

# Stereopy: modeling comparative and spatiotemporal cellular heterogeneity via multi-sample spatial transcriptomics

Received: 4 March 2024

Accepted: 5 March 2025

Published online: 21 April 2025

Check for updates

Shuangfang Fang<sup>1,2,12</sup>, Mengyang Xu<sup>2,3,12</sup>, Lei Cao<sup>1,2,12</sup>, Xiaobin Liu<sup>3,12</sup>, Marija Bezulj<sup>2,12</sup>, Liwei Tan<sup>2,12</sup>, Zhiyuan Yuan<sup>4,12</sup>, Yao Li<sup>3</sup>, Tianyi Xia<sup>1,2</sup>, Longyu Guo<sup>2</sup>, Vladimir Kovacevic<sup>2</sup>, Junhou Hui<sup>2</sup>, Lidong Guo<sup>3,5</sup>, Chao Liu<sup>2</sup>, Mengnan Cheng<sup>2,6</sup>, Li'ang Lin<sup>2</sup>, Zhenbin Wen<sup>2</sup>, Bojana Josic<sup>2</sup>, Nikola Milicevic<sup>2</sup>, Ping Qiu<sup>2</sup>, Qin Lu<sup>2</sup>, Yumei Li<sup>2</sup>, Leying Wang<sup>2</sup>, Luni Hu<sup>1,2</sup>, Chao Zhang<sup>2</sup>, Qiang Kang<sup>2</sup>, Fengzhen Chen<sup>2</sup>, Ziqing Deng<sup>1</sup>, Junhua Li<sup>2,7,8</sup>, Mei Li<sup>2</sup>, Shengkang Li<sup>2</sup>, Yi Zhao<sup>9,13</sup> , Guangyi Fan<sup>2,3,13</sup> , Yong Zhang<sup>2,10,11,13</sup> , Ao Chen<sup>2,13</sup> , Yuxiang Li<sup>2,10,11,13</sup> & Xun Xu<sup>10,13</sup>

Understanding complex biological systems requires tracing cellular dynamic changes across conditions, time, and space. However, integrating multi-sample data in a unified way to explore cellular heterogeneity remains challenging. Here, we present Stereopy, a flexible framework for modeling and dissecting comparative and spatiotemporal patterns in multi-sample spatial transcriptomics with interactive data visualization. To optimize this framework, we devise a universal container, a scope controller, and an integrative transformer tailored for multi-sample multimodal data storage, management, and processing. Stereopy showcases three representative applications: investigating specific cell communities and genes responsible for pathological changes, detecting spatiotemporal gene patterns by considering spatial and temporal features, and inferring three-dimensional niche-based cell-gene interaction network that bridges intercellular communications and intracellular regulations. Stereopy serves as both a comprehensive bioinformatics toolbox and an extensible framework that empowers researchers with enhanced data interpretation abilities and new perspectives for mining multi-sample spatial transcriptomics data.

Cells are not static. They achieve functions and form the complex architecture of multicellularity through dynamic and orderly processes, such as cellular proliferation, differentiation, and maturation, which involve spatial interactions with the microenvironment

consisting of external stimuli and neighboring cells. However, understanding the mechanisms that govern cellular processes in disease, development, and homeostasis remains an open question in scientific investigations. Such investigations often require the simultaneous

<sup>1</sup>BGI Research, Beijing, China. <sup>2</sup>BGI Research, Shenzhen, China. <sup>3</sup>BGI Research, Qingdao, China. <sup>4</sup>Institute of Science and Technology for Brain-Inspired Intelligence, MOE Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University, Shanghai, China. <sup>5</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China. <sup>6</sup>BGI Research, Hangzhou, China. <sup>7</sup>Shenzhen Key Laboratory of Unknown Pathogen Identification, BGI Research, Shenzhen, China. <sup>8</sup>BGI Research, Riga, Latvia. <sup>9</sup>Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. <sup>10</sup>BGI Research, Wuhan, China. <sup>11</sup>Guangdong Bigdata Engineering Technology Research Center for Life Sciences, BGI research, Shenzhen, China. <sup>12</sup>These authors contributed equally: Shuangfang Fang, Mengyang Xu, Lei Cao, Xiaobin Liu, Marija Bezulj, Liwei Tan, Zhiyuan Yuan. <sup>13</sup>These authors jointly supervised this work: Yi Zhao, Guangyi Fan, Yong Zhang, Ao Chen, Yuxiang Li, Xun Xu.

e-mail: [zhaoyi@ict.cn](mailto:zhaoyi@ict.cn); [fanguangyi@genomics.cn](mailto:fanguangyi@genomics.cn); [zhangyong2@genomics.cn](mailto:zhangyong2@genomics.cn); [chenao@genomics.cn](mailto:chenao@genomics.cn); [liyuxiang@genomics.cn](mailto:liyuxiang@genomics.cn); [xuxun@genomics.cn](mailto:xuxun@genomics.cn)

analysis of datasets comprising multiple samples to track the specificity and variation of cells and genes across different conditions, time points, and spatial dimensions<sup>1,2</sup>. The advent of high-resolution spatially resolved transcriptomics (SRT) technologies, such as Stereo-seq<sup>3</sup>, Slide-seq<sup>4</sup>, MERFISH<sup>5</sup>, SeqFish<sup>6</sup>, STARmap<sup>7</sup>, and Xenium<sup>8</sup>, holds the potential to generate large-scale multi-sample datasets. These progresses underscore the demand for more advanced analytical approaches that enable the exploration of molecular alterations and characteristics in various contexts—be it conditional, temporal, or spatial<sup>9</sup>. These contexts span a wide spectrum of applications, from tracking disease progression<sup>10,11</sup>, monitoring temporal cellular development<sup>12</sup>, to dissecting the intricacies of spatial organogenesis<sup>3,13</sup>.

Pioneering analysis frameworks, such as Squidpy<sup>14</sup>, Giotto<sup>15</sup>, Scanpy<sup>16</sup>, Seurat<sup>17</sup>, and scvi-tools<sup>18</sup>, have been widely employed for the analysis of spatial or single-cell omics data, enabling temporally and/or spatially resolved studies. However, they were primarily designed for single-sample analysis<sup>19</sup>. When it comes to multi-sample analyses, there is a need for tailored data containers that can efficiently organize, provide flexibility, and scale with the data. Existing data containers like AnnData (used in Scanpy<sup>16</sup>), SeuratObject (used in Seurat<sup>17</sup>), GiottoObject (used in Giotto<sup>15</sup>), and MuData (used in Muon<sup>20</sup>) have inherent limitations in managing multiple samples efficiently. Additionally, current methods<sup>21–23</sup> lack the advanced capabilities required for comprehensive multi-sample analysis. As the data cost keeps decreasing and tissue complexity grows, efficient methods for storing, integrating, and visualizing multi-sample omics data across various dimensions are urgently needed.

