Construction of a 3D spatiotemporal transcriptomic atlas of *Drosophila* embryos and larvae using high-resolution Stereo-seq

Based on the DNBSEQ sequencing technology and using the DNBSEQ-T7 sequencing platform independently developed by MGI, BGI Research has developed Stereo-seq. This spatiotemporal omics technique affords high resolution and sensitivity, and was used to construct a 3D spatiotemporal development atlas of *Drosophila* embryos and larvae. The atlas revealed the dynamic changes of *Drosophila* testis cells during development and the regulation pattern of genes.

Relevant results of this study were published in 2022 in the journal *Developmental Cell*, under the title "High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae"[1].

Recommended application: Frontier technology – spatiotemporal omics

Recommended model: DNBSEQ-T7RS

- ## A spatiotemporal transcriptomic tool with ultrahigh resolution

  The Stereo-seq spatiotemporal omics technique enables *in situ* capture of gene expression information from cells spatially resolved across a tissue using the DNBSEQ-T7 sequencing platform.

- ## Mapping a 3D spatiotemporal development atlas using Stereo-seq

  Using Stereo-Seq data from serial sections of samples, a 3D reconstruction of tissue organs and integration of gene expression information translated into the 3D transcript model of tissues in *Drosophila* embryos and larvae.

- ## Revealing the regulation pattern of genes during *Drosophila* development

  3D spatiotemporal transcriptome data from Stereo-Seq allows finely partitioning tissue areas of *Drosophila*, accurately analyzing the regulatory state of cells, and characterizing the dynamic regulatory process of gene expression in both spatial and temporal dimensions.

Background

For a long time, *Drosophila* has been a classic model organism for biologists and geneticists studying various developmental processes, such as embryogenesis, organogenesis, gametogenesis, and aging. Given its role in determining the functions of cells and tissues, the tissue-specific atlas has been extensively studied in *Drosophila* . At present, researchers have established several databases to sort out tissue-specific transcription atlases of *Drosophila* , including FlyAtlas1[2], FlyAtlas2[3], and DGET[4].

In recent years, single-cell multi-omics techniques have rapidly advanced, making it possible to analyze genome, transcription, epigenome, and proteome information at single-cell resolution. Researchers have conducted several studies on tissue-specific cells in *Drosophila* using single-cell multi-omics sequencing techniques. Spatial information of the cell transcriptome can serve as reference to study the biological functions of cells in coordinating intracellular signals, metabolism, and development. However, current technologies, many of which adopt *in situ* hybridization, cannot detect unknown transcripts and isomers, and are limited in identifying transcripts with low copy numbers and accurately quantifying gene expression levels. In addition, the spatial transcription atlas of *Drosophila* has not been completely elucidated. At present, the spatiotemporal expression patterns of over 40% of genes in *Drosophila* remain unknown[5]. In recent years, methods such as calculative strategy[6], physical division[7], and local mRNA capture and sequencing[8] have been used to analyze the spatial patterns of transcriptomes, surpassing any advantages of *in situ* hybridization and propelling progress in spatial transcriptomics methods. Nevertheless, existing DNA barcode arrays generally lack the submicron resolution and mRNA capture efficiency to study *Drosophila* embryos.

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

Study description

Based on the Stereo-seq technique[9] (Fig. 1) and using DNB microarray chips, corresponding sequencing reagents, and the high-throughput DNBSEQ-T7 sequencer, a team at BGI Research obtained the spatial transcription atlas of small samples, specifically *Drosophila* embryos and larvae. The resulting atlas featured nanoscale resolution with DNB barcode spacing of about 500 nm and high-sensitivity detection. In addition, the team reconstructed the 3D spatial transcription atlas of *Drosophila* using the spatial transcriptome data of longitudinal sections from *Drosophila*. The atlas lays a foundation to study transcription regulation of *Drosophila* at tissue and organ levels in a 3D spatial scale system.

Materials and Methods

Sample description and processing

The model used in this study was w1118 wild *Drosophila*. Late embryos were collected in 14-16 h and 16-18 h after oviposition, and called E14-16 and E16-18, respectively. Additionally, three developmental stages were sampled (which were called L1-L3). *Drosophila* tissues from different developmental stages were frozen and cut into sections with a thickness of 7 μm, except tissue sections at the L2-L3 stages which were 10 μm thick. The tissue sections were stored in Tissue-Teck O.C.T. in a -80°C refrigerator until use (Fig. 2A).

