High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae

**Graphical abstract**

**Highlights**

- Stereo-seq faithfully captures *Drosophila* spatial transcriptomes with high resolution
- Developing *Drosophila* tissues can be 3D reconstructed with Stereo-seq data
- 3D Stereo-seq data enable the identification of tissue subregions and cell state changes
- 3D Stereo-seq data uncover spatial patterns of regulon dynamics

**Authors**

Mingyue Wang, Qinan Hu, Tianhang Lv, ..., Xun Xu, Yuhui Hu, Longqi Liu

**Correspondence**

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

**In brief**

Using Stereo-seq, Wang et al. resolved the 3D spatial transcriptomes of *Drosophila* late-stage embryos and larvae with molecular and cellular dynamics patterns. This work opens the opportunity for the systematic study of gene regulatory networks during *Drosophila* development with organism-wide spatiotemporally resolved transcriptomic information.
High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae

Mingyue Wang,1,2,3 Qinan Hu,2,3,4,12 Tianhang Lv,1,5,12 Yuhang Wang,1,6,12 Qing Lan,1,12 Rong Xiang,1,12 Zhencheng Tu,1,5,12 Yanrong Wei,3 Kai Han,9 Chang Shi,1 Junfu Guo,1 Chao Liu,1 Tao Yang,11 Wensi Du,9 Yanru An,1 Mengnan Cheng,1,5 Zhencheng Tu,1,5,12 Yanrong Wei,3 Kai Han,9 Chang Shi,1 Junfu Guo,1 Chao Liu,1 Tao Yang,11 Wensi Du,9 Yanru An,1

1BGI-Shenzhen, Shenzhen 518083, China
2Shenzhen Key Laboratory of Gene Regulation and Systems Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518055, China
3Department of Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518055, China
4Department of Pharmacology, School of Medicine, Southern University of Science and Technology, Shenzhen 518005, China
5College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
6School of Biology and Biological Engineering, South China University of Technology, Guangzhou 510006, China
7BGI College & Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450000, China
8BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China
9China National Gene Bank, BGI-Shenzhen, Shenzhen 518120, China
10Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen 518120, China
11Shenzhen Key Laboratory of Single-Cell Omics, Shenzhen 518083, China
12These authors contributed equally
13Lead contact
*Correspondence: liyuxiang@genomics.cn (Y.L.), wangxiaoshan@genomics.cn (X.W.), xuxun@genomics.cn (X.X.), huyh@sustech.edu.cn (Y.H.), liulongqi@genomics.cn (L.L.)

https://doi.org/10.1016/j.devcel.2022.04.006

SUMMARY

*Drosophila* has long been a successful model organism in multiple biomedical fields. Spatial gene expression patterns are critical for the understanding of complex pathways and interactions, whereas temporal gene expression changes are vital for studying highly dynamic physiological activities. Systematic studies in *Drosophila* are still impeded by the lack of spatiotemporal transcriptomic information. Here, utilizing spatial enhanced resolution omics-sequencing (Stereo-seq), we dissected the spatiotemporal transcriptomic changes of developing *Drosophila* with high resolution and sensitivity. We demonstrated that Stereo-seq data can be used for the 3D reconstruction of the spatial transcriptomes of *Drosophila* embryos and larvae. With these 3D models, we identified functional subregions in embryonic and larval midguts, uncovered spatial cell state dynamics of larval testis, and revealed known and potential regulons of transcription factors within their topographic background. Our data provide the *Drosophila* research community with useful resources of organism-wide spatiotemporally resolved transcriptomic information across developmental stages.

INTRODUCTION

For over a century, *Drosophila* has been a fruitful 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 profiles 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 (*Drosophila* Gene Expression Tool) (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 (Karaïkos et al., 2017), imaginal disc (Arias 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., 2015), and gonad (Jeavitt et al., 2020; Rust et al., 2020; Witt et al., 2019) (reviewed in the study conducted by Li, 2020). These studies substantially expanded our knowledge of cellular diversity, functional heterogeneity, and microenvironmental regulation in Drosophila tissues. Comparing with 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., 2022), 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 spatial gene expression patterns 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 inherent drawbacks, including the 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 protein-coding genes in the Drosophila genome, over 40% of them 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 (Karaïkos et al., 2017), physical segmentation (Junker et al., 2014), and local mRNA capture and sequencing (Rodrigues 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 the study conducted by Liao et al., 2021). In an effort to capture the actual spatial transcriptomes in situ, several 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 and thus can overcome the disadvantages of in situ hybridization, make a valuable complement 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 a submicrometer level resolution or sufficient mRNA capture efficiency, impeding their application in Drosophila embryos (~500 × 100 μm in section size). Due to the small sizes of these samples, efforts in resolving their complete spatial transcriptomes remained computational (Karaïkos et al., 2017). Our recently developed SpaTial Enhanced RESolution Omics-sequencing (Stereo-seq) technique (Chen et al., 2022) can resolve spatial transcriptomes with nanometer resolution while retaining high sensitivity, providing a powerful tool to obtain spatial transcriptomes from small-sized samples like Drosophila embryos. In Stereo-seq, patterned DNA nanoballs (DNBs) with randomly barcoded sequences are photolithographically etched on modified chips and are distributed ~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 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 that take altogether ~5 days. Pupal stage follows and lasts for another ~4 days. Adult flies subsequently eclose from pupal cases. In this study, we applied Stereo-seq to late-stage embryos and all stages of larvae, generating 2D spatial transcriptomic data across these developmental stages. We also demonstrated that with 2D Stereo-seq data from all the sections of a sample, it is possible to 3D reconstruct its spatial transcriptome in silico. Unsupervised clustering and annotation could associate clusters with anatomical structures. With 3D Stereo-seq data, we detected functional subregions in embryonic and larval midgut, analyzed spatial cell state changes during larval spermatogenesis, and identified active transcription factor regulons in multiple tissues across development. Available and future Drosophila Stereo-seq data are curated in a searchable database, Flysta3D: https://db.cnbg.org/stomics/flysta3d/, for easy access. These data open exciting opportunities for Drosophila transcriptomic 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) (Figure 1A). In each experiment, 7-μm thick (10-μm thick for L2 and L3) cryosection slices from individual samples were subjected to patterned 1 × 1-cm Stereo-seq chips. Stereo-seq was performed as previously described (Chen et al., 2022). Sequencing data were processed and integrated to generate 2D spatial transcriptomes of each section. Typically, Stereo-seq captured more than 1,500 unique transcripts representing over 400 genes per merged bin.
Based on Stereo-seq data, 2D expression matrices with defined x-y coordinates and positional expression profiles of each bin were generated. We then combined all 2D expression matrices from the same sample and performed unsupervised clustering. Clusters were associated with tissue types based on previously reported marker gene profiles, followed by manual validation and annotation (Figure S1C). Clustering results in each section were consistent with anatomical structures of embryos and larvae (Figures 1B, 1C, and S2).