Developing a multi-sample analysis framework presents several complex challenges, including the establishment of a standardized framework for analysis modules and visualization functions, the design of scalable data representation for managing multi-sample data, and the provision of integrative solutions for diverse tasks. In response to these challenges, we propose Stereopy, a comprehensive toolkit that offers a complete set of extensible tools for managing, analyzing, and visualizing multi-sample spatial omics data. To facilitate the management of multiple samples in a unified and convenient manner, Stereopy incorporates flexible cross-sample storage of input data and results and prioritizes the accessibility and traceability of outcomes. Furthermore, a flexible framework is designed to enable data extraction and analysis on specific samples, effectively manage dependencies between different steps, and facilitate the transformation of single-sample results into integrative results.

The comprehensive analysis solutions provided by Stereopy improve the utilization of multi-sample information when applied to different datasets (Supplementary Note 1). In comparative studies, Stereopy enables the comparison of disturbed or disease samples with control samples, analyzes the diversity at both global and local levels in the spatial context, and identifies changes in functional mechanisms resulting from stress responses or disease perturbations. The multi-sample cell community detection (CCD) algorithm introduced in Stereopy detects variations at the cell community level in comparative samples (Supplementary Note 2.1). For temporal studies, it explores variations in cell types and gene expression over time, and captures the molecular dynamics of organismal development with spatial resolution. The proposed spatially resolved temporal gene pattern inference (TGPI) algorithm in Stereopy enhances the detection of spatiotemporal gene patterns by simultaneously considering spatial and temporal features (Supplementary Note 2.2). In three-dimensional (3D) integrative studies, Stereopy provides the NicheReg3D tool for reconstructing the cell niche and investigating the effect of intercellular signaling on intracellular regulation within spatial constraints, thus proposing an improved model on the molecular mechanisms of organogenesis (Supplementary Note 2.3). Moreover, Stereopy offers flexible data visualization techniques for both two-dimensional (2D) and 3D datasets, allowing researchers to explore complicated comparative and spatiotemporal changes in genes and cells, and accurately model underlying biological

processes across different dimensions. Stereopy can be accessed at <https://github.com/STOmics/Stereopy>. Its documentation and extensive tutorials are available at <https://stereopy.readthedocs.io/en/latest>.

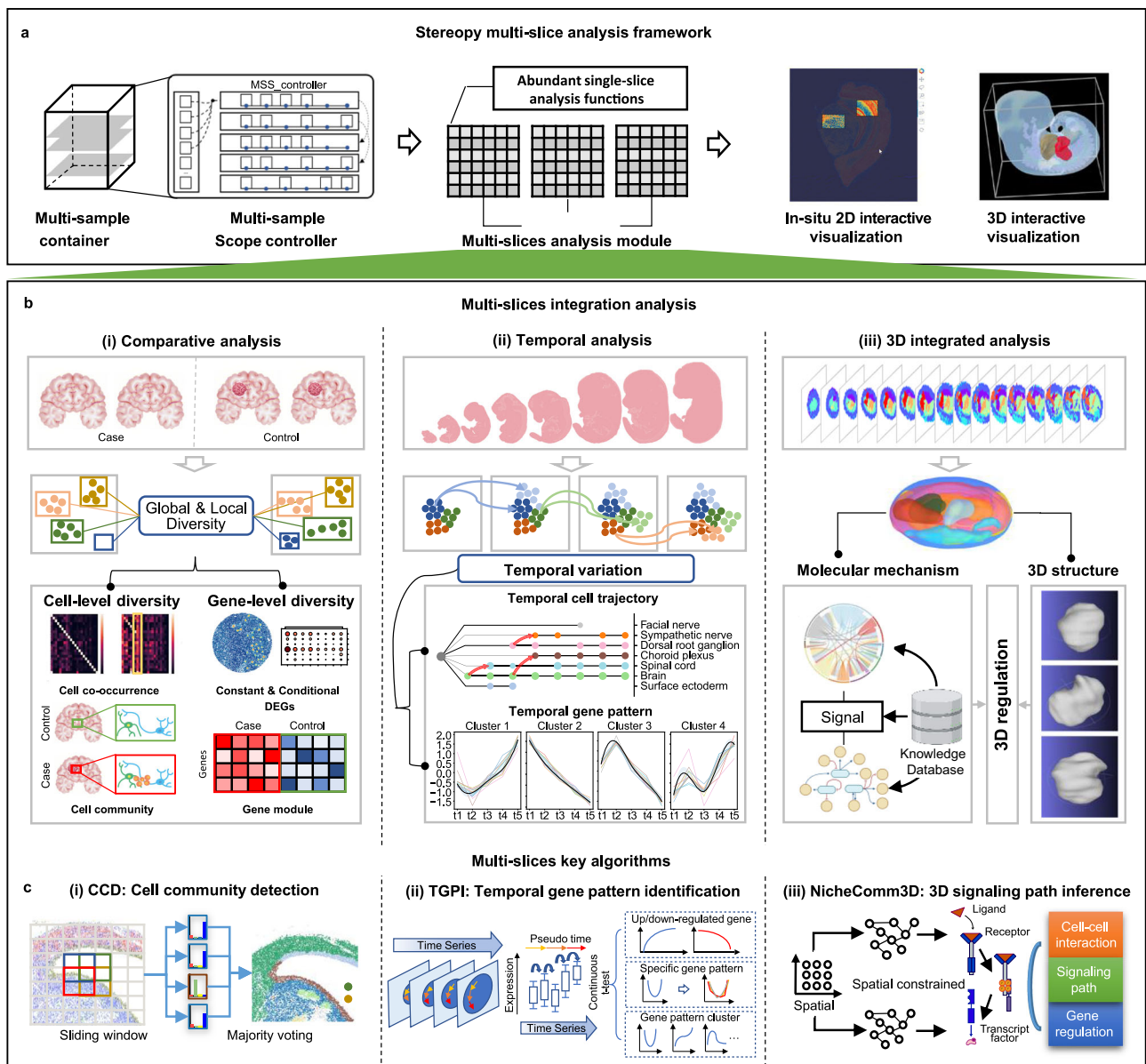
## Results

### Overview of stereopy

Stereopy provides a comprehensive and robust solution for multi-sample analysis, comprising three main components: a scalable framework, analysis modules, and visualization tools (Fig. 1a). The framework is designed to facilitate data management through a universal container for storing multimodal data, a controller for selecting and analyzing specific subsets of interest, and a transformer for integrating data. Stereopy also offers well-organized modules and key algorithms tailored for three fundamental scenarios of multi-sample spatial omics data analysis: comparative, spatiotemporal, and 3D integrative analyses (Supplementary Note 1). These analytical capabilities include the identification of specific cell communities and functional modules in comparative datasets (Fig. 1b (i) and Supplementary Note 1.1), the detection of temporal variable genes and gene patterns in time-series datasets (Fig. 1b (ii) and Supplementary Note 1.2), and the inference of complete signaling paths from cell-cell communication to gene regulation networks in 3D datasets (Fig. 1b (iii) and Supplementary Note 1.3). Stereopy's key algorithms for each data type are highlighted: 1) the cell community detection algorithm, which enables the discovery of common or specific communities between case-control samples, enhancing comparative analysis capability (Fig. 1c (i) and Supplementary Note 2.1); 2) the spatially resolved temporal gene pattern identification method, which delves into specific gene modules related to temporal development within spatial constraints (Fig. 1c (ii) and Supplementary Note 2.2); and 3) the 3D regulation mechanism inference, which uncovers comprehensive gene regulation models by mining extracellular ligand-receptor interactions, intracellular regulation networks, and signaling pathways across the entire 3D tissue level (Fig. 1c (iii) and Supplementary Note 2.3). Furthermore, Stereopy provides 2D and 3D interactive visualization capabilities for spatial omics data<sup>24</sup>, which allow researchers to generate high-quality data explorations and facilitate user-defined browsing. By providing such unique features, Stereopy proves to be an invaluable tool for researchers in the analysis and interpretation of multi-sample SRT data, with powerful functionalities that enable researchers to gain a deeper understanding of biological processes and mechanisms.