DNB microarray chips

Stereo-seq (spatial enhanced resolution omics-sequencing) is a spatial transcriptome technique with enhanced resolution independently developed by BGI Research based on DNA nanoball (DNB) array chips and *in situ* capture technology. The DNB microarray chip first deposits the DNBs containing random barcode sequences onto a photoetched modified chip. Each DNB has a diameter of about 220 nm, a ball spacing of about 500 nm or 715 nm, and captures a single DNB signal to achieve nanoscale resolution (Fig. 2B). Compared with bead-based methods, the DNB-marked random barcode produced by rolling circle amplification (RCA) has a greater spatial barcode pool. The array is sequenced to obtain the coordinate ID (CID) of each etched DNB, which corre-

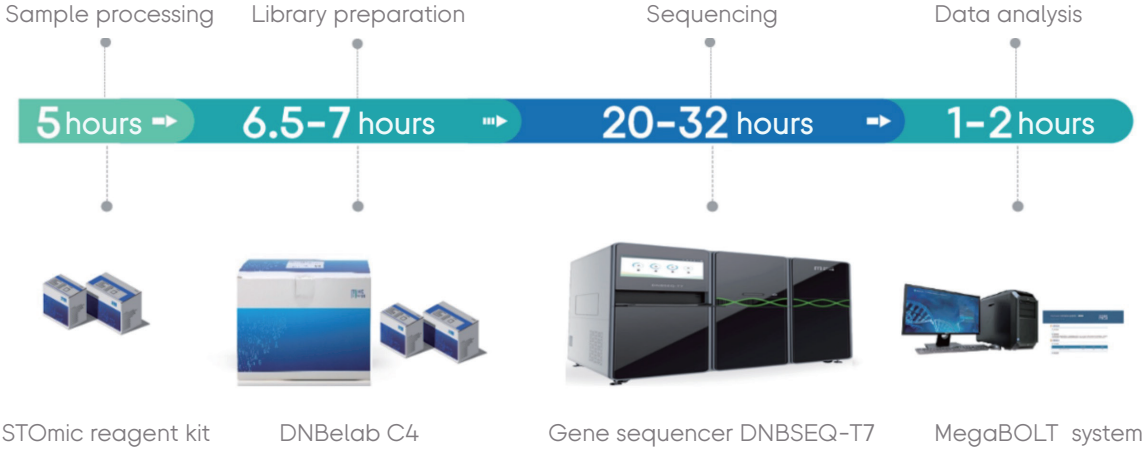

Fig. 1. Complete flowchart of spatiotemporal atlas reconstruction, outlining products from sample collection and processing, library preparation, high-throughput sequencing, and bioinformatic analysis.

sponds to its coordinate position. By mapping the relationship between CID and coordinate position, the subsequently captured mRNA is assigned to its spatial position. Meanwhile, each point includes a molecular ID (MID) to distinguish between different transcripts. The CID sequence length of the standard Stereo-seq library is 25 bp, and the MID length is 10 bp.

Library preparation and *in situ* sequencing

The process for *in situ* sequencing began with the capture of mRNA from tissue sections. The fresh nitrogen-frozen tissue sections were loaded onto the surface of the chip, preserved, and then permeabilized. Finally, reverse transcription and PCR amplification were conducted (Fig. 2C). The amplified cDNA, which served as the template for library preparation, was sequenced with a corresponding CID. The DNBSEQ-T7 sequencer independently developed by MGI was used for this sequencing process.

Sequencing data processing and reconstruction of 3D spatiotemporal transcriptome atlas

To reconstruct the 3D spatiotemporal transcriptome atlas of *Drosophila*, the researchers used the following methods (Fig. 2D): alignment of regions (PASTE algorithm[10]), determination of the location of gene clusters (binning and processed using R *Seurat*[11] and Python package), determination of the type of gene clusters (Python package scanpy[12]), statistical analysis of gene types, mapping of 2D gene spatial distribution atlas, generation of 3D spatiotemporal atlas by simulation (*skimage* algorithm[13] and 3D Slicer software[14]), and finally analysis of gene expression in the 3D atlas (Sparse VFC function of Dynamo software[15]).

