High-resolution spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae

Mingyue Wang[1,2,3*], Qinan Hu[2,3*], Tianhang Lv[1,4*], Yuhang Wang[1,5*], Qing Lan[1*], Zhencheng Tu[1,4*], Rong Xiang[1], Yanrong Wei[6], Kai Han[7], Yanru An[1], Mengnan Cheng[1,4], Jiangshan Xu[1,4], Miguel A. Esteban[1,8,9,10], Haorong Lu[11,12], Wangsheng Li[11], Shaofang Zhang[11], Ao Chen[1], Wei Chen[2,3], Yuxiang Li[1], Xiaoshan Wang[1], Xun Xu[1], Yuhui Hu[2,3] & Longqi Liu[1,13]

[1]BGI-Shenzhen, Shenzhen 518083, China.

[2]Shenzhen Key Laboratory of Gene Regulation and Systems Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518005, China.

[3]Department of Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518005, China.

[4]College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.

[5]South China University of Technology, Guangzhou 510006, China.

[6]BGI College & Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450000, China.

[7]BGI-Qingdao, BGI-Shenzhen, Qingdao, 266555, China.

[8]Laboratory of Integrative Biology, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China.

[9]CAS Key Laboratory of Regenerative Biology and Guangdong Provincial Key Laboratory of Stem Cells and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Guangzhou 510530, China.

[10]Institute of Stem Cells and Regeneration, Chinese Academy of Sciences, Beijing 100101, China.

[11]China National Gene Bank, BGI-Shenzhen, Shenzhen 518120, China.

[12]Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen, 518120, China.

[13]Shenzhen Key Laboratory of Single-Cell Omics, Shenzhen 518083, China.

[*]These authors contributed equally to this paper.

Correspondence and requests for materials should be addressed to Y.L. (liyuxiang@genomics.cn), X.W.

(wangxiaoshan@genomics.cn), X.X. (xuxun@genomics.cn), Y.H. (huyh@sustech.edu.cn) or L.L. (liulongqi@genomics.cn).

SUMMARY

Drosophila has long been a successful model organism in multiple fields such as genetics and developmental biology. *Drosophila* genome is relatively smaller and less redundant, yet largely conserved with mammals, making it a productive model in studies of embryogenesis, cell signaling, disease mechanisms, etc. Spatial gene expression pattern is critical for understanding of complex signaling pathways and cell-cell interactions, whereas temporal gene expression changes need to be tracked during highly dynamic activities such as tissue development and disease progression. Systematic studies in *Drosophila* as a whole are still impeded by lack of these spatiotemporal transcriptomic information. *Drosophila* embryos and tissues are of relatively small size, limiting the application of current technologies to comprehensively resolve their spatiotemporal gene expression patterns. Here, utilizing SpaTial Enhanced REsolution Omics-sequencing (Stereo-seq), we dissected the spatiotemporal transcriptomic changes of developing *Drosophila* with high resolution and sensitivity. Our data recapitulated the spatial transcriptomes of embryonic and larval development in *Drosophila*. With these data, we identified known and previously undetected subregions in several tissues during development, and revealed known and potential gene regulatory networks of transcription factors within their topographic background. We further demonstrated that Stereo-seq data can be used for 3D reconstruction of *Drosophila* embryo spatial transcriptomes. Our data provides *Drosophila* research community with useful resources of spatiotemporally resolved transcriptomic information across developmental stages.

INTRODUCTION

For over a century, *Drosophila* has been a productive model organism for developmental biologists and geneticists to study a diverse range of developmental processes, such as embryogenesis, organogenesis, gametogenesis, and aging. Numerous studies in *Drosophila* have led to groundbreaking discoveries, which have greatly impacted the biomedical field.

The transcriptomic profiles of cells and tissues largely determine their functions. In multicellular organisms, various types of cells with diverse transcriptomic profiles together orchestrate the functions of tissues and organs to ensure the smooth execution of biological processes. Traditionally, tissue/cell types were distinguished based on their function, anatomy, morphology, and expression of a few marker genes. As a well-established model organism, *Drosophila* has been intensively studied as to its tissue-specific transcriptomes. Marker gene expression profile in each tissue/cell type are readily available. Several databases have been established to curate the collection of tissue-specific transcriptomic profiles in *Drosophila*, including FlyAtlas1 (Chintapalli et al., 2007), FlyAtlas2 (Leader et al., 2018) and DGET (Graveley et al., 2011).

Recently, the rapid development of single-cell multi-omics technologies enables the mapping of genomic, transcriptomic, epigenomic and proteomic information at single-cell resolution. This has led to multiple studies revealing cell heterogeneity of *Drosophila* tissues by single-cell multi-omics sequencing, such as embryo (Karaiskos et al., 2017), imaginal disc (Ariss et al., 2018; Bageritz et al., 2019; Deng et al., 2019), gut (Guo et al., 2019; Hung et al., 2020), brain (Brunet Avalos et al., 2019; Davie et al., 2018) and gonad (Jevitt et al., 2020; Rust et al., 2020; Witt et al., 2019) (reviewed in Li, 2020). These studies substantially expanded our knowledge of cellular diversity, functional heterogeneity and microenvironmental regulation in *Drosophila* tissues. Comparing to traditional cell type identification methods, single-cell multi-omics sequencing measures many more dimensions of cell states and expression profiles, revealing novel cell types in multiple tissues. The Fly Cell Atlas project (Li et al., 2021), a global collaborative effort, is in progress to build comprehensive cell atlases of *Drosophila* tissues across developmental stages with single-cell multi-omics techniques.

Owing to the complex intercellular communication and coordination within or between tissues in multicellular organisms, our understanding of cellular functions greatly relies on their morphological contexts. The spatial information of cells' transcriptomes provides a wealth of insights for how cells coordinate to perform their biological functions in cell-cell signaling, metabolism and development. Currently, several databases are available for the spatial expression pattern of developing *Drosophila* embryos and larvae, such as Berkeley *Drosophila* Genome Project (BDGP) gene expression pattern studies (Tomancak et al., 2002) and Fly-FISH (Lécuyer et al., 2007). However, these databases mainly utilize *in situ* hybridization, which has several inherit drawbacks, including inability to detect unknown transcripts and isoforms, difficulty in identifying transcripts with low copy number, and lack of accuracy in quantification of gene expression levels. Moreover, the *Drosophila* spatial transcriptome is yet to be completely resolved. Within the ~14,000 genes in the *Drosophila* genome, over 40% still lack information in their spatiotemporal expression patterns (Zhou et al., 2019), and the ones with these patterns are also far from comprehensively resolved.

Recent years have witnessed the advances of spatial transcriptomic technologies, which utilize various methods to resolve the spatial patterns of transcriptomes, including computational strategies (Karaiskos et al., 2017), physical segmentation (Junker et al., 2014), local mRNA capture and sequencing (Rodriques et al., 2019) etc. These methods vary in throughput, number of genes measured and spatial resolution, and have been successfully applied to multiple tissues (reviewed in Liao et al., 2021). In an effort to capture the actual spatial transcriptomes *in situ*, some methods utilize surfaces covered with an array of uniquely DNA-barcoded beads. mRNAs from tissues (usually a section) are directly transferred to the surface and captured by the beads carrying oligo dT sequences. The spatial patterns of transcriptomes can thus be resolved after sequencing and mapping. These tools enable untargeted and comprehensive capture of cellular transcriptomic

profiles *in situ*, thus can overcome the disadvantages of *in situ* hybridization, make valuable supplement to the current spatial transcriptomics databases and facilitate discoveries of previously undetectable transcriptomic changes. They have been successfully applied to tissues such as mouse brain (Ortiz et al., 2020), human brain (Chen et al., 2020) and human heart (Asp et al., 2019).

Most of the existing DNA-barcoding array methods lack sub-micrometer level resolution or sufficient mRNA capture efficiency, impeding their application in *Drosophila* embryos (~500 μm × 100 μm in section size). Due to the small sizes of these samples, efforts in resolving their complete spatial transcriptome remained computational (Karaiskos et al., 2017). Our recently developed SpaTial Enhanced REsolution Omics-sequencing (Stereo-seq) technique (Chen et al., 2021) can resolve spatial transcriptomes with nanometer resolution while retaining high sensitivity, providing a powerful tool to obtain spatial transcriptome from small-sized samples like *Drosophila* embryos. In Stereo-seq, patterned DNA nanoballs (DNBs) with randomly barcoded sequences are photolithographically etched on a modified chip ~500 nm away from each other. This technique allows mRNA capture with high density and sensitivity, enabling spatial transcriptomic analysis at a much higher resolution. As a sequencing-based spatial transcriptomic technique with subcellular resolution, Stereo-seq makes it possible to capture the detailed spatially resolved transcriptomes in small-sized samples such as *Drosophila* embryos and larvae.

At 25 °C, the *Drosophila* life cycle starts with ~1 day of embryogenesis [can be arbitrarily divided into 17 stages based on morphology (Campos-Ortega and Hartenstein, 1997)], followed by three larval stages which take altogether ~5 days. Pupal stage follows and lasts for another ~4 days. Adult flies subsequently eclose from the pupal cases. In this study, we applied Stereo-seq to late-stage embryos and all stages of larvae, generating spatial transcriptomic data across these developmental stages. Computational analysis could correlate bin clusters with anatomical structures. With the Stereo-seq data, we also detected subregions of tissues by refined clustering, and identified active transcription factor regulatory networks in multiple tissues across embryonic and larval development. Lastly, we demonstrated that with Stereo-seq data from all the sections of an embryo, it is possible to 3D reconstruct the spatial transcriptome of an embryo *in silico*. This opens exciting opportunities for *Drosophila* research to be done in an organism-wide, 3D spatially resolved and systematic manner.

RESULTS

Application of Stereo-seq to *Drosophila* embryos and larvae

We applied Stereo-seq to the following *Drosophila* samples: late-stage embryos (14-16 h and 16-18 h after egg laying, corresponding to stage 16~17 of embryogenesis, hereafter termed E14-16 and E16-18 respectively) and all three stages of larvae (hereafter termed L1~L3) **(Fig.1A)**. In each experiment, a patterned 1cm × 1cm Stereo-seq chip was covered with 10-μm thick cryosection slices from each sample. Stereo-seq was performed as previously described (Chen et al., 2021). Sequencing

data were processed and integrated to generate 2D spatial transcriptomes of each sample. On average, Stereo-seq captured more than 1,500 unique transcripts representing over 400 genes per bin (bin 20 × 20 for E14-16, E16-18, L1 and L2, equivalent to ~10 μm × 10 μm; bin 50 × 50 for L3, equivalent to ~25 μm × 25 μm) across samples **(Fig.S1A, Table S1)**. Based on Stereo-seq data, 2D expression matrices with defined X-Y coordinates and positional expression profiles of each bin were generated. We then performed unsupervised clustering for each 2D expression matrix **(Fig.S1B)**. Clusters were associated with tissue types according to previously reported marker gene profiles, followed by manual validation and annotation. Clustering results were consistent with anatomical structures of embryos and larvae **(Fig.1B-F, Fig.S2)**.

To validate the spatial transcriptome we obtained with Stereo-seq, we compared the localization patterns of a list of transcripts that we observed in late-stage embryo samples with known *in situ* hybridization results for those transcripts, obtained from BDGP database (Tomancak et al., 2002, 2007). The spatial patterns of these transcripts in our Stereo-seq data matched well with those of *in situ* hybridization results **(Fig.S3 and S4A)**. These observations indicate that Stereo-seq effectively reflects the spatial gene expression patterns of *Drosophila* embryos. In addition, our Stereo-seq data yielded localization information for multiple transcripts that are not included in previous high-throughput *in situ* hybridization experiments from BDGP and Fly-FISH projects **(Fig.S4B)**. These data provide a valuable complement to current *in situ* hybridization databases.

Thus, Stereo-seq generated high-quality spatial transcriptomic data from *Drosophila* embryo and larva samples, identifying spatial patterns of both transcripts matching previous *in situ* hybridization data and previously undetected transcripts.

Spatiotemporal gene expression profiling reveals subregions of embryonic and larval midgut

By virtue of large number of genes captured by Stereo-seq, clusters corresponding to highly complex tissues can be further subclustered based on differences in cell subtype marker gene expression. The digestive system occupies the largest space in late-stage *Drosophila* embryos and larvae, thus correlated to the largest number of bins in our Stereo-seq data. Moreover, unsupervised clustering generated multiple distinct clusters corresponding to midgut across samples, suggesting tissue heterogeneity within embryonic and larval midgut. Thus, we extracted clusters representing midgut from all stages of investigated samples and performed refined clustering. Further unsupervised clustering of midgut clusters revealed subregions with distinct marker genes **(Fig.S5A).** We then inspected the expression patterns of marker genes in each subcluster.

