *Fgf15* taken from ABA (Lein et al., 2007) (Figure S7F). Likewise, regulon analysis of the E12.5–E16.5 developing dorsal midbrain identified putative transcription factors associated with neuroblast and glioblast lineage commitment (Figures 5J and 5K). For example, HES1 was highly active in glioblasts, and its target genes included known genes specific to the glia lineage such as *Aldh1l1* and *Nfib*. HES6 was highly active in neuroblasts, and its target genes included *Ascl1* and *Hes5* (Figures 5L and S7G). The expression patterns of *Hes1* and *Hes5* are consistent with ISH images taken from ABA (Figure S7H) (Lein et al., 2007).

We, thus, provide a comprehensive spatialy resolved transcriptomic map of progenitor cell specification in the dorsal midbrain.

Deciphering developmental disease susceptibility using MOSTA

Developmental disorders have classically been regarded as those producing gross neonatal manifestations. However, many genetic conditions that present later have a developmental, often unnoticed, component because their target genes are also expressed during organogenesis (Boycott and Ardigo, 2018; Barnat et al., 2020). Nowadays, scRNA-seq datasets are becoming widely used to predict the origin and mechanisms of genetic diseases including developmental disorders (Watanabe et al., 2019). However, the cell capture bias and lack of contextual microenvironment limits the power of this approach. We envisaged that MOSTA could be particularly useful in dissecting the developmental origins of mammalian genetic diseases. To demonstrate this, we took the 2,429 disease list (including 2,149 genes) from the developmental disorders genotype-to-phenotype database (DDG2P) (Wright et al., 2015), filtered them to select the top 1,959 genes based on the expression threshold in our dataset (Data S1), and then projected these genes onto our spatial transcriptomic maps for all developmental time points. This showed enrichment in different tissues and organs including brain, spinal cord, heart, liver, lung, connective tissue, bone, and muscle (Figure 6A). Some genes were expressed in most or all embryonic stages, whereas others had a restricted window. Similarly, although many genes were mainly enriched in one tissue (in some cases, they were confined to a specific region) or system, others were more widely distributed. As informative examples, *NKX2-1*, which is mutated in brain-lung-thyroid syndrome (Shetty et al., 2014), was mainly enriched in brain (hippocampus and subpallium) and lung and *Lmod3*, which is associated with nemaline myopathy (Berkenstadt et al., 2018), was enriched in the skeletal muscle (Figure 6B).

To illustrate that MOSTA can help identify potentially targeted cells in the affected tissues, we focused on Robinow syndrome caused by mutations in *WNT5A*, a cytokine regulating the Wnt signaling pathway (Wright et al., 2015). We observed that *Wnt5a* localized in the craniofacial region, limb, genital ridge, and brain (floor plate and neocortex) from multiple developmental stages (Figure 6B), consistent with the disease phenotype (Menezes et al., 2010). We performed cell resolution reclustering of the gene-by-cell expression matrices retrieved from the maxilla and the limb of the E16.5 embryo. This identified different cell types including mesenchymal cells, fibroblasts, endothelial cells, keratinocytes, tenocytes, myoblasts, chondrocytes, and others. We observed high *Wnt5a* expression in mesenchymal cells and fibroblasts from the maxilla, whereas in the limb, it was located mainly in mesenchymal cells (Figures 6C and 6D). This suggests that the disease mechanisms may differ in the two tissues. We then investigated *Msx1*, a homeobox gene encoding a transcription factor with known crosstalk with WNT5A (Lin et al., 2011). *Msx1* was expressed at different levels in multiple cell types of the maxilla including mesenchymal cells, but in the limb, it was rather restricted to the latter cell type only. Similarly, the MSX1 regulon was more enriched in mesenchymal cells and fibroblasts in the maxilla, and mesenchymal cells in the limb (Figures 6C and 6D). These observations are consistent with the facial and digit abnormalities present in *Msx1* knockout mice

Figure 5. Spatial heterogeneity of progenitor cell types in the developing dorsal midbrain

(A) UMAP of the segmented cells from the dorsal midbrain regions of the E12.5 (E1S3), E14.5 (E1S3), and E16.5 (E1S3, E2S6, and E2S7) stages. Cells are colored by the annotations and the stages (upper right). Endo, endothelial cell; Ery, erythrocyte; Fibro, fibroblast; GABA NeuB, GABAergic neuroblast; GABA Neu, GABAergic neuron; GlioB, glioblast; Glu NeuB, glutamatergic neuroblast; Glu Neu, glutamatergic neuron; Micro, microglia; NeuB, neuroblast; and RGC, radial glia cell.

(B) Spatial visualization of the progenitor cells identified in (A). Cells are colored by the annotations. Scale bars, 200 μm.

(C) Line chart showing the indicated gene expression gradient along the rostro-caudal axis.

(D) UMAP of the cell-cycle phases and pseudotime score of the progenitor cells identified in (A). The arrow indicates the differentiation path.

(E) Spatial visualization of the pseudotime score of the progenitor cells identified in (A). Cells are colored by the pesudotime score. Scale bars, 200 μm.

(F) Line chart showing the pseudotime score along the rostro-caudal axis as indicated in (E).

(G) Heatmaps illustrating the genes linked to cell fate divergence at the branch point of radial glia cells to glioblasts (left) and neuroblasts (right).

(H) Line chart showing the indicated gene expression gradient along the rostro-caudal axis.

(I) Spatial visualization of *Rspo1* and *Fgf15* expressions at the indicated time points of dorsal midbrain development. Scale bars, 500 μm.

(J) Heatmap showing the normalized activity score of the indicated regulons for the three progenitor cell types.

(K) Spatial visualization of the activity of the HES1 and HES6 regulons at the indicated time points of dorsal midbrain development. Scale bars, 200 μm.

(L) Gene regulatory networks of HES1 (glioblast) and HES6 (neuroblast) regulons in the developing dorsal midbrain as visualized by Cytoscape. Selected target genes were shown.

See also Figure S7.

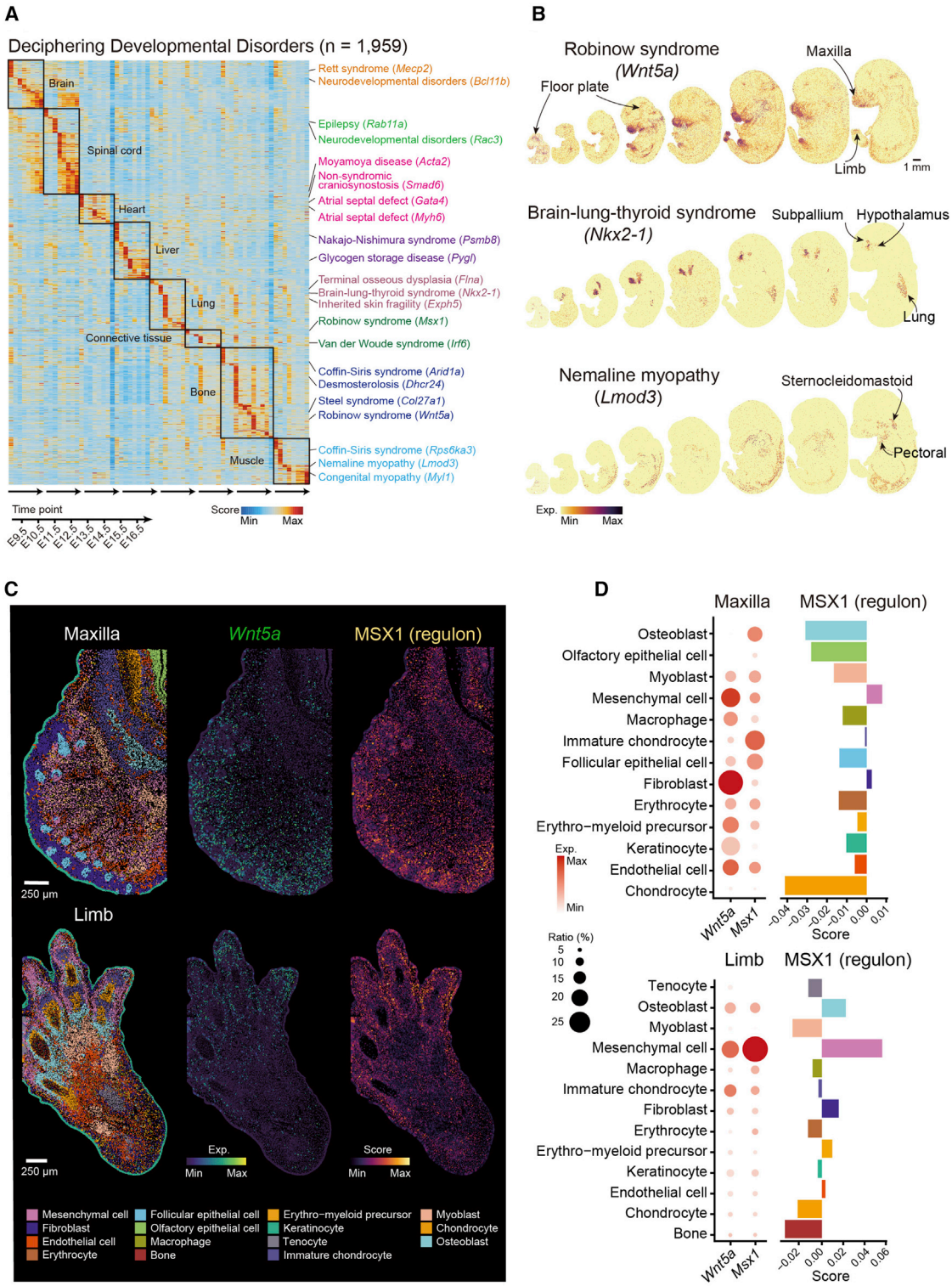


Figure 6. Association of MOSTA profiles with human developmental disorders

(A) Heatmap showing the normalized expression level of 1,959 genes selected from the developmental disorders genotype-to-phenotype database (DDG2P) in the representative anatomic regions (brain, spinal cord, heart, liver, lung, connective tissue, bone, and muscle) in embryo sections from E9.5 to E16.5. Embryo sections including E9.5 E1S1, E9.5 E2S1, E10.5 E1S1, E10.5 E2S1, E11.5 E1S1, E12.5 E1S1, E12.5 E2S1, E13.5 E1S1, E14.5 E1S1, E15.1 E1S1, and E16.5 E1S1 were used in this analysis.

(legend continued on next page)

(Jumlongras et al., 2001) and provide potential insight into the mechanisms underlying the skeletal phenotypic features of Robinow syndrome.

These results demonstrate the utility of MOSTA for defining the spatiotemporal windows of disease-related gene expression during development and for establishing potential vulnerabilities.

DISCUSSION

Topographic transcriptomic information is fundamental for dissection of the molecular events driving tissue patterning during development. Stereo-seq is a DNB-based spatially resolved transcriptomic technology with genome-wide coverage, cellular resolution, high sensitivity, and large field of view. These parameters are all critical for accurately profiling the transcriptomic heterogeneity of relatively large mid- and late-stage mammalian embryos. Moreover, Stereo-seq is an affordable technology. The current DNB capture chip costs ~220 RMB/mm^2 ($35/mm^2) or ~22,000 RMB/cm^2 ($3,500/cm^2). The high-density RNA capture with DNB chip requires a sequencing depth of around 150 gigabases/cm^2, which costs ~5,000 RMB (~$800) based on current short-read sequencing platforms.

We have used Stereo-seq to create MOSTA, an expandable panoramic and high-definition transcriptomic resource for understanding mouse organogenesis. MOSTA provides a detailed topographic information about the stepwise emergence of tissue-specific cell identities across the different stages of mouse organogenesis. Examples of the latter include the identification of neuroblast-rich and glioblast-rich domains within the ventricular zone of the developing dorsal midbrain, which will give rise to the specific patterning consisting of many cell subtypes.

