Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis

Liu, Chang; Li, Rui; Li, Young; Lin, Xiumei; Zhao, Kaichen; Liu, Qun; Wang, Shuowen; Yang, Xueqian; Shi, Xuyang; Ma, Yuting; Pei, Chenyu; Wang, Hui; Bao, Wendai; Hui, Junhou; Yang, Tao; Xu, Zhicheng; Lai, Tingting; Berbéroglu, Michael Arman; Sahu, Sunil Kumar; Esteban, Miguel A.; Ma, Kailong; Fan, Guangyi; Li, Yuxiang; Liu, Shiping; Chen, Ao; Xu, Xun; Dong, Zhiqiang; Liu, Longqi

Published in:
Developmental Cell

DOI:
10.1016/j.devcel.2022.04.009

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis

Graphical abstract

Highlights
- Stereo-seq is used to generate a ZESTA
- Co-varying spatial gene modules reveal the interactions between functional regions
- Integration of Stereo-seq and scRNA-seq builds spatial developmental trajectories
- Spatiotemporal ligand-receptor dynamics provides insights to regulatory mechanisms

Authors
Chang Liu, Rui Li, Young Li, ..., Xun Xu, Zhiqiang Dong, Longqi Liu

Correspondence
xuxun@genomics.cn (X.X.), dongz@mail.hzau.edu.cn (Z.D.), liulongqi@genomics.cn (L.L.)

In brief
Liu et al. employ Stereo-seq to dissect the spatiotemporal transcriptomic dynamics in developing zebrafish embryos. This study provides a spatially resolved resource for the research of vertebrate embryogenesis and also helps to uncover how molecular and cellular interactions contribute to zebrafish embryogenesis.
Resource

Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis

Chang Liu,1,8,13 Rui Li,1,8,13 Young Li,1,8,13 Xiumei Lin,1,4,8,13 Kaichen Zhao,2,13 Qun Liu,1,5,13 Shuowen Wang,2,3 Xueqian Yang,2 Xuyang Shi,1,4,8 Yuting Ma,1,4 Chenyu Pei,2 Hui Wang,2 Wendai Bao,2 Junhou Hui,1 Tao Yang,6 Zhicheng Xu,1 Tingting Lai,1 Michael Arman Berberoglu,6 Sunil Kumar Sahu,1 Miguel A. Esteban,1,9,10,11 Kailong Ma,1 Guangyi Fan,1,8 Xuyang Li,1 Shiting Liu,1,8 Ao Chen,1,7 Xun Xu,1,12* Zhiqiang Dong,2,3* and Longqi Liu1,8,14,*

1BGI-Shenzhen, Shenzhen 518083, China
2College of Biomedicine and Health, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, Hubei 430070, China
3Brain Research Institute, Taihe Hospital, Hubei University of Medicine, Shiyan, Hubei 442000, China
4College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
5BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China
6China National GeneBank, Shenzhen, Guangdong 518120, China
7Department of Biology, University of Copenhagen, Copenhagen 2200, Denmark
8BGI Key Laboratory of Single-Cell Omics, Shenzhen 518083, China
9Laboratory of Integrative Biology, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
10CAS 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
11Institute of Stem Cells and Regeneration, Chinese Academy of Sciences, Beijing 100101, China
12Guangdong Provincial Key Laboratory of Genome Read and Write, Shenzhen 518120, China
13These authors contributed equally
14Lead contact
*Correspondence: xuxun@genomics.cn (X.X.), dongz@mail.hzau.edu.cn (Z.D.), liulongqi@genomics.cn (L.L.)

https://doi.org/10.1016/j.devcel.2022.04.009

SUMMARY

A major challenge in understanding vertebrate embryogenesis is the lack of topographical transcriptomic information that can help correlate microenvironmental cues within the hierarchy of cell-fate decisions. Here, we employed Stereo-seq to profile 91 zebrafish embryo sections covering six critical time points during the first 24 h of development, obtaining a total of 152,977 spots at a resolution of 10 \( \times \) 10 \( \times \) 15 \( \mu \)m\(^3\) (close to cellular size) with spatial coordinates. Meanwhile, we identified spatial modules and co-varying genes for specific tissue organizations. By performing the integrated analysis of the Stereo-seq and scRNA-seq data from each time point, we reconstructed the spatially resolved developmental trajectories of cell-fate transitions and molecular changes during zebrafish embryogenesis. We further investigated the spatial distribution of ligand-receptor pairs and identified potentially important interactions during zebrafish embryonic development. Our study constitutes a fundamental reference for further studies aiming to understand vertebrate development.

INTRODUCTION

Vertebrate embryogenesis is an intricate and dynamic process with intense gene-expression changes and frequent cell-state transitions within short time windows. Extrinsic and intrinsic cues, including transcription factors (TFs), morphogens, signaling pathways, and signals from the extracellular matrix (ECM), play pivotal roles in determining different cell fates that present distinct morphologies, spatial positions, and functions (Bardot and Hadjantonakis, 2020; Marlow, 2020; Vining and Mooney, 2017). How these regulatory factors spatially interact and function together to induce a complex vertebrate embryo in a precisely controlled manner is one of the fundamental questions about embryogenesis that demands further investigation. In particular, understanding how ligand-receptor pairs spatially interact to switch on/off specific signaling pathways, such as Notch, Wnt, sonic hedgehog (SHH), and TGF-\( \beta \) during vertebrate embryogenesis is crucial but poorly documented.

The zebrafish is a widely used model organism for studying vertebrate embryonic development thanks to its fast development, embryonic transparency, and accessibility to both physical and genetic manipulation. Advances in sequencing...
technologies have made possible the assembly of single-cell atlases of model organisms during development, which enables genome-wide profiling of multimodal information including gene expression, epigenetic state, and protein levels in individual cells (Cao et al., 2017; Han et al., 2020; Karaiskos et al., 2017; Steitzer et al., 2015; The Tabula Muris Consortium et al., 2018; Trevino et al., 2020). Efforts have been made to map the gene-expression landscapes during zebrafish embryogenesis and developmental trajectories have been constructed and defined the transcriptomic states of cells as they acquire their fates (Briggs et al., 2018; Farrell et al., 2018; Wagner et al., 2018). However, how the time course transcriptomic states correlate with each other in the background of spatial localization in a complex, developing zebrafish embryo remains elusive and the spatial organization of different cell types in complex tissues remains poorly understood due to the limitation of current spatial transcriptomic technologies (Li et al., 2021; Liu et al., 2020; Rodrigues et al., 2020; Stickels et al., 2021; Vickovic et al., 2019). Based on traditional in situ hybridization (ISH) technology, a representative database of gene expression in the developing zebrafish embryos and larvae, the zebrafish information network (ZFIn) (Sprague et al., 2003) is available for scientists to examine the expression patterns of specific genes, but the global changes of transcripts cannot be explored. A spatial transcriptomic technology termed Tomo-seq has been applied in zebrafish embryos, but this technology is based on bulk sequencing of cryosections and the spatial information on the tissue sections is missing (Holler et al., 2021; Junker et al., 2014). Here, we adopted spatial enhanced resolution omics-sequencing (Stereo-seq) (Chen et al., 2021) to dissect the spatiotemporal gene-expression landscapes during the first 24 h of zebrafish embryo development. Our study provides important data resources for the research of gene expression, cellular organization, and regulatory networks during zebrafish embryogenesis.