In adult *Drosophila* midgut, recent studies characterized the molecular specificity and cellular composition of midgut subregions by bulk and single-cell RNA-seq (scRNA-seq) (Buchon et al., 2013; Hung et al., 2020; Marianes and Spradling, 2013). However, although larval midgut was reported to

contain distinct domains in an *in vivo* reporter screen decades ago (Murakami R, Matsumoto A, Yamaoka I, Tanimura T.), its spatial gene expression profile was largely unknown. Here, our Stereo-seq data validated that the embryonic and larval midgut exhibits spatially distinct subregions, indicated by different marker gene expression patterns. Specifically, genes encoding digestive enzymes (including trypsin, amylase, and chymotrypsin) that were regional markers in adult midgut, were also expressed in specific subregions along the embryonic and larval midgut **(Fig.2)**. For instance, in a representative L3 section, the expression of serine protease genes *αTry*, *βTry*, *γTry*, *Jon25Bi* and *Jon25Bii* were localized to a midgut subregion different from that of serine protease gene *CG11911* **(Fig.2E)**. In a representative L2 section, the expression of serine protease genes *βTry*, *γTry* and *Jon25Bi* were localized to a midgut subregion different from that of starch digestive enzyme genes *Amy-p* and *Amy-d,* and serine protease gene *CG11911* **(Fig.2D)**. Similarly, in a representative L1 section, the set of digestive enzyme genes were also expressed in spatially distinct midgut clusters **(Fig.2C)**. Furthermore, as early as in E14-16 and E16-18 late-stage embryos, spatial expression patterns of digestive enzyme genes were already established **(Fig.2A-B)**, suggesting the maturation of presumptive midgut during embryonic organogenesis. To further reveal the regional diversity of embryonic midgut, we performed Gene Ontology (GO) enrichment analysis on marker genes of each midgut subcluster in E14-16. Marker genes of each subcluster were enriched for distinct functions **(Fig.S5B)**, suggesting that functional regionalization in the midgut already occurs as early as stage 16 of embryogenesis.

Together, our Stereo-seq data uncovered the dynamic spatiotemporal characteristics of functional genes along the digestive tract across *Drosophila* developmental and provided clues for midgut maturation timepoint during embryogenesis.

Spatiotemporal gene expression and cell state dynamics in reproductive and neural systems

Upon further examination of spatiotemporal features of embryonic and larval tissue transcriptomes, clusters representing male reproductive organs and central nervous system (CNS) stood out as two other regions of high cell type complexity and mRNA abundance. Thus, we extended subclustering and spatial gene expression analysis to these two tissues.

Similar to midgut clusters, we performed further subclustering in clusters representing L3 testis and adjacent clusters expression male germline related genes, and were able to identify clusters corresponding to spermatocytes at different stages of development, distinguished by their marker gene expression pattern (Zhao et al., 2010). To identify the subcluster type, we compared marker gene expression in the clusters representing male reproductive organs with previously published single-cell transcriptomes of L3 gonads subpopulations (Mahadevaraju et al., 2021). Unsupervised clustering enabled identification of spermatogonia and early, mid, and late primary spermatocyte

(Fig.3A) with distinct marker genes **(Fig.S6)**. Notably, cell type specific marker genes showed clear expression patterns within testis **(Fig.3B)**, reflecting the cell states of spermatocytes at different developmental stages.

In light of the effective capture of cell states in different stages of spermatocytes, we employed RNA velocity analysis (La Manno et al., 2018) to further investigate cell state dynamics. To perform RNA velocity analysis, unspliced mRNAs captured alongside with mature mRNAs by oligo dT primers through secondary priming were first identified. Their abundance was then compared with spliced mature mRNAs and their dynamics were used to infer the directions to which transcript splicing and cell states will change. With RNA velocity analysis, we observed the spatial transcriptional dynamics of spermatocyte maturation in testis **(Fig.3C)**. We then profiled the expression patterns of genes whose mRNAs display high velocity (i.e., undergo rapid transcriptional changes). We found several highly dynamic genes expressed in mid or late spermatocytes, such as Globin gene *glob2*. *glob2* has been reported to display an unexpected high level of expression in testis, possibly to relieve oxidative stress accumulated during spermatogenesis (Gleixner et al., 2012).

We then profiled the spatial expression patterns of genes involved in Hedgehog or JAK-STAT signaling, which regulate stem cells behavior in the *Drosophila* testis (reviewed in Zhang et al., 2013). We found that transcripts encoding some inhibitors of these pathways, including *Uba3* (Du et al., 2011), *Pp1-13C* (Su et al., 2011), and *EloB* (Stec et al., 2013), were mainly enriched in spermatocytes **(Fig.3D)**, possibly to inhibit the activity of stem cell regulatory pathways in the differentiated spermatocytes.

We also inspected the spatial expression patterns of marker genes in cell subtypes in CNS. The spatial expression patterns of neuronal and glial cell marker genes largely overlap, with glial cell markers showing broader expression patterns, consistent with the spatial distribution of their anatomical structures **(Fig.S7)**. Moreover, we looked at the spatial expression of marker genes of peptidergic and dopaminergic neurons in L2 and found that their spatial patterns can be clearly distinguished **(Fig.S7)**.

To summarize, the high sensitivity of Stereo-seq enabled delicate subclustering of complex tissues to identify cell subtypes in their morphological background. Spatial RNA velocity analysis could provide evidence for activation timing of genes with spatiotemporally restricted regulatory functions.

Spatial gene regulatory networks during *Drosophila* development

One major driving force of cellular heterogeneity during development is the changes in gene regulatory networks (GRNs), mediated by the interactions of transcription factors (TFs), co-factors and their downstream target genes. Spatial transcriptomic data from Stereo-seq provides a unique

opportunity to analyze GRNs in their morphological contexts, allowing GRN identification and analysis in a spatially resolved tissue-specific manner.

To profile the spatial GRNs across *Drosophila* development, we employed single-cell regulatory network inference and clustering (SCENIC) protocol, which were developed to reconstruct GRNs and identify cell states from scRNA-seq data (Aibar et al., 2017). To find regulons of a TF of interest, SCENIC first identifies genes co-expressed with the TF, then enriches for direct-binding downstream target genes with *cis*-regulatory motif analysis. Finally, SCENIC scores the activity of regulons in each cell to identify the cells with high sub-network activity.

SCENIC analysis of Stereo-seq spatial transcriptomic data revealed GRNs with spatial distribution patterns. Across embryonic and larval samples, we identified multiple GRNs mediated by previously reported TFs. Some of the GRNs are tissue-specific, and their corresponding TFs are known to play regulatory roles in the tissues where the GRNs are located, including *GATAe* in midgut (Okumura et al., 2005, 2016), *Pdp1* in muscle (Lin et al., 1997), and *l(3)neo38* (Schneiderman et al., 2010), *kn* (Seecoomar et al., 2000) and *jim* (Iyer et al., 2013) in CNS **(Fig.4)**. Other GRNs are not restricted to certain tissues, but display similar patterns over development, such as *ham*, *srp*, *grh* and *Hr78* **(Fig.4)**. The spatial GRN patterns of all these TFs correlate with those of late-stage embryo *in situ* hybridization results **(Fig.4)**, suggesting that SCENIC analysis of Stereo-seq data effectively recapitulated the spatial GRN patterns of these TFs. Among the identified TFs, *ham* was known for its regulatory role in central and peripheral nervous system, and external sensory neurons (Moore et al., 2002, 2004). Some of the examined slices in which *ham* GRNs were identified did not contain CNS, but *ham* GRNs consistently mapped to midgut in late-stage embryos and larvae **(Fig.4)**. It is known that in larvae, *ham* also shows expression enrichment in midgut in addition to CNS (Leader et al., 2018). This suggest that *ham* may play additional regulatory roles in the digestive system.

Thus, SCENIC analysis of Stereo-seq data from developing *Drosophila* identifies both known and potential GRNs, indicating that Stereo-seq can effectively capture transcription regulation networks in the spatial contexts of TFs.

3D reconstruction of spatial transcriptome of *Drosophila* late-stage embryos

With the efficient and faithful capture of 2D spatial transcriptome of *Drosophila* embryos by Stereo-seq, we wondered if it is possible to 3D reconstruct the spatial transcriptome of the whole embryo. As a proof of principle, we collected all the 7-μm thick cryosection slices (14 slices obtained) of one entire E16-18 embryo sectioned along the left-right axis (so that each section is a sagittal plane) and performed Stereo-seq with them. We then combined all 2D expression matrices of these samples, performed clustering and annotation individually in each slice **(Fig.S8)**. We developed a pipeline for 3D reconstruction (see **Methods**). Briefly, we extracted 2D sample regions from visualized

Stereo-seq expression matrices and aligned them based on shape and transcriptome similarities, so that each bin was assigned an x-y-z 3D coordinate. We then minimized batch effects of each section and merged identical clusters in adjacent slices. The contour and position of 3D clusters resembled the anatomical structures of embryos **(Fig.5A and Supplemental Data 1)**. The marker genes of each tissue displayed expected spatial distribution around their corresponding tissues **(Fig.5B)**. In the 3D reconstructed model, various tissues, represented by clusters aligned across sections, display intra-tissue structural continuity as well as inter-tissue gene expression heterogeneity **(Fig.5C and Supplemental Data 2)**. Moreover, the 3D clusters representing complex tissues such as foregut and midgut can also be further divided into detailed structures based on subclustering and marker gene identification **(Fig.5D and Supplemental Data 3)**.

In brief, Stereo-seq is capable of efficient recapitulation of the 3D spatial transcriptome of *Drosophila* embryos. With future optimization of Stereo-seq, we aim to establish an organism-wide, high-resolution, and high-sensitivity 3D spatiotemporal transcriptomic atlas of all stages of *Drosophila* embryos and larvae. We anticipate that with modification, the 3D transcriptome reconstruction strategies we developed here can be applied to additional *Drosophila* adult tissues to resolve their spatial transcriptomes as well.

DISCUSSION

Spatial transcriptomics techniques greatly expanded our knowledge of gene expression within topographical contexts. In *Drosophila*, existing databases of spatial gene expression patterns were mostly generated from results of *in situ* hybridization analyses. However, *in situ* hybridization techniques are highly limited by probe design strategy and cannot be easily multiplexed. Thus, the gene expression patterns covered by *in situ* hybridization databases are biased. Consequently, previous spatial mapping of *Drosophila* transcriptome and regulatory network is largely based on *in situ* hybridization data and/or computation, which lack the capability for unbiased and global gene expression profiling within intact spatial background.

We present spatiotemporal transcriptomic maps in *Drosophila* from late-stage embryos to 3rd instar larvae based on our Stereo-seq platform. With its unique combination of DNA nanoball patterned arrays and unbiased *in situ* mRNA capture, Stereo-seq provides significantly enhanced resolution and sensitivity compared to other spatial transcriptomics techniques, allowing its application to small-sized tissue sections such as *Drosophila* embryos to simultaneously resolve their spatial transcriptome and morphology. The highly customizable sizes of Stereo-seq chips enable simultaneous mapping of multiple tissue sections, avoiding batch effects introduced by separate experimental runs.

From embryos to larvae, *Drosophila* tissues undergo fundamental diversification in morphology and functions as they develop, which are dictated by spatial (within or between tissue primordia) and temporal (across developmental time) heterogeneity in gene expression profiles. With spatiotemporally resolved gene expression patterns generated by Stereo-seq, expression profiles of tissues across embryonic and larval stages can be dissected in both dimensions to comprehensively reveal their changes during development. Recently, the rapid development of scRNA-seq techniques and the joint efforts of *Drosophila* research community have discovered many new and unanticipated cell types across various tissues during *Drosophila* development (Li et al., 2021). Our Stereo-seq platform can help quickly establish the positions of these newly discovered cells *in vivo*, without the need for *in situ* hybridization with multiple probes. With its high resolution and sensitivity, Stereo-seq itself may also lead to the discovery of new cell types within the spatial contexts in tissues of interest.

We show that Stereo-seq data from *Drosophila* embryos and larvae effectively reflected their tissue anatomical structure and spatial heterogeneity of gene expression. The developmental transcriptomic maps we generated complements current *in situ* hybridization data with quantitative spatial information of both known and previously undetected transcripts. Additionally, the high-resolution and high-sensitivity spatiotemporal transcriptomic data allowed RNA velocity and SCENIC analysis with actual instead of computed spatial background, uncovering the dynamics of known and potential spatially defined tissue-specific gene regulatory networks.

Our results demonstrated the capability of Stereo-seq in resolving the spatial transcriptomes of small-sized samples like *Drosophila* embryos. To our best knowledge, the Stereo-seq data we generated here produces the first actual organism level spatial transcriptomic maps of *Drosophila* late-stage embryos and larvae, which can be combined with scRNA-seq data and provide valuable insights for the systematic study of tissue formation paradigms and regulatory network changes during development. With future optimization of Stereo-seq, we aim to expand the current transcriptomic atlas to generate complete organism-wide 3D spatial transcriptomic profiles covering the entire lifespan of developing *Drosophila*. Application of Stereo-seq to additional genotypes and stages of *Drosophila* will establish an encyclopedic spatial transcriptomic database that will be of great interest to the *Drosophila* research community.