To illustrate the utility of MOSTA for inferring the regulatory mechanisms behind developmental processes, we have systematically dissected the spatially resolved gene regulons driving organogenesis. This identified tissue- and area-specific transcription factors and their associated targets across different embryonic stages. Although the function of many of the transcription factors controlling these regulons has been previously reported, we have identified the spatial domains in which they act. Importantly, Stereo-seq also captures intronic transcripts, making it possible to calculate spatial RNA velocity with single-cell resolution, which as in scRNA-seq studies (La Manno et al., 2018; Qiu et al., 2022b), can facilitate the study of cell fate transitions. Using spatial RNA velocity, we have generated transcriptomic maps of tangentially migratory interneurons in the developing cerebral cortex. In addition, we have projected the expression of developmental disease loci (including monogenic diseases involving transcription factors or morphogens) onto our spatial maps, highlighting a potentially relevant functional connection between *Wnt5* and *Msx1* in different cell types but especially in mesenchymal cells in Robinow syndrome. Mice and humans differ in multiple aspects of normal physiology and disease; hence, there are caveats, but this approach clearly has the potential to uncover disease mechanisms. Systematic independent explorations using our interactive portal will provide users with insights to unlock the black boxes of mouse embryogenesis and developmental disorders. In this regard, expansion with datasets from genetically engineered mice that bear lineage tracing cassettes, mimic developmental disorders, or cause alteration of developmental pathways will be useful to test and refine any models of normal and abnormal development generated with MOSTA.

Larger Stereo-seq chips will enable the profiling of embryos from other mammalian species including primates. The integrated knowledge from multiple species could be instrumental in discerning the determinants of spatiotemporal differences in tissue patterning. Forthcoming optimizations of Stereo-seq will enhance gene capture, further facilitating the assignment of individual cell identities and states in these panoramic maps. Similarly, oligonucleotide-labeled antibodies (Stoeckius et al., 2017) could be incorporated into the DNB for simultaneous protein detection of surface markers, which would be useful for easier identification of specific small cell types (*e.g.*, infiltrating immune cells). Besides the study of development, Stereo-seq and its future refinements have the ability to transform multiple other research fields. Due to its characteristics and applicability, Stereo-seq also has the potential to move into routine clinical practice as an extraordinary diagnostic tool complementary to medical imaging and histopathology data.

Limitations of the study

The high resolution of Stereo-seq allows efficient image-based cell segmentation, but in some cases, where multiple cell types are close to each other, particularly smaller cell types like immune cells, cell segmentation may be imperfect. Similarly, Stereo-seq has a genome-wide coverage, but due to limitations in capture, relevant genes with low expression may be missed. This is mitigated because the large field-of-view facilitates the profiling of enough cell numbers to power the statistical analysis. Importantly, we have profiled a total of 53 sagittal sections corresponding to an overall area of 23.4 cm^2, but more sections may still be necessary to provide adequate coverage of all regional tissue types. The incomplete sampling could also affect the inference of relationship of anatomic regions and developmental trajectories. Future expansions of MOSTA will help address this caveat, and the addition of enough sections will facilitate 3D transcriptomic rendering that is not possible with the current dataset. Meanwhile, despite the remarkable utility for unveiling the molecular logic of embryogenesis through analysis of genes, gene modules, or regulons, special attention

(B) Spatial visualization of the expression of the indicated genes related to the selected human developmental disorders (Robinow syndrome, brain-lung-thyroid syndrome, and nemaline myopathy) in embryos from E9.5 to E16.5. Disease-related anatomic regions are annotated in the figures. Representative sections including E9.5 E1S1, E10.5 E1S1, E11.5 E1S1, E12.5 E1S1, E13.5 E1S1, E14.5 E1S1, E15.1 E1S1, and E16.5 E1S1 were shown. Scale bars, 1 mm.
(C) Spatial visualization of cell types (left), *Wnt5a* expression (middle), and MSX1 regulon activity (right) at the maxilla (upper) and limb (bottom) from E16.5 (E1S3). Scale bars, 250 μm.
(D) Left: bubble plot showing the normalized expression of *Wnt5a* and *Msx1* in the indicated cell types in maxilla (upper) and limb (bottom). Right: barplot showing the activity of the MSX1 regulon in the indicated cell types of the maxilla (upper) and limb (bottom) from E16.5 (E1S3).

should be paid to validations including functional assays. In the case of regulons, integration with scATAC-seq datasets can help address this issue, but we anticipate that future modifications of Stereo-seq will allow direct spatial measurements of chromatin accessibility. Finally, although we have developed algorithms including the SCC for the analyses in this study, others may be needed for a broader range of applications in various tissues from different species.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal
- METHOD DETAILS
  - Stereo-seq chip preparation
  - Stereo-seq library preparation and sequencing
  - In situ hybridization
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Stereo-seq raw data processing
  - Image-based single cell segmentation and diffusion analysis
  - Stereo-seq benchmarking analysis and clustering
  - Spatially resolved gene regulatory networks
  - Identification of spatially auto-correlated gene or regulon modules
  - Cell type mapping
  - Alignment of single-cell transcriptome with Stereo-seq by Tangram
  - Reanalysis of scATAC-seq data
  - Inference of developmental trajectories
  - Spatial gene enrichment analysis of genes associated with human developmental disorders

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.04.003.

ACKNOWLEDGMENTS

This work is part of the "SpatioTemporal Omics Consortium" (STOC) paper package. A list of STOC members is available at: http://sto-consortium.org. We would like to thank the MOTIC China Group, Rongqin Ke (Huaqiao University, Xiamen, China), Jiazuan Ni (Shenzhen University, Shenzhen, China), Wei Huang (Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China), and Jonathan S. Weissman (Whitehead Institute, Boston, USA) for their help. This work was supported by the grant of Top Ten Foundamental Research Institutes of Shenzhen, the Shenzhen Key Laboratory of Single-Cell Omics (ZDSYS20190902093613831), and the Guangdong Provincial Key Laboratory of Genome Read and Write (2017B030301011); Longqi Liu was supported by the National Natural Science Foundation of China (31900466) and Miguel A. Esteban's laboratory at the Guangzhou Institutes of Biomedicine and Health by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16030502), National Natural Science Foundation of China (92068106), and the Guangdong Basic and Applied Basic Research Foundation (2021B1515120075).

AUTHOR CONTRIBUTIONS

J.W., X.X., A.C., L.L., and Y. Li conceived the idea; J.W., X.X., A.C., L.L., Y. Li, and M.A.E. supervised the work; A.C., S. Liao, M.C., and L.L. designed the experiment; S. Liao and J.Y. generated the Stereo-seq chip with the help from X.C. and X.H.; M.C., J.X., Y.J., H.Z., and P.G. performed the majority of the experiments; K.M., L. Wu, X.L., Huifang Lu, H.X., B.C., J.H., and M.L. performed data preprocessing and quality evaluation; L. Wu, Y. Lai, X.Q., S.H., X. Wang, Xiaoyu Wei, W.F., Y. Liu, Y. Yang, and Xiaofeng Wei analyzed the data; Z.L., Y.H., J.P., Shuai Liu, M.S., C.L., Q.L., Y. Yuan, Zhifeng Wang, Zhaohui Wang, B.Q., G.L., F.Z., X.D., S.W., H.K., L. Wang, J.-F.F., O.W., Haorong Lu, and B.W. provided technical support; L.H., P.M.-C., P.H.M., J.P.T., Q.-F.W., Shiping Liu, Y.G., M.N., W.Z., F.M., Y. Yin, H.Y., M.L., R.J.C., J.M., and M.U. gave the relevant advice; A.C., M.C., L. Wu, Y. Lai, X.Q., M.A.E., L.L., X.X., and J.W. wrote the manuscript.

DECLARATION OF INTERESTS

The chip, procedure, and applications of Stereo-seq are covered in pending patents. J.W. and H.Y. are founders of BGI-Shenzhen, and employees of BGI have stock holdings in BGI.

Received: August 6, 2021
Revised: January 24, 2022
Accepted: April 1, 2022
Published: May 4, 2022

REFERENCES

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086.

Arimura, N., Dewa, K.I., Okada, M., Yanagawa, Y., Taya, S.I., and Hoshino, M. (2019). Comprehensive and cell-type-based characterization of the dorsal midbrain during development. Genes Cells *24*, 41–59.

Barnat, M., Capizzi, M., Aparicio, E., Boluda, S., Wennagel, D., Kacher, R., Kassem, R., Lenoir, S., Agasse, F., Braz, B.Y., et al. (2020). Huntington's disease alters human neurodevelopment. Science *369*, 787–793.

Berkenstadt, M., Pode-Shakked, B., Barel, O., Barash, H., Achiron, R., Gilboa, Y., Kidron, D., and Raas-Rothschild, A. (2018). LMOD3-associated nemaline myopathy: prenatal ultrasonographic, pathologic, and molecular findings. J. Ultrasound Med. *37*, 1827–1833.

Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, Å., Zhang, M., et al. (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat. Methods *18*, 1352–1362.

Boycott, K.M., and Ardigó, D. (2018). Addressing challenges in the diagnosis and treatment of rare genetic diseases. Nat. Rev. Drug Discov. *17*, 151–152.

Broudy, V.C. (1997). Stem cell factor and hematopoiesis. Blood *90*, 1345–1364.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

Cho, C.S., Xi, J., Si, Y., Park, S.R., Hsu, J.E., Kim, M., Jun, G., Kang, H.M., and Lee, J.H. (2021). Microscopic examination of spatial transcriptome using Seq-Scope. Cell *184*, 3559–3572.e22.

Codeluppi, S., Borm, L.E., Zeisel, A., La Manno, G., van Lunteren, J.A., Svensson, C.I., and Linnarsson, S. (2018). Spatial organization of the somatosensory cortex revealed by osmFISH. Nat. Methods *15*, 932–935.

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

Di Bella, D.J., Habibi, E., Stickels, R.R., Scalia, G., Brown, J., Yadollahpour, P., Yang, S.M., Abbate, C., Biancalani, T., Macosko, E.Z., et al. (2021). Molecular logic of cellular diversification in the mouse cerebral cortex. Nature *595*, 554–559.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dries, R., Zhu, Q., Dong, R., Eng, C.L., Li, H., Liu, K., Fu, Y., Zhao, T., Sarkar, A., Bao, F., et al. (2021). Giotto: a toolbox for integrative analysis and visualization of spatial expression data. Genome Biol. *22*, 78.

Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G., et al. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science *327*, 78–81.

Goff, L.A., Groff, A.F., Sauvageau, M., Trayes-Gibson, Z., Sanchez-Gomez, D.B., Morse, M., Martin, R.D., Elcavage, L.E., Liapis, S.C., Gonzalez-Celeiro, M., et al. (2015). Spatiotemporal expression and transcriptional perturbations by long noncoding RNAs in the mouse brain. Proc. Natl. Acad. Sci. USA *112*, 6855–6862.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29.

He, P., Williams, B.A., Trout, D., Marinov, G.K., Amrhein, H., Berghella, L., Goh, S.T., Plajzer-Frick, I., Afzal, V., Pennacchio, L.A., et al. (2020). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. Nature *583*, 760–767.

Hinze, C., Ruffert, J., Walentin, K., Himmerkus, N., Nikpey, E., Tenstad, O., Wiig, H., Mutig, K., Yurtdas, Z.Y., Klein, J.D., et al. (2018). GRHL2 is required for collecting duct epithelial barrier function and renal osmoregulation. J. Am. Soc. Nephrol. *29*, 857–868.

Hu, J., Li, X., Coleman, K., Schroeder, A., Ma, N., Irwin, D.J., Lee, E.B., Shinohara, R.T., and Li, M. (2021). SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. Nat. Methods *18*, 1342–1351.

Jumlongras, D., Bei, M., Stimson, J.M., Wang, W.F., DePalma, S.R., Seidman, C.E., Felbor, U., Maas, R., Seidman, J.G., and Olsen, B.R. (2001). A nonsense mutation in MSX1 causes Witkop syndrome. Am. J. Hum. Genet. *69*, 67–74.

La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A.R., Dratva, L.M., et al. (2021). Molecular architecture of the developing mouse brain. Nature *596*, 92–96.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

Lebrigand, K., Bergenstråhle, J., Thrane, K., Mollbrink, A., Barbry, P., Waldmann, R., and Lundeberg, J. (2020). The spatial landscape of gene expression isoforms in tissue sections. bioRixv. https://doi.org/10.1101/2020.08.24.252296.

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature *445*, 168–176.