RESULTS

Generation of the spatiotemporal transcriptomic data by Stereo-seq in zebrafish embryos

Zebrafish embryos of different stages, including 3.3, 5.25, 10, 12, 18, and 24 h post-fertilization (hpf), were harvested for spatial transcriptomic analysis (Figure 1A, top). Embryos were dechondionated and embedded in optimal cutting temperature (OCT) compound for the preparation of sagittal cryosections with a thickness of 15 μm (close to cellular size). High-resolution Stereo-seq chips (spot size, 220 nm; center-to-center size, 715 nm; chip size, 1 cm²) were used for in situ capture of RNA transcripts (Figure 1A, bottom). In order to avoid incomplete sampling of cell types and batch effects, we placed multiple sections from one single embryo at each stage for Stereo-seq library preparation. Taking a 24-hpf embryo as an example, 17 sections of the same embryo were attached onto two 1-cm² chips and subjected to Stereo-seq (Figures 1B and 1C). All the sections analyzed in this study were numbered at each developmental stage. The position of each section and the total numbers of sections processed at each stage are shown in Figures S1A and 1G.

The brightfield image of each developmental stage embryo used for sagittal cryosections, as well as nucleic-acid dye-staining image of each section can be viewed in our data portal (https://db.cngb.org/stomics/zesta). We applied a grid-based strategy, which aggregates a number of DNA nanoball (DNB) spots, to cluster the spatial signature into substructures (see STAR Methods). As expected, by comparison of spatial clustering with different grid (bin) sizes ranging from 143 μm (bin 200, 200 × 200 DNB) to 10 μm (bin 15, 15 × 15 DNB, which is equivalent to ~1 cell diameter), we found more refined clustering with increased cluster numbers and higher accuracy (Figure 1D). Meanwhile, the number of transcripts and genes detected among different sections of the 24-hpf embryo was comparable (Figures 1E and 1F), with an average of 2,576 unique molecular identifiers (UMIs) and 786 genes at bin 15 (Figure 1G). The number of UMIs or genes of sections at each time point were overall sufficient for further analysis across the six developmental stages (Figure 1G).

Thus, Stereo-seq, combined with a multiple-section strategy, enables the spatiotemporal transcriptomic analysis of small-size samples such as zebrafish embryos. This allowed us to create a high-quality zebrafish embryogenesis spatiotemporal transcriptomic atlas (ZESTA).

Spatial clustering and molecular characterization of the zebrafish embryogenesis spatiotemporal transcriptomic atlas

We pooled all the sections from each stage for unsupervised clustering (see STAR Methods), revealing spatial heterogeneity in each embryo section and regional specification during development (Figure 2A). As Stereo-seq captured a high density of signals, we were able to identify more elaborate structures with deeper clustering (Figures 2B, 2D, S1B, and S1D–S1H; Table S1). Because the cell-type signatures vary across sections (Table S2), the sections were selected to display according to the needs of each analysis. To validate the sensitivity and accuracy of Stereo-seq for transcript capture in zebrafish embryos, we randomly selected 100 genes with ISH results in ZFIN and displayed the spatial expression patterns at 24-hpf (Data S1); 95 genes could be detected in at least one of the sections from
this stage. Of note, the spatial patterns were more obvious in genes with higher average-expression levels and became less clear as the expression level declined. We finally identified 83 genes showing similar patterns with ISH data in at least one of the sections and 12 genes with very low expression showing inconsistent patterns (Data S1). This result indicates that Stereo-seq could accurately reflect expression patterns of genes with high or decent expression level.

At 3.3-hpf, three cell types with different spatial distributions were identified: marginal blastomere, superficial blastomere, and deep blastomere (Figures 2B and S1B). As disclosed by the pseudotime analysis using Monocle 2 (Figure S1C; Qiu et al., 2017), the differentiation sequence was from marginal blastomere to deep blastomere and superficial blastomere, which is consistent with a previous study (Kimmel et al., 1995). Moreover, we found some cell-type-specific genes, such as mad2li1 and thy1, that have not been reported as marker genes of superficial blastomere at this stage. We further validated the gene-expression patterns by ISH on the classic marker gene grhl3 (Miles et al., 2017) and the new marker gene thy1 for superficial blastomere (Figure 2C). Interestingly, more detailed structures could be identified with clustering for specific regions. For example, the pronephros of the 24-hpf embryo can be divided into anterior pronephros and posterior pronephros (Figure 2E, top). The spinal cord, floor plate, and notochord are interacting regions and could be separated precisely in the embryo sections, as shown by specific expression of marker genes including pdm8, ntn1b, and ndt5, respectively (Figure 2E, bottom).

To build a spatial map of the nervous system, we extracted the neural components from the Stereo-seq data of the 18- and 24-hpf embryos for more detailed clustering (Figures 2F, 2G, and S1B; Table S3). We observed subclusters corresponding to the features of different cell types distributed in different anatomical regions such as telencephalon, midbrain, and hindbrain. For example, regions showing signatures corresponding to neural stem cells (NSCs) and neural intermediate progenitor cells (nIPCs) (Figures 2F and 2G). In the 24-hpf hindbrain, we observed that the NSCs were mainly enriched at the dorsal side, while the nIPCs were distributed at the ventral side, which was consistent with previous reports (Figure 2H; Tambalo et al., 2020). To study the spatial heterogeneity of differentiation potential of each progenitor cell type, we performed pseudotime analysis which identified two branch points of differentiation, one from early NSCs and the other from nIPCs (Figure 2I). Consistent with a previous report (Paradaen and Huttner, 2014), the early NSCs (state 1) transition to more differentiated states, i.e., NSCs (state 2) and nIPCs (state 3). Interestingly, the state 1 NSCs were mostly located at anterior and dorsal regions, while the state 2 NSCs located at posterior and ventral regions (Figure 2J). Differentially expressed gene (DEG) analysis revealed that genes related to the cell cycle and stemness (e.g., pcna and notch3) are largely enriched in state 1 NSCs (Figure 1K). State 2 NSCs start to express genes related to the ribosomal proteins involved in mRNA translation (e.g., rps21 and rpl35a), suggesting a more differentiated state compared with state 1 NSCs (Figure 1K). Taken together, our analyses indicate that the hindbrain NSCs show nonsynchronous differentiation in both anterior versus posterior and dorsal versus ventral directions.

Our Stereo-seq constituted a unique resource of ZESTA with unprecedented resolution, and the publicly available interactive data portal can be accessed at https://db.cngb.org/stomics/zesta/.

Spatial gene modules uncover the interaction of different spatial regions

The spatial patterns of gene expression depict different functional regions across embryo sections. In order to identify co-varying genes showing similar spatial distribution, we applied Hotspot (DeTomaso and Yosef, 2021) to identify gene modules from our Stereo-seq dataset. We identified 19 spatial modules for the 24-hpf embryo sections. Gene ontology (GO) enrichment analysis on the spatially correlated genes of each module confirmed the functions of these spatial modules (Figure 3A), which also showed consistency with the spatial clusters

**Figure 2. A spatial transcriptomic atlas of the developing zebrafish embryo**

(A) Unsupervised clustering of the zebrafish embryo section across sequential developmental stages analyzed by Stereo-seq at bin-15. Bins are colored by different regions. Scale bars, 0.25 mm. In this study, all sections from 10- to 24-hpf are displayed with head on the upper left and tail on the bottom right. The “#number” next to the sections represents the number given to each section, and all the sections in the following figures are marked with the “#number.”

(B) Unsupervised clustering of the 3.3-hpf zebrafish embryo sections at bin-15. Bins are colored by spatial identities inferred from expressed marker genes. Scale bars, 0.25 mm.