### Stereopy develops an efficient multi-sample data analysis framework

Stereopy has developed an efficient multi-sample data analysis framework that includes the MsData (multi-sample data) container, MSS (multi-sample scope) controller, and multi-sample analysis transformer. The MsData container extends the AnnData format to incorporate additional features applicable to multiple samples while preserving single-sample dependencies (Fig. 2a). This allows users to access the entire dataset and individual samples through a single handler, facilitating flexible analysis across multiple samples (Supplementary Figs. 1, 2). The MSS controller manages result storage, tracks analysis dependencies, and visualizes outcomes (Fig. 2b). By adjusting scope parameters in each MsData function, users can associate meta-information and results with corresponding samples for subsequent association analysis. Stereopy's multi-sample transformer supports customized analysis of multi-sample datasets with diverse demands (Fig. 2c), providing functions to integrate single-sample results into the multi-sample context or reversibly split multi-sample data for single-sample analysis. These transformations are particularly useful for modules such as clustering and annotation, which may involve manual curation or calculation comparisons using different algorithms. The framework facilitates parallel or integrative analysis across multiple samples (Fig. 2d and Supplementary fig. 3a), enabling comprehensive multi-sample joint



**Fig. 1 | Overview of stereopy. a** Stereopy provides solutions for multi-sample analysis, including a multi-sample data container and framework, multi-sample data modules, and multi-sample interactive visualization. **b** Stereopy offers key analysis modules for three main multi-sample data analysis scenarios: (i) Comparative studies: Stereopy provides functions at both the cell level and gene level to infer the global and local similarity and diversity for comparative SRT datasets. DEGs: Differentially expressed genes. (ii) Temporal studies: Stereopy offers temporal trajectory analysis and spatially resolved temporal gene pattern analysis to phase the temporal variable datasets. (iii) 3D integrative studies: Stereopy enables

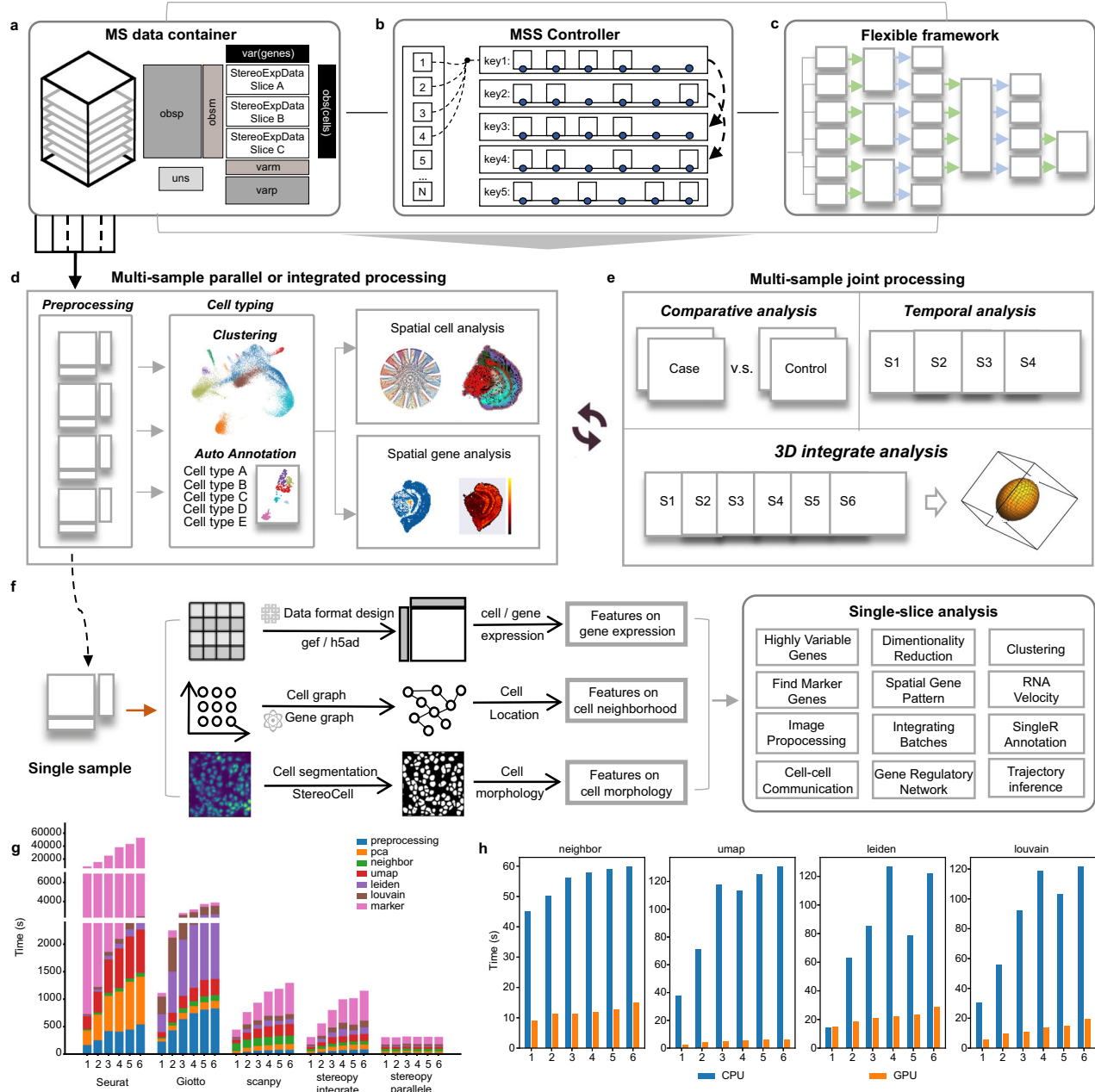
3D data reconstruction and 3D signaling path identification to explore regulation mechanisms. **c** Stereopy contributes key algorithms for the three analysis scenarios: (i) CCD algorithm detects cell communities in single/multi-sample datasets, with a focus on finding common and specific communities. (ii) TGPI algorithm identifies temporal variable gene patterns with spatial restriction, enabling the discovery of gene patterns related to development or temporal variation. (iii) NicheReg3D algorithm investigates inter- and intracellular regulation mechanisms from the 3D aspect. Source data are provided as a Source Data file.

analyses and interactive visualization of multi-sample data (Fig. 2e and Supplementary Fig. 3b, c). In addition, Stereopy is a powerful tool for single-sample spatial omics data analysis by sharing several key functions and modules (Fig. 2f and Supplementary Fig. 3).