To validate the spatial transcriptomes that we obtained with Stereo-seq, we compared the spatial expression patterns of top marker genes of embryo tissue clusters with known *in situ* hybridization results for these genes in embryos at comparable stages, obtained from BDGP database (Tomancak et al., 2002, 2007). To better reflect the spatial expression patterns of each gene in the entire embryo, their expression matrices in all sections of a sample were combined and projected to a single 2D graph. The spatial patterns of these transcripts in our Stereo-seq data matched well with those of *in situ* hybridization results (Figure 2A). To further explore the correlation between our Stereo-seq data and BDGP *in situ* hybridization results, we assigned each gene in embryo samples to a dominant tissue cluster based on spatial enrichment and compared these enrichment identities with BDGP tissue annotations. We found significant overlap between Stereo-seq tissue enrichment and BDGP tissue annotations (Figure 2B; Data S1). These observations indicate that Stereo-seq effectively reflects the spatial gene expression patterns of *Drosophila* embryos. In addition, our Stereo-seq data generated spatial information for multiple transcripts that are not included in BDGP database (Figure 2C; Data S1). These patterns can be reproducibly detected in late-stage embryos. 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.

**3D reconstruction of spatial transcriptomes of *Drosophila* embryos and larvae**

With the efficient and faithful capture of 2D spatial transcriptomes of *Drosophila* embryos and larvae by Stereo-seq, we wondered if it is possible to 3D reconstruct the spatial transcriptome of the
Figure 2. Comparisons of spatial gene expression patterns between Stereo-seq data and in situ hybridization

(A) Spatial expression patterns of top marker genes of different tissue clusters in Stereo-seq. For each gene, spatial expression patterns from Stereo-seq data of E14–16 or E16–18 samples are shown on the left. Expression matrices from all sections of the entire embryo were combined and projected along the z axis (along which cryosection was performed) to generate 2D graphs of spatial expression patterns; in situ hybridization results of embryos at stage 13–16 from BDGP database are shown on the right.

(B) Venn diagrams showing overlaps between BDGP tissue annotation and genes assigned to each tissue cluster based on spatial enrichment of E14–16 and E16–18 Stereo-seq data.

(C) Spatial expression patterns of genes with no available in situ hybridization results from BDGP. These patterns are reproducibly detected in both E14–16 and E16–18 Stereo-seq data. A full list of such patterns can be found in Data S1. Genes mentioned in this and following figures are listed in Table S2 with FlyBase IDs for ID matching and validation. All Stereo-seq samples are in near lateral view. Viewpoints of each in situ hybridization image are labeled under each panel. Scale bars, 100 μm. A-P, anterior-posterior.
entire sample. We collected all the cryosection slices of the 5 studied stages (every 4th slice for L3) sectioned along the left-right axis (so that each section is a sagittal plane) and performed Stereo-seq with them. After mapping, clustering, and annotation, we developed a pipeline for sample 3D reconstruction (see STAR Methods). Briefly, we extracted 2D graphic 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 clusters with the same annotation in adjacent slices. The contour and position of 3D clusters resembled the anatomical structures of embryos and larvae (Figure 3A; Data S2). 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. Representatively, the marker genes of each tissue displayed expected spatial distribution around their corresponding tissues in the E14–16 embryo (Figure 3B; Data S2). Moreover, 3D clusters representing complex tissues such as foregut and midgut can be further divided into detailed structures based on marker gene identification (Figure 3C; Data S2).

In brief, Stereo-seq is capable of efficient recapitulation of the 3D spatial transcriptome of Drosophila embryos and larvae. We anticipate that with modification, the 3D transcriptome reconstruction strategies we developed here can be applied to additional stages of Drosophila embryos and adult tissues to resolve their spatial transcriptomes as well.

Regionalization of embryonic and larval midgut
In the recent decade, studies have recognized that adult fly midgut contains subregions with distinct cell morphologies and physiological capacities. Various genetic and genomic techniques, including enhancer trap reporter screens, transcriptomic profiling of hand-dissected midgut domains, and single-cell RNA-seq (scRNA-seq) of midgut cells were used to characterize the regionalization of adult midgut (Buchon et al., 2013; Guo...
et al., 2019; Hung et al., 2020; Marianes and Spradling, 2013). In contrast, studies on regionalization of embryonic and larval midgut have been lacking. A study decades ago revealed specific reporter patterns in larval midgut (Murakami et al., 1994), but the complexity of developing embryonic and larval midgut remains to be explored.

To explore the regionalization of Drosophila embryonic and larval midgut, we employed hotspot analysis, which identifies modules of correlated genes by integrating single-cell gene expression profiles and other multimodal data such as spatial position information (DeTomaso and Yosef, 2021). By applying hotspot analysis to Stereo-seq data, we identified spatially distinct and functionally diverse gene modules across 5 stages of samples (Figures 4A and S3). Inspecting these modules, we observed that as the organism matures, expression levels of gene families encoding chymotrypsin, trypsin, and amylase, which are required for protein or starch digestion, increase during embryonic and larval development (Figure 4B).