(C) The in situ hybridization images and the expression patterns of marker genes grhl3 and thy1 for superficial blastomere at 3.3-hpf. Scale bars, 0.25 mm.

(D) Unsupervised clustering of the 24-hpf zebrafish embryo sections at bin-15. Bins are colored by spatial identities inferred from expressed marker genes. Scale bars, 0.25 mm.

(E) Spatial visualization of indicated areas of the 24-hpf embryo on the left: detailed anatomical structures identified in the pronephros, and a combined structure including the spinal cord, the floor plate, and the notochord. The expression patterns of marker genes for each anatomical structure are shown on the right. Scale bars, 0.1 mm.

(F) Unsupervised subclustering of the 24-hpf zebrafish nervous system. Bins are colored by spatial identities and cell types inferred from expressed marker genes. The subcluster colors are the same as the font colors of cluster names in (G). Scale bars, 0.25 mm.

(G) Heatmap shows the mean expression level of marker genes between the indicated clusters of the zebrafish nervous system.

(H) Spatial visualization of the hindbrain at 24-hpf, from left to right including indicated cell types and the expression of marker genes of different cell types. Scale bars, 0.25 mm.

(I) Pseudotime analysis of hindbrain-related clusters as performed by Monocle 2. Bins are colored by cell types on the top (color legends of cell types are the same as those in Figure 2H) and colored by pseudotime stages on the bottom.

(J) Spatial visualization of detailed structures of pseudotime states identified by Monocle 2 in the hindbrain on different sections at 24-hpf. Scale bars, 0.25 mm.

(K) Heatmap shows the mean expression level of the top 30 DEGs different pseudotime states 1, 2, and 3. Representative genes related to each state are highlighted by red color.
Spatial visualization of modules showed that the distribution of spatial modules reasonably matched their region-specific biological functions (Figures 3A and 3C). For example, module 10 (M10, blood vasculature) gathered genes involved in sprouting angiogenesis, and M8 (eye) was abundant in genes related to eye development. For the other developmental stages, we identified 2 spatial modules for 3.3-, 4 for 5.25-, 13 for 10-, 17 for 12-, and 14 for 18-hpf (Figure S2). Similarly, the GO function enrichment and spatial location reasonably matched with each other for most spatial modules (Figure S2). Interestingly, some clusters contained multiple gene-expression modules; for example, at 10-hpf, we could divide anterior and posterior neural keel (Figure S1E) into subspatial modules including proliferative neural keel (M6), optic vesicle (M9), neural crest (M11), midbrain-hindbrain boundary neural keel (M12), and anterior neural keel (M13) (Figures S2E and S2F).

In contrast, we found that a single gene module could correspond to multiple subregions. For instance, M12 in 24-hpf embryo spread across various clusters including pigment cells, erythroid lineage cells, and sensory organs (otic vesicle and eye) (Figures 3B and 3D). GO enrichment of this module revealed functions related to the development of pigment cells, mesenchymal cells, and the peripheral nervous system (PNS), and to the proliferation and migration of cells, which were characterized as different differentiation states of the early neural crest (Figure 3E). In order to explore the connection among these four cell types/organisms, the search tool for retrieval of interacting genes (STRING) was employed to seek potential gene interaction which was visualized with Cytoscape (Figure 3F; Shannon et al., 2003). Among these genes, we highlighted TFs in the network. Intriguingly, we found that the TF gene sox10 intensively interacts with genes from these cell types and organs, consistent with the reported function of sox10 (Rocha et al., 2020). As expected, projection of specific TFs from pigment cells and mesenchymal cells on the embryo sections revealed consistent spatial distributions with these cell-fate destinations (Figure 3G). These results suggested a complex regulatory network underlying the differentiation states of the neural crest. For example, β-catenin (ctnmb1) was enriched in the esenchyma-related submodule, and it showed extensive interactions with other submodules (Figure 3F), consistent with previous reports on Wnt signaling regulating the differentiation of neural crest (Hari et al., 2012).

Secreting SHH, which is produced by notochord and floor plate, has been reported to function as a key morphogen regulating organogenesis during early embryonic development (Male et al., 2020; Murtaugh et al., 1999). To investigate the interaction of shh-secreting tissues with the neighboring regions, we constructed the interaction network of shha/b in M13 (shh-secret tissues located in floor plate and notochord, Figures 3A and 3H, right) and the neighboring modules, M1 and M5, which correspond to the musculature system and the spinal cord, respectively (Figures 3A and 3H, left and middle). Within the potential interacting genes revealed by STRING, we found that myod1 and hhatab in M1 showed interaction with shha and shhb in M13 (Figure 3I). Previous studies have found that shh plays a role in muscle development and adult myogenesis by regulating the activity of myod1 (Voronova et al., 2013), and hhatab also has been reported to be essential for muscle development (Van et al., 2015). While hhatab has been rarely reported to interact with shh to play a role in muscle development, the interaction of hhatab and shh disclosed by our spatial module analysis could potentially be a novel regulatory mechanism of muscle development. In contrast, fgf13 and two other NSCs maker genes (gfp and pou4f2) in M5 show interaction with shha/b (Figure 3I) which was consistent with the previously reported role of SHH in regulating the development of the spinal cord (Yu et al., 2013). The interaction of these genes was further confirmed by their tightly adjacent characters of spatial distribution (Figure 3J).

In summary, spatial modules depict the functional regions or subregions in an organism, provide a powerful tool to explore the gene interaction between different spatial regions, and serve as a reliable resource to discover novel functions as well as unknown interactions of genes.

**Construction of spatially resolved developmental trajectories by integrated analysis of scRNA-seq and Stereo-seq data**

Both our Stereo-seq data and previously published scRNA-seq data indicate that developmental time is a strong source of variation due mainly to active transition between different cell states during embryogenesis. To construct the spatiotemporal developmental trajectories of embryonic cells, we further performed droplet-based scRNA-seq (Liu et al., 2019; see STAR Methods) with zebrafish embryos at the same developmental time points.
Developmental Cell

Resource

A

B

C

D

Branch: Central Nervous System Module

Gene: mael

E

Branch: Pigment Cells Module

Gene: tfap2a

F

3.3 hpf 5.25 hpf 10 hpf 12 hpf 18 hpf 24 hpf

G

Late Blastodisc

H

Lateral Plate Meosderm

I

Neural Crest

scRNA-seq developmental trajectory

SPOTlight spatial distribution

TFs target expression pattern

Developmental Cell 57, 1284–1298, May 23, 2022

1291
as in the Stereo-seq profiling (Figure 4A). We obtained 86,307 cells from embryos of the 6 developmental stages and identified 71 cell types, which are mostly consistent with previous studies (Farrell et al., 2018; Wagner et al., 2018; Figures 4A, S3A, and S3B; Table S4). To introduce spatial information to the cell clusters of scRNA-seq, we conducted an integrated analysis of the scRNA-seq and Stereo-seq data by applying SPOTlight (Elosua-Bayes et al., 2021) to calculate the cell composition of each bin in the Stereo-seq data (Figure 4A). All the scRNA-seq and integrated analysis data are also deposited to the ZESTA database (https://db.cngb.org/stomics/zesta/).

At 3.3-hpf, the emergence of two cell types was observed (Figure S3C), for which we used Monocle 2 to undertake pseudotime analysis to determine their differential identities. Based on the pseudotime relationship, we determined the two cell types to be early blastodisc and late blastodisc respectively (Figure S3D). The results of SPOTlight showed that early blastodisc was distributed on the ventral side of the embryo, while late blastodisc was mainly on the dorsal side (Figure S3E), which was consistent with the spatial location of the 3.3-hpf clusters from the Stereo-seq data (Figure S1B), suggesting that the integrated analysis could accurately endow the cell clusters of scRNA-seq with spatial information. To construct the developmental trajectory, we integrated scRNA-seq data of adjacent time points using Harmony (Korsunsky et al., 2019) and applied KNN mapping to predict the developmental fate of each cell of the earlier time point (see STAR Methods). We then projected the developing trajectory using the Sankey diagram (Figure S3F; Table S4).