Stereopy accelerates multi-sample analysis through algorithmic and parallel computing approaches. By implementing parallel analysis for dependent functions including preprocessing, cell clustering, and annotation, Stereopy reduces the overall processing time. Notably, Stereopy's common SRT analysis modules outperform existing tools such as Giotto, Scanpy, and Seurat in terms of processing time for both parallel and integrative analysis across different numbers of samples (Fig. 2g). Meanwhile, Stereopy leverages GPU acceleration to enhance

the performance of time-consuming but necessary functions such as dimensionality reduction, neighborhood searching, Leiden<sup>25</sup> / Louvain<sup>26</sup> clustering, and SingleR annotation<sup>27</sup> (re-implemented in Python as a part of Stereopy). The GPU-accelerated functions demonstrate a substantial improvement in execution time compared to their CPU counterparts (Fig. 2h).

**Stereopy unveils cell and gene diversity in comparative analysis** Comparing samples that are disturbed or affected by disease with control samples allows researchers to understand changes in functional mechanisms at both local and global levels. Stereopy incorporates cell-level and gene-level modules, supported by our in-house

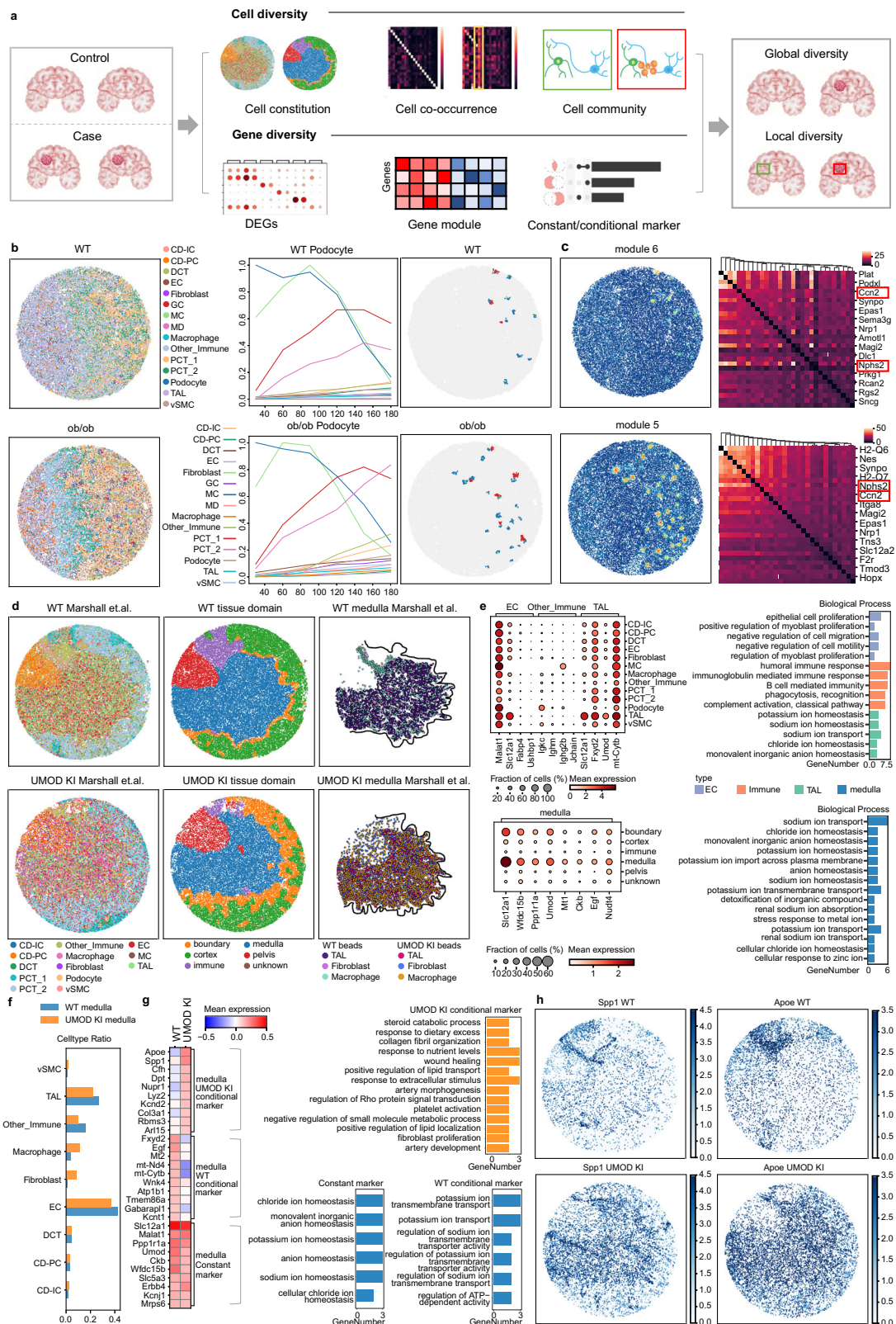


**Fig. 2 | Stereopy designs a flexible multi-sample data analysis framework and accelerates multi-sample analysis.** **a** Stereopy introduces MSData container, **b** MSS controller, and **c** multi-sample analysis transformer to support scalable analysis of multi-sample SRT data. MS: Multi-sample, MSS: Multi-sample scope. **d** Stereopy provides the option of multi-sample parallel or integrative processing and **e** joint multi-sample processing for comparative, temporal, and 3D integrative studies. **f** Stereopy facilitates the restoration and processing of gene expression,

spatial information, and corresponding image features for each sample in spatial omics analysis. **g** Comparison of the execution time of Stereopy, Seurat, Giotto, and Scanpy in basic processes, including preprocessing, PCA, finding neighbors, UMAP, Leiden clustering, Louvain clustering, and finding marker genes. **h** Comparison of the execution time of basic processes, including finding neighbors, UMAP, Leiden clustering, and Louvain clustering with or without GPU mode. Source data are provided as a Source Data file.

algorithms, to identify and analyze global and local diversities in comparative samples (Supplementary Note 1.1). The cell-level analysis focuses on exploring cell diversity in terms of cell type, cell co-occurrence, and cell community via multi-sample comparisons. To enhance the detection of cell communities, we have developed the multi-sample CCD algorithm (Supplementary Fig. 4, Supplementary Note 2.1 and Methods). At the gene level, Stereopy investigates the gene diversity within specific cell types and communities (Fig. 3a), and introduces the concept of constant (remain unaffected by disturbances) and conditional markers (respond to disturbances and contribute to functional changes).