We then visualized these gene modules in 3D reconstructed midgut model and observed 3–4 distinct modules distributed along digestive tracts in each stage of samples (Figure 4C; Data S3). Across gene modules and developmental stages, genes encoding digestive enzymes displayed spatially distinct and temporally dynamic patterns (Figure 4D). For example, maltose alpha-glucosidase gene Mal-A3 is enriched in posterior midgut of embryos (E14–16 module 2/E16–18 module 1) and early larvae (L1 module 1E3/L2 module 1), whereas Mal-A6 is enriched in anterior late larval midgut (L3 module 5); trypsin family gene Try29F is specifically expressed in posterior embryonic midgut (E14–16 module 2/E16–18 module 1), whereas other trypsin family genes such as alphaTry and betaTry are enriched in anterior late larval midgut (L3 module 5&6) (Figure 4D). These dynamic regional expression patterns within gene families further indicated the functional regionalization dynamics of midgut during embryonic and larval stages. Gene ontology (GO) analysis also supported that these spatial gene modules represent functional domains of embryonic and larval midgut, as each module is enriched for distinct functions in nutrient digestion (Figure 4E).

Together, our Stereo-seq data suggested the dynamic spatiotemporal characteristics of functional genes along the digestive tract across Drosophila development and provided clues for midgut regionalization changes during embryogenesis. **Spatial cell state dynamics in larval testes**

Upon further examination of spatiotemporal features of embryonic and larval tissue transcriptomes, clusters representing testis stood out as another region of high cell type complexity and mRNA abundance. Drosophila testis contains continuously differentiating germ cells, providing an excellent model to study in vivo cell states and their transitions. Thus, we performed subclustering and spatial gene expression analysis on one transverse section of L3 testes.

After further subclustering of clusters representing testes, we determined the cell types that each merged bin (bin 15 × 15 DNBs) represents by mapping Stereo-seq data to an annotated larval testis scRNA-seq dataset (Mahadevaraju et al., 2021) (Figure 5A). We generated prediction scores for germ line cell types for each bin and identified germ cell bins with high prediction scores. We determined bins representing somatic cells with the same method (Figure 5B). We identified various cell types during spermatogenesis, including spermatogonia and primary spermatocyte of different stages (Figures 5A and 5B). The spatial distribution of these cell types along the differentiation progression matches the actual anatomical structure of larval testes (Figure 5C). Marker gene expression profiles of these identified cell types also coincide with previous reports (Figure S4A; Table S3) (Mahadevaraju et al., 2021; Zhao et al., 2010).

Notably, we observed mRNA abundance variation in differentiating germ cells (Figure S4B). UMAP projection of three germ line cell types showed directional transition of transcriptome (Figure S4C). GO analysis also uncovered functional pathway enrichment shifts along differentiation (Figure S4D). All these observations reflected cell state changes of germ cells at different developmental stages. Thus, we employed RNA velocity analysis (La Manno et al., 2018) to the developing germ cells we captured in this single section at a single time point. With RNA velocity analysis, which demonstrates that the balance of captured unspliced and spliced mRNA abundance is an indicator of future cell states, we can visualize spatial transcriptional dynamics of spermatogenesis in larval testes. We observed robust transition from spermatogonia to early primary spermatocyte and then to late primary spermatocyte, as indicated by the flow of velocity vectors (Figure 5D). Spermatogonia displayed higher spatial velocity compared with primary spermatocytes, as indicated by thicker vectors, suggesting stronger transcriptional dynamics (Figure 5D). Pseudotemporal ordering of developing germ cells identified genes with high velocity potential (i.e., undergo rapid transcriptional changes) during spermatogenesis (Figure 5E). We then profiled the spatial expression patterns of genes whose mRNAs display a high velocity (Figure 5F). Many of these genes (165/283) was previously reported to be specifically enriched in adult male testis (Leader et al., 2018; Zhao et al., 2010). GO analysis of genes with a high spatial velocity showed enrichment of various pathways critical for germ line development (Figure 5G).

To summarize, the high sensitivity of Stereo-seq enabled delicate subclustering of complex tissues to identify cell types in their morphological background. Spatial RNA velocity analysis...
Figure 5. Cell type subclustering and cell state changes of *Drosophila* larval testis

(A) Subclustering results of clusters representing testes in a representative single L3 transverse section. Each merged bin (dot) represents 15 × 15 DNBs. Scale bars, 100 μm.
could provide clues for location and timing of activation of genes with spatiotemporally restricted regulatory functions.

**Spatial patterns of regulon activities 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 provide a unique opportunity to analyze regulons of GRNs in their morphological contexts, allowing regulon identification and analysis in a spatially resolved tissue-specific manner.

To profile the spatial regulons across Drosophila development, we employed single-cell regulatory network inference and clustering (SCENIC) analysis, which was developed to reconstruct regulons and identify cell states from scRNA-seq data by combining transcriptomic information of TFs and their co-expressed downstream targets (Aibar et al., 2017).

SCENIC analysis of Stereo-seq data revealed regulons with specific spatial patterns. Across embryo and larva samples, we identified multiple regulons mediated by previously reported TFs. Some of the regulons are tissue specific, and their corresponding TFs are known to play regulatory roles in the tissues where the regulons are located, such as Rbp6 in CNS (central nervous system) (Siddall et al., 2012), grh in epidermis (Wang et al., 2009), srp in fat body (Abel et al., 1993), kay in midgut (Souid and Yanicostas, 2003), and Mef2 in muscle (Bour et al., 1995) (Figure 6). The spatial patterns of these regulons in embryos coincide with those of late-stage embryo in situ hybridization results (Figure 6), suggesting that SCENIC analysis of Stereo-seq data effectively recapitulated the spatial regulon patterns of these TFs. Among the TFs of regulons we identified, functions of CG16779 are largely unknown. Its expression was only detected in adult CNS (Leader et al., 2016) but not in embryos (Tomancak et al., 2002) previously. Spatial SCENIC analysis of our Stereo-seq data identified CNS-specific CG16779 regulon activity across embryo and larva samples, implying its potential regulatory functions in CNS development.