Although single-cell studies have mapped the cell types and their molecular features during vertebrate embryogenesis, the dynamics of the spatial organization of cells are still poorly understood (Farrell et al., 2018; Wagner et al., 2018). To introduce spatial information to the cell-fate lineages, SPOTlight was applied to project the developmental trajectories to the embryo sections (Figures 4B, 4C, S4A, and S4B). These results revealed the spatial dynamics of different developmental trajectories. Taking the two developmental branches of the presumptive ectoderm as an example, both the central nervous system branch and the pigment cells branch developed from the presumptive ectoderm which appeared at 5.25-hpf. The cell fates started to be determined toward two directions at 5.25-hpf, and the tendency became even more obvious at 10-hpf, as shown by the separated-count distribution curves (Figure 4C; see also STAR Methods). We calculated the DEGs of each branch at each developmental stage. The expression patterns of these gene modules and representative genes were consistent with the spatial distribution of corresponding branches (Figures 4D, 4E, and S4C; see also STAR Methods). For all branches, spatial distributions of SPOTlight projections, gene modules and individual gene-expression patterns were uploaded to the ZESTA database (https://db.cngb.org/stomics/zesta/).

A prominent phenomenon of the developmental trajectory is that some cell types bifurcate into multiple branches at certain time points (Figure 4F). We noticed that the distributions of some developmental directions seem to be spatially separated and distinguishable in the progenitor cell type at an early stage (Figures 4G–4I), because we observed that the count distribution peaks of different branches can be distinguished early (Figures 4G–4I, left). We then integrated the scRNA-seq data and applied the single-cell regulatory network inference and clustering (SCENIC) algorithm (Aibar et al., 2017) to predict TFs with high activity at each developmental direction (see STAR Methods). Moreover, we visualized the regulatory activities of these key TFs by displaying the expression of target genes on embryo sections, which are largely consistent with the spatial distribution of the corresponding cell type (Figures 4G–4I, right). For example, different TFs were identified to regulate the segregated developmental directions of the 12-hpf lateral plate mesoderm, which gives rise to the cardiovascular system, erythroid lineage cells, and endothelial cells. The nkh2.5 gene is predicted as a crucial TF for the development of the cardiovascular system, as cdx4 is predicted for erythroid lineage cells (Figure 4H, right). Similarly, different TFs were identified for the separate developmental lineages of the 3.3-hpf late blastodisc and 18-hpf neural crest (Figures 4G and 4I, right).

Interestingly, the 18- and 24-hpf nervous systems could be divided into subgroups (Figures 2F, 2G, and S1I), which allowed us to construct the spatially resolved trajectory between these subclusters (Figure S4D). Notably, we identified a NSCs cluster from Stereo-seq data (labeled as Stereo-NSCs-1) (Figure S4E, left) as well as a similar cluster from scRNA-seq data (labeled as sc-NSCs-1) at 18-hpf (Figure S4F). By applying SPOTlight to project cluster “sc-NSCs-1” to the embryo sections (Figure S4G), we found that these two clusters from different data-sets are highly correlated. Intriguingly, the “Stereo-NSCs-1”
Dynamic spatiotemporal distribution of ligand-receptor pairs during zebrafish embryogenesis

The extrinsic cues, including the interaction between ligands and receptors, are crucial for the regulation of cell fate during embryogenesis. We therefore calculated the relative distance of different ligand-receptor pairs at all six developmental stages with the distance of adjacent bins (Figures 5A and S5A). To validate whether these potential interacting ligand-receptor pairs tend to be expressed close to each other than genes with random distribution, we compared the distances between ligand and receptor-expressing bins with those of randomly selected bins (see STAR Methods). As expected, the receptor-expressing bins were closer to the ligand-expressing bins than those randomly selected bins, especially at later time points (Figure S5B). To find the ligand-receptor pairs that are essential for zebrafish development, we regarded the ligand-receptor pairs with an average relative distance less than 5 bins in at least one time point as having strong potential interaction. Then we tested the correlation between receptors and ligands across spatial clusters and only kept the ligand-receptor pairs that correlated significantly (Pearson’s r > 0.5, p value < 0.05, Figures S5C and S5D). We found 21 ligand-receptor pairs that passed these strict filtering, including some Notch ligands and receptor-expressing bins with a relative distance of less than 2 bins (Figure 5C, middle), and illustrated the corresponding spatial clusters (Figures S5F–S5I). We further analyzed target genes of the Notch pathway and found consistent spatial Notch activities with the spatial expression dynamics of Notch ligand-receptor pairs (Figure 5C, bottom). Our results revealed a highly dynamic spatial distribution of Notch components expression and Notch activity during zebrafish embryogenesis, which are consistent with previous reports that Notch might control somite formation by regulating the synchronous development of somite cells before 18-hpf (ÖZbudak and Lewis, 2008). Along with the development progressing, the Notch signal tends to be highly enriched in the nervous system at 18- and 24-hpf and play an important role in the maintenance and differentiation of NSCs (Mayoshi et al., 2010).

We further explored gene modules that are spatially correlated with Notch signaling pathway using hotspot analysis on the Stereo-seq data of 10-, 12-, 18-, and 24-hpf zebrafish embryos (Figures 5D–5G and S5J–S5L; see also STAR Methods). At 10-hpf, one main Notch-correlated spatial module with a specific spatial pattern is identified with distribution in the somite-forming region and contains genes such as *msgn1*, *myf5*, and *cdx4* in M1, which are functionally associated with muscle development (Figure 5D). At 12-hpf, in addition to M3 gathering in the somitogenesis region, M1 and M2 started to enrich in the nervous system, and they comprised genes that were important in neurogenesis and brain development, i.e., *egr2b*, *irx7*, and *neurog1* (Figures 5E and S5J). More modules related to nervous system development showed up at later stages. For example, modules enriched at 18-hpf were mainly related to the development of the brain and the spinal cord (Figures 5F and S5K). At 24-hpf, the Notch-correlated spatial modules were also enriched in the nervous system (Figures 5G and S5L). M1 contains genes such as *pcna* and *mcm5*, which are related to cell cycle, RNA splicing, and DNA metabolism, indicating that the dominant role of this...
module is to regulate the development of neural progenitor cells (Ino and Chiba, 2000; Yano et al., 2010). In order to identify the dynamic changes of candidate genes that directly or indirectly interact with the Notch pathway at each developmental stage, we further characterized the co-varying tendency of module genes relative to the Notch pathway from 10- to 24-hpf (Figure S5H). Therefore, the spatial gene modules showed co-varying expression patterns and functional correlation with the Notch signaling pathway at specific time points, which indicates that the genes involved can very likely interact with Notch signaling, either directly or indirectly, to play an important role in the development of defined cell types.