Lim, L., Mi, D., Llorca, A., and Marín, O. (2018). Development and functional diversification of cortical interneurons. Neuron *100*, 294–313.

Lin, M., Li, L., Liu, C., Liu, H., He, F., Yan, F., Zhang, Y., and Chen, Y. (2011). Wnt5a regulates growth, patterning, and odontoblast differentiation of developing mouse tooth. Dev. Dyn. *240*, 432–440.

Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C.C., Tebaldi, T., Zhang, D., Kim, D., Bai, Z., et al. (2020). High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. Cell *183*, 1665–1681.e18.

Lohoff, T., Ghazanfar, S., Missarova, A., Koulena, N., Pierson, N., Griffiths, J.A., Bardot, E.S., Eng, C.L., Tyser, R.C.V., Argelaguet, R., et al. (2022). Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. Nat. Biotechnol. *40*, 74–85.

Menezes, R., Letra, A., Kim, A.H., Küchler, E.C., Day, A., Tannure, P.N., Gomes da Motta, L., Paiva, K.B., Granjeiro, J.M., and Vieira, A.R. (2010). Studies with Wnt genes and nonsyndromic cleft lip and palate. Birth Defects Res. A Clin. Mol. Teratol. *88*, 995–1000.

Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., and Virshup, I.J.b. (2021). Squidpy: a scalable framework for spatial single cell analysis. bioRixv. https://doi.org/10.1101/2021.02.19.431994.

Peng, G., Suo, S., Cui, G., Yu, F., Wang, R., Chen, J., Chen, S., Liu, Z., Chen, G., Qian, Y., et al. (2019). Molecular architecture of lineage allocation and tissue organization in early mouse embryo. Nature *572*, 528–532.

Pijuan-Sala, B., Wilson, N.K., Xia, J., Hou, X., Hannah, R.L., Kinston, S., Calero-Nieto, F.J., Poirion, O., Preissl, S., Liu, F., and Göttgens, B. (2020). Single-cell chromatin accessibility maps reveal regulatory programs driving early mouse organogenesis. Nat. Cell Biol. *22*, 487–497.

Qiu, C., Cao, J., Martin, B.K., Li, T., Welsh, I.C., Srivatsan, S., Huang, X., Calderon, D., Noble, W.S., Disteche, C.M., et al. (2022). Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nature Genetics *54*, 328–341.

Qiu, X., Zhang, Y., Martin-Rufino, J.D., Weng, C., Hosseinzadeh, S., Yang, D., Pogson, A.N., Hein, M.Y., Min, J., Wang, L., et al. (2022). Mapping transcriptomic vector fields of single cells. Cell *185*, 690–711, e45.

Rao, A., Barkley, D., França, G.S., and Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. Nature *596*, 211–220.

Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat. Methods *14*, 975–978.

Shetty, V.B., Kiraly-Borri, C., Lamont, P., Bikker, H., and Choong, C.S. (2014). NKX2-1 mutations in brain-lung-thyroid syndrome: a case series of four patients. J. Pediatr. Endocrinol. Metab. *27*, 373–378.

Srivatsan, S.R., Regier, M.C., Barkan, E., Franks, J.M., Packer, J.S., Grosjean, P., Duran, M., Saxton, S., Ladd, J.J., Spielmann, M., et al. (2021). Embryo-scale, single-cell spatial transcriptomics. Science *373*, 111–117.

Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science *353*, 78–82.

Stickels, R.R., Murray, E., Kumar, P., Li, J., Marshall, J.L., Di Bella, D.J., Arlotta, P., Macosko, E.Z., and Chen, F. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat. Biotechnol. *39*, 313–319.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868.

Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat. Methods *18*, 1333–1341.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233.

Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., and Yu, T.J.P. (2014). scikit-image: image processing in Python. PeerJ *2*, e453.

Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., Äijö, T., Bonneau, R., Bergenstråhle, L., Navarro, J.F., et al. (2019). High-definition spatial transcriptomics for in situ tissue profiling. Nat. Methods *16*, 987–990.

Watanabe, K., Umićević Mirkov, M., de Leeuw, C.A., van den Heuvel, M.P., and Posthuma, D. (2019). Genetic mapping of cell type specificity for complex traits. Nat. Commun. *10*, 3222.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

Wright, C.F., Fitzgerald, T.W., Jones, W.D., Clayton, S., McRae, J.F., van Kogelenberg, M., King, D.A., Ambridge, K., Barrett, D.M., Bayzetinova, T., et al. (2015). Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet *385*, 1305–1314.

Yao, Z., Liu, H., Xie, F., Fischer, S., Booeshaghi, A.S., Adkins, R.S., Aldridge, A.I., Ament, S.A., Pinto-Duarte, A., Bartlett, A., et al. (2020). An integrated transcriptomic and epigenomic atlas of mouse primary motor cortex cell types. bioRxiv. https://doi.org/10.1101/2020.02.29.970558.

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular architecture of the mouse nervous system. Cell *174*, 999–1014.e22.

Zhao, E., Stone, M.R., Ren, X., Guenthoer, J., Smythe, K.S., Pulliam, T., Williams, S.R., Uytingco, C.R., Taylor, S.E.B., Nghiem, P., et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. Nat. Biotechnol. *39*, 1375–1384.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-DIG-AP fab antibody	Roche	Cat#11093274910; RRID:AB_2734716
Biological samples		
Adult mouse olfactory bulb	This study	N/A
Adult mouse brain (12 weeks)	This study	N/A
E9.5 mouse embryos	This study	N/A
E10.5 mouse embryos	This study	N/A
E11.5 mouse embryos	This study	N/A
E12.5 mouse embryos	This study	N/A
E13.5 mouse embryos	This study	N/A
E14.5 mouse embryos	This study	N/A
E15.5 mouse embryos	This study	N/A
E16.5 mouse embryos	This study	N/A
Chemicals, peptides, and recombinant proteins		
Tissue-Tek OCT	Sakura	Cat#4583
T4 ligase	NEB	Cat#M0202V
AMPure XP Beads	Vazyme	Cat#N411-03
Nucleic Acid Dye	Thermo	Cat#Q10212
20 × SSC	Thermo	AM9770
Pepsin	Sigma	P7000
RNase inhibitor	NEB	M0314L
SuperScript II	Invitrogen	Cat#18064-014
Exonuclease I	NEB	M0293L
KAPA HiFi Hotstart Ready Mix	Roche	KK2602
Qubit™ dsDNA Assay Kit	Thermo	Q32854
TRIzol™ Reagent	INVITROGEN	Cat#15596-026
Paraformaldehyde	SIGMA	Cat#158127
Formamide	SIGMA	Cat#F9037
EDTA	INVITROGEN	Cat#15575-038
5 M NaCl	AMBION	Cat#AM9759
RNase A	Sigma	Cat#R4642
blocking buffer	Roche	Cat#11096176001
BM purple	Roche	Cat#11442074001
Tris-HCL pH-7.5	VWR	Cat#100216-512
$MgCl_2$	Ambion	AM9530G
Tween 20	Sigma	P9416
Deposited data		
HDST data of mouse olfactory bulb	GEO	GEO: GSE130682
Seq-Scope data of liver	GEO	GEO: GES169706
DBiT-Seq data of mouse embryo	GEO	GEO: GSE137986
SLIDE-seqV2 data of mouse olfactory bulb	Broad	Stickels et al., 2021
Visium data of mouse olfactory bulb	GEO	GEO: GSE153859
Single-cell data of adult mouse brain	Linnarsson lab	http://mousebrain.org/adolescent/
Single-cell data of mouse developing cortex	GEO	GEO: GSE153164

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Single-cell data of mouse developing mouse brain	Linnarsson lab	http://mousebrain.org/development/
smFISH data of adult mouse brain	Linnarsson lab	http://linnarssonlab.org/osmFISH/availability/
ABA ISH data	Allen Institute	https://mouse.brain-map.org/
Developmental Disorders Genotype-to-Phenotype database (DDG2P)	Wellcome Trust Sanger Institute	https://panelapp.genomicsengland.co.uk/panels/484/
Raw data of Stereo-seq	This study	CNGB: CNP0001543
Processed data of Stereo-seq	This study	https://db.cngb.org/stomics/mosta/
Oligonucleotides		
DNB library oligo1: TGTGAGCCAAGGAGTTGAACTGCTGACGTACTGAGAGGCATGGCGACCTTATCAGNNNNNNNNNNNNNNNNNNNNNNNNNTTGTCTTCCTAAGACCG	Sangon	N/A
DNB library oligo2: /5phos/CTTGGCCTCCGACTTAAGTCGGATCGTAGCCATGTCGTTC	Sangon	N/A
Splint oligo: TCGGAGGCCAAGCGGTCTTAGGAA	Sangon	N/A
CID sequencing primer: CTGCTGACGTACTGAGAGGCATGGCGACCTTATCAG	Sangon	N/A
DNB library F: /5phos/TGTGAGCCAAGGAGT	Sangon	N/A
DNB library R: GAACGACATGGCTA	Sangon	N/A
Capture oligo: /5phos/TTGTCTTCCTAAGACNNNNNNNNNNTTTTTTTTTTTTTTTTTTTTTTV	Sangon	N/A
Stereo-seq-TSO: CTGCTGACGTACTGAGAGGC/rG//rG//iXNA_G/	Sangon	N/A
cDNA PCR primer: CTGCTGACGTACTGAGAGGC	Sangon	N/A
Stereo-seq-library-F: /5phos/CTGCTGACGTACTGAGAGG*C*A	Sangon	N/A
Stereo-seq-library-R: GAGACGTTCTCGACTCAGCAGA	Sangon	N/A
Stereo-seq-library-splint-oligo: GTACGTCAGCAGGAGACGTTCTCG	Sangon	N/A
Stereo-seq-read1: CTGCTGACGTACTGAGAGGCATGGCGACCTTATCAG	Sangon	N/A
Stereo-seq-MDA-primer: TCTGCTGAGTCGAGAACGTC	Sangon	N/A
Stereo-seq-read2: GCCATGTCGTTCTGTGAGCCAAGGAGTT	Sangon	N/A
Shh_F: GATGAGGAAAACACGGGAGC	Sangon	N/A
Shh_R: ACCCGGTTGATGAGAATGGT	Sangon	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Kitl_F: GGGCTTCATTTGCTGTCTGT	Sangon	N/A
Kitl_R: CTTTGCGGCTTTCCCTTTCT	Sangon	N/A
Sox2_F: TGATGGAGACGGAGCTGAAG	Sangon	N/A
Sox2_R: TGGAGTGGGAGGAAGAGGTA	Sangon	N/A
Pantr1_F: GCCACGCGAGGTATTTGAAA	Sangon	N/A
Pantr1_R: GCACCATTTCATCACATCAGC	Sangon	N/A
Software and algorithms		
SAW	https://github.com/BGIResearch/SAW	V2.1.0
pySCENIC	https://github.com/aertslab/pySCENIC/	V0.11.2
Hotspot	http://www.github.com/Yoseflab/Hotspot	V0.9.1
Monocle3	https://github.com/cole-trapnell-lab/monocle3	V1.0.0
Tangram	https://github.com/broadinstitute/Tangram	V1.0.2
Squidpy	https://github.com/theislab/squidpy	V1.1.2
Scikit-image	https://scikit-image.org/	V0.18.1
Seurat	https://satijalab.org/seurat/	V4.0.5
Scanpy	https://github.com/theislab/scanpy/	V1.8.2
Dynamo	https://github.com/aristoteleo/dynamo-release	V1.0.0
OpenCV	https://github.com/skvark/opencv-python	V4.5.4
STAR	https://github.com/alexdobin/STAR	V2.7.10a
Rearrr	https://github.com/LudvigOlsen/rearrr	V0.3.0
Signac	https://satijalab.org/signac/	V1.5.0
ChromVAR	https://github.com/GreenleafLab/chromVAR	V1.16.0

RESOURCE AVAILABILITY

Lead contact

Further information and requests for the resources and reagents may be directed to the corresponding author Xun Xu (xuxun@genomics.cn)

Material availability

All materials used for Stereo-seq are commercially available.