Stereopy-CCD demonstrates superior performance compared to other existing methods such as Giotto<sup>15</sup>, SpaGCN<sup>28</sup>, GraphST<sup>23</sup>, BASS<sup>22</sup>, and PRECAST<sup>21</sup> in both single-sample scenarios (e.g., whole mouse embryo brain) and multi-sample scenarios (e.g., continuous mouse brain and mouse kidney) (Supplementary Fig. 5-7, Supplementary Tables 1, 2, and Methods). Meanwhile, Stereopy-CCD performs comparably to BANKSY<sup>29</sup> in terms of accuracy but with the added benefits of lower time and memory consumption. In the analysis of the whole mouse embryo brain dataset, Stereopy-CCD is capable of identifying cell communities or domains that align with existing knowledge (Supplementary Fig. 5). In continuous adult mouse brain, Stereopy-



CCD detects common cell communities across three 2D slices (Supplementary Fig. 6). Furthermore, in our analysis of mouse kidney samples, which include a diabetic sample (UMOD KI-homozygous gene UMOD-C125R knock-in mice with monogenic disorder) and a WT sample<sup>30</sup>, the Stereopy-CCD algorithm identifies a central community present in both samples (Supplementary Fig. 7). The central community closely corresponds to the region annotated as medulla in a

previous study conducted by Marshall et al.<sup>30</sup> (Fig. 3d and Supplementary Fig. 7).

To assess the efficacy of Stereopy in detecting global diversity, we applied it to comparative mouse kidney datasets<sup>30</sup>. Specifically, we analyzed a pair of Slide-seq v2 samples: wild-type (WT) and diabetic (*ob/ob* genetic model of early diabetic kidney disease) (Supplementary Fig. 8). Using the co-occurrence calculations developed in Stereopy,

**Fig. 3 | Stereopy facilitates comparative analysis of spatial transcriptomes with multiple samples.** **a** Graphical abstract showcasing Stereopy's capabilities for comparative multi-sample analysis. Stereopy offers functions for analyzing cell constitution diversity, co-occurrence, and cell communities at the cell level, as well as differential expression gene, spatial gene modules, and constant/conditional markers at the gene level. **b** Co-occurrence result for a BTBR kidney sample. Left: spatial map of WT and *ob/ob* kidney samples. Middle: line plot showing co-occurrence of podocyte cells with other cell types. Right: spatial map confirming the co-occurrence of podocyte with GCs and MCs. Upper and lower parts represent WT and *ob/ob* samples, respectively. **c** Left column: spatial gene modules corresponding to podocyte location. Right column: local autocorrelation of corresponding gene module. Upper and lower parts represent WT and *ob/ob* samples, respectively. **d** Spatial map showing cell type annotation, tissue domain identified

by Stereopy-CCD algorithm, and medulla defined by Marshall et al. for WT and UMOD KI kidney samples. Left, middle and right parts represent cell type annotation, tissue domain annotation, and medulla defined by Marshall et al., respectively. Upper and lower represent WT and UMOD KI samples, respectively. **e** Left: differentially expressed genes for medulla in WT sample and its composing cell types such as EC, TAL, and other immune cells. Right: GO enrichment analysis for these cell types. **f** Cell type constitution and proportion for medulla. **g** Constant and conditional marker for medulla in WT and UMOD KI samples. Left shows the heatmap of constant and conditional markers. Conditional markers are genes with high expression only under certain conditions. Right shows GO enrichment analysis for each group of genes. UMOD KI conditional markers (orange) are enriched in GO terms related to wound healing and other processes. **h** Spatial heatmap of *Spp1* and *ApoE* in WT and UMOD KI samples. Source data are provided as a Source Data file.

we confirmed Marshall's previous findings<sup>30</sup> regarding the co-occurrence of Podocytes with GC cells (Fig. 3b). Moreover, our analysis inferred a higher co-occurrence of these cell types in the *ob/ob* sample compared to the WT sample. Importantly, Stereopy demonstrated a greater significance in detecting the co-occurrence of Podocytes with GCs when compared to Squidpy's co-occurrence algorithm<sup>14</sup> (Supplementary Fig. 9). Subsequently, gene modules were identified in both samples, revealing the co-expression of *Nphs2* (a Podocyte marker) and *Ccn2* (a Podocyte injury marker) in the WT and *ob/ob* samples (Fig. 3c). Additionally, local autocorrelation analysis indicated a stronger correlation between *Nphs2* and *Ccn2* in *ob/ob* sample (Fig. 3c). The analysis of differentially expressed genes (DEGs) provided evidence of *Ccn2*'s higher rank among Podocytes markers (Supplementary Fig. 10).

To assess the capability of Stereopy in detecting local diversity, we conducted additional analysis on the cell communities identified in the mouse kidney samples using CCD. These communities were annotated according to their anatomical structures, including kidney cortex, medulla, the boundary of cortex and medulla, pelvis, and immune region (Fig. 3d). As mentioned earlier, the medullary region identified by Stereopy-CCD closely aligns with the region annotated in the study conducted by Marshall et al., in which the marker genes of each region are discernible (Fig. 3d, Supplementary Figs. 7 and 11). The medulla community exhibited similar proportions of thick ascending limb (TAL), endothelial cells (EC), and other immune cell types, despite variations in the distribution of cell types (Supplementary Fig. 12). The UMOD KI sample showed increased percentages of fibroblast and macrophages in the medulla compared to WT sample (Fig. 3f and Supplementary Fig. 12), consistent with Marshall's findings<sup>30</sup>. To analyze markers of both tissues and cell types, we calculated DEGs and enriched gene ontology (GO) terms specifically for the medulla and its constituent cell types, such as TAL, EC, and other immune cell types (Fig. 3e). The marker genes exhibited greater significance, and the enriched GO terms in the renal medulla were highly relevant to renal function, including sodium ion transport, potassium ion transmembrane transport, and chloride ion homeostasis. This highlights the functional relevance of the tissue compared to individual cell types. Our findings confirm that examining the gene divergence of cell communities provides deep insights into tissue function. To comprehensively assess the response of different regions to the UMOD KI disturbance, we analyzed the number of conditional markers. Our analysis revealed a significant increase in the number of top DEGs in the medulla region compared to other regions, indicating a greater diversity of cell types (Supplementary Fig. 13). These results suggest that studying the overall differences in cell communities and conditional markers may yield more meaningful biological discoveries than focusing solely on individual cell types. Specifically, we consistently observed a marker related to renal function, including sodium, potassium, and chloride ion homeostasis, in the renal medulla of both healthy and disease samples. However, the UMOD KI sample exhibited conditional markers involved in renal function damage, such as

response to nutrient level, wound healing, and response to extracellular stimulus (Fig. 3g and Supplementary Table 3). Notably, *Spp1* emerged as a significant conditional marker (Fig. 3h), which has been identified as the top hub gene associated with kidney stones<sup>31</sup>. Further analysis revealed that the risk of renal stones persisted when both *Spp1* and *Umod* had variants, indicating the importance of these two genes in the development of kidney disease<sup>32</sup>. Another conditional marker, *ApoE*, has been reported to be associated with glomerular disorders due to its central role in lipoprotein metabolism. The increased abundance of macrophages in the UMOD KI sample is consistent with the hyperactivity of macrophages involved in *ApoE*-related glomerular disorders<sup>33</sup>.