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

**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 with 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 (Chen et al., 2022) enable simultaneous transcript capture and 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 course) 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., 2022). 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 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 cell state changes and tissue-specific regulons.

Our results demonstrated the capability of Stereo-seq in resolving the spatial transcriptomes of small-sized samples like Drosophila embryos. To the best of our 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. To facilitate access to our Stereo-seq data, we established a searchable database, Flysta3D: https://db.cngb.org/stomics/flysta3d/, to curate current and future Stereo-seq data generated from Drosophila samples. Users can visualize tissue-specific 3D spatial expression patterns of genes of interest in this database. A detailed tutorial of our database can be found in Figure S5 and in the database instruction page: https://db.cngb.org/stomics/help. 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 samples will establish an encyclopedic spatial transcriptomic database that will be of great interest to the Drosophila research community.

Limitations of the study
This study generated organism-wide 3D spatiotemporal transcriptomic data in Drosophila from late-stage embryos to 3rd instar larvae, but there were still some limitations in experimental and analysis procedures. First, although our Stereo-seq data reconstructed tissues based on spatial heterogeneity of gene expression, dissimilarities remain between 3D reconstructed tissue models and their actual locations and boundaries. It would be more ideal to obtain the actual anatomical structures of tissues and use them as guidance for tissue-specific spatial transcriptomic analysis. In the future, this can be achieved by integrating imaging information of the cryosection slice and Stereo-seq data from the same slice. Second, capture of rare cell types with Stereo-seq remains challenging, such as intestinal stem cells and enteroendocrine cells. These types of cells

Figure 6. Spatial patterns of regulons identified in Stereo-seq data of Drosophila late-stage embryos and larvae

Representative tissue-specific spatial regulon patterns identified by SCENIC. Regulon activity patterns from all sections of entire embryo or larva samples were combined and projected along the z axis (along which cryosection was performed) to generate 2D graphs of spatial activity patterns of regulons. In situ hybridization results with viewpoints and tissue annotations of the TFs in stage 13-16 embryos from BDGP database are shown for comparison. Stereo-seq samples are presented in near lateral view. A-P, anterior-posterior.
are of few numbers and/or scattered distribution with low expression levels of marker genes, resulting in underrepresentation in spatial transcriptomic data. In our further study, we will obtain scRNA-seq data of embryos and larval tissues and integrate them with Stereo-seq data of the same fly line to improve resolution of spatial transcriptomic data. Third, it is difficult to construct tissue development trajectories with only late-stage embryo and larva samples, as most fundamental changes in organogenesis occur at earlier stages. We will include more stages of embryo data for complete and detailed transcriptomic trajectories for tissue development across stages in future studies. Finally, insights generated by Stereo-seq and spatiotemporal transcriptomic analyses, such as midgut regionalization and testsis spatial RNA velocity, need to be verified with in vivo experiments like mutant analysis. As we continue to expand our database, we will generate transgenic or mutant fly lines to better address the biological questions raised by Stereo-seq data analyses.

**STAR METHODS**

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

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Fly strain maintenance
  - Sample preparation
- **METHOD DETAILS**
  - Stereo-seq
  - Stereo-seq data analysis
  - Hotspot analysis
  - Mapping of larval testes spatial transcriptomic data to scRNA-seq reference
  - RNA velocity analysis
  - SCENIC analysis
  - Gene Ontology enrichment analysis
- **ADDITIONAL RESOURCES**

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.devcel.2022.04.006.

**ACKNOWLEDGMENTS**

This work is part of the SpatioTemporal Omics Consortium (STOC) paper package. A list of STOC members is available at http://www.sto-consortium.org. This work was supported by Shenzhen BGI Research Institute, the grant of Top Ten Fundamental Research Institutes of Shenzhen. This work was supported by Shenzhen Key Laboratory of Gene Regulation and Systems Biology (grant no. ZDSYS20200811144002008, China) to W.C. and Y.H.), 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.), National Natural Science Foundation of China (grant no. 32100684 to Q.H. and grant no. 81773881 to Y.H.), 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 GenBank (CNGB). We thank Dr. Mariana F. Wolfer for helpful comments on the manuscript. Schemes in the graphical abstract, Figures 1A, and 5C were 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 experiments. 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. C.S., F.G., C.L., T.Y., and W.D. helped establish the online database. Q.H., Q.L., and M.W. 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.

**REFERENCES**

Abel, T., Michelson, A.M., and Maniatis, T. (1993). A Drosophila GATA family member that binds to Adh regulatory sequences is expressed in the developing fat body. Development 119, 623–633.

Albar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Huiselmann, 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.

Aris, M.M., Islam, A.B.M.M.K., 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, 5024.

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wardell, E., Custodio, J., Reimegård, J., Salmén, F., et al. (2019). A spatiotemporal organism-wide gene expression and cell atlas of the developing human heart. Cell 179, 1647–1660.e19.

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

Bour, B.A., O’Brien, M.A., Lockwood, W.L., Goldstein, E.S., Bodmer, R., Taghert, P.H., Abmayr, S.M., and Nguyen, H.T. (1995). Drosophila MEF2, a transcription factor that is essential for myogenesis. Genes Dev. 9, 730–741.

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 Lemaître, 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).

Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball patterned arrays. Cell. https://doi.org/10.1016/j.cell.2022.04.003.