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.

Ariss, M.M., Islam, A.B., Critcher, M., Zappia, M.P., and Frolov, M. V (2018). Single cell

RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. Nat. Commun. *9*, 1–13.

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. Cell *179*, 1647–1660.

Bageritz, J., Willnow, P., Valentini, E., Leible, S., Boutros, M., and Teleman, A.A. (2019). Gene expression atlas of a developing tissue by single cell expression correlation analysis. Nat. Methods *16*, 750–756.

Brunet Avalos, C., Maier, G.L., Bruggmann, R., and Sprecher, S.G. (2019). Single cell transcriptome atlas of the *Drosophila* larval brain. Elife *8*, e50354.

Buchon, N., Osman, D., David, F.P., Fang, H.Y., Boquete, J.P., Deplancke, B., and Lemaitre, B. (2013). Morphological and molecular characterization of adult midgut compartmentalization in *Drosophila*. Cell Rep *3*, 1725–1738.

Campos-Ortega, J.A., and Hartenstein, V. (1997). The Embryonic Development of *Drosophila melanogaster* (Springer-Verlag Berlin).

Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Yang, J., Li, W., Xu, J., Hao, S., et al. (2021). Large field of view-spatially resolved transcriptomics at nanoscale resolution. BioRxiv 2021.01.17.427004.

Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I., et al. (2020). Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell *182*, 976–991.

Chintapalli, V.R., Wang, J., and Dow, J.A.T. (2007). Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. *39*, 715–720.

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., et al. (2018). A Single-Cell Transcriptome Atlas of the Aging *Drosophila* Brain. Cell *174*, 982–998.

Deng, M., Wang, Y., Zhang, L., Yang, Y., Huang, S., Wang, J., Ge, H., Ishibashi, T., and Yan, Y. (2019). Single cell transcriptomic landscapes of pattern formation, proliferation and growth in *Drosophila* wing imaginal discs. Development *146*, dev179754.

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.

Du, J., Zhang, J., Su, Y., Liu, M., Ospina, J.K., Yang, S., and Zhu, A.J. (2011). In vivo RNAi screen reveals neddylation genes as novel regulators of Hedgehog signaling. PLoS One *6*, e24168.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., and Sonka, M. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn. Reson. Imaging *30*, 1323–1341.

Gleixner, E., Herlyn, H., Zimmerling, S., Burmester, T., and Hankeln, T. (2012). Testes-specific hemoglobins in *Drosophila*evolved by a combination of sub- and neofunctionalization after gene duplication. BMC Evol. Biol. *12*, 1–12.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. Nature *471*, 473–479.

Guo, X., Yin, C., Yang, F., Zhang, Y., Huang, H., Wang, J., Deng, B., Cai, T., Rao, Y., and Xi, R. (2019). The cellular diversity and transcription factor code of *Drosophila* enteroendocrine cells. Cell Rep. *29*, 4172–4185.

Hung, R.-J., Hu, Y., Kirchner, R., Liu, Y., Xu, C., Comjean, A., Tattikota, S.G., Li, F., Song, W., Ho Sui, S., et al. (2020). A cell atlas of the adult *Drosophila* midgut. Proc. Natl. Acad. Sci. *117*, 1514–1523.

Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. *9*, 90–95.

Iyer, E.P.R., Iyer, S.C., Sullivan, L., Wang, D., Meduri, R., Graybeal, L.L., and Cox, D.N. (2013). Functional genomic analyses of two morphologically distinct classes of *Drosophila* sensory neurons: post-mitotic roles of transcription factors in dendritic patterning. PLoS One *8*, e72434.

Jevitt, A., Chatterjee, D., Xie, G., Wang, X.-F., Otwell, T., Huang, Y.-C., and Deng, W.-M. (2020). A single-cell atlas of adult *Drosophila* ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. PLoS Biol. *18*, e3000538.

Junker, J.P., Noël, E.S., Guryev, V., Peterson, K.A., Shah, G., Huisken, J., McMahon, A.P., Berezikov, E., Bakkers, J., and van Oudenaarden, A. (2014). Genome-wide RNA Tomography in the Zebrafish Embryo. Cell *159*, 662–675.

Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., Kocks, C., Rajewsky, N., and Zinzen, R.P. (2017). The *Drosophila* embryo at single-cell transcriptome resolution. Science. *358*, 194–199.

Kibirige, H. (2017). plotnine: a grammar of graphics for Python. Graph. Softw.

Leader, D.P., Krause, S.A., Pandit, A., Davies, S.A., and Dow, J.A.T. (2018). FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. Nucleic Acids Res. *46*, D809–D815.

Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell *131*, 174–187.

Li, H. (2020). Single-cell RNA sequencing in *Drosophila*: Technologies and applications. WIREs Dev. Biol. e396.

Li, H., Janssens, J., De Waegeneer, M., Kolluru, S.S., Davie, K., Gardeux, V., Saelens, W., David, F., Brbić, M., Leskovec, J., et al. (2021). Fly Cell Atlas: a single-cell transcriptomic atlas of the adult fruit

fly. BioRxiv 2021.07.04.451050.

Liao, J., Lu, X., Shao, X., Zhu, L., and Fan, X. (2021). Uncovering an Organ's Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. Trends Biotechnol. *39*, 43–58.

Lin, S.C., Lin, M.H., Horvath, P., Reddy, K.L., and Storti, R. V (1997). PDP1, a novel *Drosophila* PAR domain bZIP transcription factor expressed in developing mesoderm, endoderm and ectoderm, is a transcriptional regulator of somatic muscle genes. Development *124*, 4685–4696.

Mahadevaraju, S., Fear, J.M., Akeju, M., Galletta, B.J., Pinheiro, M.M.L.S., Avelino, C.C., Cabral-de-Mello, D.C., Conlon, K., Dell'Orso, S., Demere, Z., et al. (2021). Dynamic sex chromosome expression in *Drosophila* male germ cells. Nat. Commun. *12*, 892.

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.

Marianes, A., and Spradling, A.C. (2013). Physiological and stem cell compartmentalization within the *Drosophila* midgut. Elife *2*, e00886.

Moore, A.W., Jan, L.Y., and Jan, Y.N. (2002). hamlet, a Binary Genetic Switch Between Single- and Multiple- Dendrite Neuron Morphology. Science (80-.). *297*, 1355–1358.

Moore, A.W., Roegiers, F., Jan, L.Y., and Jan, Y.-N. (2004). Conversion of neurons and glia to external-cell fates in the external sensory organs of *Drosophila* hamlet mutants by a cousin-cousin cell-type respecification. Genes Dev. *18*, 623–628.

Murakami, R., Shigenaga, A., Kawano, E., Matsumoto, A., Yamaoka, I., and Tanimura, T. (1994). Novel tissue units of regional differentiation in the gut epithelium of Drosopbila, as revealed by P-element-mediated detection of enhancer. Roux's Arch. Dev. Biol. *203*, 243–249.

Okumura, T., Matsumoto, A., Tanimura, T., and Murakami, R. (2005). An endoderm-specific GATA factor gene, dGATAe, is required for the terminal differentiation of the *Drosophila* endoderm. Dev. Biol. *278*, 576–586.

Okumura, T., Takeda, K., Kuchiki, M., Akaishi, M., Taniguchi, K., and Adachi-Yamada, T. (2016). GATAe regulates intestinal stem cell maintenance and differentiation in *Drosophila* adult midgut. Dev. Biol. *410*, 24–35.

Ortiz, C., Navarro, J.F., Jurek, A., Märtin, A., Lundeberg, J., and Meletis, K. (2020). Molecular atlas of the adult mouse brain. Sci. Adv. *6*, eabb3446.

Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2021). Squidpy: a scalable framework for spatial single cell analysis. BioRxiv 2021.02.19.431994.

Qiu, X., Zhang, Y., Hosseinzadeh, S., Yang, D., Pogson, A.N., Wang, L., Shurtleff, M., Yuan, R., Xu,

S., Ma, Y., et al. (2021). Mapping Transcriptomic Vector Fields of Single Cells. BioRxiv 696724.

Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science (80-.). *363*, 1463–1467.

Rust, K., Byrnes, L.E., Yu, K.S., Park, J.S., Sneddon, J.B., Tward, A.D., and Nystul, T.G. (2020). A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. Nat. Commun. *11*, 1–17.

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. *15*, 2247–2276.

Schneiderman, J.I., Goldstein, S., and Ahmad, K. (2010). Perturbation analysis of heterochromatin-mediated gene silencing and somatic inheritance. PLoS Genet *6*, e1001095.

Seecoomar, M., Agarwal, S., Vani, K., Yang, G., and Mohler, J. (2000). knot is required for the hypopharyngeal lobe and its derivatives in the *Drosophila* embryo. Mech. Dev. *91*, 209–215.

Sievert, C. (2020). Interactive web-based data visualization with R, plotly, and shiny (CRC Press).

Stec, W., Vidal, O., and Zeidler, M.P. (2013). *Drosophila* SOCS36E negatively regulates JAK/STAT pathway signaling via two separable mechanisms. Mol. Biol. Cell *24*, 3000–3009.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902.

Su, Y., Ospina, J.K., Zhang, J., Michelson, A.P., Schoen, A.M., and Zhu, A.J. (2011). Sequential Phosphorylation of Smoothened Transduces Graded Hedgehog Signaling. Sci. Signal. *4*, ra43.

Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E., et al. (2002). Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. Genome Biol *3*, 1–14.

Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2007). Global analysis of patterns of gene expression during *Drosophila* embryogenesis. Genome Biol. *8*, 1–24.

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

Wickham, H. (2011). ggplot2. Wiley Interdiscip. Rev. Comput. Stat. *3*, 180–185.

Witt, E., Benjamin, S., Svetec, N., and Zhao, L. (2019). Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational bias in *Drosophila*. Elife *8*, e47138.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 1–5.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing

biological themes among gene clusters. Omi. a J. Integr. Biol. *16*, 284–287.

Zeira, R., Land, M., and Raphael, B.J. (2021). Alignment and Integration of Spatial Transcriptomics Data. BioRxiv 2021.03.16.435604.

Zhang, Z., Pan, C., and Zhao, Y. (2013). Hedgehog in the *Drosophila* testis niche: what does it do there? Protein Cell *4*, 650–655.

Zhao, J., Klyne, G., Benson, E., Gudmannsdottir, E., White-Cooper, H., and Shotton, D. (2010). FlyTED: the *Drosophila* Testis Gene Expression Database. Nucleic Acids Res *38*, D710–D715.

Zhou, J., Schor, I.E., Yao, V., Theesfeld, C.L., Marco-Ferreres, R., Tadych, A., Furlong, E.E.M., and Troyanskaya, O.G. (2019). Accurate genome-wide predictions of spatio-temporal gene expression during embryonic development. PLoS Genet. *15*, e1008382.

ACKNOWLEDGEMENTS

This work was supported by the Shenzhen Key Laboratory of Gene Regulation and Systems Biology (Grant No. ZDSYS20200811144002008, China) (to W.C. and Y.H.), the Shenzhen Science and Technology Program (Grant No. KQTD20180411143432337, China) (to W.C. and Y.H.), Shenzhen Key Laboratory of Single-Cell Omics (Grant No. ZDSYS20190902093613831, China) (to L.L.) and Guangdong Provincial Key Laboratory of Genome Read and Write (Grant No. 2017B030301011, China) (to H.L.). This work was also supported by China National GeneBank (CNGB). We thank Dr. Mariana Wolfner for helpful comments on the manuscript. Fig.1A was created with Biorender.com.

AUTHOR CONTRIBUTIONS

M.W., Q.H. and X.W. conceived the idea; Y.L., X.X., W.C., Y.H. and L.L. supervised the work; M.W., Q.H. and Q.L. designed the experiment; M.W., Q.H., Q.L., Y.W. and Y.A. performed the experiments; T.L., Y.W., Z.T., R.X., K.H. and X.W. processed and analyzed the data; M.C., J.X. and A.C. helped with Stereo-seq library preparation; H.L., W.L. and S.Z. helped with sequencing; Q.H. and Q.L. wrote the manuscript; all authors read and edited the manuscript.

DECLARATION OF INTERESTS

The chip, procedure, and application of Stereo-seq are covered in pending patents. Employees of BGI have stock holdings in BGI.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for the resources and reagents may be directed to the corresponding author Longqi Liu (liulongqi@genomics.cn).

Material availability

All materials used for Stereo-seq are commercially available.

Data and code availability

All raw data generated by Stereo-seq and custom codes using open-source software to support this study are provided in supplementary materials (**Supplemental Data 4**) or deposited to CNGB Nucleotide Sequence Archive database with accession code: CNP0002189 (https://db.cngb.org/search/project/CNP0002189). All data were analyzed with standard programs and packages, as detailed in **Methods**.

METHODS

Fly strain maintenance

All embryo and larva samples in this study were from *Drosophila* strain *w1118* (Tsinghua Fly Center). Flies were maintained on cornmeal-sucrose-agar media (Hopebio, HB8590) in a 25 °C incubator (Laifu, PGX-280A-3H) on a 12 h/12 h light/dark cycle.