Besides well-studied pathways, our data also revealed some novel working models of ligand-receptor interactions during embryogenesis. For example, Midkine (mdk) has been reported to be essential in neurogenesis during embryonic development by promoting cell growth, survival, and migration (Winkler et al., 2003). We firstly uncovered that the mdk family ligand mdka could potentially interact with receptors such as sdc4, c1d, and lrp2a in the nervous system at different developmental stages (Figures SS5M–SS5O). Secondly, the interactions of mdka with the two known co-receptors (sdc4 and c1d) (Sarrazin et al., 2011; Zamir et al., 1997) began to appear at 12-hpf with the similar spatial interaction patterns (Figures SS5M and SSN), which suggested that sdc4 and c1d might cooperate with mdka to participate in neurogenesis regulation in early nervous system development. Additionally, the interactions between mdka and lrp2a were mainly concentrated in the spinal cord and the brain at 24-hpf, while few interactions were found in early development stages (3.3- to 18-hpf) (Figure S5O), which suggested that the regulation of the nervous system by mdk and lrp2a might be mainly after the onset of the interactions of mdka with sdc4 and c1d. Furthermore, we also found a few interactions of mdka and lrp2a in pronephric at 24-hpf (Figure S5O).

Therefore, our study revealed important spatiotemporal interactions of ligand-receptor pairs during zebrafish embryogenesis and provided clues to investigate potential developmental regulatory mechanisms.

We thereby composed a spatiotemporal expression atlas of various ligand-receptor pairs and provided a useful database for the investigation of the dynamics of different signaling pathways as exemplified by the comprehensive analysis of the Notch pathway.

**DISCUSSION**

The mapping of a spatial transcriptomic landscape is essential for understanding vertebrate embryogenesis; however, this has never been achieved due to the limitation of current spatial transcriptomic technologies. Our Stereo-seq now enables the spatial transcriptomic landscape of zebrafish embryogenesis (Chen et al., 2021). In addition, the flexible size of the Stereo-seq chip allowed us to simultaneously profile multiple sections on one chip, which can greatly eliminate incomplete sampling of cell types and the batch effects introduced by separate experimental runs. The present study employed the Stereo-seq for the first time and drafted the dynamic spatiotemporal landscape of gene expression as well as spatial regulatory factors at 10 × 10 × 15-μm³ resolution (close to cellular size) during the first 24 h of zebrafish embryonic development. Our Stereo-seq data successfully distinguished anatomical structures such as the hypoblast, enveloping layer (EVL), hatching gland, and spinal cord at different developmental stages (Figure S1) and identified organ subdivisions such as the telencephalon, midbrain, midbrain-hindbrain boundary, and hindbrain of the brain (Figures 2F and 2G). Furthermore, Stereo-seq data could help us address questions that cannot be answered by single-cell sequencing data. For example, using our spatial atlas, we have identified two regional NSCs with different stemness and differentiation potential in the dorsal hindbrain (Figures 2H–2K). We have also identified a cluster of NSCs at 18-hpf that show the spatially distinct potential of differentiation toward distinct lineages, which indicates that spatial patterning by microenvironment of the NSCs is crucial for cell-type diversification (Figures S4D and S4E). Besides, our zebrafish embryogenesis atlas data can be searched by gene and anatomical regions for all sections using our interactive data portal (https://db.cngb.org/stomics/zebra/). However, we should be cautious of patterns in genes with low expression levels. For example, among 100 randomly selected genes, although 83 genes showed similar patterns with ISH results, we found discrepancies in 12 low-expressed genes and 5 undetected genes (Data S1).

Taking advantage of our Stereo-seq data, we identified spatial modules at different developmental stages and integrated the spatial modules with spatial regions to investigate key co-varying genes. Nevertheless, how the spatially correlated gene sets interact and mutually regulate each other both inside and outside the modules to form a regulatory network remains elusive, and unraveling it would improve our understanding of the molecular mechanisms of vertebrate embryogenesis. For example, the notochord/floor plate provides directional signals to the surrounding tissue during embryogenesis, in which SHH signal plays a critical role. We therefore constructed the potential interacting network of shta/shhb by analyzing the interaction of spatial module genes (Figures 3H and 3I), which could provide the rationale for further mechanistic investigations on the function of shta/shhb during organogenesis.

To accurately map the cell-type determination and developmental trajectory reconstruction, we have constructed the spatiotemporal developmental trajectories of the zebrafish embryo by combining the scRNA-seq and Stereo-seq data to delineate cell-state transitions with spatial coordinates. The developmental trajectories showed clear and specific spatiotemporal characteristics, which are consistent with previously known facts about zebrafish embryonic development. Besides illuminating the spatial separation of closely correlated cell types, our spatiotemporal developmental trajectories resolved some previously unanswered questions. For instance, it has been unclear when the lateral plate mesoderm starts to differentiate into different tissues. Our results disclosed that the fate of this structure was already decided at 12-hpf to differentiate into three different directions, namely, the cardiovascular system, erythroid lineage cells, and endothelial cells (Figures 4H and S4A).

TFs, together with their downstream gene networks, are one of the key factors that drive cell-fate transition during
embryogenesis. We investigated candidate TFs that potentially dominate the cell types at different time points in our spatiotemporal developmental trajectories. Our data identified known TFs, whose spatiotemporal distribution was consistent with their reported functions. The TF gene *cdx4* has been shown to be involved in the development of hematopoietic stem cells (Davidson and Zon, 2008) and our analysis showed that *cdx4* regulates the cell-fate transition from lateral plate mesoderm to erythroid lineage cells (Figure 4H, right). We also discovered a potential new role of certain TFs such as *foxn4* in the differentiation of the presumptive ectoderm from the blastodisc, while *foxn4* was previously known to function in the regulation of NSC division (Misra et al., 2014; Figure 4G, right).

Based on our high-resolution Stereo-seq data, we were able to precisely calculate the spatial distance between different ligands and receptors at embryo scale in zebrafish. Besides the Notch components, we uncovered interesting dynamics of many other ligand-receptor pairs. The lipid-metabolism-related ligand-receptor pairs interacted as early as 3.3-hpf such as *lrp5-aopoe* and *scarb1-aopoe*, while the pairs related to the development of the nervous system and ECM began to interact at a later stage such as *mdka-c1d* at 10-hpf and *coll1a1a-itga5* at 18-hpf (Figure 5B). The dynamic interactions of paired ligands and receptors are consistent with the roles of these genes in organogenesis regulation at different development stages. In contrast, the highly dynamic spatiotemporal changes of ligand-receptor expression suggested that different ligand-receptor pairs are precisely programed to coordinate the cell-fate determination and organogenesis during zebrafish embryonic development (Figures 5C and S5M–S5O). Through ligand-receptor analysis, we provided some hints to novel mechanisms of organogenesis regulation. Taking the *mdka* and *lrp2a*, for example, their interaction mainly enriched in the spinal cord and the brain at 24-hpf (Figure S5O), the function of which in the development of the nervous system is unclear. *Lrp1* has been reported to be expressed in the developing brain and spinal cord and participated in cell-fate specification and migration (Auderset et al., 2016). Our findings suggested that *lrp2a* may play a similar role to *lrp1*, and interact with *mdka* to regulate the neurogenesis of the spinal cord and the brain. Meanwhile, we also found the interaction of *mdka* and *lrp2a* in pronephric at 24-hpf (Figure S5O). *Lrp2a* has been reported in the pronephric epithelial cells (Elmonem et al., 2018) and *mdka* has been reported to play a role in epithelial-mesenchymal transition during organ development (Muramatsu, 2002). Therefore, we speculated that the interaction between *mdka* and *lrp2a* might mediate epithelial-mesenchymal transition, which is essential for the development of the renal system.