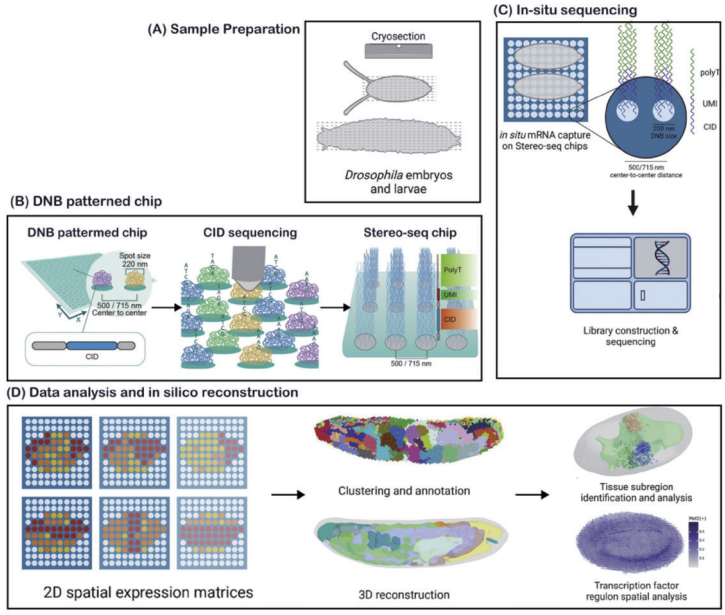

Fig. 2. Method of constructing the 3D spatiotemporal development atlas of *Drosophila* using DNB etched microarray chips and *in situ* sequencing.This figure is derived from Figure 1 in Chen et al., 2022[16] and Figure 1 in Wang et al., 2022[1].

Results

Improvement of *Drosophila* gene expression spatial position information in BDGP database

First, the researchers extracted gene cluster information from the data obtained with Stereo-seq and mapped the 2D gene spatial expression atlas of *Drosophila* embryos and larvae (Fig. 3A). Then, they compared and analyzed the known gene in situ hybridization information in the 2D atlas and BDGP database (Berkeley *Drosophila* Genome Project[17,18]) (Fig. 2B-C). They found that genes captured by Stereo-seq were highly consistent with the anatomical structure information expressed by genes in the BDGP, proving that Stereo-seq accurately reflected the spatial gene expression of *Drosophila* late embryos. Furthermore, the researchers also found that Stereo-seq captured large amounts of gene spatial position information detected in the BDGP database, and the expression patterns of these genes could be detected in the duplicate samples of late embryos (Fig. 3C). Therefore, Stereo-seq captured high-quality transcription expression information of *Drosophila* late embryos and larvae in a high-throughput way and filled in gaps of the BDGP *in situ* hybridization database.