Chen, H., and Boutros, P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 12, 35.

Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Sala Frigerio, C., Corthout, N., Qian, X., Laliková, J., Kuhнемund, M., Votyuk, I., et al. (2020). Spatial
transcriptomics and in situ sequencing to study Alzheimer's disease. Cell 182, 976–991.e19.

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, I., Alibar, 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.e20.

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.

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

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.

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

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artergi, 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.e5.

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

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95.

Jeavitt, A., Chatterjee, D., Xie, G., Wang, X.-F., Otwell, T., Huang, Y.-C., and Jevitt, A., Chatterjee, D., Xie, G., Wang, X.-F., Otwell, T., Huang, Y.-C., and Yan, Y. (2019). Single cell transcriptomic landscapes of pattern formation, proliferation and growth in Drosophila wing imaginal discs. Development 146, dev179754.

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.

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.

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

Murakami, R., Shigenaga, A., Matsumoto, A., Yamaoka, I., and Tanimura, T. (1994). Novel tissue units of regional differentiation in the gut epithelium of Drosophila, as revealed by P-element-mediated detection of enhancer. Rouxs Arch. Dev. Biol. 203, 243–249.

Ortiz, C., Navarro, J.F., Jurek, A., Martin, 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., Ibara, I.L., Holmberg, O., Virshup, I., et al. (2021). Squidpy: A scalable framework for spatial single cell analysis. Preprint at bioRxiv. https://doi.org/10.1101/2021.02.19.431994.

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

Rodrigues, 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 363, 1463–1467.

Rust, K., Bymes, 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, 5628.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Siddall, N.A., Calicina, M., Johanson, T.M., Monk, A.C., Casagrande, F., Been, R.P., McLaughlin, E.A., and Hume, G.R. (2012). Drosophila Rbp6 is an orthologue of vertebrate Msi-1 and Msi-2, but does not function redundantly with dMsi to regulate germline stem cell behaviour. PLoS One 7, 1–13.

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

Souid, S., and Yanicostas, C. (2003). Differential expression of the two Drosophila fos/kayak transcripts during oogenesis and embryogenesis. Dev. Dyn. 227, 150–154.

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.e21.

Tomancak, P., Beaton, A., Weissmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2002). Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol. 3, RESEARCH0088.

Tomancak, P., Berman, B.P., Beaton, A., Weissmann, 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, R145.

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmann, G., Alibar, 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.

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.

Wang, S., Tsrhouhas, V., Xylouridis, N., Sabri, N., Tiklová, K., Nautil, N., Gallo, M., and Samakovlis, C. (2009). The tyrosine kinase Sticher activates
Grainy head and epidermal wound healing in *Drosophila*. Nat. Cell Biol. 11, 890–895.

Wickham, H. (2011). ggplot2. WIREs. Comp. 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, 15.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (N Y) 2, 100141.

Zeira, R., Land, M., and Raphael, B.J. (2021). Alignment and integration of spatial transcriptomics data. Preprint at bioRxiv. https://doi.org/10.1101/2021.03.16.435604.

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-Ferrer, 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.
## STAR METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Chemicals, peptides, and recombinant proteins | | |
| Agar | Vetec | V900500 |
| Propionic acid | Aladdin | P110444 |
| Phosphate acid | LingFeng, Shanghai | - |
| Bromophenol blue | Macklin | B802656 |
| Tissue-Teck OCT | Sakura | 4583 |
| Pepsin | Sigma | P7000 |
| 20X Saline-sodium citrate (SSC) | Thermo | AM9770 |
| RNase inhibitor | NEB | M0314L |
| Critical commercial assays | | |
| Cornmeal-sucrose-agar Drosophila media | Hopebio | HB8590 |
| Drosophila incubator | Laifu | PGX-280A-3H |
| Cryostat | Leica | CM1950 |
| Stere-seq chips and apparatus | BGI Shenzhen | - |
| KAPA Hotstart Ready Mix | Roche | KK2602 |
| SuperScript reverse transcription mix | Invitrogen | 18064-014 |
| VAHTS DNA clean beads | Vazyme | N411-03 |
| Deposited data | | |
| Sequencing data generated by Stereo-seq | This study | https://db.cngb.org/stomics/flysta3d/ |
| Experimental models: Organisms/strains | | |
| Drosophila melanogaster w1118 | Tsinghua Fly Center | - |
| Software and algorithms | | |
| STAR | Dobin et al., 2013 | https://github.com/alexdobin/STAR |
| PASTE | Zeira et al., 2021 | https://github.com/raphael-group/paste |
| Seurat | Stuart et al., 2019 | https://github.com/satijalab/seurat/ |
| Scanpy | Wolf et al., 2018 | https://scanpy.readthedocs.io/en/stable/ |
| Squidpy | Palla et al., 2021 | https://squidpy.readthedocs.io/en/stable/ |
| Skimage | Van der Walt et al., 2014 | https://github.com/scikit-image/scikit-image |
| 3D slicer | Fedorov et al., 2012 | https://www.slicer.org/ |
| Plotly | Sievert, 2020 | https://github.com/plotly/plotly.py |
| Plotnine | Kibirige, 2017 | https://plotnine.readthedocs.io/en/stable/ |
| VennDiagram | Chen and Boutros, 2011 | https://cran.r-project.org/web/packages/VennDiagram/index.html |
| Dynamo | Qiu et al., 2022 | https://github.com/aristoteleo/dynamo-release |
| Hotspot | DeTomaso and Yosef, 2021 | https://github.com/YouLab-Hotspot |
| Velocyto | La Manno et al., 2018 | http://velocyto.org/velocyto.py |
| pySCENIC | Van de Sande et al., 2020 | https://github.com/aertslab/pySCENIC |
| clusterProfiler | Wu et al., 2021 | https://github.com/YuLab-SMU/clusterProfiler |
| Other | | |
| Custom codes generated using open-source software | This study | https://db.cngb.org/search/project/CNP0002189 |
| Custom scripts | This study | Data S4 |
RESOURCES AVAILABILITY

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

Materials availability
All materials used for Stereo-seq are commercially available.