Data and code availability

All raw data generated by Stereo-seq have been deposited to CNGB Nucleotide Sequence Archive (accession code: CNP0001543 (https://db.cngb.org/search/project/CNP0001543). All data were analyzed with standard programs and packages, as detailed above. Custom code supporting the current study is available at https://github.com/BGIResearch/SAW. Additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal

All relevant procedures involving animal experiments presented in this study are compliant with ethical regulations regarding animal research and were conducted under the approval of the Animal Care and Use committee of the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences (license number IACUC2021002). Mouse olfactory bulb and brain were dissected from 12-week-old C57BL/6J female mice. E9.5, E10.5, E11.5, E12.5, E13.5, E14.5, E15.5 and E16.5 embryos were collected from pregnant C57BL/6J female mice. After collection, tissues were snap-frozen in liquid nitrogen prechilled isopentane in Tissue-Tek OCT (Sakura, 4583) and transferred to a -80°C freezer for storage before the experiment. Cryosections were cut at a thickness of 10 μm in a Leika CM1950 cryostat. Adult mouse olfactory bulb and mouse hemibrain were cut coronally, mouse embryos were cut sagittally.

METHOD DETAILS

Stereo-seq chip preparation

Generation of Stereo-seq chips

To generate the patterned array, we synthesized two oligo sequences: one containing 25 random deoxynucleotides DNB library oligo1 and the other a fixed sequence with 5' phosphorylated DNB library oligo2. These two oligos were ligated with splint oligo1 at 37°C for 2 hours using T4 ligase (NEB; 1U/μl T4 DNA ligase and 1× T4 DNA ligation buffer). The products were purified using the AMPure XP Beads (Vazyme, N411-03) and then PCR amplified with the following steps: 95°C for 5 minutes, 12 cycles at 98°C for 20 seconds, 58°C for 20 seconds, 72°C for 20 seconds and a final incubation at 72°C for 5 minutes. The PCR products were purified using the AMPure XP Beads. DNB were then generated by rolling circle amplification and loaded onto the patterned chips according to the MGI DNBSEQ-Tx sequencer manual. Next, to determine the distinct DNB-CID sequences at each spatial location, single-end sequencing was performed using a CID sequencing primer in MGI DNBSEQ-Tx sequencer with SE25 sequencing strategy. After sequencing, the capture oligo including 22 nt poly-T and 10 nt UMI was hybridized with the DNB in 5 × SSC buffer at 37°C for 30 minutes, and then incubated with T4 ligase (NEB, 1 U/μl T4 DNA ligase, 1 × T4 DNA ligation buffer and 0.5% PEG2000) at 37°C for 1 hour. This produces capture probes containing a 25 nt CID barcode, a 10 nt UMI and a 22 nt poly-T ready for poly-A RNA capture. A detailed step-by-step protocol describing this procedure is included in the MOSTA website (https://db.cngb.org/stomics/mosta).

Calling of CID

CID sequences together with their corresponding coordinates for all DNB were determined using a base calling method according to manufacturer's instruction of MGI DNBSEQ-Tx sequencer. After sequencing, the capture chip was split into smaller size chips (5 mm × 10 mm, 10 mm × 10 mm, 10 mm × 20 mm) ready for use. At this stage, we filtered out all duplicated CID that correspond to non-adjacent spots.

Stereo-seq library preparation and sequencing

Tissue processing

Tissue sections were adhered to the Stereo-seq chip surface and incubated at 37°C for 3-5 minutes. Then, the sections were fixed in methanol and incubated for 40 minutes at -20°C before Stereo-seq library preparation. Where indicated, the same sections were stained with nucleic acid dye (Thermo fisher, Q10212) and imaging was performed with a Ti-7 Nikon Eclipse microscope prior to *in situ* capture at the channel of FITC.

In situ reverse *transcription*

After washed with 0.1 × SSC buffer (Thermo, AM9770) supplemented with 0.05 U/μl RNase inhibitor (NEB, M0314L), tissue sections placed on the chip were permeabilized using 0.1% pepsin (Sigma, P7000) in 0.01 M HCl buffer, incubated at 37°C for 12 minutes and then washed with 0.1 × SSC buffer (Thermo, AM9770) supplemented with 0.05 U/μl RNase inhibitor (NEB, M0314L). RNA released from the permeabilized tissue and captured by the DNB was reverse transcribed overnight at 42°C using SuperScript II (Invitrogen, 18064-014, 10 U/μl reverse transcriptase, 1 mM dNTPs, 1 M betaine solution PCR reagent, 7.5 mM $MgCl_2$, 5 mM DTT, 2 U/μl RNase inhibitor, 2.5 μM Stereo-seq-TSO and 1 × First-Strand buffer). After reverse transcription, tissue sections were washed twice with 0.1 × SSC buffer and digested with Tissue Removal buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 37°C for 30 minutes. cDNA-containing chips were then subjected to Exonuclease I (NEB, M0293L) treatment for 1 hour at 37°C and were finally washed once with 0.1x SSC buffer.

Amplification

The resulting cDNAs were amplified with KAPA HiFi Hotstart Ready Mix (Roche, KK2602) with 0.8 μM cDNA-PCR primer. PCR reactions were conducted as follows: incubation at 95°C for 5 minutes, 15 cycles at 98°C for 20 seconds, 58°C for 20 seconds, 72°C for 3 minutes and a final incubation at 72°C for 5 minutes.

Library construction and sequencing

The concentrations of the resulting PCR products were quantified by Qubit™ dsDNA Assay Kit (Thermo, Q32854). A total of 20 ng of DNA were then fragmented with *in-house* Tn5 transposase at 55°C for 10 minutes, after which the reactions were stopped by the

addition of 0.02% SDS and gently mixing at 37°C for 5 minutes after fragmentation. Fragmented products were amplified as described below: 25 μl of fragmentation product, 1 × KAPA HiFi Hotstart Ready Mix and 0.3 μM Stereo-seq-Library-F primer, 0.3 μM Stereo-seq-Library-R primer in a total volume of 100 μl with the addition of nuclease-free H_2O. The reaction was then run as: 1 cycle of 95°C 5 minutes, 13 cycles of 98°C 20 seconds, 58°C 20 seconds and 72°C 30 seconds, and 1 cycle of 72°C 5 minutes. PCR products were purified using the AMPure XP Beads (0.6× and 0.15×), used for DNB generation and finally sequenced on MGI DNBSEQ-Tx sequencer with read length indicated in Table S1.

In situ hybridization

Preparation of probes and sections

Total RNA was prepared using Trizol from the whole embryos of E9.5 and E11.5 respectively and were further used as template for preparing probes. After reverse transcription using polyT, target genes were amplified using oligonucleotides harboring T7 promoter. The PCR products were used as template to synthesize Digoxin-labelled antisense RNA probes by in vitro transcription. Cryosections were sagittally cut into 10 μm thickness of slides in a Leika CM1950 cryostat ready to use for hybridization.

Hybridization

Slides were incubated at 37°C for 3 minutes, then were fixed in 4% paraformaldehyde and incubated for 1 hour at room temperature. Slides were washed by PBS supplemented with 0.1% Tween for 5 minutes for three times and permeabilized in PBS supplemented with 0.3% Triton for 20 minutes at room temperature. For pre-hybridization, slides were placed into the hybridization buffer (10% dextran, 5 × SSC, 50% formamide, 0.1% Tween, 1 mg/ml yeast RNA, 100 μg/ml heparin, 1 × Denhardt's solution, 0.1% CHAPS and 5 mM EDTA) at 60°C for 1 hour. After pre-hybridization, slides were hybridized with 500 ng/ml RNA probes in hybridization buffer overnight at 60°C. The next day, slides were washed by gradient SSC buffer containing 0.1% Tween 20 at 60 °C, which include twice in 5 × SSC for 30 minutes, 2 × SSC for 30 minutes, 0.2 × SSC for 30 minutes, and finally cooling to room temperature with 0.2 × SSC for 30 minutes.

Staining

Next, slides were washed twice with TNE buffer (10 mM Tris-pH7.5, 500 mM NaCl, 1 mM EDTA) for 10 minutes, and further subjected to RNase A (Sigma, R4642) treatment in TNE buffer for 1 hour at room temperature. After washing twice with TNE buffer for 10 minutes, slides were washed three times with MAB buffer (100 mM maleic acid, 150 mM NaCl, 0.1% Tween) for 5 minutes, and were further treated in blocking buffer (Roche, 11096176001) at room temperature for 1 hour. After blocking, slides were incubated with Anti-DIG-AP fab antibody (Roche, 11093274910), and washed five times with MAB buffer for 10 minutes at room temperature, followed by washing twice with AP buffer (100 mM Tris pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl and 0.1% Tween) for 10 minutes. Finally, slides were incubated in BM purple (Roche, 11442074001) for staining. The resultant stained slides were imaged with a Ti-7 Nikon Eclipse microscope.

QUANTIFICATION AND STATISTICAL ANALYSIS

Stereo-seq raw data processing

Fastq files were generated using a MGI DNBSEQ-Tx sequencer. CID and MID are contained in the read 1 (CID: 1-25 bp, MID: 26-35 bp) while the read 2 consist of the cDNA sequences. CID sequences on the first reads were first mapped to the designed coordinates of the *in situ* captured chip achieved from the first round of sequencing, allowing 1 base mismatch to correct for sequencing and PCR errors. Reads with MID containing either N bases or more than 2 bases with quality score lower than 10 were filtered out. CID and MID associated with each read were appended to each read header. Retained reads were then aligned to the reference genome (mm10) using STAR (Dobin et al., 2013) and mapped reads with MAPQ > 10 were counted and annotated to their corresponding genes). UMI with the same CID and the same gene locus were collapsed, allowing 1 mismatch to correct for sequencing and PCR errors. Finally, this information was used to generate a CID-containing expression profile matrix. The whole procedure was integrated into a publicly available pipeline SAW available at https://github.com/BGIResearch/SAW.

Image-based single cell segmentation and diffusion analysis

Cell segmentation with nucleic acid *staining*

We leveraged nucleic acid staining from the same section to segment cells by projecting the staining image to the Stereo-seq chips. To achieve this, we summed the total UMI in each DNB spot which harbors a specific spatial coordinate to generate a spatial density matrix, then converted the matrix into an image where each pixel corresponds to one DNB and total UMI of DNB spot the grayscale of the pixel. We then registered the DNB image with nucleic acid staining image manually. After alignment, we applied Scikit-image package (V0.18.1) to perform cell segmentation analysis (Van der Walt et al., 2014). Briefly, the background of the staining image was removed with a global threshold approach, then watershed algorithm was applied to obtain single cell segmentations. The number of markers required for the watershed algorithm were obtained through Gaussian-weighted local threshold binarization with block size of 41 and offset of 0.003. We then exacted Euclidean distance transformation (with distance of 13 or 15) from the background removed images. For each of the segmented cell, UMI from all DNB within the corresponding segmentation were aggregated per-gene and then summed to generate a cell by gene matrix for downstream analysis. The centroid of each cell was determined using rearrr (https://github.com/LudvigOlsen/rearrr).

Diffusion analysis

We assessed the lateral diffusion in Stereo-seq based on the spatial distribution of three transcript sets: (1) known nuclear localized long non-coding RNAs *Malat1* and *Neat1*; (2) Mitochondrial RNAs and (3) *Vip* transcripts that specifically expressed in *Vip*$^+$ interneurons. For (1) and (2), we sampled 100 cells from all the segmented cells, and calculated the mean distance between each transcript to the centroid of each cell. After repeating 50 times, the distribution of (1) and (2) were plot. For (3), we compared the lateral diffusion of Stereo-seq with published smFISH data downloaded from Linnarsson Lab website (http://linnarssonlab.org/osmFISH/). *Vip*$^+$ interneurons were selected either based on cell segmentation-based clustering (Stereo-seq) or the reported annotations (smFISH). The diffusion distance was further determined based on the distance between *Vip* transcripts with the centroid of each *Vip*$^+$ interneurons.