Our study demonstrates that Stereopy provides a systematic analysis of cell-level and gene-level similarity and diversity between case and control samples, yielding results of high biological significance. The application of the Stereopy-CCD algorithm and the identification of conditional markers contribute to our understanding of tissue structure and function in comparative analyses.

### Stereopy identifies spatiotemporal variation in time-series analysis

The growth and development of organisms involve complex biological processes characterized by variations in cell types and gene expression over time. These temporal variation capture the dynamic molecular changes occurring during development. To investigate temporal variations in time-series datasets, Stereopy emphasizes detecting dynamic changes in both the spatial and temporal dimensions (Supplementary Note 1.2). In terms of cell type changes, Stereopy adopts a manifold partitions-based method<sup>34</sup> to preserve the global topology and infer the trajectory of cell types across different samples, which provides a visual representation of cell trajectories and changes in cell numbers across different time points (Supplementary Note 3). Meanwhile, Stereopy proposes a spatially resolved TGPI method for identifying genes with similar temporal expression changes, including continuously up- or downregulated genes, as well as other complex patterns observed in real time and pseudotime (Fig. 4a and Supplementary Note 2.2).

The false positive risk score (FPR score), which we have introduced in Stereopy-TGPI, is a metric used to detect continuously up- and down-regulated genes by merging *p*-values. To assess the effectiveness of the FPR score, we conducted an evaluation using mouse embryo forebrain datasets. These datasets consist of three distinct cell types and were collected at 7, 5, and 3 time points, respectively. The results demonstrate that FPR score is a stable and reliable approach for identifying genuine up- and down-regulated genes with continuous changes in gene expression across multiple time points (Supplementary Figs. 14–16 and Methods). Stereopy-TGPI serves as a valuable tool not only for identifying temporal gene up- and downregulation but also for elucidating intricate temporal or pseudotime expression patterns. It simultaneously considers the consistency of gene expression in both temporal and spatial aspects (Supplementary Figs. 17, 18 and



Methods). In comparison to Mfuzz<sup>35</sup>, Stereopy-TGPI's identification was more correlated to real and pseudotime tendencies and capable of enriching significant GO terms relevant to neuron development in the time-series whole mouse brain (Supplementary Fig. 19 and Methods).

We further investigated the trajectory and temporal gene pattern of Stereo-seq mouse embryos across eight time-point samples, ranging from E9.5 to E16.5<sup>3</sup>. We inferred the trajectory of the integrative

dataset from eight time points and visualized the cell type development using a tree plot (Fig. 4b, c). Next, the flexibility of Stereopy's data container enabled manual clustering and annotation of the brains in each sample independently (Fig. 4d, Supplementary Fig. 20a). Pseudotime analysis<sup>34</sup> confirmed a gradual increase in pseudotime, with higher pseudotime values in the forebrain region indicating later development (Fig. 4d, Supplementary Fig. 20b). Additionally, we

**Fig. 4 | Stereopy enables temporal analysis of spatial transcriptomes with multiple samples.** **a** Graphical abstract depicting the pipeline for time series analysis. **b** Spatial trajectory visualization of mouse embryo multi-sample transcriptomes from E9.5 to E16.5. **c** Tree plot indicating the development of mouse embryo ectoderm. The x-axis represents time point, the dot size represents the cell number, and the red arrow indicates the PAGA trajectory. **d** Manual annotation and pseudotime assignment for time-series mouse brain samples. **e** Development tree for cell types in the time series. The x-axis represents time point, and the height of Sankey represents the cell number of a certain cell type at a particular time point. **f** PAGA graph for mouse brain trajectory inference. Red arrow points at the cell types selected for downstream analysis. **g** Up and downregulated genes for the mouse forebrain trajectory and GO enrichment analysis. The cell numbers for each cell type are 3972, 524, 2827, 2838, 3629, respectively. The box extends from the first quartile (Q1) to the third quartile (Q3) and the line inside represents the

median. The whiskers extend from the box to the farthest data point lying within 1.5 times the inter-quartile range from the box. Flier points are those beyond the end of the whiskers. **h** F-score among time points and correlation with pseudotime of the top 1000 gene for each cluster using Stereopy-TGPI (blue) and Mfuzz (yellow) methods. **i** A temporal gene pattern identified by Stereopy-TGPI for mouse forebrain trajectory. **j** A temporal gene pattern identified by Stereopy-TGPI for mouse forebrain time series datasets. The upper and lower edges of the error band represent Q1 and Q3 of the gene expression levels of the gene group in *i* and *j*, respectively. **k** Gene expression of *Tead1* in each cell type at each time point. **l** Spatial heatmap for AUC scores of *Tead1*(+) regulons and corresponding GO terms. **m** Time-series gene network for *Tead1*. The radial line represents a group of genes, and the points on it indicate the time points when these genes are expressed. The point size indicates the gene number. Source data are provided as a Source Data file.

provided statistics on the cell number of each cell type across eight time points and inferred the cell trajectory (Fig. 4e, f and Supplementary Fig. 20c).

Our focus then shifted to the temporal gene pattern in the forebrain trajectory series, encompassing the forebrain progenitor, cortical hem, dorsal forebrain, forebrain intermediate progenitor, and forebrain cortical glutamatergic stages. By utilizing Stereopy-TGPI to calculate gene up- and downregulation, we identified *Foxg1* as the top-ranked temporal upregulated gene. This key transcription factor (TF) showed gradual upregulation along the forebrain trajectory, consistent with its known role in regulating forebrain development<sup>36</sup> (Fig. 4f–h and Supplementary Fig. 21). Conversely, *Hes5*, a gradually downregulated gene, exhibited high expression in embryonic neural precursor cells and played a crucial role in negatively regulating neural and oligodendrocyte differentiation<sup>37</sup>. We also employed Stereopy-TGPI to identify cell-type-specific expression patterns along the forebrain trajectory. Among these patterns, we observed a distinctive gene pattern characterized by an upregulation trend prior to the cortical hem stage, followed by continuous downregulation. In this dataset, the cortical hem was exclusively present before E14.5, which aligned with previous research indicating the emergence of Cajal-Retzius neurons, the constituent cell type of the cortical hem, during early developmental stages<sup>32</sup>. To explore the functions of the cortical hem and the key factors contributing to its disappearance, we examined temporal gene patterns related to time points where the cortical hem consistently occurred in the forebrain. We observed a distinct temporal gene pattern, with peak expression levels during the developmental stages from E11.5 to E14.5, coinciding with the presence of the cortical hem, followed by a noticeable decrease thereafter (Fig. 4i). By intersecting genes from the cell-type-trajectory gene pattern and the temporal gene pattern, we identified *Tead1* as a key TF that exhibited high expression in the cortical hem and low expression after its disappearance (Fig. 4j). *Tead* TFs have been previously implicated in regulating cortical development<sup>33</sup>. Leveraging the interactive visualization capabilities of Stereopy, we performed gene regulatory network (GRN) analysis on the forebrain of each sample (Supplementary Fig. 22). Our investigation revealed a notable decrease in the number of genes regulated by *Tead1*, from 338 at E12.5 to 7 at E13.5 (Fig. 4k). Furthermore, we found that the enriched GO terms at E11.5 and E12.5 were similar and primarily associated with forebrain development. However, at E13.5, the enriched GO terms were more related to neuroblast proliferation and forebrain neuron generation. For example, *Tcf4*, a target gene (TG) regulated by *Tead1* from E11.5 to E13.5, controls the positioning of cortical projection neurons<sup>38</sup>. Interestingly, by E14.5, GO terms were no longer related to neurons or the brain (Fig. 4l). This observation coincided with the appearance of forebrain cortical glutamatergic cells, suggesting that *Tead1* and cortical hem had completed their neurogenesis function by this stage.