Sample preparation

Embryos were collected at two-hour intervals on a grape juice plate [2.15% w/v agar (Vetec, V900500), 49% v/v grape juice, 0.2% v/v propionic acid (Aladdin, P110444), 0.02% phosphate acid (LingFeng, Shanghai)] from a population cage and aged to desired stages. Larvae of desired stages were isolated from the same population cage. Samples were incubated in phosphate buffer saline (PBS) containing 0.5 mg/mL bromophenol blue (Macklin, B802656) for 10 min for staining and better visualization during cryosection. Samples were then rinsed in PBS to remove excessive dye before embedding. To compare the efficiency of mRNA capture and 3D reconstruction, we attempted to cut the samples in two directions (along the left-right axis for sagittal section, or along the dorsal-ventral axis for transverse section) during cryosection. The exact orientations of each sample used in this study were indicated in the figure legends. Samples were oriented so that they were most likely to be sectioned in the desired axis. Orientated samples were immobilized with double-sided tapes to prevent disturbance from flowing embedding media. 6 samples of the same stage (2 samples for L2

549 and L3) were embedded and sectioned together. Sample were embedded with Tissue-Teck OCT
550 (Sakura, 4583) and transferred to a −80 °C freezer for storage until used. Cryosection was performed
551 with indicated slice thickness in a Leica CM1950 cryostat. Sample sections were applied to Stereo-seq
552 chips immediately after cryosection.
553
554 **Stereo-seq**
555 Stereo-seq library preparation and sequencing were performed as previously described (Chen et
556 al., 2021). Briefly, embryo and larva sections on Stereo-seq chips were fixed in pre-chilled methanol
557 at -20 °C for 40 min. After removal of methanol, sections were permeabilized on chip with 100 µl
558 0.1% pepsin (Sigma, P7000) in 0.01 M HCl at 37 °C for 5 min. Permeabilization solution was then
559 removed and sections were washed with 100 µl 0.1× saline-sodium citrate (SSC, Thermo, AM9770)
560 buffer supplemented with 0.05 U/µl RNAase Inhibitor (MGI). mRNAs captured by DNBs on the chip
561 were reverse transcribed with SuperScript II reverse transcription (RT) mix (Invitrogen, 18064-014, 10
562 U/µL reverse transcriptase, 1 mM dNTPs, 1 M betaine solution PCR reagent, 7.5 mM $MgCl_2$, 5 mM DTT,
563 2 U/µL RNase inhibitor, 2.5 µM Stereo-seq template switch oligo and 1× First-Strand buffer) at 42 °C
564 for 1 h. RT mix was then removed. The chip was washed with 0.1× SSC and incubated in tissue
565 removal buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 37 °C for 30 min. Tissue
566 removal buffer was removed and the chip was washed twice with 0.1× SSC. cDNAs on the chip were
567 amplified with KAPA HiFi Hotstart Ready Mix (Roche, KK2602) with 0.8 µM cDNA-PCR primer.
568 Sequencing libraries were prepared with PCR products undergoing the following steps: fragmentation
569 (in-house Tn5 transposase), amplification (KAPA HiFi Hotstart Ready Mix), and purification (Vazyme,
570 N411-03). Final libraries were sequenced on a MGI DNBSEQ-Tx sequencer.
571
572 **Stereo-seq data analysis**
573 **Raw data processing.** Stereo-seq raw data processing and unsupervised clustering were
574 performed as previously described (Chen et al., 2021). Briefly, CID sequences were first mapped to
575 the designed coordinates on chip with 1 base mismatch tolerance. MID sequences with quality score
576 lower than 10 were filtered out. cDNA sequences were aligned to the reference genome (*Dm3*) by
577 STAR (Dobin et al., 2013). Expression profile matrices with CID were generated based on the
578 information above.
579 **Unsupervised clustering.** The expression profile matrices of sections from embryos and L1/L2
580 were divided into bins with 20 × 20 DNBs, while those from L3 were divided into bins with 50 x 50
581 DNBs. After obtaining the expression profile matrices through the in-house processing software, we
582 used *Seurat* package to perform data normalization and unsupervised clustering (Stuart et al., 2019).
583 Briefly, *SCTransform* function was first applied to normalize and identify highly variable genes. After

584 dimensionality reduction with *runPCA* function, *runUMAP* function was used to perform a
585 two-dimensional projection on the data with the following parameter: *dims = 6*. The identification of
586 each cluster was then performed with *FindNeighbors*, *FindClusters*, and *FindAllMarkers* functions in
587 *Seurat* with the following parameters: *dims = 6, verbose = FALSE, resolution = 0.6, only.pos = TRUE,*
588 *min.pct = 0.25, logfc.threshold = 0.25*.

589 **Cluster annotation.** For the list of marker genes of each cluster in the unsupervised clustering
590 results, we extracted the top 30 genes with the most significant *p* values. To infer specific tissue types
591 of each cluster, we queried these 50 genes in the publicly available databases, including BDGP,
592 FlyBase, FlyAtlas1, FlyAtlas2 and FlyMine, as well as other published results. We assigned the tissue
593 type(s) that match the most marker genes of the cluster.

594

595 **RNA velocity analysis**

596 Spliced and unspliced RNA were counted for all detected genes (~400 genes per bin in the male
597 reproductive organ clusters) by every bin 50 × 50 using Velocyto command line interface (La Manno
598 et al., 2018), with about 1.3% of genes per bin were detected as unspliced type. The generated splice
599 and unspliced expression matrices were then used to estimate the differentiation dynamics of cell
600 lineages with dynamo software (Qiu et al., 2021), the transitioning between gene expression states
601 was modeled and inferred, with lineage relationship of early primary spermatocyte to mid primary
602 spermatocyte set as confident cell velocities, to determine the expression dynamics of testis cell.
603 Genes with high correlation with the transition were also determined in the same process.

604

605 **SCENIC analysis**

606 *pySCENIC* pipeline (Van de Sande et al., 2020) was used to predict activity scores of TFs on each
607 section. *pySCENIC* uses TFs' motif and downstream gene expression to predict TFs' activity scores at
608 single cell level. *pySCENIC* pipeline was first used to generate AUC matrices of TFs, with rows
609 representing bins and columns representing TFs. Values in each cell represent TFs' activity scores.
610 AUC values in each section were then visualized with *ggplot2* (Wickham, 2011).

611

612 **Gene Ontology enrichment analysis**

613 R package *clusterProfiler* (Yu et al., 2012) was used to identify enriched Gene Ontology terms.
614 Marker gene lists from each cluster were inputted with default parameters.

615

616 **Imputation of spatial gene expression**

617 For better visualization of spatial gene expression patterns, *SparseVFC* function of dynamo
618 software was used to impute gene expression by 2D or 3D coordinates to render spatial patterns of

gene expression. This method was used in **Fig. 2**, **Fig.3C** and **Fig.S7**.

3D Reconstruction of embryo transcriptome

Stereo-seq data preprocessing. 2D expression matrices of all 14 sections of an entire E16-18 embryo were subjected to preprocessing. Specifically, for each section, Python package *scanpy* (Wolf et al., 2018) was used to normalize expression values for total UMI counts per bin 20 × 20. *calculate_qc_metrics* function was used for data quality control, with the following non-default parameters: *percent_top = None, log1p = False, inplace = True*. Highly variable genes were identified by function *highly_variable_genes* with parameters: *flavor = seurat, n_top_genes = 2000, inplace=True*. Other functions in preprocessing used standard methods and default processes of *scanpy*, including total-count normalization and principal component analysis.

Unsupervised clustering and cluster annotation. For each section, PCA algorithm was used for data dimensionality reduction with parameter *n_comps = 20*. Python package *squidpy* (Palla et al., 2021) algorithm was used for unsupervised clustering. Specifically, after using *neighbors* function in *scanpy* to find the neighbors of each bin 20 × 20 based on gene expression information with parameter *n_neighbors = 30*, *spatial_neighbors* function in *squidpy* was used to find the spatial neighbors of each spot with parameter *n_neigh = 6*. The neighbor information obtained through the two algorithms were then merged. Based on the merged neighbor information, *leiden* algorithm in *scanpy* was used for unsupervised clustering. After individual testing, optimal resolution was determined for each section. Clusters in the optimal resolution of each section were annotated as described above.

Section alignment and 3D modeling. For two adjacent sections, PASTE (Zeira et al., 2021) algorithm was used to align them along the z axis based on both gene expression similarities and spatial coordinates. For all the 2D expression matrices of 14 sections, *pairwise_align* function in PASTE was run sequentially along the z axis, and each bin was assigned an x-y-z 3D coordinate. After that, on 2D graphics of all sections, bins with the same annotation were assigned the same color code. The aligned 3D coordinates and color codes representing different tissues were integrated into 3D graphics in tiff format with *skimage* (Van der Walt et al., 2014) algorithm. *3D Slicer* (Fedorov et al., 2012) software was used to transform tiff data into a smooth 3D model of the embryo. Major smoothing and 3D modeling methods used were *Margin*, *Closing*, *Opening*, *Median*, *Gaussian*, and *Joint smoothing*. Spatial distribution of 3D clusters and gene expression patterns were visualized with *scanpy* and *plotly* (Sievert, 2020). Python tools *matplotlib* (Hunter, 2007) and *plotnine* (Kibirige, 2017) were used to draw 2D graphics, and *plotly* was used to display 3D graphics and models.

FIGURE LEGENDS

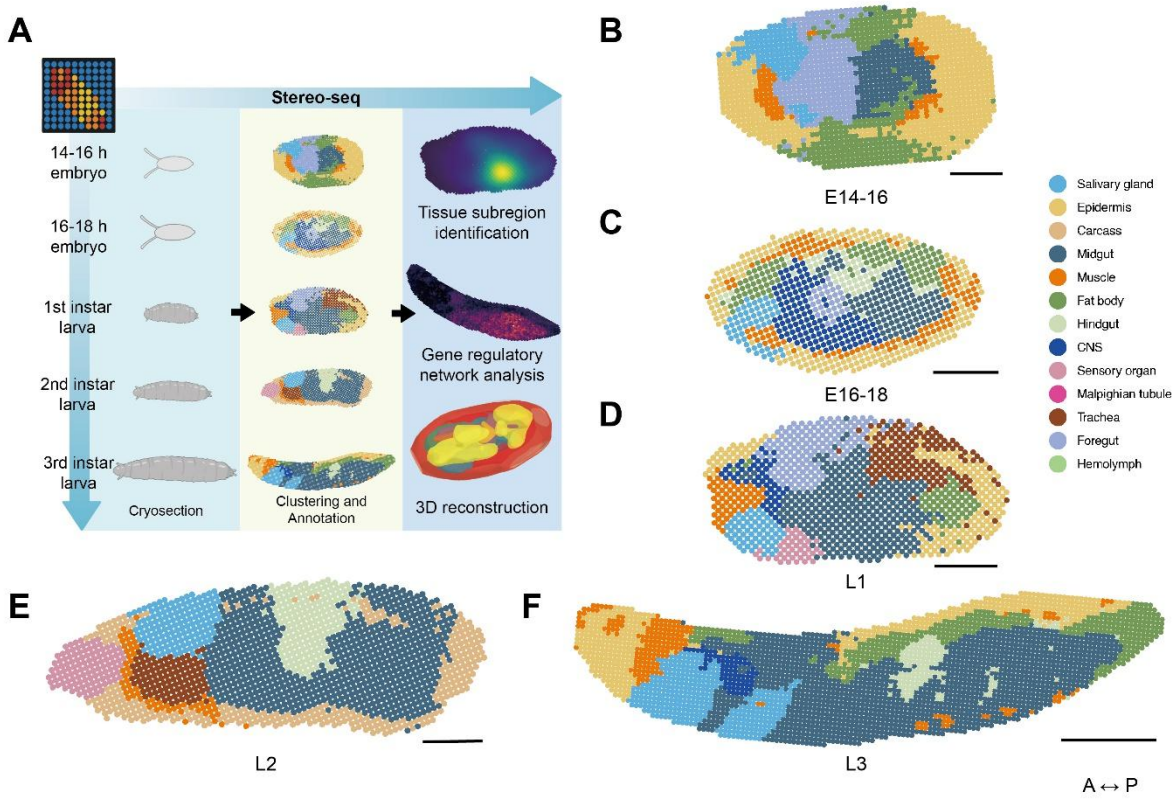
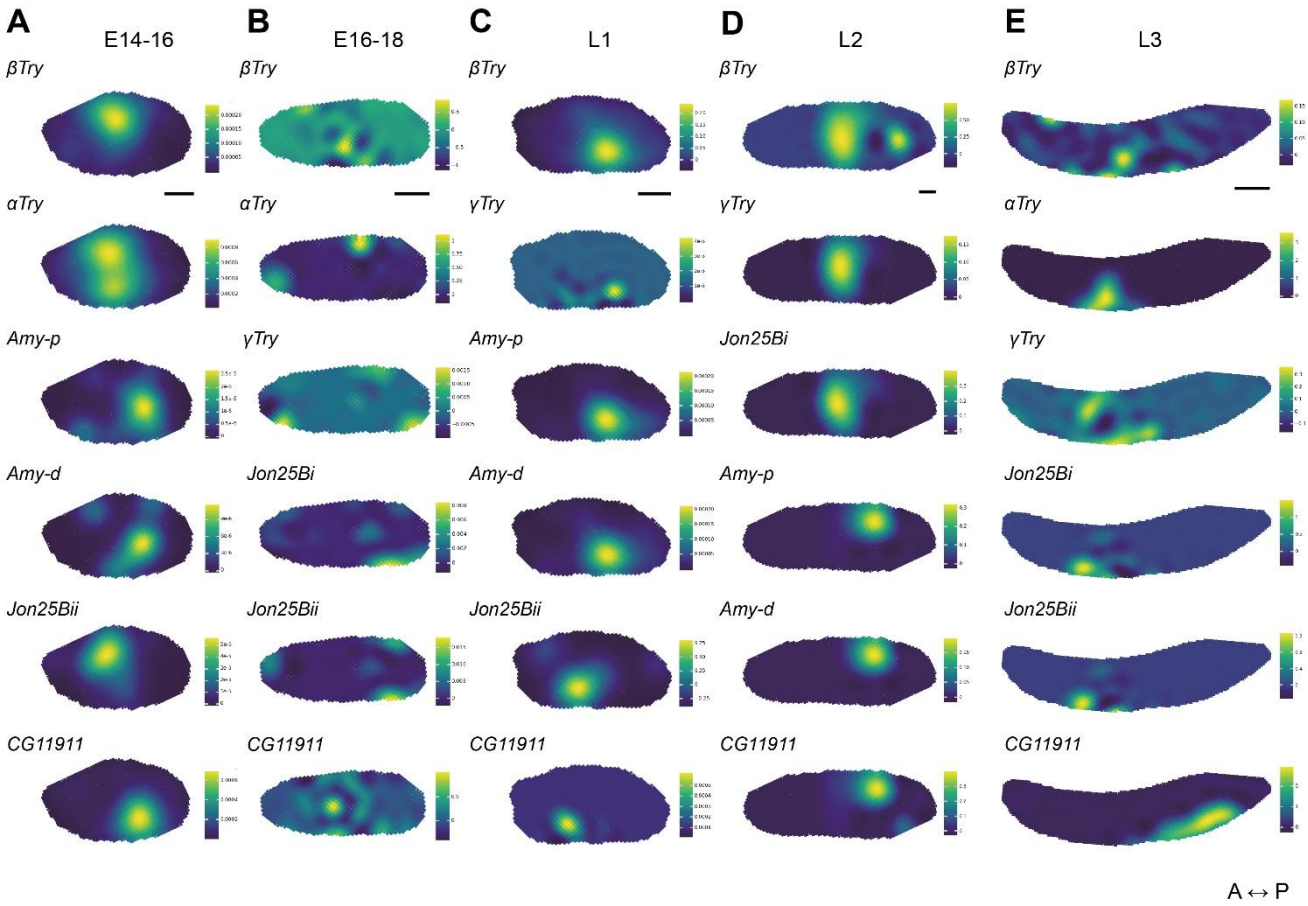