Data and code availability
Raw data generated by Stereo-seq in this study and associated analysis protocols and software can be accessed in our online database, Flysta3D: https://db.cngb.org/stomics/flysta3d/. All data were analyzed with standard programs and packages, as detailed in method details. All custom codes using open-source software to support this study are provided in supplemental information (Data S4). Any additional information required to re-analyze the data reported in this study is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Fly strain maintenance
All embryo and larva samples in this study were from Drosophila strain w1118. Flies were maintained on cornmeal-sucrose-agar media in a 25 °C incubator 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, 49% v/v grape juice, 0.2% v/v propionic acid, 0.02% phosphate acid) from a population cage and aged to desired stages. Unsexed 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 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 main text and figure legends. Samples were oriented so that they were most likely to be sectioned along 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 and L3) were embedded and sectioned together. Sample were embedded with Tissue-Tek OCT and transferred to a −80 °C freezer for storage until used. Cryosection was performed with indicated slice thickness in a Leica CM1950 cryostat. Sample sections were applied to Stereo-seq chips immediately after cryosection.

METHOD DETAILS

Stereo-seq
Stereo-seq library preparation and sequencing were performed as previously described (Chen et al., 2022). See additional resources for a detailed protocol for Stereo-seq. Briefly, embryo and larva sections on Stereo-seq chips were fixed in pre-chilled methanol at -20 °C for 40 min. After removal of methanol, sections were permeabilized on chip with 100 μL 0.1% pepsin in 0.01 M HCl at 37 °C for 5 min. Permeabilization solution was then removed and sections were washed with 100 μL 0.1× saline-sodium citrate (SSC) buffer supplemented with 0.05 U/μL RNase Inhibitor. mRNAs captured by DNBS on the chip were reverse transcribed with SuperScript II reverse transcription (RT) mix (10 U/μL reverse transcriptase, 1 mM dNTPs, 1 M betaine solution PCR reagent, 7.5 mM MgCl2, 5 mM DTT, 2 U/μL RNase inhibitor, 2.5 μM Stereo-seq template switch oligo and 1× First-strand buffer) at 42 °C for 1 h. RT mix was then removed. The chip was washed with 0.1× SSC and incubated in tissue removal buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 37 °C for 30 min. Tissue removal buffer was removed and the chip was washed twice with 0.1× SSC. cDNAs on the chip were amplified with KAPA HiFi Hotstart Ready Mix with 0.8 μM cDNA-PCR primer. Sequencing libraries were prepared with PCR products undergoing the following steps: fragmentation (in-house Tn5 transposase), amplification (KAPA HiFi Hotstart Ready Mix), and purification (VAHTS DNA clean beads). Final libraries were sequenced on a MGI DNSEQ-Tx sequencer.

Stereo-seq data analysis
Raw data processing
Stereo-seq raw data processing and unsupervised clustering were performed as previously described (Chen et al., 2022). Briefly, CID sequences were first mapped to the designed coordinates on chip with 1 base mismatch tolerance. UMI sequences with quality score lower than 10 were filtered out. cDNA sequences were aligned to the reference genome (Dm6) by STAR (Dobin et al., 2013). Expression profile matrices with CID were generated based on the information above.

Section alignment
For two adjacent sections, PASTE (Zeira et al., 2021) algorithms were used to align them along the z axis based on both gene expression similarities and spatial coordinates. For all the 2D expression matrices of every embryo, pairwise_align function in PASTE was run sequentially along the z axis, and each bin was assigned an x-y-z 3D coordinate.
Unsupervised clustering
The expression profile matrices of sections from E14-16, E16-18, L1 and L2 samples were divided into bins with 20 x 20 DNBs, while those from L3 samples were divided into bins with 50 x 50 DNBs. After obtaining the expression profile matrix through the in-house processing software, data was normalized and integrated with R package Seurat (Stuart et al., 2019) and unsupervised clustering with spatial constraints was performed. In brief, SCTransform was first applied to normalize and identify highly variable genes. RPCA-based integration was used to integrate all sections of every embryo. To account for both spatial and transcriptome information, a k-nearest neighbor graph based on transcriptomic data in Python package Scanpy (Wolf et al., 2018) and a spatial k-nearest neighbor graph based spatial information in Python package Squidpy (Palla et al., 2021) were generated respectively. The union of two graphs was then used as input for leiden clustering.

Cluster annotation
For the list of marker genes of each cluster in unsupervised clustering results, top 50 genes with the most significant $p$ values were extracted. To infer specific tissue types of each cluster, these genes were queried in the publicly available databases, including BDGP, FlyBase, FlyAtlas1, FlyAtlas2 and FlyMine, and other published results. Tissue type(s) that match the most marker genes of the cluster were assigned to that cluster. After annotation, scanpy.pl.rank_genes_groups_heatmap function of Python package scanpy (Wolf et al., 2018) was used to generate heatmaps with the following parameters: n_genes=10, groupby = ‘annotation’, vmin = 0, vmax = 3.

3D modeling
After cluster annotation, 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) algorithms. 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 package matplotlib (Hunter, 2007) and plotnine (Kibirige, 2017) were used to draw 2D graphics, and plotly was used to display 3D graphics and models.

Tissue enrichment gene statistics
After cluster annotation, Seurat was used to identify genes specifically enriched to each tissue cluster of Stereo-seq data, using functions FindMarkers and FindAllMarkers with the following parameters: avg_log2FC >= 0.75, min.pct = 0.1, test.use = “wilcox”. R package VennDiagram (Chen and Boutros, 2011) were used to draw Venn diagrams showing overlaps between genes assigned to each tissue cluster based on spatial enrichment of E14-16 and E16-18 Stereo-seq data and BDGP tissue annotation.