Stereopy offers comprehensive solutions for temporal multi-sample analysis, with a particular emphasis on temporal gene patterns.

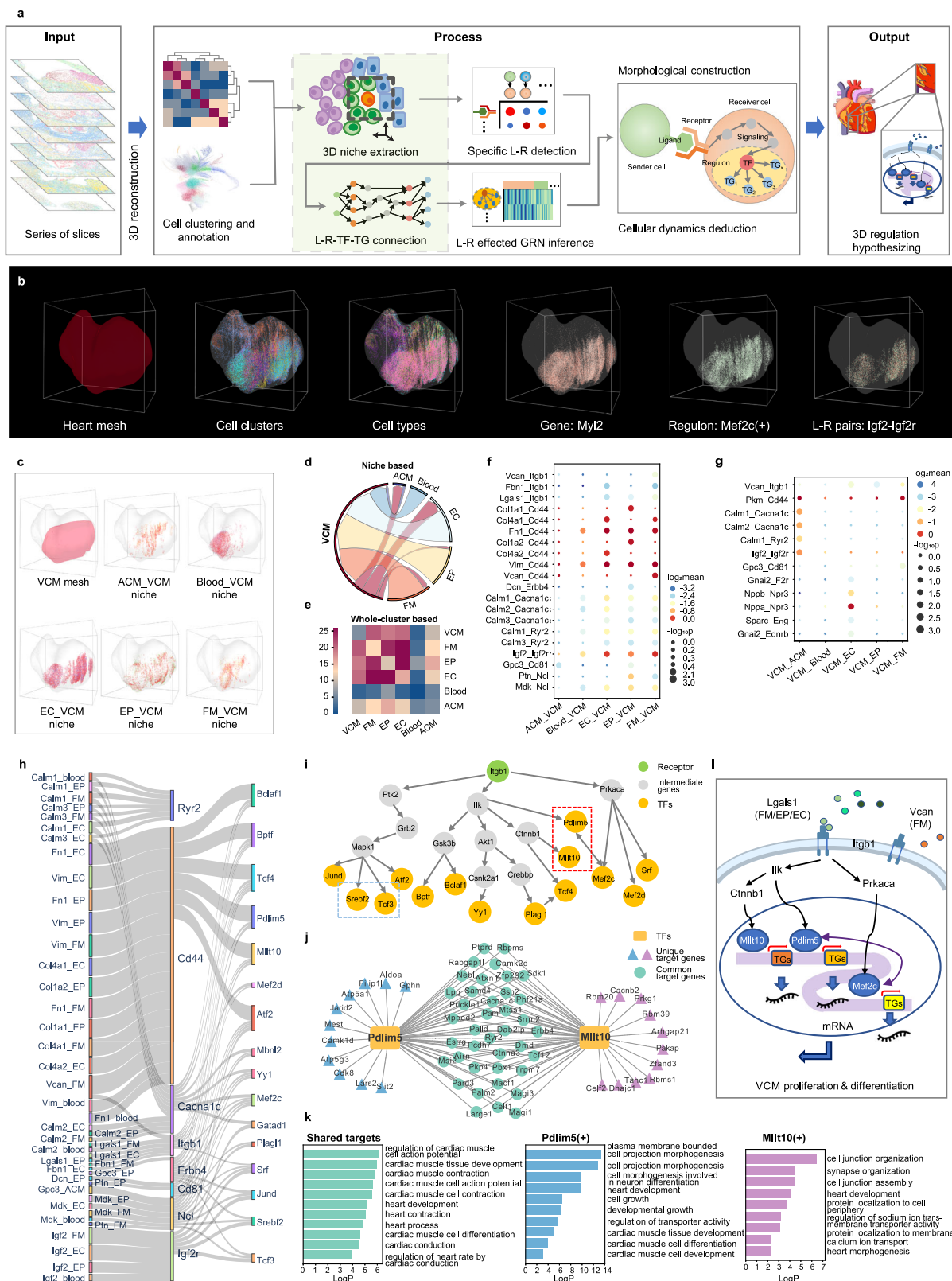
TGPI, a key feature of Stereopy, enables the detection of temporal gene patterns within time-series datasets, revolutionizing our understanding of temporal dynamics in biological systems.

### Stereopy reveals niche-mediated regulations in 3D analysis

In multicellular organisms, cells and tissues are organized within a 3D structure, leading to complex cellular interactions that cannot be adequately captured in 2D culture<sup>39</sup>. Conventional analytical approaches are confined to 2D methodologies, resulting in the loss of crucial interaction information along the z-axis. Stereopy's NicheReg3D pipeline addresses this limitation by precisely characterizing the cellular constitution of 3D niches and facilitating the exploration of intercellular and intracellular interactions (Supplementary Note 1.3). It combines data preprocessing, 3D alignment and reconstruction, 3D cellular niche confinement, cell-niche communication, ligand-receptor (L-R)-TF-TG pathway inference, and intracellular TF-centered regulatory network prediction (see Methods). The core algorithms underpinning 3D joint analysis and the underlying 3D regulation model are elucidated in Fig. 5a. We applied NicheReg3D to the well-studied system of the mouse cortical region, which was sequenced using the BARseq technique, a high-throughput in situ sequencing method<sup>40</sup>. Our results demonstrated that 3D niches, composed of accurately defined cells from diverse cell types after 3D reconstruction of consecutive 2D slices, outperformed 2D niche composition analysis of individual slice. This improved analysis of 3D niches benefited downstream analyses, such as the predictive identification of cortical areas (Supplementary Fig. 23).

In the 3D context, cellular heterogeneity is not only governed by the intracellular GRN but also influenced by the extracellular micro-environment, collaboratively accomplishing various biological tasks<sup>41,42</sup>. However, current computational methods for modeling both interactions simultaneously are insufficient<sup>43</sup>. Our approach provides unique insights into the intracellular regulation mediated by biochemical signals of intercellular crosstalk in multiple dimensions in a 3D spatial multiple-sample setting. To showcase its effectiveness in exploring niche-mediated regulations, we applied the Stereopy-NicheReg3D pipeline to analyze the cardiac development of a mouse embryo sequenced by Stereo-seq<sup>3</sup>. We extracted 59 10-μm-thick 2D serial cryosections at a distance of 10 μm, covering the entire mouse embryonic heart. For the SRT data of 90,411 high-quality segmented cells with 30,254 genes inferred from subcellular spots, we performed unsupervised clustering analysis for each sample and identified six cardiac cell clusters (Supplementary Fig. 24 and Supplementary Table 4). The 3D reconstructed model provided a multi-hierarchical transcriptomic architecture, ranging from organ meshes, cell types and clusters, spatially variable genes, spatially specific regulons, and niche-specific L-R pairs (Fig. 5b).