Figure 1 Clustering and annotation of Stereo-seq generated spatial transcriptomes of *Drosophila* late-stage embryos and larvae. (A) Schematic workflow of Stereo-seq data acquisition and analysis in this study. **(B-F)** Unsupervised clustering and annotation of representative single sections of **(B)** E14–16, **(C)** E16-18, **(D)** L1, **(E)** L2, and **(F)** L3 samples. Clusters with the same annotation are merged and assigned the same color code. Scale bars = 100 μm for E14-16, E16-18, L1 and L2, and 500 μm for L3. A-P: anterior-posterior.



660 **Figure 2 Spatial heterogeneity of digestive enzyme gene expression in midgut across**
661 ***Drosophila* development.** Spatial expression patterns of genes encoding digestive enzymes (trypsin,
662 amylase, and chymotrypsin) in clusters annotated as midgut in representative single sections of **(A)**
663 E14-16, **(B)** E16-18, **(C)** L1, **(D)** L2, and **(E)** L3. All samples are presented in near dorsal/ventral view.
664 Scale bars = 100 µm for E14-16, E16-18, L1 and L2, 500 µm for L3.

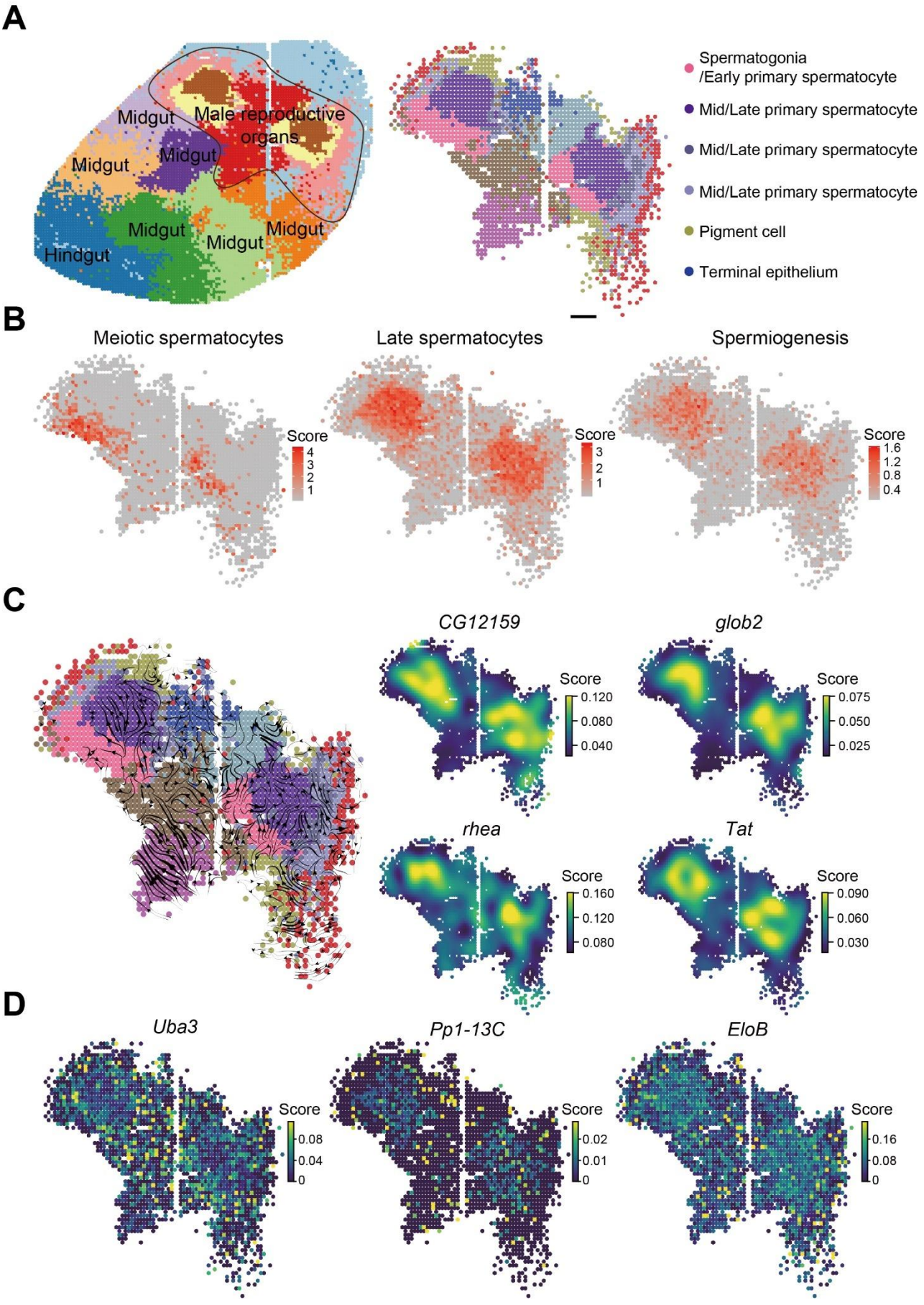


Figure 3 Cell type subclustering and cell state change tracking of *Drosophila* 3rd instar larval male reproductive organs. (A) Subclustering results of clusters annotated as testis and adjacent clusters expression male germline related genes in a representative L3 single slice. Scale bar = 100 µm. **(B)** Spatial expression patterns of marker gene sets representing cells at different stages of spermatogenesis (meiotic spermatocytes: *Rbp4*, *Taf12*, *blanks*; late spermatocytes: *tHMG1*, *CG2955*, *CG13476*, *CG4375*, *Pen*, *SAK*, *lobo*, *CG16719*; spermiogenesis: *p-cup*, *Pif2*, *whip*, *CG4073*, *CG31128*, *boly*, *CG14294*, *hale*, *Glut3*, *c-cup*, *CG12126*, *CG15059*, *CG6333*, *CG31797*, *CG30099*, *CG31404*, *CG13337*, *CG31815*, *CG10512*, *CG14546*, *CG8851*, *CG30110*) (Zhao et al., 2010). **(C)** RNA velocity analysis of clusters representing male reproductive organs and expression patterns of representative genes with high velocity. **(D)** Spatial expression patterns of genes encoding inhibitors of Hedgehog and JAK-STAT pathways in male reproductive organs.