In conclusion, we demonstrated a spatiotemporal landscape of the transcriptional dynamics in the developing zebrafish embryo and provided a useful resource for studying the cellular and molecular mechanisms of germ-layer specification, organogenesis, and cell-fate determination. Future perspectives can include the construction of the spatiotemporal zebrafish developmental atlas at a longer time window with shorter intervals, a 3D reconstruction based on multiple sectioning strategy, and further exploration of the regulatory networks by integrating more spatial omics data such as the spatial genomics and chromatin accessibility.
support. T.Y., Z.X., and T.L. helped establish the online database. A.C. and Yuxiang Li gave the relevant advice. M.A.E., M.A.B., S.K.S., S.L., and X.X. participated in the manuscript editing and discussion. C.L., Z.D., and L.L. wrote the manuscript. All authors edited and approved the manuscript.

DECLARATION OF INTERESTS

The chip, procedure, and applications of Stereo-seq are covered in pending patents. Employees of BGI have a stock holding in BGI.

Received: September 24, 2021
Revised: February 6, 2022
Accepted: April 5, 2022
Published: May 4, 2022

REFERENCES


# STAR METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-DIG-AP Fab Fragments</td>
<td>Roche</td>
<td>Cat#11093274910; RRID: AB_2734716</td>
</tr>
<tr>
<td><strong>Chemicals, peptides, and recombinant proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7 transcriptase</td>
<td>Roche</td>
<td>Cat#10881767001</td>
</tr>
<tr>
<td>Paraformaldehyde</td>
<td>Sigma</td>
<td>Cat#P7629</td>
</tr>
<tr>
<td>Tricine solution</td>
<td>Biomebio</td>
<td>Cat#886-86-2</td>
</tr>
<tr>
<td>20X Saline-sodium citrate (SSC)</td>
<td>Thermo</td>
<td>Cat#AM9770</td>
</tr>
<tr>
<td>DIG-RNA labeling mix</td>
<td>Roche</td>
<td>Cat#1277073</td>
</tr>
<tr>
<td>Pronase E</td>
<td>Sigma</td>
<td>Cat#P5147-1G</td>
</tr>
<tr>
<td>Hi-Fetal bovine serum</td>
<td>Biological Industries</td>
<td>Cat#04-001-1ACS</td>
</tr>
<tr>
<td>Trypsin-EDTA (1×)</td>
<td>BioSharp</td>
<td>Cat#BL512A</td>
</tr>
<tr>
<td>Trypan blue</td>
<td>Thermo</td>
<td>Cat#15250061</td>
</tr>
<tr>
<td>Bovine serum albumin</td>
<td>Sigma</td>
<td>Cat#A8022-50G</td>
</tr>
<tr>
<td>Phosphate buffered saline</td>
<td>Meilunbio</td>
<td>Cat#MA0015</td>
</tr>
<tr>
<td>Tissue-Tek OCT</td>
<td>Sakura</td>
<td>Cat#4583</td>
</tr>
<tr>
<td>Nucleic acid dye</td>
<td>Thermo</td>
<td>Cat#Q10212</td>
</tr>
<tr>
<td><strong>Critical commercial assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOmics Gene Expression kit S1</td>
<td>BGI</td>
<td>Cat#1000028493</td>
</tr>
<tr>
<td>The DNBelab C Series High-throughput Single-Cell Library Preparation Kit</td>
<td>MGI</td>
<td>Cat#940-000047-00</td>
</tr>
<tr>
<td>AMPure XP beads</td>
<td>Vazyme</td>
<td>Cat#N411-03</td>
</tr>
<tr>
<td>Qubit ssDNA Assay Kit</td>
<td>Thermo</td>
<td>Cat#Q10212</td>
</tr>
<tr>
<td><strong>Deposited data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>In situ</em> hybridization data of zebrafish embryos</td>
<td>ZFIN</td>
<td><a href="http://zfin.org">http://zfin.org</a></td>
</tr>
<tr>
<td>Raw data of scRNA-seq</td>
<td>This paper</td>
<td>CNGB Nucleotide Sequence Archive: CNP0002220</td>
</tr>
<tr>
<td>Raw data of Stereo-seq</td>
<td>This paper</td>
<td>CNGB Nucleotide Sequence Archive: CNP0002220</td>
</tr>
<tr>
<td><strong>ZESTA</strong></td>
<td>This paper</td>
<td><a href="https://db.cngb.org/stomics/zesta/">https://db.cngb.org/stomics/zesta/</a></td>
</tr>
<tr>
<td><strong>Experimental models: Organisms/strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zebrafish embryos</td>
<td>China Zebrafish Resource Center</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Software and algorithms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seurat</td>
<td>Stuart et al., 2019</td>
<td><a href="https://github.com/satijalab/seurat">https://github.com/satijalab/seurat</a></td>
</tr>
<tr>
<td>Hotspot</td>
<td>DeTomaso and Yosef, 2021</td>
<td><a href="https://yoseflab.github.io/Hotspot/">https://yoseflab.github.io/Hotspot/</a></td>
</tr>
<tr>
<td>Cytoscape</td>
<td>Shannon et al., 2003</td>
<td><a href="https://cytoscape.org/">https://cytoscape.org/</a></td>
</tr>
<tr>
<td>pySCENIC</td>
<td>Abar et al., 2017</td>
<td><a href="https://github.com/aertslab/pySCENIC">https://github.com/aertslab/pySCENIC</a></td>
</tr>
<tr>
<td>Harmony</td>
<td>Korsunsky et al., 2019</td>
<td><a href="https://github.com/immunogenomics/harmony">https://github.com/immunogenomics/harmony</a></td>
</tr>
<tr>
<td>SPOTlight</td>
<td>Elosua-Bayes et al., 2021</td>
<td><a href="https://github.com/MarcElosua/SPOTlight">https://github.com/MarcElosua/SPOTlight</a></td>
</tr>
<tr>
<td>ggplot2</td>
<td>Wickham, 2011</td>
<td><a href="https://github.com/tidyverse/ggplot2">https://github.com/tidyverse/ggplot2</a></td>
</tr>
</tbody>
</table>
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 and droplet based scRNA-seq are commercially available.

Data and code availability
All raw data generated by Stereo-seq and scRNA-seq have been deposited to CNGB Nucleotide Sequence Archive: CNP0002220 (https://db.cngb.org/search/project/CNP0002220/). Additional data, including processed H5ad data, the brightfield images of each embryo, nucleic acid dye staining images of each section, the spatially resolved developmental trajectories data can be accessed from https://db.cngb.org/stomics/zesta/.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All relevant procedures involving animal experiments presented in this paper were approved by Animal Care and Use Committee of Huazhong Agriculture University (HZAUFI-2021-0001). Zebrafish embryos from AB wild-type crosses were collected at 3.3, 5.25, 10, 12, 18, and 24 h post-fertilization. After being dechorionated manually with forceps, embryos were anesthetized by soaking in tricaine solution (Bomeibio, 886-86-2). Embryos were then placed in cryomold with OCT (Sakura, 4583), and extra egg water was removed using a glass pipette. The embryos were oriented with a blunt metal needle to the right position. After the OCT freezed solidly on the flat surface of a block of dry ice, cryosections were cut at a thickness of 15 μm using a Leika CM1950 cryostat.

METHOD DETAILS

In situ hybridization
Probes of grhl3 and thy1 were synthetized by T7 transcriptase (Roche, 10881767001), and labeled by digoxigenun-RNA labeling mix (Roche, cat. no. 1277073). The whole-mount in suit hybridization was performed as described by Luo et al (He et al., 2020). In brief, 3.3-hpf embryo were careful dechorionated with fine forceps and fixed overnight at 4°C in 4% paraformaldehyde (Sigma, P7629). After that, these embryos were dehydrated with methanol and stored at -20°C for at least 2 h. Embryos were hybridized with probes in hybridization buffer, followed by blocking and incubation with anti-digoxigenun-AP fab fragments (Roche, 11093274910) in blocking buffer. Staining was developed using NBT/BCIP solution. Images of stained embryos. Images of stained embryos were obtained by Leica M205 fluorescence stereomicroscope.