Stereo-seq benchmarking analysis and clustering

Comparison of Stereo-seq with published methods

Expression profile matrix of mouse olfactory bulb was divided into non-overlapping bins covering an area of X × X DNB, with X ∈ (*3, 14, 140*) and the transcripts of the same gene were aggregated within each bin. HDST data were taken from GSE130682 (Vickovic et al., 2019), SLIDE-seqV2 data from the Single Cell Portal of the Broad Institute (Stickels et al., 2021), DBiT-seq data from GSE137986 (Liu et al., 2020), Visium data from GSE153859 (Lebrigand et al., 2020), and Seq-Scope data from GSE169706 (Cho et al., 2021). For Figure 1C, to ensure that proper comparisons were made, the data of Stereo-seq were binned into bin 3 (3 × 3 DNB, ~2 μm), bin 14 (14 × 14 DNB, ~10 μm) or bin 140 (140 × 140 DNB, ~100 μm).

Spatially constrained clustering (SCC)

Expression profile matrix of adult mouse olfactory bulb were binned into bin 14, mouse brain or mouse embryo were binned into bin 50 (50 × 50 DNB) and the transcripts of the same gene were aggregated within each bin. Next, data were log-normalized in dynamo and SCC was performed using dyn.tl.scc and dyn.tl.purity functions. In brief, to include spatial information during clustering, we built a spatial k-nearest neighbor graph $G^{k_1}_{spatial}$ (k_1 is by default set to be 8 as each bin has 8 nearest spatial neighbors) using Squidpy (Palla et al., 2021) and then took the union with the k-nearest neighbor graph $G^{k_2}_{expression}$ based on transcriptomic data (k_2 is by default set to be 30). The combined graph ($G_{combined} = G^{k_1}_{spatial} \cup G^{k_2}_{expression}$) was then used as input for leiden clustering. Further, each cluster was annotated based on the cluster specific markers identified by the rank_genes_groups function of SCANPY (Wolf et al., 2018) using default parameters as well as the anatomic annotation based on eHistology Kaufman Annotations (http://www.emouseatlas.org/emap/home.html) or Allen Brain Atlas (http://mouse.brain-map.org/). Where indicated, the selected organs were further subjected to recluster using SCC.

Unsupervised clustering of the segmented cells

For the adult mouse hemibrain, cells containing a total number of detected genes < 200, and those with a total number of detected genes ranking in the top 0.1% were filtered out. The resulting cells were further processed by Seurat (Hao et al., 2021) followed by SCTransform, scaling, feature genes selection, PCA dimension reduction and clustering with resolution parameter set at 3.8. For the E16.5 whole mouse embryo, cells with a total number of detected genes < 200 were filtered out, and further processed by Seurat followed by same procedure described above but with resolution parameter set at 2.5. The data of the E16.5 brain was processed likewise but with resolution parameter 5.0. For the mouse developing dorsal midbrain, segmented cell data from E12.5 (E1S3), E14.5 (E1S3), E16.5 (E1S3, E2S6, E2S7) were subjected to the same procedure with resolution parameter set at 0.5.

Spatially resolved gene regulatory networks

The analysis of regulon activity was performed by following the standard SCENIC pipeline (Aibar et al., 2017). Binned (bin 50) or segmented cell expression data matrix was used as the input to SCENIC. The expression matrix was subjected to GENIE3 algorithm analysis to reconstruct the co-expressed gene network for each transcription factor. Transcription factors co-expression modules were then analyzed by RcisTarget and their potential targets were further filtered with default parameters. The filtered potential targets were used to build the regulons. The regulon activity (area under the curve) was analyzed with AUCell and the active regulons were determined with AUCell default threshold. The activity of regulons for each bin was then mapped to the physical space. The gene network for Figure 5G was constructed by selecting the target genes of corresponding transcription factors related to gliogenesis or neurogenesis, and further visualized by Cytoscape.

Identification of spatially auto-correlated gene or regulon modules

Spatially auto-correlated gene or regulon modules were identified using Hotspot (DeTomaso and Yosef, 2021). The expression matrix for the top 5,000 variable genes and all regulon activity matrix of each embryo were used as the input. For the gene modules, the data were normalized by the total UMI number of each bin, from which a k-nearest neighbor (knn) graph of genes was created using the *create_knn_graph* function with the parametes: *n_neighbors = 30* (for regulon, *n_neighbors = 10*), then genes or regulons with significant spatial autocorrelation (*FDR < 0.05*) were kept for further analysis. The modules were identified using the *create_modules* with the parameters: *min_gene_threshold = 20* and *fdr_threshold = 0.05* (for the regulon: *min_gene_threshold = 5* and *fdr_threshold = 0.05*).

Cell type mapping

To validate the cell type identity characterized in our adult mouse hemibrain section, annotated scRNA-seq profiles of the adult mouse brain downloaded from Linnarsson Lab website (http://mousebrain.org/adolescent/) (Zeisel et al., 2018) were used for comparison. In brief, cells from Stereo-seq and scRNA-seq were subjected to SCTransform, FindIntegrationAnchors and IngegrateData using reciprocal principal component in Seurat, and were further co-embedded in low-dimension space via UMAP. After co-embedding, a knn based method was used to connect cell clusters between Stereo-seq and scRNA-seq. Briefly, for a random subset of 80% of cells in each cluster of the Stereo-seq data, we identified five closest neighbor cells from the adult mouse hemibrain scRNA-seq data in the co-embedding. We next calculated the cell type proportion of such neighbors mapped to each cell clusters of the scRNA-seq data. After repeating 500 times, the median cell type proportions were then used as cell-type probability score between Stereo-seq clustering and scRNA-seq clustering results. For Figure S1I, cell types with probability < 0.05 were filtered out. For Figure S1J, cell types with probability < 0.1 were filtered out.

Alignment of single-cell transcriptome with Stereo-seq by Tangram

To validate the cell type distribution patterns derived from Stereo-seq, we used Tangram to map the annotated scRNA-seq mouse embryonic brain data downloaded from Linnarsson Lab (http://mousebrain.org/development/) (La Manno et al., 2021), or annotated scRNA-seq mouse developing cortex from GSE153164. In brief, marker genes in the single cell clusters were determined by FindAllMarkers function in Seurat, the top 20 DEG based on fold change were selected as training genes for Tangram to project the single cell to Stereo-seq. Then, the normalized cell type probabilities were visualized in Stereo-seq data.

Reanalysis of scATAC-seq data

scATAC-seq data of mouse developing cortex were retrieved from GSE153164 (Di Bella et al., 2021). Cell type annotation for scATAC-seq profiled were transfered from scRNA-seq via the TransferData function in Seurat. Then, Signac (Stuart et al., 2021) was employed to analyze the scATAC-seq data using standard process. The transcription factor deviation score was calculated using ChromVAR (Schep et al., 2017).

Inference of developmental trajectories

Embryo developmental trajectory analysis

To construct embryo-wide developmental trajectory during organogenesis, we integrated and co-embedded bins from representative sections (E9.5 E1S1, E9.5 E2S1, E10.5 E1S1, E10.5 E2S1, E11.5 E1S1, E11.5 E1S2, E12.5 E1S1, E12.5 E2S1, E13.5 E1S1, E13.5 E1S2, E14.5 E1S1, E14.5 E1S2, E15.5 E1S1, E15.5 E1S2, E16.5 E1S1, E16.5 E1S2) at consecutive time points using Seurat with FindIntegrationAnchors and IntegrateData functions with default settings. After integration, edge weight between organ across embryonic stages were calculated as described in TOME (Qiu et al., 2022a). At last, continuous trajectory with edges weight > 0.2 from the UMAP embedding were retained for the resulting directed graph.

Spatial RNA velocity analysis

Interneuron RNA velocity analysis (Figure 4) was performed using dynamo (Qiu et al., 2022b) following the tutorial (https://dynamo-release.readthedocs.io/). Unspliced and spliced RNA for each cell were extracted from the E16.5 telencephalon with SAW. Inhibitory neurons were extracted according to the tangram score of forebrain GABAergic neuroblasts or forebrain GABAergic neurons. The data matrix was then processed by dynamo to normalize the expression, select feature genes and perform PCA dimension reduction, followed by default parameters to estimate the kinetic parameters and gene-wise RNA velocity vectors that were then projected to the physical space. Specifically, the ''Fokker-Planck'' kernel implemented in dynamo was used for projection of high-dimensional RNA velocity vectors to physical space. Streamlines of only inhibitory neurons were visualized. To facilitate the understanding of gene expression dynamics over space, we reconstructed the continuous vector field in the PCA space, and then utilized hodge decomposition which takes the simplicial complexes (a sparse directional graph) constructed based on the learned vector field function to infer the vector field based pseudotime (vf pseudotime). At last, we also illustrated the kinetics of all high PCA loading genes along the vf pseudotime.

Monocle 3 analysis

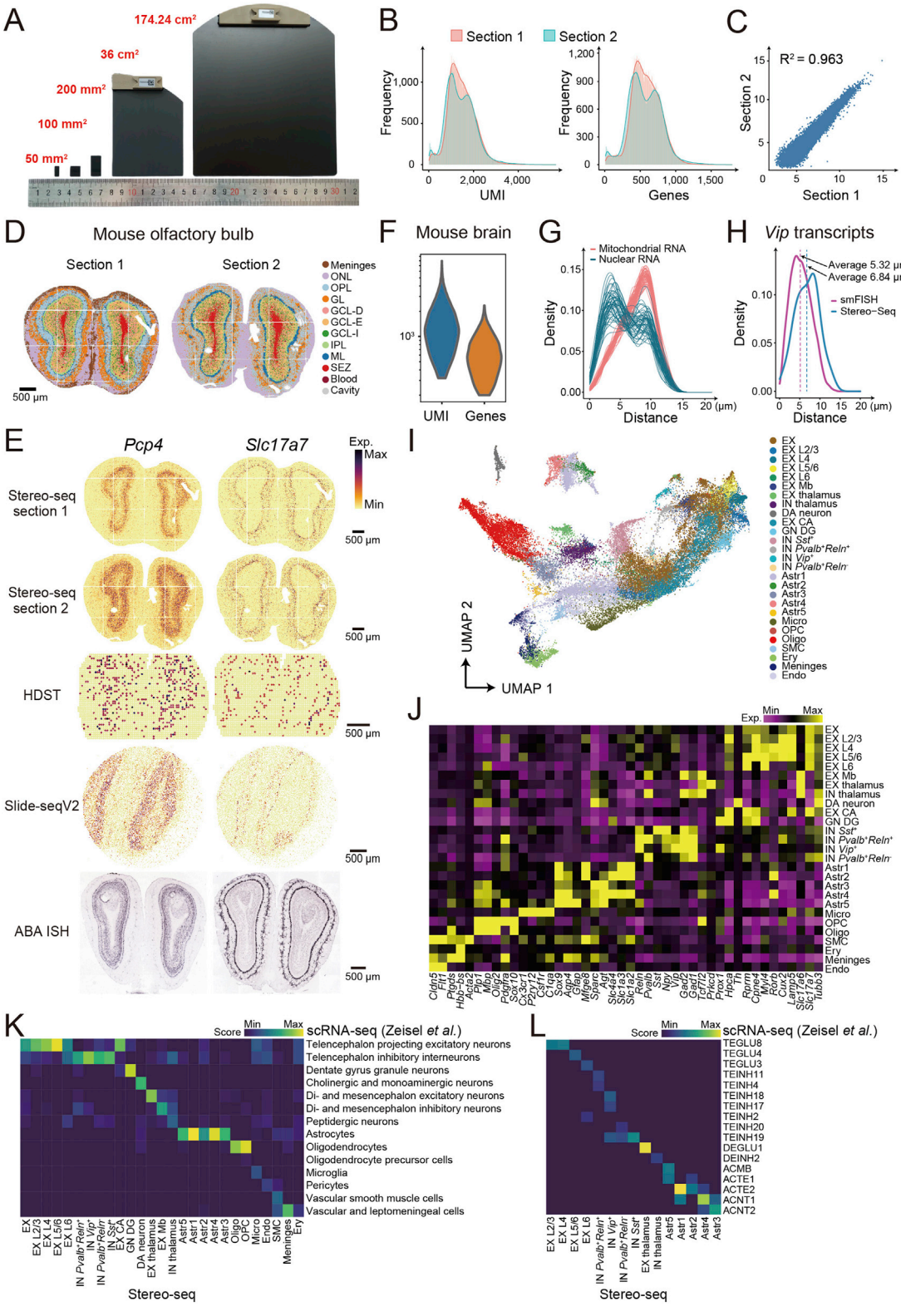
For midbrain developmental trajectory (Figure 5), after initial data processing, DEG identification and cell type annotation, the putative radial glia cell, glioblast and neuroblast clusters were subjected to pseudotime analysis by monocle 3 (Cao et al., 2019). DEG of radial glia cell, glioblast and neuroblast were visualized by heatmap ordered by the pseudotime trajectory.