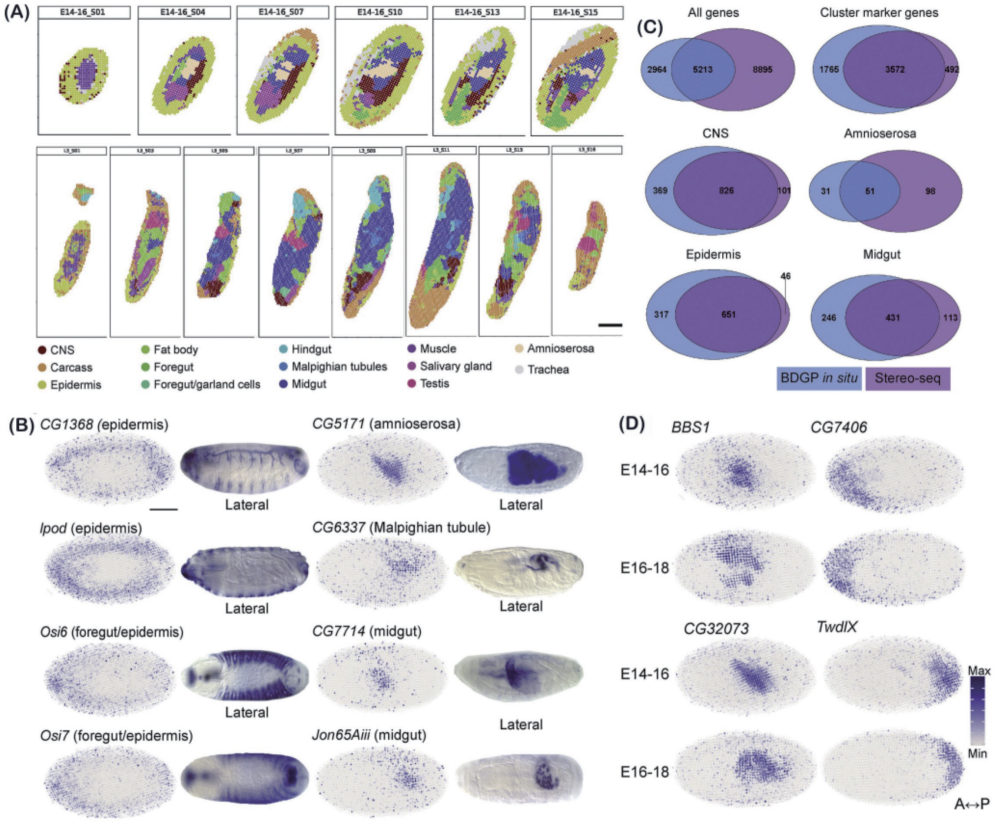

Fig. 3. The degree of matching between the gene spatial expression atlas generated by Stereo-seq and known genes in the BDGP database.

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

3D reconstruction of a spatial transcriptome of *Drosophila* embryos and larvae

The researchers collected samples from five stages of *Drosophila* embryos, including late and larval stages, and performed Stereo-seq sequencing on all frozen sections. From the results, they developed a scheme to reconstruct 3D transcriptome information. First, the researchers extracted 2D regions of each section of the sample in the sequencing matrix and implemented 3D registration according to the shape of the sample section and similarity of the transcriptome. Then, after quality control, clustering and annotation, they obtained the 3D reconstruction model of the sample (Fig. 4A). The researchers also checked the reconstructed 3D *Drosophila* model, finding that the structure of the model was highly consistent with the known anatomical structure of *Drosophila*, and the marker gene of the corresponding tissue was consistent with the results of previous research.