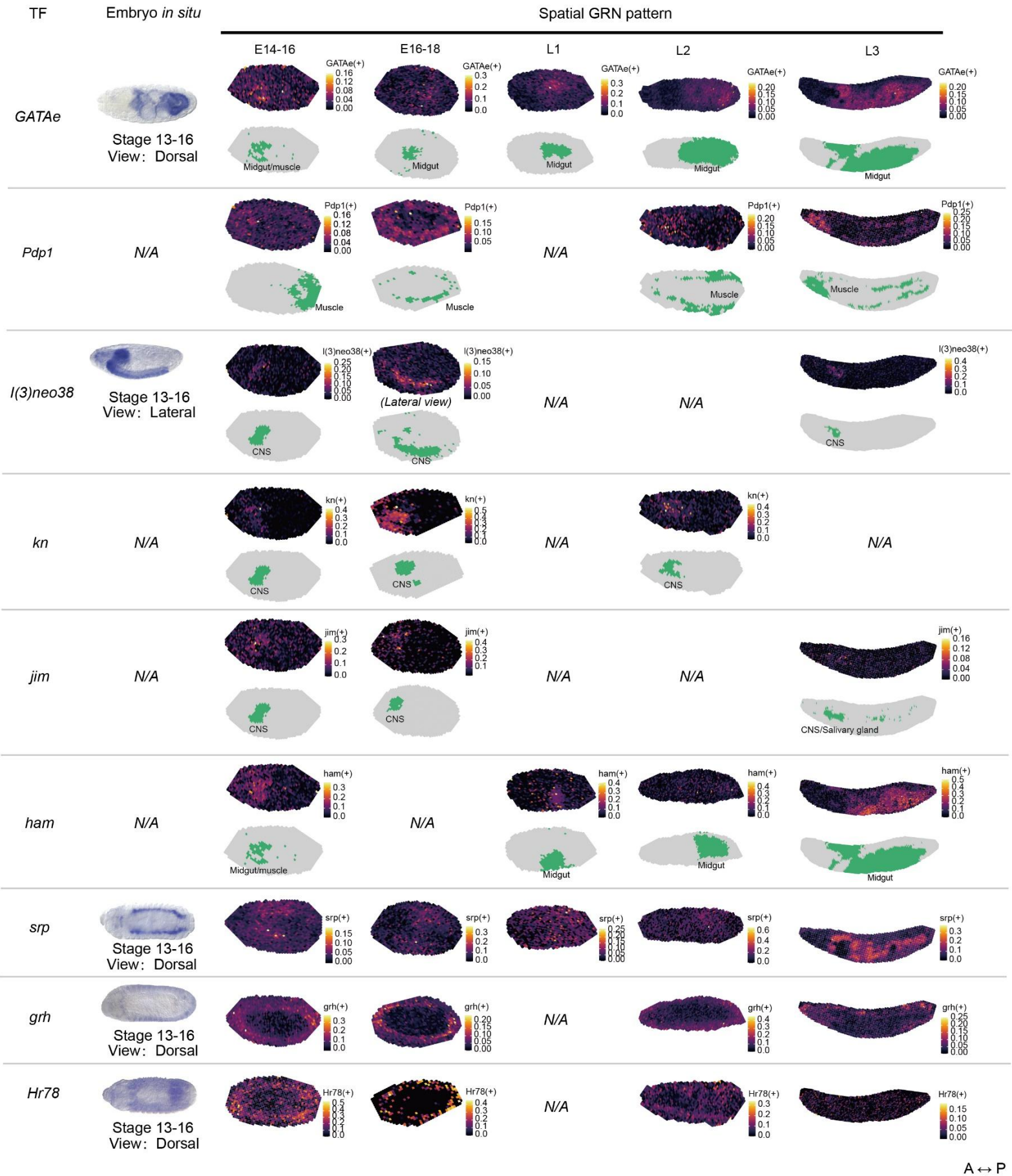
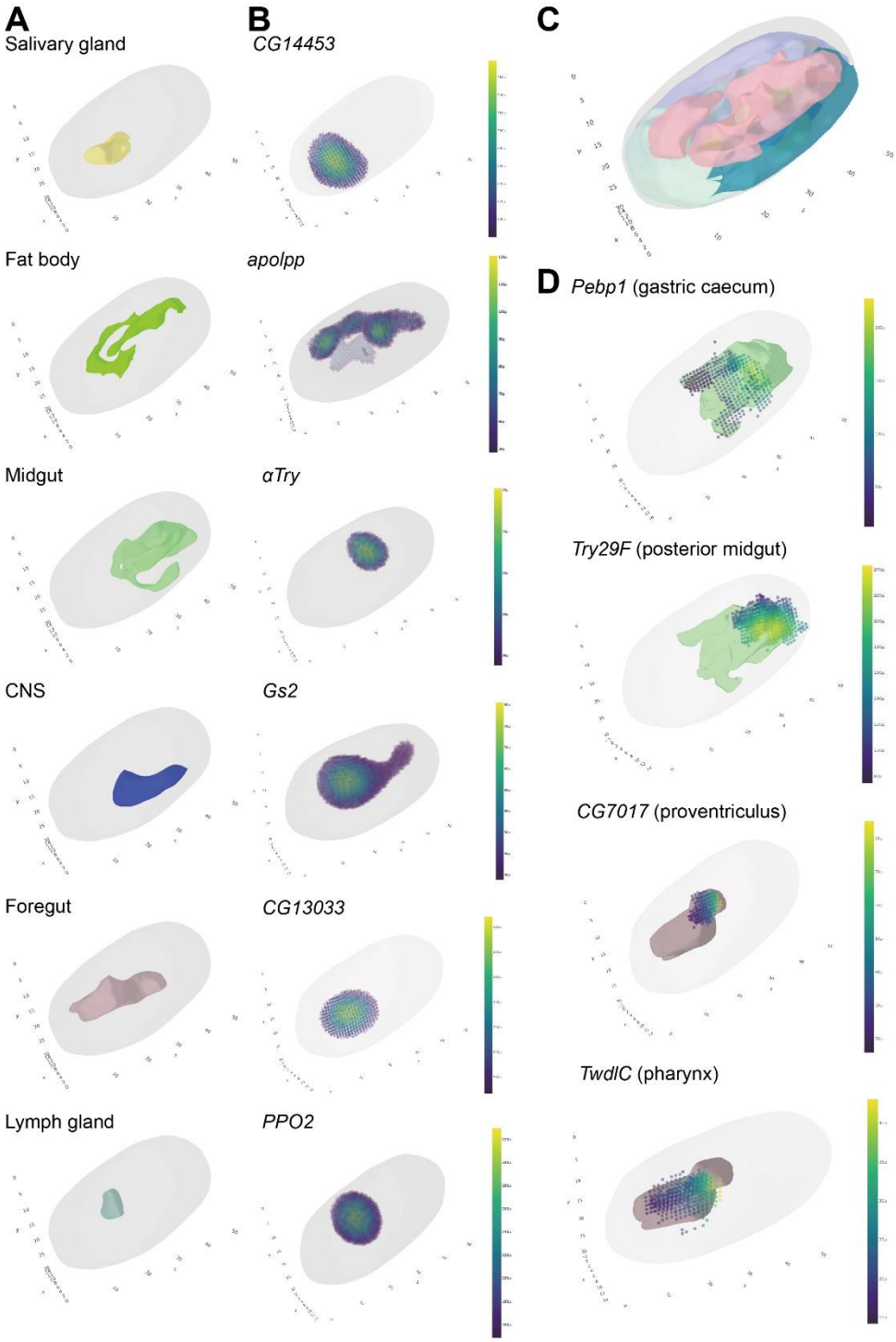


Figure 4 Spatial gene regulatory networks identified in Stereo-seq data. Spatial GRN patterns identified by SCENIC. For each row with two sub-rows, the upper sub-row shows GRNs mapped to spatial transcriptomic profiles generated by Stereo-seq, and the lower sub-row shows the positions and annotations of the clusters co-colocalized with the GRN in the same sample. The left-most column shows *in situ* hybridization results of the TFs from BDGP database (if available). Stages and viewpoints of *in situ* hybridization results are labeled but do not completely match those of our embryo samples. Unless otherwise labeled, Stereo-seq samples are presented in near dorsal/ventral view. N/A indicates that spatial GRN patterns of the TF were not identified in any examined samples of the indicated stage, likely due to incomplete spatial transcriptome capture resulting from cryosection position and orientation. A-P: anterior-posterior.



688 **Figure 5 3D reconstruction of spatial transcriptomes of a late-stage *Drosophila* embryo. (A)** 3D
689 reconstruction, clustering, and annotation of Stereo-seq generated spatial transcriptome of an entire
690 E16-18 embryo. Individual 3D clusters representing specific tissues are shown. **(B)** 3D spatial
691 distribution of representative marker genes of tissues presented in **(A). (C)** The integrated 3D model
692 of the entire embryo. **(D)** 3D clusters representing foregut and midgut, and representative marker
693 genes of their detailed anatomical structures. For scales, each unit in x axis is equivalent to 7 μm (the
694 thickness of a single cryosection slice), and each unit in y and z axis is equivalent to 9.15 μm (the
695 length of 20 DNBs in bin 20 × 20 analysis).

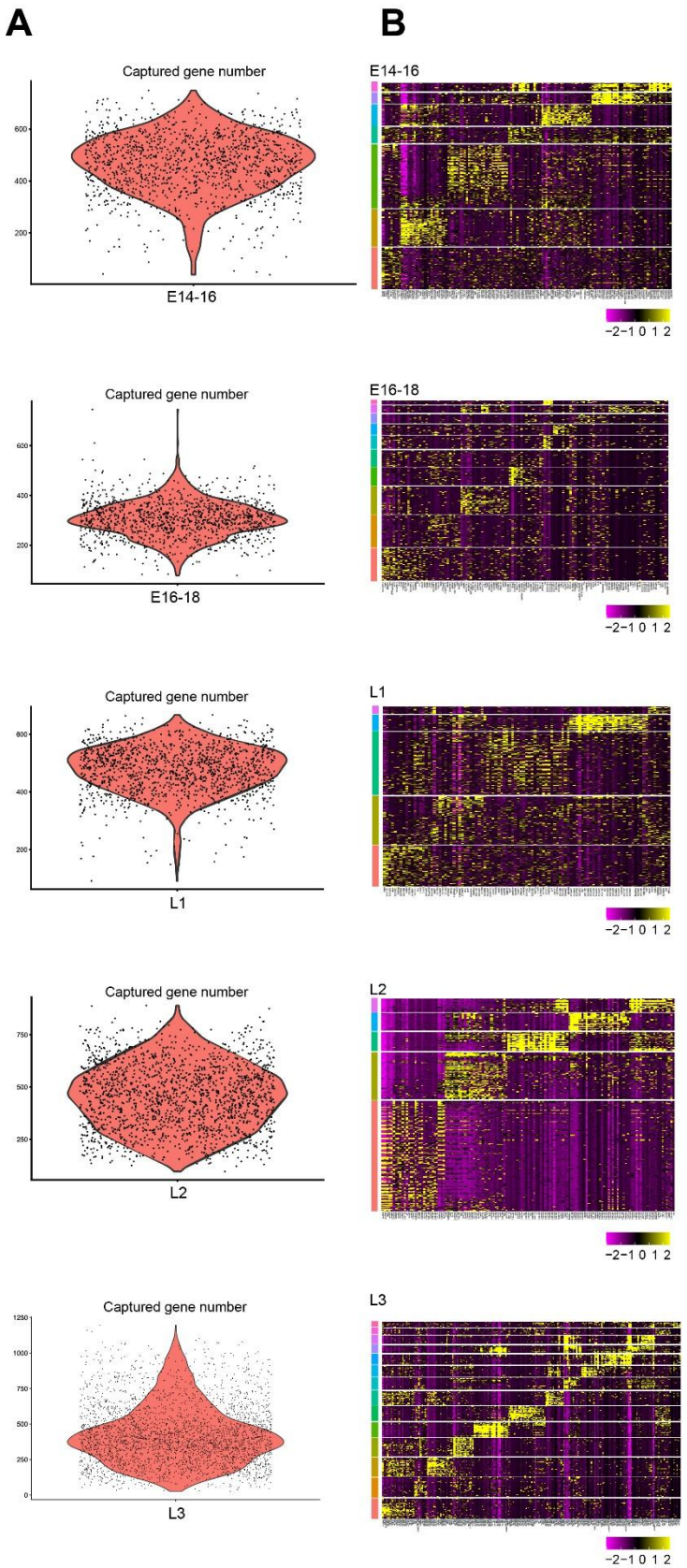


Figure S1 Quality control of Stereo-seq data from samples of each examined stages. (A) Numbers of genes captured by Stereo-seq in representative samples, each dot represents one bin (bin 20 × 20 for E14-16, E16-18, L1 and L2, bin 50 × 50 for L3). **(B)** Heatmap of expression levels of top marker genes of each cluster after unsupervised clustering in the same samples as **(A)**.

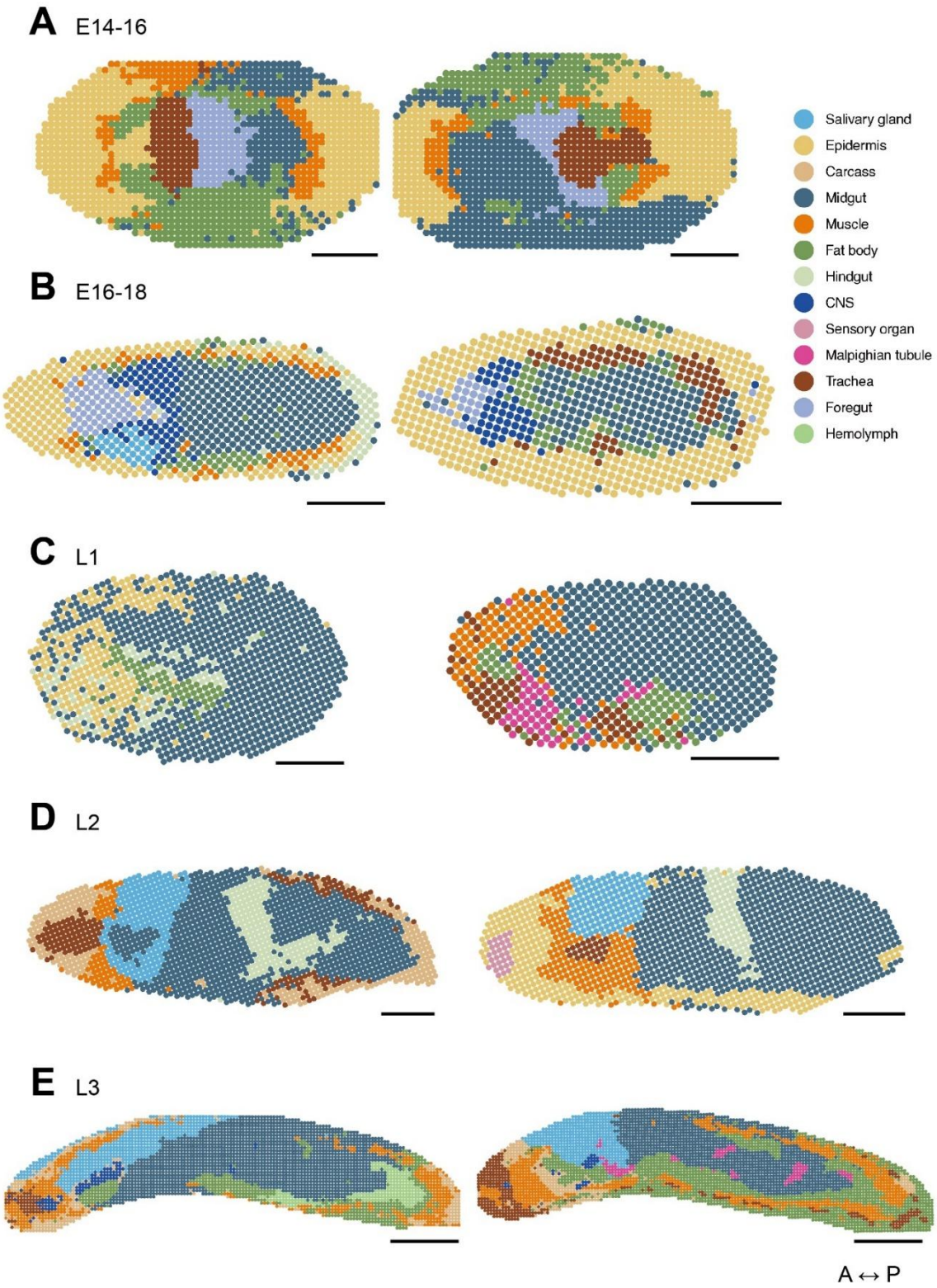


Figure S2 Biological replicates of spatial transcriptomes of *Drosophila* embryos and larvae generated by Stereo-seq. (A-E) Unsupervised clustering and annotation of single sections of **(A)** E14−16, **(B)** E16-18, **(C)** L1, **(D)** L2, and **(E)** L3 samples. Clusters with the same annotation are merged and assigned the same color code. Scale bars = 100 μm for E14-16, E16-18, L1 and L2, and 500 μm for L3. A-P: anterior-posterior.