Spatial gene enrichment analysis of genes associated with human developmental disorders

Monogenic disease information was retrieved from DDG2P (v2.29) (Wright et al., 2015). After filtering out low-expressed genes with counts per million less than 1 in all anatomic regions, a list of total 1,919 genes were kept for further analysis. The expression level of indicated organs at selected time point were aggregated, *Z* score normalized, and further visualized by as scatterplots with the physical information.

CellPress
OPEN ACCESS

Supplemental figures



(legend on next page)

Figure S1. Performance of stereo-seq, related to Figures 1 and 2

(A) Stereo-seq chips of different sizes ranging from 50 mm^2 to 174.24 cm^2.

(B) Distribution of UMI (left) and genes (right) from two adjacent mouse olfactory bulb sections.

(C) Pearson correlation coefficient (R^2 = 0.9666) of pseudo-bulk profiles from the two mouse olfactory bulb sections.

(D) Unsupervised SCC of the two mouse olfactory bulb sections analyzed by Stereo-seq data at bin 14 resolution. Bins are colored by their annotation. ONL, olfactory nerve layer; OPL, outer plexiform layer; GL, glomerular layer; GCL-D, granular cell zone deep; GCL-E, granular cell layer externa; GCL-I, granular cell layer internal; IPL, internal plexiform layer; ML, mitral layer; and SEZ, subependymal zone. Scale bars, 500 μm.

(E) Spatial visualization of *Pcp4* and *Slc17a7* expressions of the two mouse olfactory bulb sections and reported HDST (Vickovic et al., 2019), Slide-seqV2 (Stickels et al., 2021), and ISH images of the adult mouse olfactory bulb taken from ABA (Lein et al., 2007). Scale bars, 500 μm.

(F) Violin plots showing the number of detected UMI (left) and genes (right) for the segmented cells of adult mouse hemibrain.

(G) Distribution of nuclear localized RNA (*Malat1* and *Neat1*) and mitochondria RNA (cytoplasmic) of the sampled cells, each line representing each sampling.

(H) Comparison of lateral diffusion distance for *Vip* transcripts to the centroid of Vip^+ GABAergic interneurons between the adult mouse hemibrain Stereo-seq data and a reported mouse somatosensory cortex smFISH data (Codeluppi et al., 2018).

(I) UMAP visualization of the segmented cells from the adult mouse hemibrain section. Cells were colored by their annotation. EX, excitatory glutamatergic neuron; IN, GABAergic interneuron; DA, dopaminergic neuron; GN DG, granule cell of dentate gyrus; Astr, astrocyte; Micro, microglia; OPC, oligodendrocyte precursor cell; Oligo, oligodendrocyte; SMC, smooth muscle cell; Ery, erythrocyte; and Endo, endothelial cell.

(J) Heatmap showing the normalized expression of selected marker genes for the indicated cell types of the mouse hemibrain section shown in Figure 2C.

(K) Correspondence between the indicated cell types identified by Stereo-seq and the cell clusters identified by reported scRNA-seq (Zeisel et al., 2018).

(L) Correspondence between the indicated cell subtypes identified by Stereo-seq and the cell subtypes identified by reported scRNA-seq (Zeisel et al., 2018).

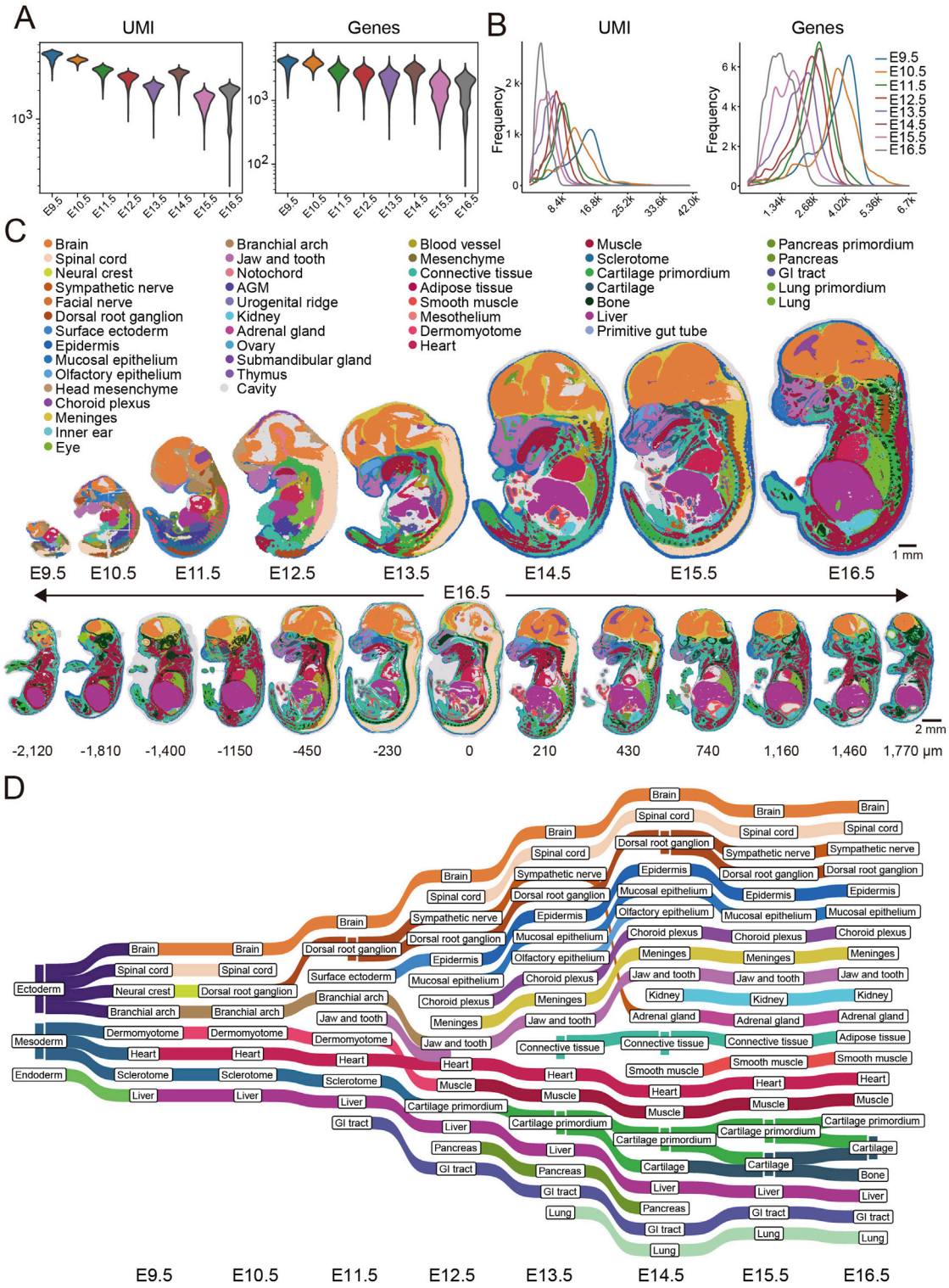


Figure S2. Spatially resolved transcriptomic atlas of mouse organogenesis, related to Figure 3

(A) Violin plots showing the number of detected UMI (left) and gene (right) per bin 50 for the E9.5–E16.5 sections shown in Figure 3A.

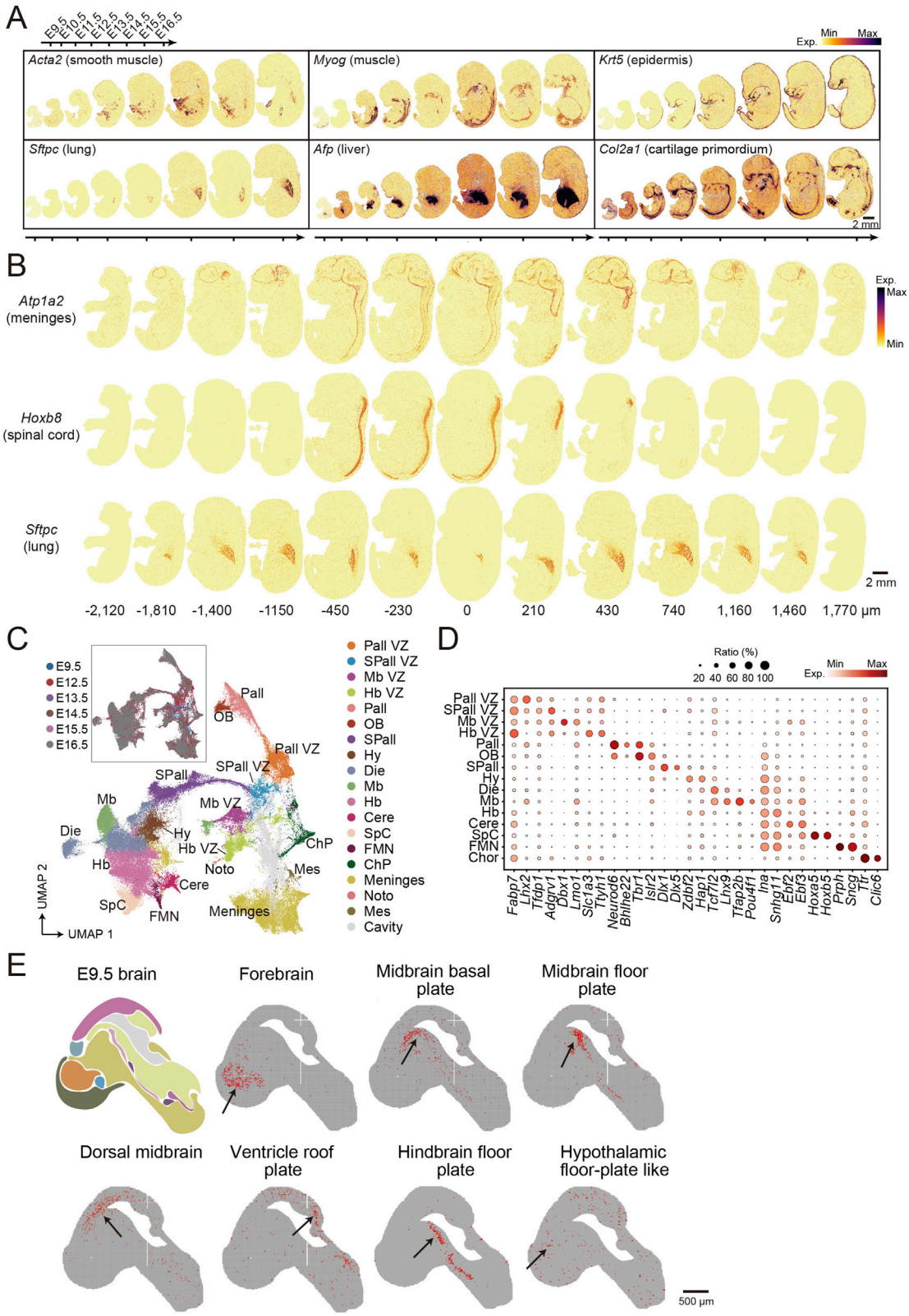
(B) Distribution of number of detected UMI (left) and gene (right) per bin 50 for the E9.5–E16.5 sections shown in Figure 3A.

(C) Upper: unsupervised SCC of mouse embryo sections across E9.5–E16.5. Embryo sections including E9.5 E2S1, E10.5 E2S1, E11.5 E1S2, E12.5 E2S1, E13.5 E1S2, E14.5 E1S2, E15.5 E1S2, and E16.5 E1S2 were shown. Bins are colored by their annotation. Bottom: unsupervised SCC of the mouse embryo sections at

(legend continued on next page)

E16.5. Thirteen sections from the same E16.5 embryo including E2S1–E2S13 are shown. Bins are colored by their annotation. The distance of each section from the midline is indicated. Scale bars, 1 mm.