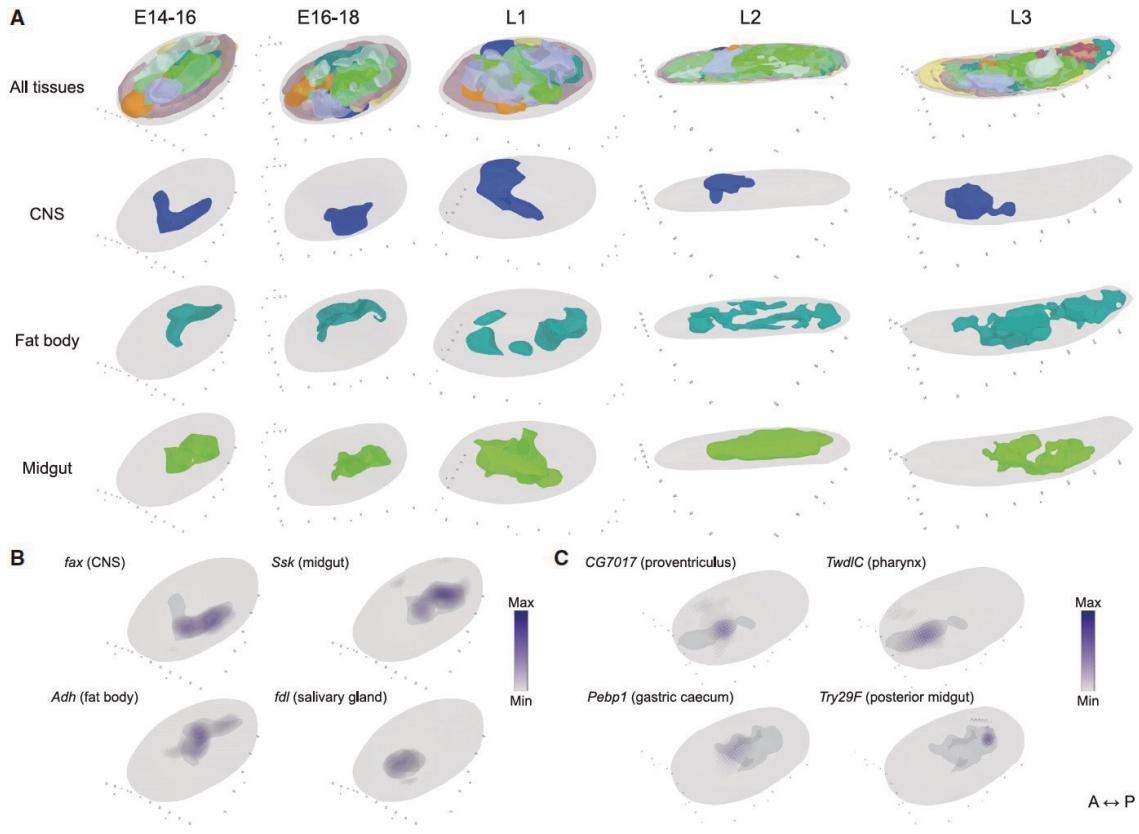

Fig. 4. 3D reconstruction of a spatial transcriptome of *Drosophila* late embryos and larvae.

Identification and division of midgut subregions in *Drosophila* late embryos and larvae

The researchers conducted a hotspot analysis to integrate gene expression profiles of single cells and other multi-modal data (such as spatial position information[19]) to identify the module information of related genes, and then identified spatially and functionally different gene modules in samples during five developmental stages of *Drosophila* (Fig. 4A). By analyzing these different gene modules, researchers found that the expression levels of the gene families encoding chymotrypsin, trypsin, and amylase increased during embryonic and larval development as the organism matured (Fig. 4B). By further visualizing the 3D reconstructed midgut model, the researchers found that 3-4 different gene modules were distributed in the digestive tract of the samples at each stage (Fig. 5C). For example, the genes encoding digestive enzymes exhibited different dynamic patterns in time and space (Fig. 5D). These findings suggest that Stereo-seq data reveal the spatiotemporal dynamic features of the functional genes in the digestive tract during the development of *Drosophila* and provide clues to the changes in midgut regionalization during embryonic development.

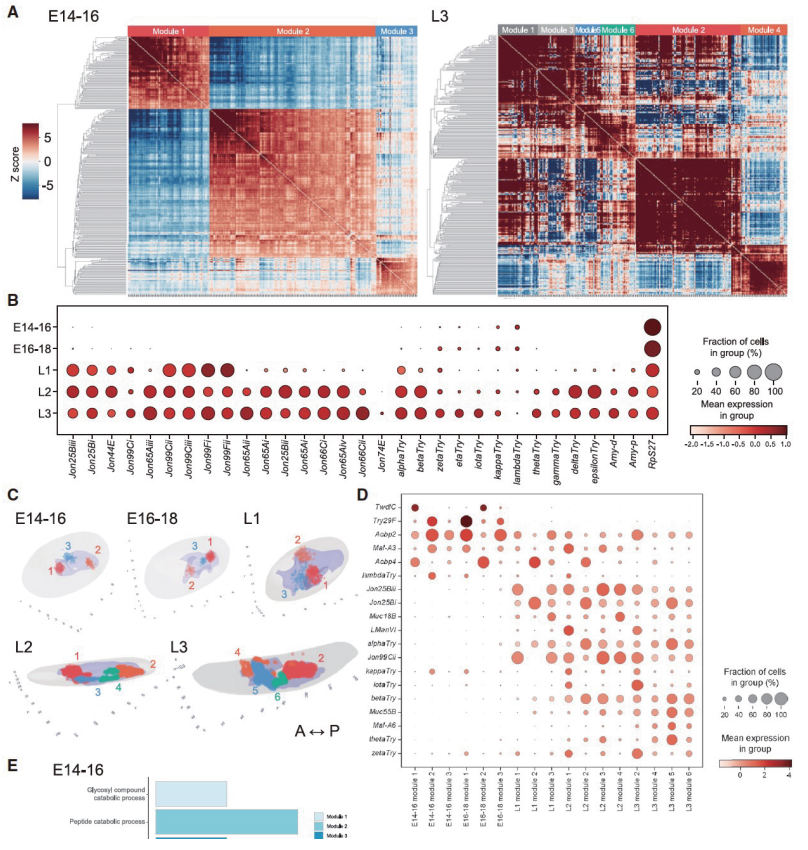

Fig. 5. Spatial subregions of midgut in *Drosophila* embryos and larvae.