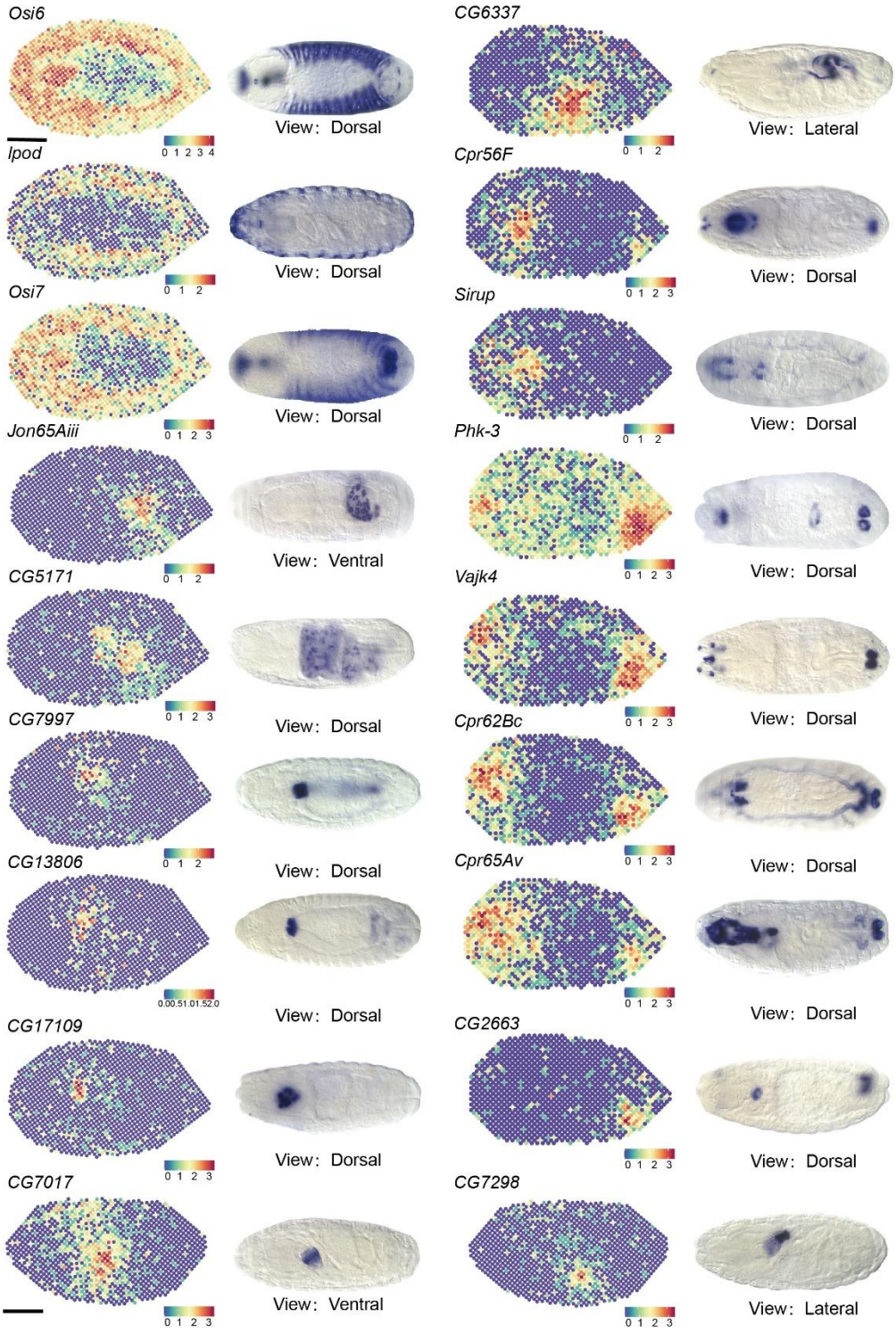
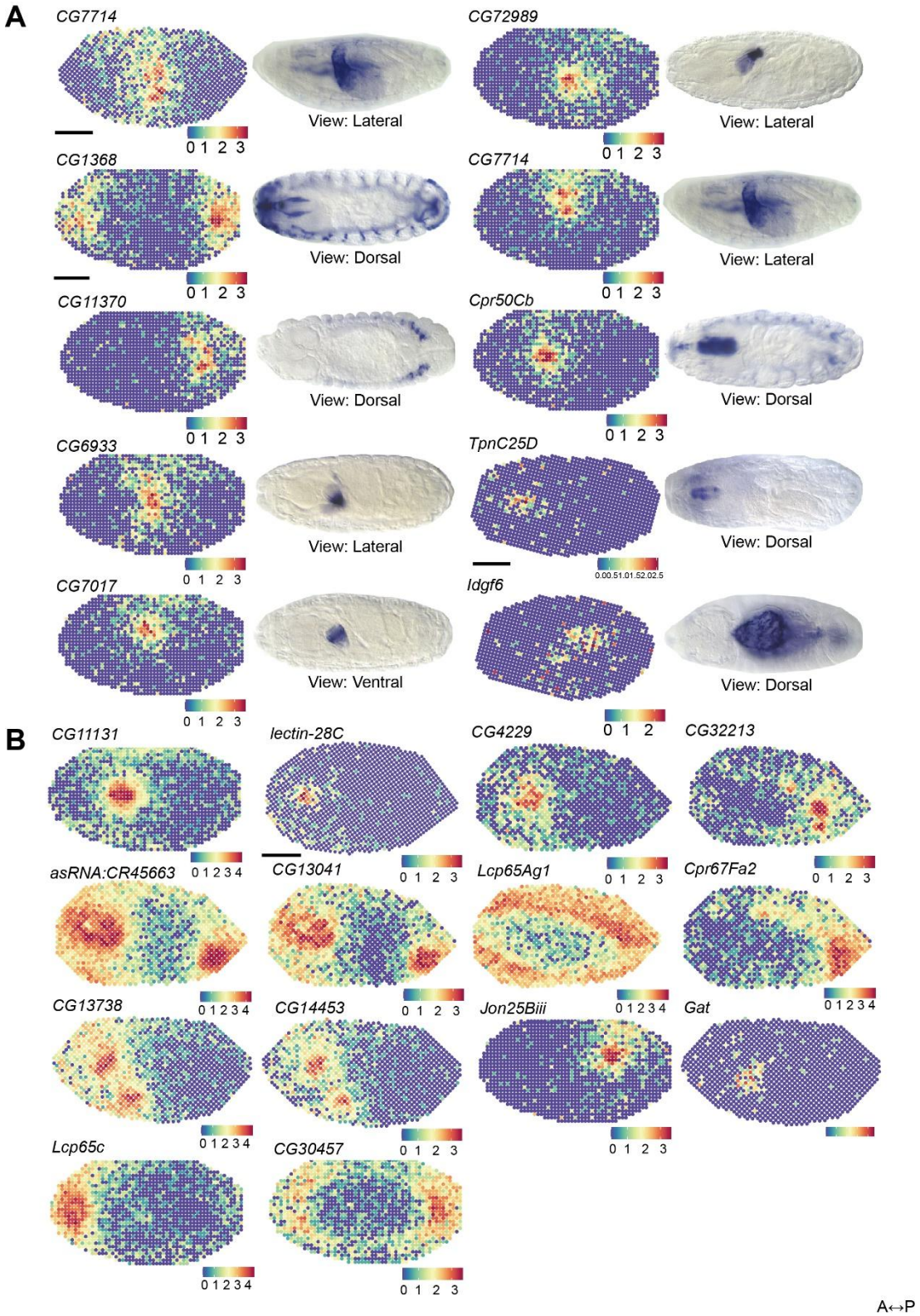
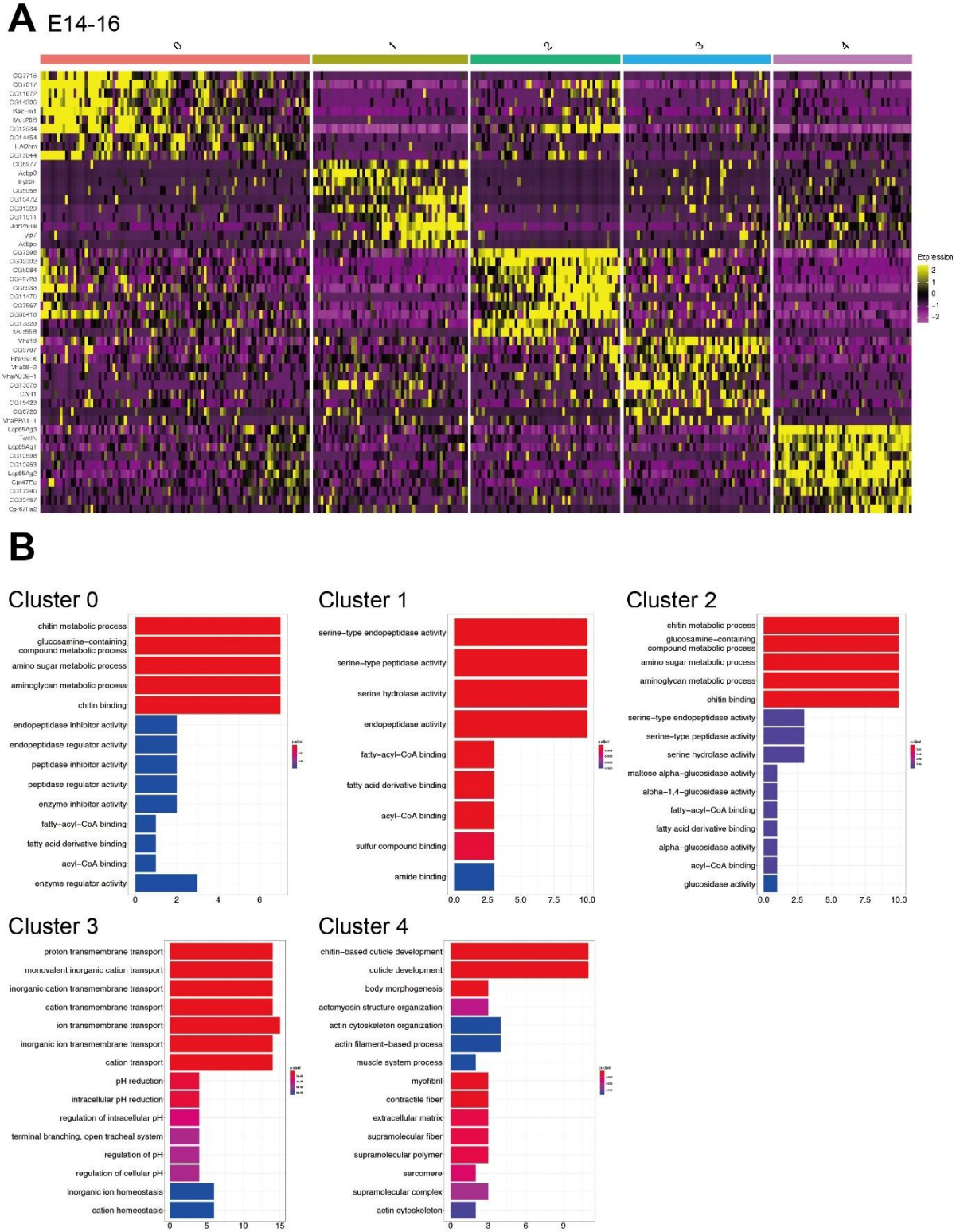