(D) Directed acyclic graph showing inferred relationships between anatomic regions across sections from E9.5 to E16.5 shown in Figures 3A and S2C. Each row corresponds to one of the anatomic regions shown in Figures 3A and S2C, and the columns to the developmental stages spanning E9.5 to E16.5.



(legend on next page)

Figure S3. Spatiotemporal dynamics of regional specification, related to Figure 3

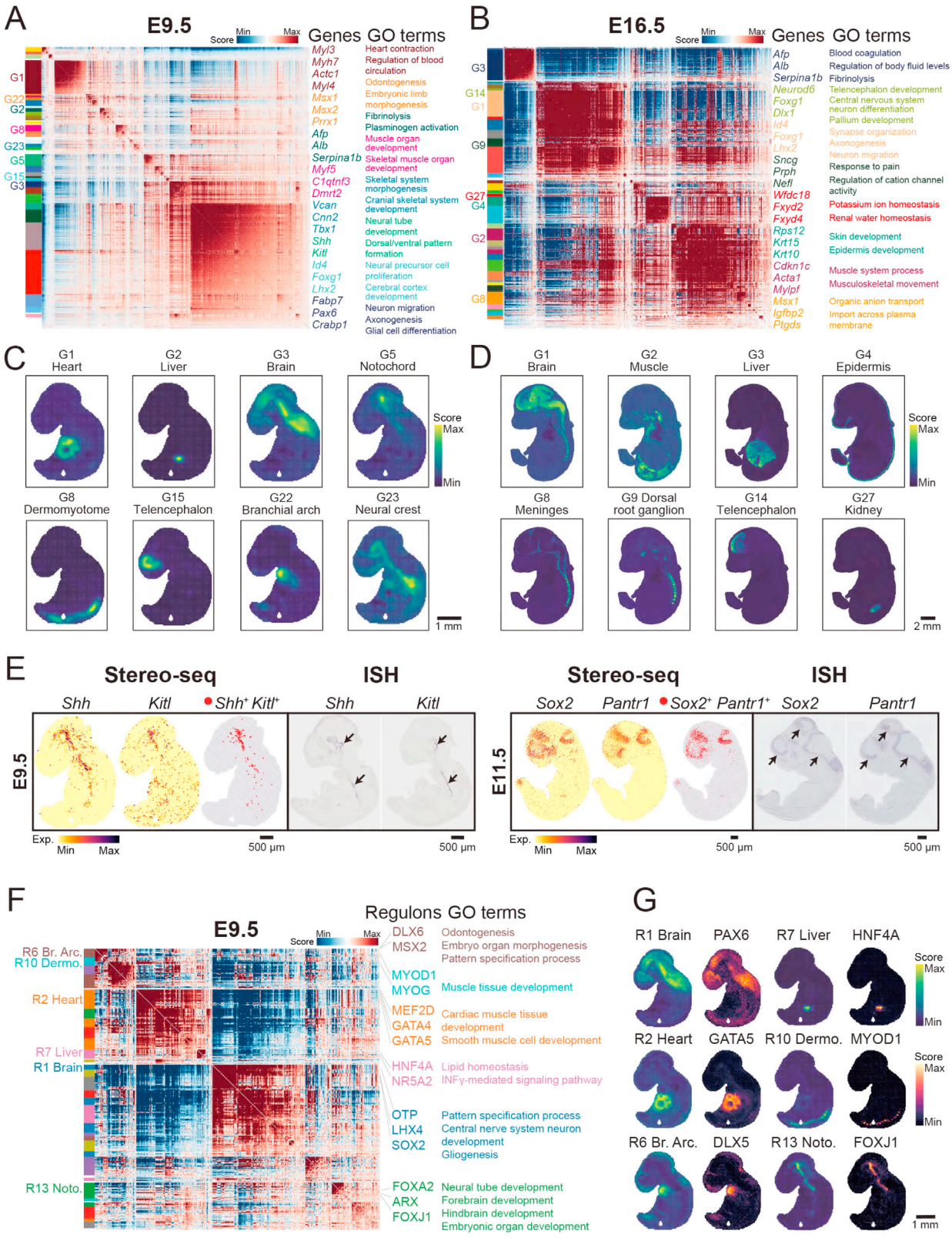
(A) Spatial visualization of the indicated gene expression for smooth muscle (*Acta2*$^+$), muscle (*Myog*$^+$), epidermis (*Krt5*$^+$), lung (*Sftpc*$^+$), liver (*Afp*$^+$), and cartilage primordium (*Col2a1*$^+$) in the sections shown in Figure 3A. Scale bars, 2 mm.

(B) Spatial visualization of the indicated gene expression for meninges (*Atp1a2*$^+$), spinal cord (*Hoxb8*$^+$), and lung (*Sftpc*$^+$) in 13 sections of E16.5 shown in (C). Scale bars, 2 mm.

(C) UMAP visualization of the brain retrieved from the E9.5 and E12.5–E16.5 sections shown in Figure 3C. Bins are colored by stage (upper left) and the annotations (bottom). MZ, mantle zone; VZ, ventricular zone; Pall, pallium; SPall, subpallium; Die, diencephalon, Hy, hypothalamus; Mb, midbrain; SpC, spinal cord; Hb, hindbrain; Cere, cerebellum; OB, olfactory bulb; ChP, choroid plexus; Noto, notochord; FMN, facial motor nucleus; Mes, mesenchyme.

(D) Bubble plot showing the expression levels of the specific markers in each of the brain regions shown in Figure S3C.

(E) Tangram inferred spatial distributions of different types of radial glia cells for an E9.5 section (E1S1). A reported single-cell mouse developing brain data (La Manno et al., 2021) was used for the analysis. Scale bars, 500 μm.



(legend on next page)

Figure S4. Modules of gene expression and gene regulatory networks in the developing embryos, related to Figure 3

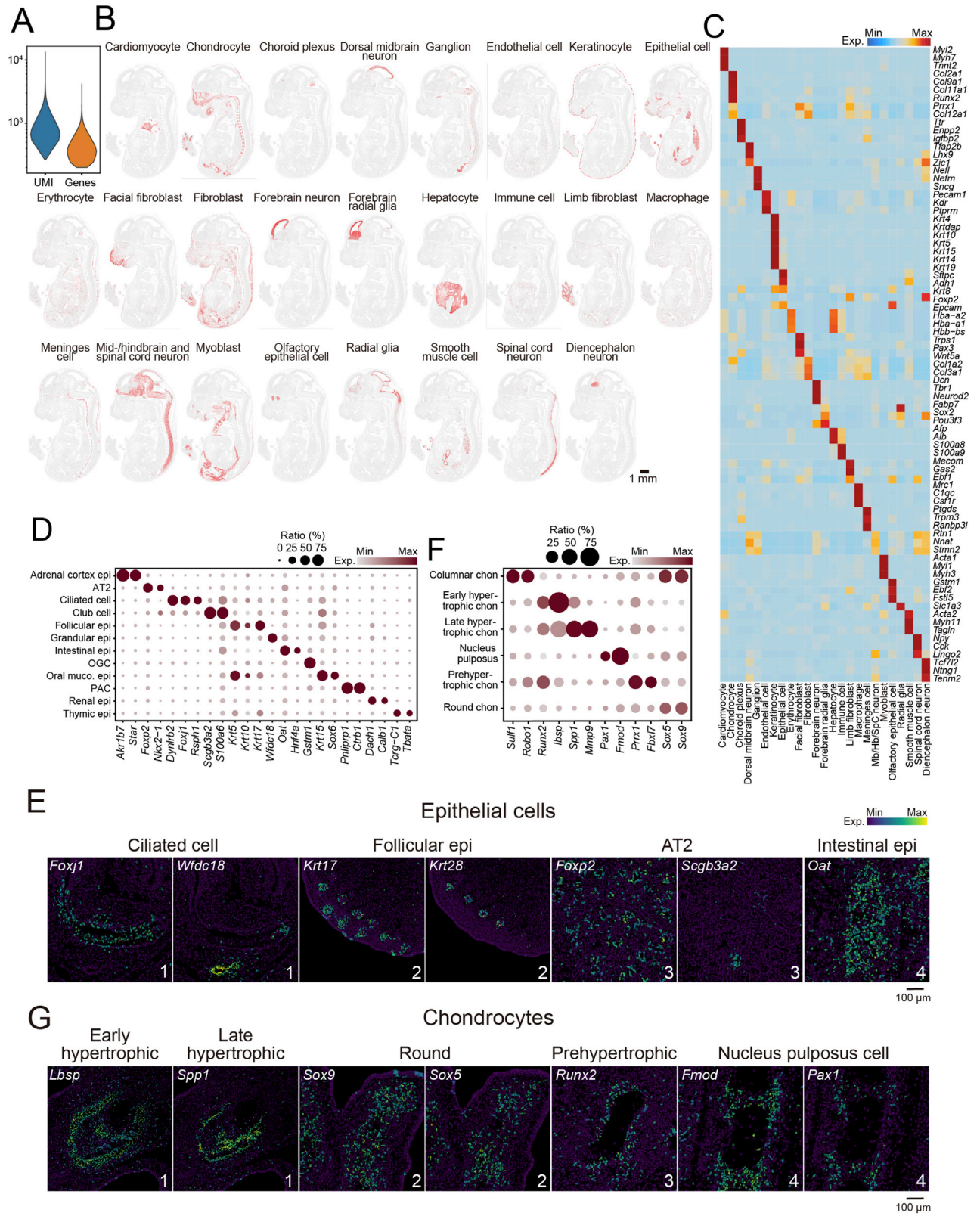
(A and B) Heatmaps showing the genes with significant spatial autocorrelation grouped into different gene modules based on pairwise spatial correlations of E9.5 (left, E1S1) and E16.5 (right, E1S1) sections. Selected genes and GO terms related to representative gene modules are highlighted on the right side of each heatmap.

(C and D) Spatial visualization of the example gene modules related to the indicated organs on E9.5 (left, E1S1) and E16.5 (right, E1S1) sections. Scale bars, 1 mm (left) and 2 mm (right).

(E) Spatial visualization of gene expression, coexpression, and ISH images for *Shh* and *Kitl* in an E9.5 section (E1S1, left), and *Sox2* and *Pantr1* in an E11.5 section (E1S4, right). Scale bars, 500 μm.

(F) Heatmap showing the regulons with significant spatial autocorrelation grouped into different modules based on pairwise spatial correlations of an E9.5 section (E1S1). Selected regulons and their corresponding GO terms related to representative regulon modules are highlighted on the right side.

(G) Spatial visualization of example regulon modules and their representative regulon for the indicated organs. Scale bars, 1 mm.



(legend on next page)

Figure S5. Spatial heterogeneity of cell types and cell states in the E16.5 embryo, related to Figure 4

(A) Violin plot showing the number of detected UMI and genes per segmented cell in the E16.5 section (E1S3) shown in Figure 4A.
(B) Spatial visualization of the indicated cell types shown in Figure 4A. Scale bars, 1 mm.
(C) Heatmap showing the normalized expression of selected markers for the indicated cell types from E16.5 section (E1S3) shown in Figure 4A.
(D) Bubble plot showing the normalized expression of selected markers for the indicated subtypes of the epithelial cell clusters shown in Figure 4A.
(E) Spatial visualization of the expression of the indicated genes in the regions squared in Figure 4B. Scale bars, 100 μm.
(F) Bubble plot showing the normalized expression of selected markers for the indicated cell states of the chondrocyte clusters shown in Figure 4A.
(G) Spatial visualization of the expression of the indicated genes in the regions squared in Figure 4C. Scale bars, 100 μm.