Dynamic spatial changes of testis cells in *Drosophila* larva

To study the state and transformation features of cells in *Drosophila*, the researchers analyzed the subclustering and gene spatial expression in the cross-section of testis cells of larvae at L3 stage. By comparing and integrating Stereo-seq data and an scRNA-seq dataset[20], the researchers identified various types of germ cells, including spermatogonia and primary spermatocytes at different stages (Fig. 6A and 6B). The researchers also mapped the dynamic spatial transcription atlas of spermatogenesis in larvae testis via RNA rate analysis (Fig. 6D). Through pseudo-time and pseudo-space enrichment analysis on developing germ cells, the researchers identified the genes with high speed potential during spermatogenesis. The researchers mapped the spatiotemporal expression pattern during spermatogenesis, from spermatogonia, to early primary spermatocytes, and to late primary spermatocytes (Fig. 5E-G).

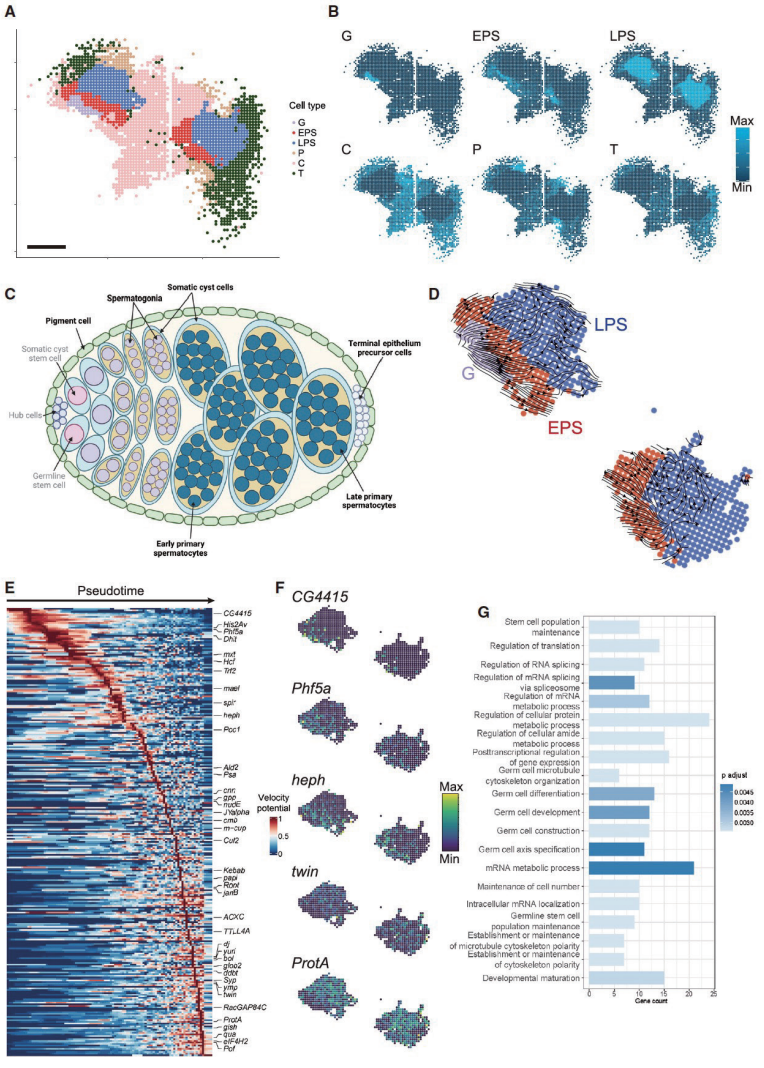

Fig. 6. Subtype differentiation and dynamic changes of testis cells in *Drosophila* larvae.

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

Construction of the spatiotemporal regulation pattern of genes during development of *Drosophila*

To study the spatial regulation patterns of genes during the development of *Drosophila*, the researchers analyzed the transcription factors (TFs) in spatial transcriptome data using single-cell regulatory network inference and clustering (SCENIC) (Fig. 7). They found that the spatial patterns of TFs, like srp[21] in fat body and Mef2[22] in muscle, were consistent with the *in situ* hybridization results of late embryos (Fig. 7). This analysis also revealed potential regulatory functions of some unknown TFs (such as CG16779, which is specifically expressed in the central nervous system).