Stereo-seq library preparation and sequencing
The STOmics Gene Expression kit S1 (BGI, 1000028493) was utilized according to the standard protocol V1.1 with minor modification (Chen et al., 2021). Tissue sections were adhered to the Stereo-seq chip, and incubated in -20°C methanol for 30 min fixation, followed by nucleic acid dye staining (Thermo fisher, Q10212) and imaging (Ti-7 Nikon Eclipse microscope). For permeabilization, tissue sections before 18-hpf were permeabilized at 37°C for 3 minutes, while tissue sections of 18-hpf and 24-hpf were permeabilized at 37°C for 5 minutes. The cDNA was purified using AMPure XP beads (Vazyme, N411-03). The indexed single-cell RNA-seq libraries were constructed according to the manufacturer’s protocol. The sequencing libraries were quantified by Qubit ssDNA Assay Kit (Thermo Fisher Scientific, Q10212). DNA nanoballs (DNBs) were loaded into the patterned Nano arrays and sequenced on MGI DNBSEQ-Tx sequencer (50 bp for read 1, 100 bp for read 2).

Zebrafish embryo collection and single cell isolation for scRNA-seq
The protocol for zebrafish embryos dissociation and single cell suspension isolation was adapted according to published protocol with a few modifications (Manoli and Driever, 2012). Briefly, 250–1000 embryos were collected and transferred into a Petridish containing Pronase E protease solution (Sigma-Aldrich, P5147-1G). When 20–30% embryos were hatched, 1 ml 56°C pre-heated Hi-Fetal Bovine Serum (Biological Industries, 04-001-1ACS) was added. Then embryos were washed once with 0.5X Danieau’s solution containing 10% Hi-FBS, and thrice with 0.5X Danieau’s solution. Then deyolking buffer was added and pipetted up and down until only the bodies of the embryos were visible. After dissociation with 1 x trypsin-EDTA solution (Bio­sharp, BL512A), incubation was stopped by adding Hi-FBS to a final concentration of 5%. Finally, single cells were resuspended in 0.04% BSA (Sigma, A8022-50G) in PBS (Meilunbio, MA0015).

Single cell RNA-seq library construction and sequencing
The DNBelab C Series High-throughput Single-Cell Library Preparation Kit (MGI, 940-000047-00) was utilized according to the manufacturer’s protocol (Liu et al., 2019). In brief, single cell suspensions were loaded into the chip for droplet generation. Then the droplets were gently removed from the collection and incubated at room temperature for 20 minutes to capture mRNA release from cells.
And then emulsion breakage and beads collection were performed. After reverse transcription, and cDNA amplification, purified PCR products were used for DNB generation and finally sequenced on an MGI DNBSEQ-Tx using the following read length: 41bp for read1, 100bp for read2, and 10bp for sample index.

QUANTIFICATION AND STATISTICAL ANALYSIS

Binning data of spatial Stereo-seq data
Raw data were processed in the same procedure as previous work (Chen et al., 2021). Transcripts captured by 15*15 DNBs were merged as one bin 15. We treated the bin 15 as the fundamental analysis unit. Bin IDs were synthesized by their spatial coordination (spatial_x and spatial_y) at the capture chip. In specific, the DNB at the left bottom of bin 15 was selected to represent the location of bin 15. Sample contours were manually drawn to exclude bins that were not from the tissue samples. For comparison with other resolutions, we also binned DNBs at 24-hpf in other bin sizes, i.e., bin 50, 100, 150, 200, and 200, and performed unsupervised clustering. The thresholds of gene number used for filtering low-quality bins were: 350 for bin 15, 1800 for bin 50, 5000 for bin 100, 8000 for bin 150 and 10000 for bin 200 respectively.

Unsupervised clustering of Stereo-seq data
Data normalization, scaling, and bins clustering were processed using the R package Seurat (v4.0.5, https://github.com/satijalab/seurat). Sections from the same embryo were pooled together as a data set for analysis. Cell identities of clusters were annotated using marker genes. Gene number captured was used for quality control, with 200 genes for 5.25-, 10- and 12-hpf, 300 genes for 3.3- and 18-hpf, and 350 genes for 24-hpf, and bins with a lower number of genes captured were filtered out. Bins located in the yolk with low gene numbers were also filtered out in this procedure. To get more detailed cell types, bins belonging to a specific cell type, or groups of relevant cell types were further clustered and annotated.

Single-cell RNA-seq data processing: cell clustering and identification of cell types
Raw sequencing reads of each sample were processed using DNBelab_C_Series_HT_scRNA-analysis-software (https://github.com/MGI-tech-bioinformatics/DNBelab_C_Series_HT_scRNA-analysis-software) which include alignment, primary filtering, and gene expression generation of each cell. We merged all data sets for each time point using the merge function and used Seurat (V4.0.3) to filter the low-quality cells by gene number of each cell and reads mapped to mitochondrial genes. The data were then log normalized using ‘NormalizeData’ function with the scaled factor set to the default value of 10,000. We scale the data using ‘ScaleData’ function and identified the 2,000 high variable features using ‘FindVariableFeatures’ function. All these variable genes were used to do principal-component analysis (PCA). The cells were clustered using ‘FindNeighbors’ function using the first 20 principal components, followed by ‘FindClusters’ function. All cell clusters were identified by using known cell type–specific markers for each cluster.

Marker genes of clusters and GO an KEGG gene enrichment analysis
Marker genes of different clusters were found by ‘FindMarkers’ of Seurat R package, with parameter (min.pct = 0.25, logfc.threshold = 0.25) and filtered by pvalue.adj<0.05. Gene set enrichment analysis, were processed with function ‘enrichGO’ using parameters (OrgDb=org.Dr.e finished gene expression generation of each cell. We merged all data sets for each time point using the merge function and used Seurat (V4.0.3) to filter the low-quality cells by gene number of each cell and reads mapped to mitochondrial genes. The data were then log normalized using ‘NormalizeData’ function with the scaled factor set to the default value of 10,000. We scale the data using ‘ScaleData’ function and identified the 2,000 high variable features using ‘FindVariableFeatures’ function. All these variable genes were used to do principal-component analysis (PCA). The cells were clustered using ‘FindNeighbors’ function using the first 20 principal components, followed by ‘FindClusters’ function. All cell clusters were identified by using known cell type–specific markers for each cluster.

Marker genes of clusters and GO an KEGG gene enrichment analysis
Marker genes of different clusters were found by ‘FindMarkers’ of Seurat R package, with parameter (min.pct = 0.25, logfc.threshold = 0.25) and filtered by pvalue.adj<0.05. Gene set enrichment analysis, were processed with function ‘enrichGO’ using parameters (OrgDb=org.Dr.eg.db, pAdjustMethod="BH", pvalueCutoff=0.01, qvalueCutoff = 0.05) and function ‘enrichKEGG’ with parameter (organism='dre', pvalueCutoff=0.05) of R package clusterProfiler (V3.16.1, https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html) (Wu et al., 2021).