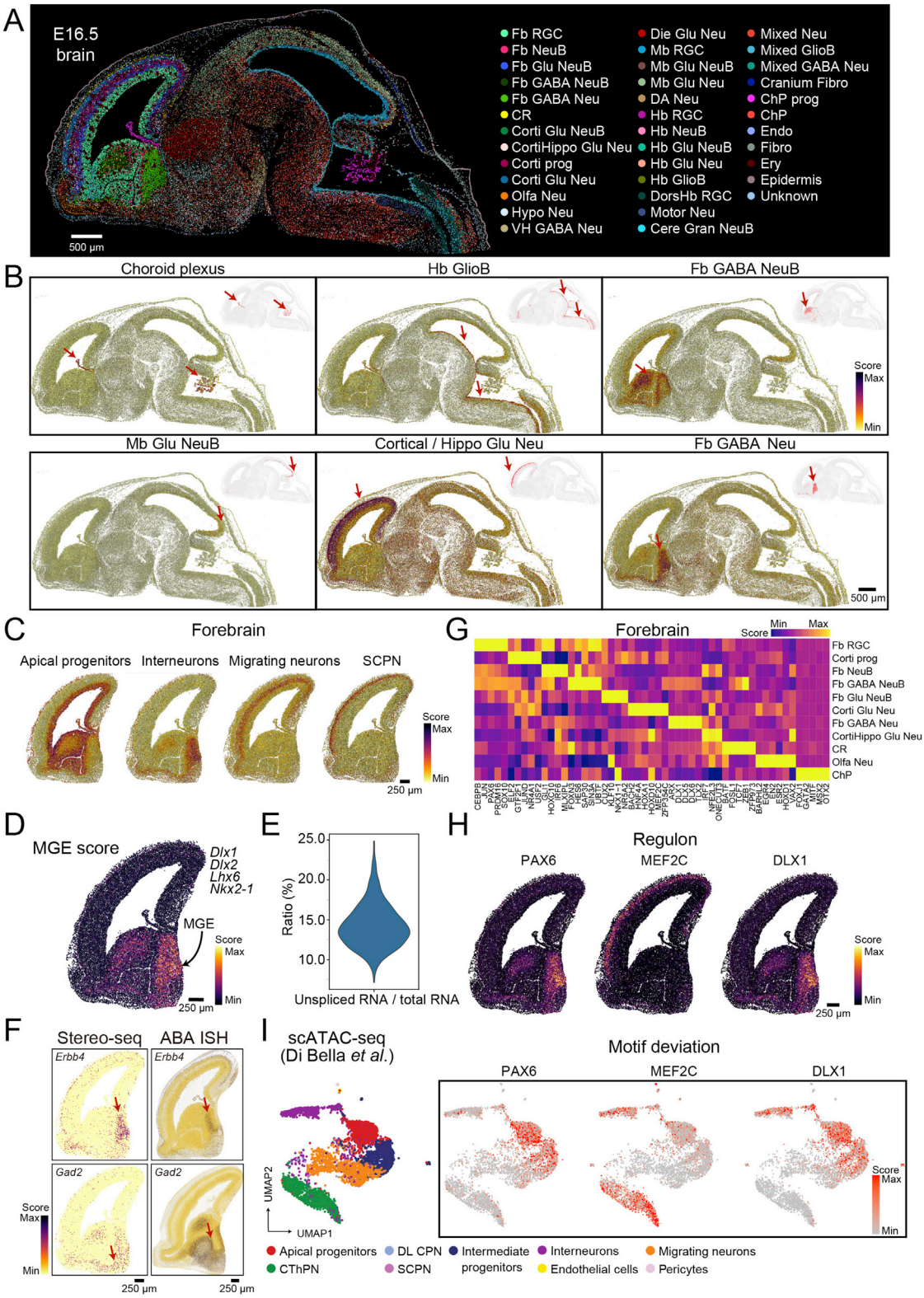


Figure S6. Spatial heterogeneity of cell types in the E16.5 brain, related to Figure 4

(A) Spatial visualization of cell types in the E16.5 brain. Cells are colored by the annotations. Fb RGC, forebrain radial glia cell; Fb NeuB, forebrain neuroblast; Fb Glu NeuB, forebrain glutamatergic neuroblast; Fb GABA NeuB, forebrain GABAergic neuroblast; Fb GABA Neu, forebrain GABAergic neuron; CR, Cajal-Retzius

(legend continued on next page)

cell; Corti Glu NeuB, cortical glutamatergic neuroblast; CortiHippo Glu Neu, cortical or hippocampal glutamatergic neuron; Corti prog, cortical intermediate progenitor; Corti Glu Neu, cortical glutamatergic neuron; Olfa Neu, olfactory neuron; Hypo Neu, hypothalamus neuron; VH GABA Neu, ventromedial hypothalamus GABAergic neuron; Die Glu Neu, diencephalon glutamatergic neuron; Mb RGC, midbrain radial glia cell; Mb Glu NeuB, midbrain glutamatergic neuroblast; Mb Glu Neu, midbrain glutamatergic neuron; DA Neu, dopaminergic neuron; Hb RGC, hindbrain radial glia cell; Hb NeuB, hindbrain neuroblast; Hb Glu NeuB, hindbrain glutamatergic neuroblast; Hb Glu Neu, hindbrain glutamatergic neuron; Hb GlioB, hindbrain glioblast; DorsHb RGC, dorsal hindbrain radial glia cell; Motor Neu, motor neuron; Cere Gran NeuB, cerebellar granule neuroblasts; Mixed Neu, mixed region neuron; Mixed GlioB, mixed region glioblast; Mixed GABA Neu, mixed region GABAergic neuron; Cranium Fibro, cranium fibroblast; ChP, choroid plexus; Endo, endothelial cell; Fibro, fibroblast; Ery, erythrocyte. Scale bars, 500 μm.

(B) Tangram inferred spatial distributions of the indicated cell types at E16.5 brain (E1S3). A reported scRNA-seq of mouse developing brain data (La Manno et al., 2021) was used for the integration analysis. Scale bars, 500 μm.

(C) Tangram inferred spatial distributions of the indicated cell types at E16.5 embryo brain (E1S3). A reported scRNA-seq of mouse developing cortex data (Di Bella et al., 2021) was used for the integration analysis. Scale bars, 250 μm.

(D) Spatial visualization of the MGE module score. Genes used for MGE module score calculation are indicated on the right side. Scale bars, 250 μm.

(E) Violin plot showing the ratio of unspliced RNA versus total RNA in E16.5 GABAergic neurons/neuroblasts (E1S3).

(F) Left: spatial visualization of the expression of the indicated genes in the E16.5 telencephalon (E1S3). Scale bars, 250 μm. Right: ISH for the indicated genes of the E15.5 embryo taken from ABA. Scale bars, 250 μm.

(G) Heatmap showing the activity of regulons for the indicated cell types from the E16.5 telencephalon shown in Figure 4D.

(H) Spatial visualization of the activity for the representative regulons at the E16.5 telencephalon (E1S3). Scale bars, 250 μm. (I) UMAP visualization of a reported scATAC-seq data of the mouse developing cortex (Di Bella et al., 2021). Cells are colored by the annotations (left), or the motif deviation score for the indicated transcription factor motifs (right).

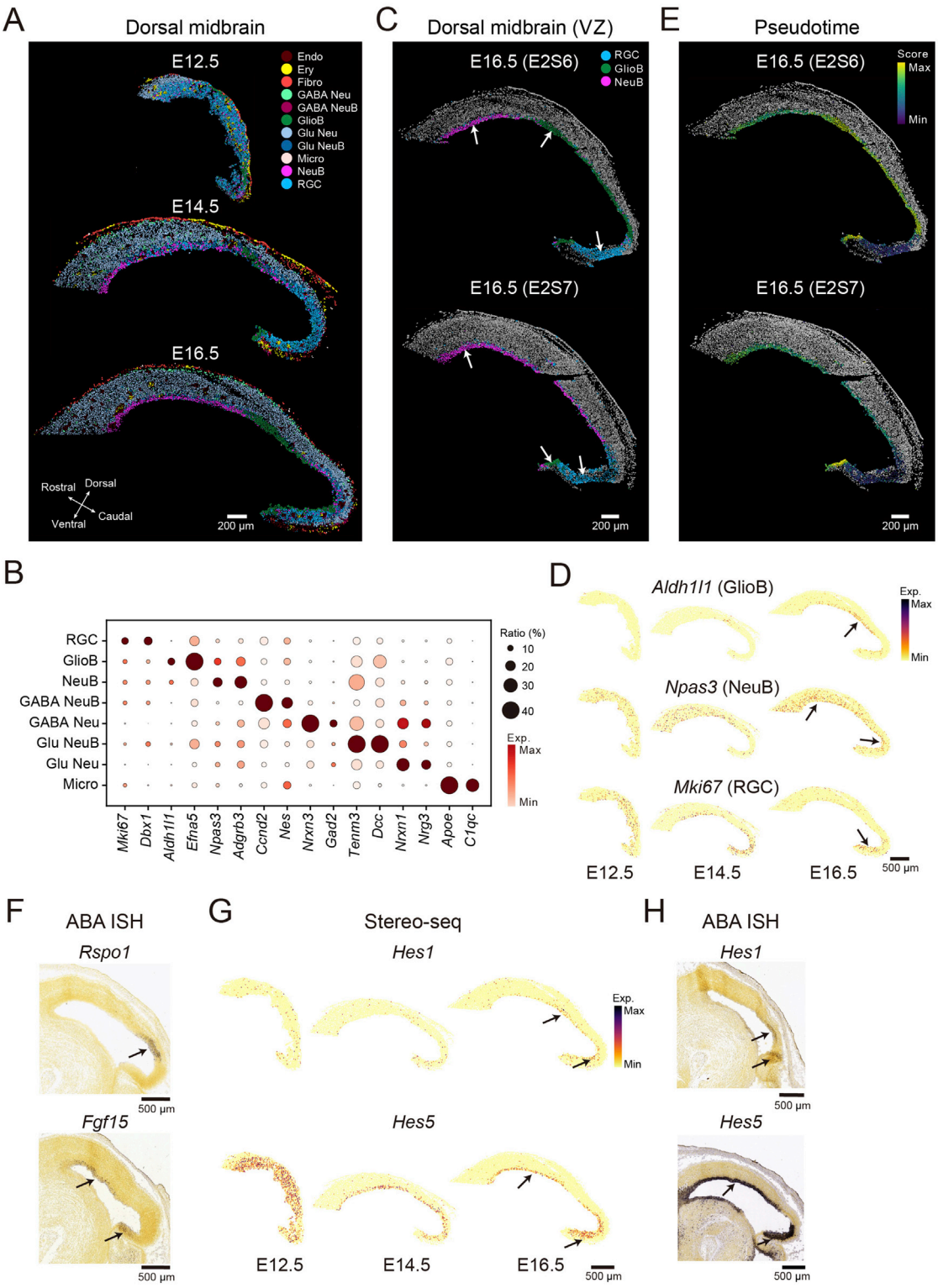


Figure S7. Spatial diversity of cell types in the developing dorsal midbrain, related to Figure 5

(A) Spatial visualization of cell types in the dorsal midbrain at the indicated regions of the E12.5 (E1S3), E14.5 (E1S3), and E16.5 (E1S3, E2S6, E2S7). Cells are colored by the annotations. Endo, endothelial cell; Ery, erythrocyte; Fibro, fibroblast; GABA Neu, GABAergic neuron; GlioB, glioblast; Glu Neu, glutamatergic neuron; Glu NeuB, glutamatergic neuroblast; Micro, microglia. NeuB, neuroblast; RGC, radial glia cell. Scale bars, 200 μm.

(B) Bubble plot showing the normalized expression of selected markers for the indicated cell types shown in Figure 5A.

(legend continued on next page)

(C) Spatial visualization of progenitor cell types in the ventricular zone of the dorsal midbrain in sections from another embryo (E16.5 E2S6 and E2S7). Scale bars, 200 μm.
(D) Spatial visualization of *Aldh1l1*, *Npas3*, and *Mki67* expression in the dorsal midbrain at the indicated time points (E12.5 E1S3, E14.5 E1S3, and E16.5 E1S3). Scale bars, 500 μm.
(E) Spatial visualization of the pseudotime score of progenitor cells in the ventricular zone of the dorsal midbrain in sections from another embryo (E16.5 E2S6 and E2S7). Scale bars, 200 μm.
(F) ISH images of the E15.5 embryo for the indicated genes taken from ABA. Scale bars, 500 μm. (G) Spatial visualization of the expression of the indicated genes in the dorsal midbrain at the indicated time points (E12.5 E1S3, E14.5 E1S3, and E16.5 E1S3) and the corresponding ISH images of the E15.5 embryo taken from ABA (H). Scale bars, 500 μm.