Summary

As a spatially resolved transcriptome technique independently developed by BGI Research, Stereo-seq has nanoscale resolution, centimeter-level field of view, high sensitivity, and very high repeatability. It is a reliable tool for analyzing the transcriptome of complex tissues and organisms. This application briefly have summarized the results of a study that used Stereo-seq to reconstruct the 3D spatiotemporal transcriptome atlas of *Drosophila* late embryos and larvae. The study also analyzed the dynamic changes of testis cells and the regulation pattern of genes during the development of *Drosophila*. In the future, Stereo-seq can complement medical images and histopathological data to generate comprehensive atlases of health, disease, evolution and organ development, bringing about great changes in foundational and clinical research.

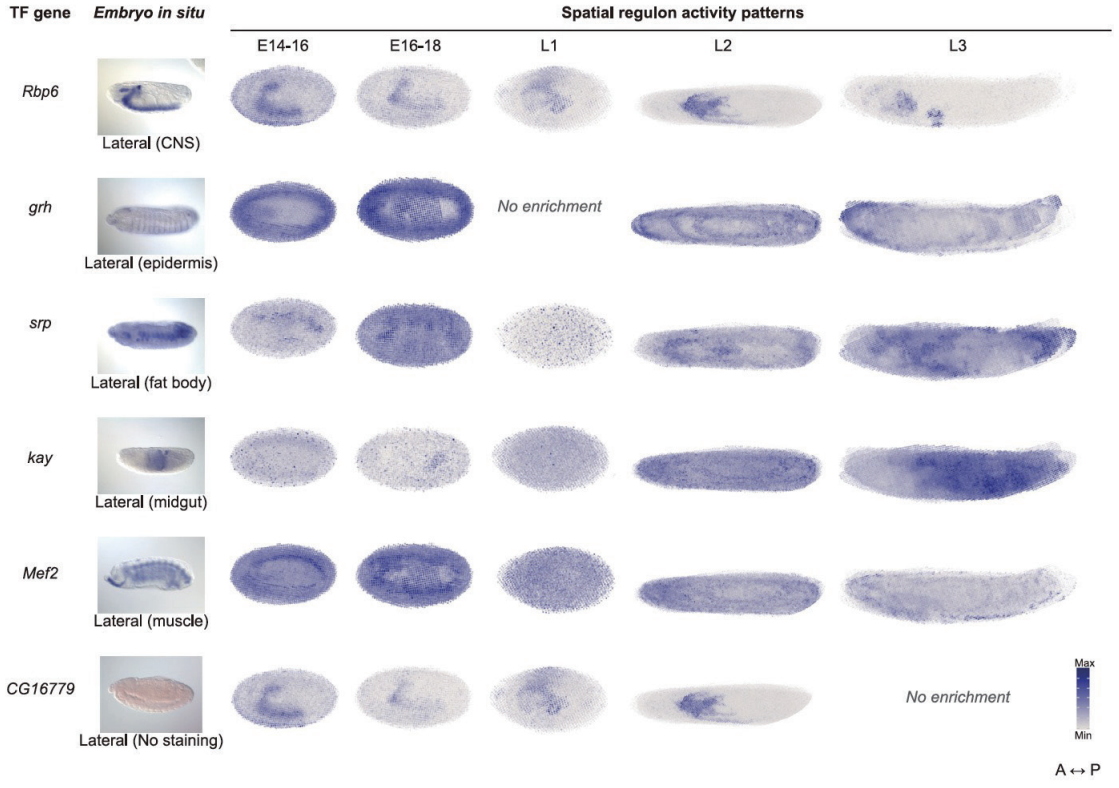

Fig. 7. Spatial dynamic patterns of key transcription factores in *Drosophila* embryos and larvae.