2D gene expression patterns
After preprocessing, gene expression patterns were visualized with plotnine. Sections were aligned and projected along the z axis. For better visualization, opacity of each bin was controlled by its gene expression level with parameter alpha = gene expression value. The range of opacity was set to [0,1], consistent with the minimum and maximum values of gene expression.

Imputation of spatial gene expression
For better visualization of spatial gene expression patterns, SparseVFC function of Dynamo software (Qiu et al., 2022) was used to impute gene expression by 2D or 3D coordinates to render spatial patterns of gene expression.

3D gene expression patterns
After imputation of spatial gene expression patterns, plotly was used to visualize 3D models with parameter opacity = 0.1 and spatial gene expression patterns with parameter opacity = 0.2. Spatial gene expression portions with low expression levels were filtered out to reduce background noise. Spatial locations of gene expression patterns were adjusted according to 3D models.

Hotspot analysis
Hotspot analysis
Python package Hotspot (DeTomaso and Yosef, 2021) was used to identify midgut functional gene modules with default parameters by combining expression matrices and 3D coordinate matrices of each sample. Identified gene modules were then annotated according to known midgut functional genes.

Expression patterns of midgut subregion functional genes
scanpy.pl.dotplot function from scanpy was used to visualize differential expression of midgut functional genes in gene modules of embryo and larva samples.

Heatmap of hotspot subregions
scanpy.pl.heatmap function from scanpy was used to visualize genes expression patterns in midgut subregion gene modules with midgut marker genes as input genes.

Mapping of larval testes spatial transcriptomic data to scRNA-seq reference
Reference larval testes scRNA-Seq data was downloaded from Gene Expression Omnibus with accession number GSE125947 (Mahadevaraju et al., 2021). Seurat was used for label transfer from scRNA-seq dataset to our Stereo-seq data. Briefly, scRNA-seq reference data and spatial transcriptome query were first pre-processed. Label transfer was then performed, generating a probabilistic classification score for each bin for each cell type. Based on the predicted scores, each bin was assigned a most likely cell
Unsupervised clustering results of spatial transcriptome data were used to correct predicted annotation. For each spatial transcriptome data cluster, the predicted cell type that best defines that cluster was identified.

**RNA velocity analysis**

Spliced and unspliced RNA were counted for all detected genes (~500 genes per bin in the male reproductive organ clusters) by every merged bin of 15 × 15 DNBs using *Veloccyto* command line interface (La Manno et al., 2018), with about 1.3% of genes per bin were detected as unspliced type. The generated splice and unspliced expression matrices were then used to estimate the differentiation dynamics of cell lineages with *Dynamo* software, the transitioning between gene expression states was modeled and inferred, with lineage relationship of early primary spermatocyte to late primary spermatocyte set as confident cell velocities, to determine the expression dynamics of germ cells. Genes with high correlation with the transition were also determined in the same process.

**SCENIC analysis**

*pySCENIC* pipeline (Van de Sande et al., 2020) was used to predict activity scores of TFs on each section. *pySCENIC* uses TFs’ motif and downstream gene expression to predict TFs’ activity scores at single cell level. *pySCENIC* pipeline was first used to generate area under curve (AUC) matrices of TFs, with rows representing bins and columns representing TFs. Values in each cell represent TFs’ activity scores. AUC values in each section were then visualized with *ggplot2* (Wickham, 2011). Gene regulatory networks was constructed and then visualized using Cytoscape (Shannon et al., 2003).

**Gene Ontology enrichment analysis**

R package *clusterProfiler* (Wu et al., 2021) was used to identify enriched Gene Ontology terms. Marker gene lists from indicated clusters were inputted with default parameters.

**ADDITIONAL RESOURCES**

- Detailed Stereo-seq protocols: https://db.cngb.org/stomics/flysta3d/ (navigate to the “Stereo-seq” tab on the webpage).
- Raw data, experiment and analysis protocols, and analysis software in this study: https://db.cngb.org/stomics/flysta3d/ (navigate to the “Resource” tab on the webpage).
- Relevant SpatioTemporal Omics Consortium studies using Stereo-seq: https://db.cngb.org/stomics/databases.
Supplemental information

High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae

Mingyue Wang, Qinan Hu, Tianhang Lv, Yuhang Wang, Qing Lan, Rong Xiang, Zhencheng Tu, Yanrong Wei, Kai Han, Chang Shi, Junfu Guo, Chao Liu, Tao Yang, Wensi Du, Yanru An, Mengnan Cheng, Jiangshan Xu, Haorong Lu, Wangsheng Li, Shaofang Zhang, Ao Chen, Wei Chen, Yuxiang Li, Xiaoshan Wang, Xun Xu, Yuhui Hu, and Longqi Liu
Figure S1 Quality control of Stereo-seq data, related to Figure 1. (A-B) Numbers of (A) unique transcripts and (B) genes per merged bin (bin 20 × 20 DNBs for E14-16, E16-18, L1 and L2, bin 50 × 50 DNBs for L3) captured by Stereo-seq in all sections of late-stage embryo and larva samples. (C) Heatmap of expression levels of top 10 marker genes of tissue clusters after unsupervised clustering, annotation and merging of clusters annotated as the same tissues.
Figure S2 Clustering and annotation of Stereo-seq generated 2D spatial transcriptomes of Drosophila late-stage embryos and larvae, related to Figure 1. (A-C) Unsupervised clustering and annotation of all the cryosection slices of (A) E16-18 (B) L1 and (C) L2 samples. Clusters with the same annotation are merged and assigned the same color code. Scale bars = 100 μm for E16-18 and L1, and 500 μm for L2.
Figure S3 Hotspot analysis of embryonic and larval midgut subregions, related to Figure 4. (A) Correlation heatmap of functional gene modules (top color bars) identified by Hotspot analysis in midgut of E16-18, L1 and L2 samples. Each row and each column represent a module marker gene, and Z-score indicates correlation between module marker genes. (B) Heatmap of top marker genes in each gene module in 5 stages of samples.
Figure S4 Quality control of subclusters of testes in the single L3 transverse section, related to Figure 5. (A) Top marker genes of each cell type subclusters in Fig.5A. (B) Number of unique transcripts per merged bin (bin 15 × 15 DNBs), in clusters representing germline cell types. (C) UMAP results of transcriptomic profiles of germline cell types. (D) GO analysis of functional pathway enrichment of germline cell types. G: spermatogonia; EPS/LPS: early-/late-primary spermatocyte; C: somatic cysts cell; T: terminal epithelium precursor cell; P: pigment cell.
High-resolution spatiotemporal transcriptomic maps of developing Drosophila embryos and larvae