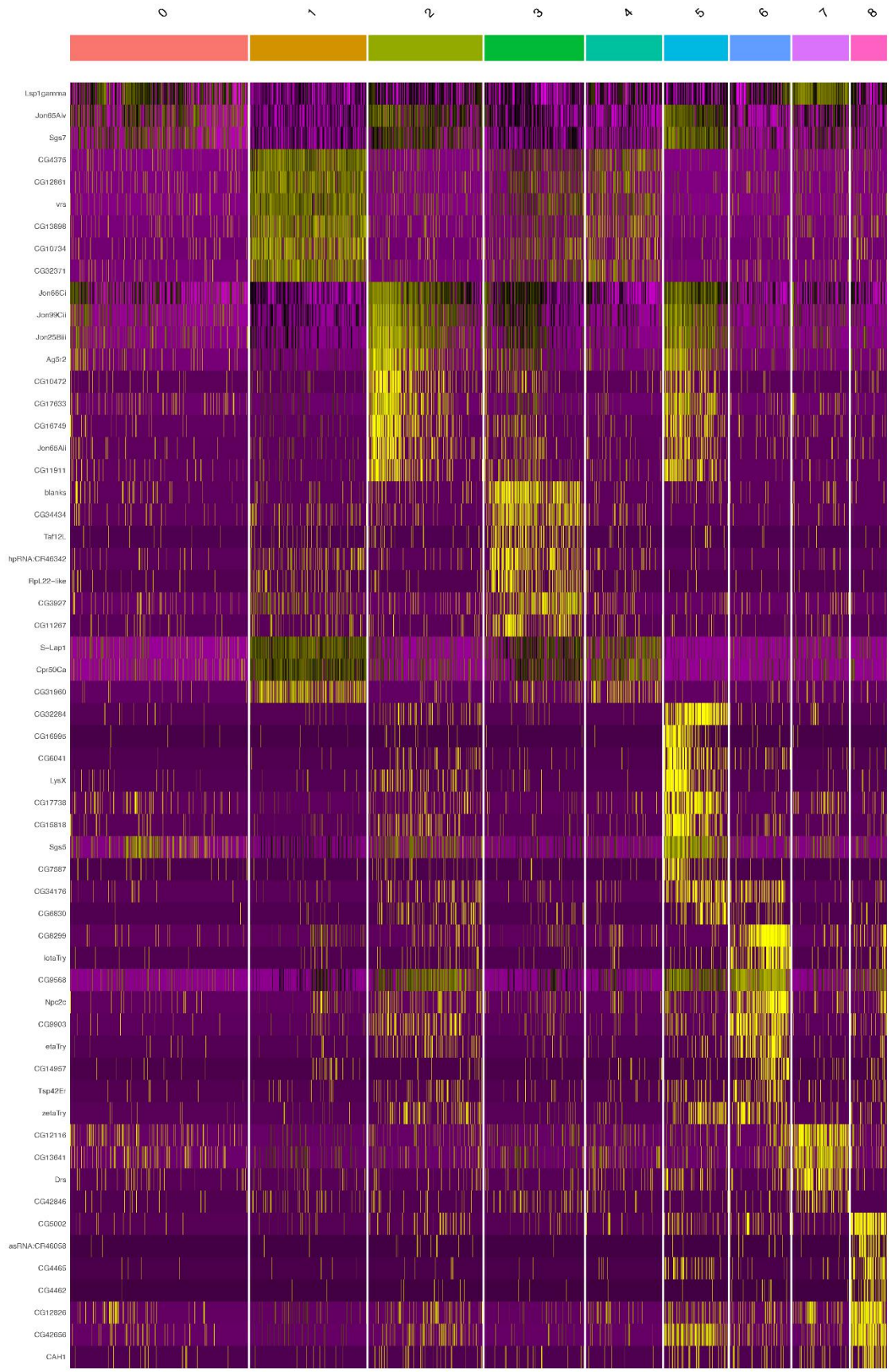
Figure S3 Comparisons of spatial gene expression patterns between Stereo-seq and *in situ* hybridization. For each gene panel, left shows spatial expression patterns generated from Stereo-seq data of representative single sections of E14-16 or E16-18 embryo samples; right shows *in situ* hybridization results of embryos at similar stages from BDGP database. All Stereo-seq samples are in near dorsal/ventral view. Viewpoints of each *in situ* hybridization image are labeled under each panel. Scale bars = 100 μm. A-P: anterior-posterior.



711 **Figure S4 Comparisons of spatial gene expression patterns between Stereo-seq and *in situ***
712 **hybridization. (A)** Continued from **Figure S3**. **(B)** Spatial expression pattern of genes with no available
713 clear *in situ* hybridization patterns from BDGP or Fly-FISH. All Stereo-seq samples are in
714 near-dorsal/ventral view. Viewpoints of each *in situ* hybridization image are labeled under each panel.
715 Scale bars = 100 μm. A-P: anterior-posterior.

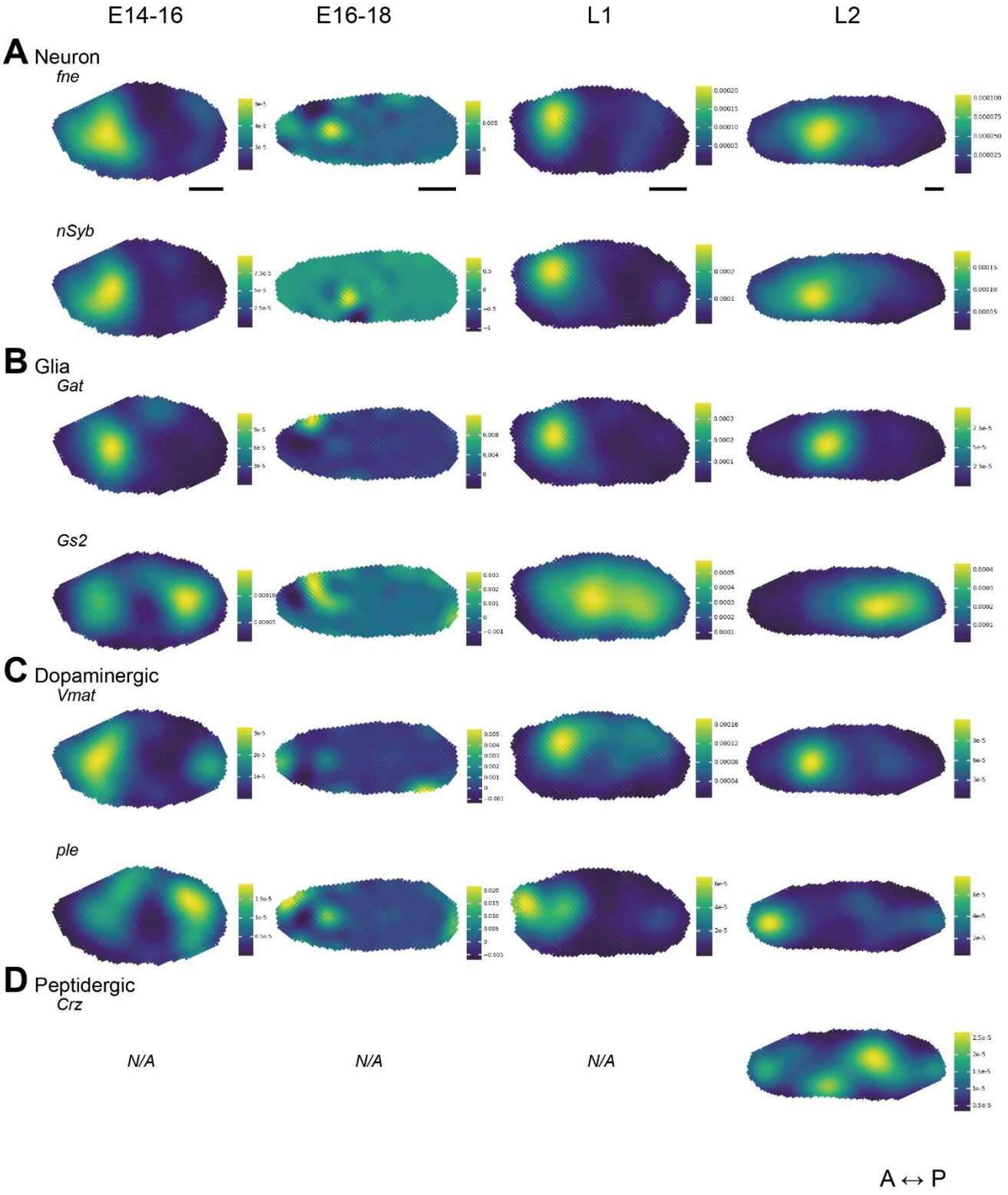


716 **Figure S5 Analysis of marker genes of embryonic and larval midgut subregions. (A)** Heatmap of
717 the top marker genes of midgut subclusters in the single section from E14-16 in **Figure 2**. **(B)** GO
718 ontology analysis of marker genes of each subcluster in **(A)**.

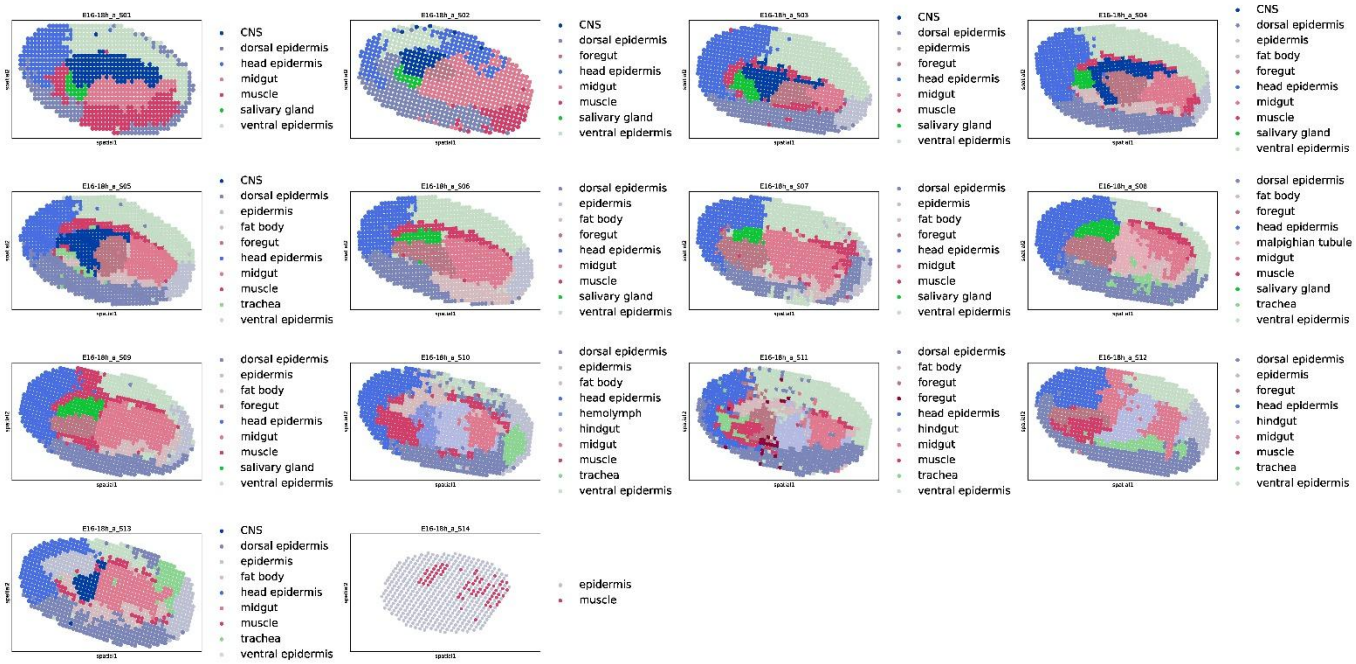


719 **Figure S6 Heatmap of the top marker genes of subclusters of male reproductive organs in the**

720 **single L3 section in Figure 3.**



721 **Figure S7 Spatial expression patterns of marker genes of embryonic and larval CNS. (A-D)**
722 Spatial expression patterns of marker genes of **(A)** neural, **(B)** glial cells, **(C)** peptidergic, and **(D)**
723 dopaminergic neurons in late-stage embryos and larvae. N/A indicates that spatial gene expression
724 patterns were not identified in any examined samples of the indicated stage, likely due to incomplete
725 spatial transcriptome capture resulting from cryosection position and orientation. Scale bars = 100
726 µm. A-P: anterior-posterior.



727 **Figure S8 Clustering and annotation of Stereo-seq generated spatial transcriptome of all**

728 **sections of an E16-18 embryo used for 3D reconstruction in Figure 5.**

Sample	Bin size (number of DNBs) used for analysis	Number of unique transcripts per bin	Number of unique genes per bin
E14-16	20 × 20	1325.2 ± 347.0	437.2 ± 81.3
E16-18	20 × 20	2969.1 ± 913.9	425.7 ± 78.9
L1	20 × 20	1602.1 ± 368.4	477.9 ± 82.4
L2	20 × 20	1928.3 ± 923.7	453.9 ± 151.4
L3	50 × 50	1768.4 ± 1737.7	423.8 ± 208.5

729 **Table S1 Numbers of transcripts and genes captured in samples shown in Figure 1.** Numbers
730 are shown in mean ± standard deviation.

731 **Supplemental Data 1** 3D spatial patterns of representative tissues in an entire E16-18 embryo
732 shown in **Figure 5**.
733 **Supplemental Data 2** 3D spatial transcriptome reconstruction model of an entire E16-18 embryo
734 shown in **Figure 5**.
735 **Supplemental Data 3** 3D spatial patterns of representative marker genes of detailed anatomical
736 structures in foregut and midgut in an entire E16-18 embryo shown in **Figure 5**.
737 **Supplemental Data 4** Custom scripts used in this study.