Reference

1. Wang, M. *et al.* High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae. *Dev.Cell* 57, 1271-1283 e1274, doi:10.1016/j.devcel.2022.04.006 (2022).
2. Chintapalli, V. R., Wang, J. & Dow, J. A. Using FlyAtlas to identify better *Drosophila* melanogaster models of human disease. *Nature Genet.* 39, 715-720, doi:10.1038/ng2049 (2007).
3. Leader, D. P., Krause, S. A., Pandit, A., Davies, S. A. & Dow, J. A. T. FlyAtlas 2: a new version of the *Drosophila* melanogaster expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic Acids Res.* 6, D809-D815,doi:10.1093/nar/gkx976 (2018).
4. Graveley, B. R. *et al.* The developmental transcriptome of *Drosophila* melanogaster. *Nature* 471, 473-479, doi:10.1038/nature09715 (2011).
5. Zhou, J. *et al.* Accurate genome-wide predictions of spatio-temporal gene expression during embryonic development. *PLoS Genet.* 15, e1008382, doi:10.1371/journal.pgen.1008382 (2019).
6. Karaiskos, N. *et al.* The *Drosophila* embryo at single-cell transcriptome resolution. *Science* 358, 194-199, doi:10.1126/science.aan3235 (2017).
7. Junker, J. P. *et al.* Genome-wide RNA Tomography in the zebrafish embryo. *Cell* 159, 662-675, doi:10.1016/j.cell.2014.09.038 (2014).
8. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 363, 1463-1467, doi:10.1126/science.aaw1219 (2019). H0088 (2002).
9. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185, 1777-1792 e1721, doi:10.1016/j.cell.2022.04.003 (2022).
10. Zeira, R., Land, M., Strzalkowski, A. & Raphael, B. J. Alignment and integration of spatial transcriptomics data. *Nat. Methods* 19, 567-575,doi:10.1038/s41592-022-01459-6 (2022).
11. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).
12. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* 19,doi:10.1186/s13059-017-1382-0 (2018).
13. Van Der Walt, S. *et al.* Scikit-image: Image processing in python. *PeerJ* 2014,doi:10.7717/peerj.453 (2014).
14. Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* 30, 1323-1341, doi:10.1016/j.mri.2012.05.001 (2012).
15. Qiu, X. *et al.* Mapping transcriptomic vector fields of single cells. *Cell* 185, 690-711.e645,doi:10.1016/j.cell.2021.12.045 (2022).
16. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185, 1777-1792.e1721, doi:10.1016/j.cell.2022.04.003 (2022).
17. Tomancak, P. *et al.* Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol.* 3, RESEARCH0088 (2002).
18. Lécuyer, E. *et al.* Global Analysis of mRNA Localization Reveals a Prominent Role in Organizing Cellular Architecture and Function. *Cell* 131, 174-187, doi:10.1016/j.cell.2007.08.003 (2007).
19. DeTomaso, D. & Yosef, N. Hotspot identifies informative gene modules across modalities of single-cell genomics. *Cell Syst.* 12, 446-456 e449, doi:10.1016/j.cels.2021.04.005 (2021).
20. Mahadevaraju, S. *et al.* Dynamic sex chromosome expression in *Drosophila* male germ cells. *Nat. Commun.* 12, 892,doi:10.1038/s41467-021-20897- y (2021).
21. Abel, T., Michelson, A. M. & Maniatis, T. A *Drosophila* GATA family member that binds to Adh regulatory sequences is expressed in the developing fat body. *Development* 119, 623-633,doi:10.1242/dev.119.3.623 (1993).
22. Bour, B. A. *et al. Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes Dev.* 9, 730-741, doi:10.1101/gad.9.6.730 (1995).

FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.

Recommended Ordering Information

Category	Product	Cat. NO.
Instruments	DNBSEQ-T7RS Genetic Sequencer	900-000128-00
Software	Data Center Appliance	900-000444-00
Library Prep	DNBelab C Series High-throughput Single-cell RNA Library Preparation Set V2.0	940-000519-00
Sequencing Reagents	DNBSEQ-T7RS High-throughput Sequencing Set (FCL PE100) V2.0	1000028455

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com

The copyright of this brochure is solely owned by MGI Tech Co. Ltd.. The information included in this brochure or part of, including but not limited to interior design, cover design and icons, is strictly forbidden to be reproduced or transmitted in any form, by any means (e.g. electronic, photocopying, recording, translating or otherwise) without the prior written permission by MGI Tech Co., Ltd.. All the trademarks or icons in the brochure are the intellectual property of MGI Tech Co., Ltd. and their respective producers.

*1. For StandardMPS and CoolMPS: Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, Spain, UK, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, Portugal, Austria and Romania. Unless otherwise informed, StandardMPS sequencing reagents, and sequencers for use with such reagents are not available in Hong Kong. No purchase orders for StandardMPS products will be accepted in the USA until after January 1, 2023.

2. For HotMPS sequencers: This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.

3. For HotMPS reagents: This sequencing reagent is only available in selected countries.