The pseudotime analysis using Monocle 2
The pseudotime trajectories were reconstructed by Monocle2 package (V2.22.0, https://www.bioconductor.org/packages/release/bioc/html/monocle.html) (Qiu et al., 2017). The raw count was first converted to CellDataSet object by importCDS function, then dimension reduction clustering and trajectory inference were performed with the reduce Dimension function and the order cells function with default parameters respectively. Differentially expressed genes of different pseudoequivalences were found by ‘Find-Markers’ of Seurat R package, with parameters (min.pct = 0.25, logfc.threshold = 0.25).

Identification of spatially auto-correlated gene modules
Modules of spatially correlated genes were identified using Hotspot (V0.9.1, https://yoseflab.github.io/Hotspot/) (DeTomaso and Yosef, 2021). The expression matrices of genes with minimum UMI number (30 for 3.3- 5.25- and 24-hpf, 50 for 10-, 12-, and 18-hpf) of all embryo sections from 6 development stages were used as inputs. The data were first normalized by the total UMIs number of these genes of each bin, then K-Nearest Neighbors (KNN) graph of genes was created using the ‘create_knn_graph’ function with the parameters: n_neighbors = 8, then genes with significant spatial autocorrelation (FDR < 0.05) were kept for further analysis. The modules were identified using the ‘create_modules’ with the parameters: min_gene_threshold (15 genes for 3.3-hpf, 20 genes for 5.25-hpf, 10-hpf and 12-hpf, and 100 genes for 18-hpf and 24-hpf) and fdr_threshold = 0.05. For the cell type enrichment analysis for each module, we calculated the cell type composition of bins with high module scores (with the criteria that module scores bigger than 3, the method of calculation of the module score see the Hotspot reference; DeTomaso and Yosef, 2021).
To find genes that spatially correlated with Notch pathway, we further conducted module analysis at 10-, 12-, 18- and 24-hpf (with minimum UMIs number for gene selection as above). We first calculated gene set expression score (Notch-score) of Notch pathway genes (genes with very low expression were removed from the Notch pathway gene list) using function ‘AddModuleScore’ of R package Seurat (V4.0.5). Then we added the Notch-score into the gene expression matrix and calculated the spatial correlations as above. Genes with Z-scores for the significance of the correlation with Notch-score bigger than 5 were chosen for further analysis. Modules of these Notch pathway correlated genes were identified using ‘create_modules’, with parameter min_gene_threshold, 20, 20, 30, 100 for 10-, 12-, 18-, 24-hpf respectively. Notch genes used for analysis included: dlc, dla, dld, dlb, dll4, jag1a, jag1b, jag2b, notch1a, notch1b, notch2, notch3, her6, her9, her12, her15.1, her15.2, her4.2, her2, her4.4, her5, her3, her11, her13, her8.2, her8a, her4.1, her4.3, her4.5, her1, hey1, hey2.

**Protein-protein interaction (PPI) network analysis of spatial modules**
We applied the search tool for the retrieval of interacting genes (STRING) (https://string-db.org) database to seek potential interactions among genes in one/multiple modules. Active interaction sources including co-expression as well as species limited to “*Danio rerio*”, and an interaction score >0.4 were applied to construct the PPI networks. The PPI networks were visualized by Cytoscape software 3.8.2 (https://cytoscape.org/) (Shannon et al., 2003). In these networks, the nodes correspond to the proteins and the node size is proportional to the relative connectivity in each network.

**Construction of single-cell developmental trajectory**
We integrated scRNA-seq data from adjacent time points using Harmony (lambda = 0.1, epsilon.harmony = -Inf) (Harmony, V1.0, https://portals.broadinstitute.org/harmony/articles/quickstart.html) (Korsunsky et al., 2019). Then we calculated the 10 nearest neighbors in the latter time point for each spot of the earlier time point using K-Nearest Neighbors algorithm. We took the most frequent clusters of the 10 nearest spots as the target states that the earlier spots would develop into. If two or more target clusters had the same frequencies in the nearest neighbors, the cluster with the shortest weighted distance was used. We repeated this procedure on each pair of adjacent time points from 3.3- to 24-hpf. While constructing the developmental trajectory between each two adjacent time points, we reserved only the connections with a proportion of source cells more than 30% of its type.

**Trajectory branch marker genes**
We calculated the marker genes of the nodes within each branch and then overlapped them with the cell type marker genes of the nodes within the single-cell RNA data of each time point.

**Deconvolution of cell types**
We divided single-cell clusters into subtypes according to the single-cell developmental trajectory. Then we calculated the cell composition of each spatial data bin using SPOTlight (V1.0.6, https://github.com/MarcElosua/SPOTlight) (Elosua-Bayes et al., 2021) with scRNA-seq data as a reference.

**Quantification of cell type distribution at embryo sections**
To better quantify the spatial distribution of different cell types at 6 developmental stages in each trajectory, we first placed embryo sections in the spatial coordinate, then calculated and plotted bin counts of these cell types distribution on the x and y axis coordinate of each embryo sections. The cell type of bin 15 was assigned with the reference cluster with the biggest value of ratio of SPOTlight deconvolution results. The distribution curves were smoothed by loess regression when plotted using ggplot2 (V3.3.5, https://github.com/tidyverse/ggplot2) (Wickham, 2011).

**TF regulation activity prediction**
We downloaded the motif database of zebrafish from the CIS-BP Database (http://cisbp.ccbr.utoronto.ca/bulk.php) and constructed the cisTarget databases for zebrafish according to the SCENIC tutorial instructions (Albar et al., 2017) (https://github.com/aertslab/create_cisTarget_databases). Then we calculated the single-cell TF regulation activity using pySCENIC (rank_threshold = 10000, auc_threshold = 0.1, nes_threshold = 2.5) (pySCENIC, V0.11.2, https://github.com/aertslab/pySCENIC). We used the weighted average expression of predicted target genes to illustrate the regulation activity of each transcription factor in the spatial transcriptomic data.

**Ligand-receptor analysis**
Ligand-receptor pairs were extracted from the LRBase.Dre.eq.db database (http://bioconductor.org/packages/release/data/annotation/html/LRBase.Dre.eq.db.html). For each bin expressing a receptor, we calculated its Euclidian distances to all the bins expressing the corresponding ligand and take the nearest distance as the distance of this ligand-receptor pair in this bin. Then the distances were scaled regarding the distance of adjacent cells as 1. The average distance of each ligand-receptor pair in each time point was taken as the ligand-receptor distance for that time point. Ligand-receptor pairs with an average distance less than 5 in one or more time points were kept for analysis. The distance score was calculated as:

\[
\text{Distance\_score} = \text{Minimum}(\log_2(1 + \frac{1}{\text{distance}}), 5)
\]
We calculated the expression score of each ligand-receptor pair as:

$$\text{Expression\_score} = \log_2(1 + (E_r \times E_l))$$

Where $E_r$ and $E_l$ represent the average expression of the receptor and the ligand respectively.

For ligand-receptor interaction of notch signal, we normalized the interaction frequency by dividing the bin number of each cell type.

To test the significance of the distances between ligands and receptors, for each ligand-receptor pair, we randomly selected the same number of bins as receptor-expressing bins, and calculated the average distance between these random bins to ligand-expressing bins. After repeating the random sampling test for 50 times, we took the mean value of the 50 tests as the average distance of random bins to ligand-expression bins. Then we compared the average distances between (1) randomly sampled bins and ligand-expressing bins; (2) receptor-expressing and ligand-expressing bins. The comparison between the distances of ligand-receptor pairs and random bins to ligand-expression bins was performed with paired t-test.

To calculate the correlation of ligand-receptor pairs, we grouped the spatial bins by previously defined clusters and calculated the correlation of average expression of ligands and receptors across all spatial clusters for each time point.