Dataset ID: STD50000060 | 155,684 Spots | 12,850 Genes

Summary
This database is intended to curate 3D spatial transcriptomes of all stages of Drosophila embryos and larvae generated by Stereo-seq. Drosophila melanogaster strain w1118 embryos were collected at two late stages (14-15 h and 16-18 h after egg laying, corresponding to stage 16-17 of embryogenesis) and all three stages were also collected. These samples were subjected to cryosection to generate 7-10 μm thick slices. All slices of each sample were applied to Stereo-seq chips to capture their 2D spatial transcriptomes. All the 2D spatial transcriptomes of each sample were combined to recreate their 3D spatial transcriptomes. More samples of different stages will be added in the future. With these data, one could visualize and analyze spatial expression patterns of genes of interest, 3D reconstruct tissue-specific spatial transcriptomes by clustering and annotation, simulate tissue developmental trajectory across development, identify cell signaling pathways and gene regulatory networks, examine gene functions in their intact spatial context etc.

Sections:
- stomics/E14-16h_a_count_normal_stereoseq_h5ad
- stomics/L3_b_count_normal_stereoseq_h5ad
- stomics/L1_a_count_normal_stereoseq_h5ad
- stomics/L2_b_count_normal_stereoseq_h5ad
- stomics/E16-18h_a_count_normal_stereoseq_h5ad
Figure S5 Navigation instructions of our online database website, related to all figures. Current and future Stereo-seq data on *Drosophila* are curated in our online database, Flysta3D: https://db.cngb.org/stomics/flysta3d/. Access to raw data and visualization of 3D models can be found in the “Resource” tab of the database, or by visiting the China National Gene Bank datasets: https://db.cngb.org/stomics/datasets/STDS0000060. The tabs on the datasets page allow users to navigate among: ① Brief summary of current datasets; ② 3D visualization of genes and tissues of interest (detailed instructions of this web app can be found in https://db.cngb.org/stomics/help). The “Sections” drop down list (⑤) can be used to switch between samples; ③ Access of unprocessed Stereo-seq data; ④ Available analysis results of Stereo-seq data.
| Sample | Bin size (number of DNBs) used for analysis | Number of unique transcripts per bin | Number of genes per bin |
|--------|------------------------------------------|------------------------------------|------------------------|
| E14-16 | 20 × 20                                   | 2300.8 ± 786.2                     | 634.5 ± 154.0          |
| E16-18 | 20 × 20                                   | 1553.6 ± 736.7                     | 400.5 ± 88.9           |
| L1     | 20 × 20                                   | 1063.9 ± 607.0                     | 358.0 ± 128.8          |
| L2     | 20 × 20                                   | 922.7 ± 641.3                      | 286.6 ± 132.2          |
| L3     | 50 × 50                                   | 7290.7 ± 6165.5                    | 1337.5 ± 634.6         |

Table S1 Numbers of unique transcripts and genes captured in Stereo-seq samples in this study, related to Figure 1. Numbers are shown in mean ± standard deviation.
| Gene symbol | Flybase ID |
|-------------|------------|
| CG1368      | FBgn0030539|
| Ipod        | FBgn0030187|
| Osi6        | FBgn0027527|
| Osi7        | FBgn0037414|
| Cpr62Bc     | FBgn0035281|
| Cpr56F      | FBgn0034499|
| CG7298      | FBgn0036948|
| CG7017      | FBgn0036951|
| CG5171      | FBgn0031907|
| CG6337      | FBgn0033873|
| CG7714      | FBgn0038645|
| Jon65Aiii   | FBgn0035665|
| CG6933      | FBgn0036952|
| CG17109     | FBgn0039051|
| Idgf6       | FBgn0013763|
| CG2663      | FBgn0037323|
| BBS1        | FBgn0035741|
| CG7406      | FBgn0030980|
| CG14191     | FBgn0030981|
| CG32073     | FBgn0052073|
| CREG        | FBgn0025456|
| Tpc2        | FBgn0035078|
| TwdlV       | FBgn0037227|
| TwdlX       | FBgn0052571|
| Fax         | FBgn0014163|
| Adh         | FBgn0000055|
| Ssk         | FBgn0036945|
| Fdl         | FBgn0045063|
| CG7017      | FBgn0036951|
| TwdlC       | FBgn0039469|
| Pebp1       | FBgn0038973|
| Try29F      | FBgn0038973|
| CG4415      | FBgn0031296|
| Phf5a       | FBgn0031822|
| hepH        | FBgn0011224|
| twin        | FBgn0011725|
| ProtA       | FBgn0013300|
| Rbp6        | FBgn0260943|
| Gene | FlyBase ID |
|------|------------|
| grh  | FBgn0259211|
| srp  | FBgn0003507|
| kay  | FBgn0001297|
| Mef2 | FBgn0011656|
| CG16779 | FBgn0037698 |

Table S2 List of genes mentioned in the main text and figures with their FlyBase IDs, related to Figures 2 to 6.