Based on the reconstructed 3D murine heart data, our investigation focused on the development of ventricular cardiomyocytes (VCMs) and their interaction with the cardiac niche, which is known to



play a significant role in cardiogenesis<sup>44</sup>. The VCM niche encompassed five other cell types: approximately 28% atrial cardiomyocytes (ACMs), 27% blood cells, 23% endocardial cells (ECs), 13% epicardial cells (EPs), and 9% fibro-mesenchymal cells (FMs) (Fig. 5c). Interestingly, within a 25- $\mu$ m physical distance in 3D, we observed that VCMs were the primary recipients of signals from surrounding cells (Fig. 5d). This contradicted the conventional understanding that communication

activities were specific to whole cell clusters<sup>45</sup> (Fig. 5e), underscoring the essentiality of 3D niche pruning. Our computational analysis predicted a significant molecular interaction between the ligand *Vcan* and its receptor *Itgb1* in FM-VCM cells (communication score = 0.293,  $p=0.00$ ), with moderate presence in other niches (Fig. 5f). This observation aligns with previous studies that have emphasized the critical role of *Vcan* in the extracellular matrix for supporting and

**Fig. 5 | Stereopy integrates spatial multi-sample data and reveals 3D regulatory mechanisms related to cardiac development.** **a** Stereopy-NicheReg3D workflow. This approach focuses on identifying regulatory networks that are responsive to cell-niche interactions in a spatial manner, through extracting the cellular niche in 3D and connecting intracellular regulations with intercellular communications (highlighted in green). **b** Stereopy-NicheReg3D illustrating multi-hierarchically transcriptomic architecture, ranging from heart organ meshes, heart cell types and clusters, spatially variable genes (*MyL2*), spatially specific regulons (*Mef2c*(+)), and niche-specific L-R pairs (*Igf2-Igf2r*) from left to right. **c** Spatial distribution of VCM's niche compositions including neighboring ACM, blood, EC, EP, and FM cells in the boundary. **d** Circos plot showing bidirected cell-cell interactions in five niches. The width of an arrow correlates with the number of significant L-R pairs. **e** Heatmap showing the CCC intensities without niche restriction, different from (**d**). **f, g** Bubble plots demonstrating cell-type-specific L-R pairs: **f** niches to VCM. **g** VCM

to niches. Circle color indicates the communication score of each L-R pair, while circle size indicates its *p*-value of permutation test. **h** Sankey plot connecting intercellular L-R interactions from sender niche cells to receiver VCM cells to VCM intracellular downstream TFs via deductive signaling pathways. Bandwidth indicates the mean expression of the two genes at both ends. **i** Regulatory network showing inferred intracellular signaling paths from receptor *Itgb1* to downstream TFs. **j** Shared and specific TGs in *Pdlim5*(+) and *Mlt10*(+) regulons showing the 3D co-regulation function. **k** GO enrichment analysis indicating the collective function of shared targets of regulons (shared, *Pdlim5*(+), and *Mlt10*(+) from left to right). **l** The 3D regulation model of extracellular signaling to intracellular gene regulatory network. The protein structure of RBPMS (PDB code: 5CYJ) was downloaded from Protein Data Bank and rendered in PyMOL. Source data are provided as a Source Data file.

remodeling VCMs<sup>46,47</sup>. Moreover, we found that the four distinct niche compositions, apart from ACMs, collectively influenced VCM gene expression through the same sets of L-R pairs (*Vim-Cd44*, *Calm1-Ryr2*, *Igf2-Igf2r*...), many of which have been implicated in the regulation of CM proliferation, migration, and differentiation<sup>48,49</sup> (Supplementary Figs. 25, 26). In comparison to other state-of-the-art cell-cell communication (CCC) tools, including single-cell CellPhoneDB<sup>45</sup> and spatially resolved NICHES<sup>50</sup> (Supplementary Table 5), Stereopy achieved the most complete identification of specific L-R pairs that covered nearly all those derived by other tools, thanks to its precise niche extraction (Supplementary Fig. 27). The majority of these L-R pairs are involved in mammalian cardiac growth and development (Supplementary Table 6). Additionally, our analysis revealed that VCM also exerted a reverse influence on the cell state or function of the cell micro-environment through another set of L-R pairs (Fig. 5g).

In addition, we inferred the specifically expressed GRNs on VCM cells within the niches (Fig. 5h, Supplementary Fig. 28 and Supplementary Table 7). Through enrichment analysis, we identified a set of candidate core TFs and their corresponding regulons, suggesting their potential susceptibility to cell-niche communications and warranting further inspection of their regulatory effects. We then established deductive signaling paths to connect intercellular signaling activities from niche cells and intracellularly influenced TFs. To simplify the complex network, we retained connections between receptors and TFs involving a maximum of two intermediate genes (Fig. 5h). Among them, *Cd44* emerged as the primary recipient of extracellular signaling, stimulated by specific ligands (*Col1a1*, *Col4a1/2*, *Vcan*) or collectively expressed from different niches (*Fnl1*, *Vim*). This signaling pathway could up- or downregulate various TFs such as *Tcf4*. Previous studies have elucidated the ability of *Cd44* to activate the canonical Wnt/ $\beta$ -catenin signaling pathway, impacting the expression of *Tcf4* and downstream genes<sup>51</sup>, thereby exerting temporal and spatial control over heart maturation<sup>52</sup>. *Igf2-Igf2r* also collectively regulated VCM proliferation and differentiation by activating PI3K/Akt pathways, as previously reported<sup>53,54</sup>. Moreover, the shared *Calm1/3-Cacna1c* family displayed potential regulation of *Mef2c/d* expression, which has been linked to excitation-contraction coupling in VCM function through calmodulin-dependent signaling pathways<sup>55,56</sup>. Our framework additionally facilitated the investigation of detailed GRNs for each user-defined receptor in the same cell. For instance, Fig. 5i depicted the GRN of the *Itgb1* receptor as a directed graph, encompassing various modes, including directed acyclic (such as *Srebf2* and *Tcf3*) and bidirected acyclic (such as *Pdlim5* and *Mlt10*). Importantly, the inferred GRN, extended to downstream TGs, highlights the potential for different intercellular communication to collectively regulate the same set of genes (Fig. 5j and Supplementary Fig. 29). For example, the *Itgb1*-related CCC might modulate both *Pdlim5* and *Mlt10* through Ilk-related pathways. GO enrichment analysis using Metascape<sup>57</sup> (<https://metascape.org/>), indicated that their shared TGs jointly managed cardiac muscle development and contraction (Fig. 5k), corroborating

prior findings<sup>58,59</sup>. In comparison to other tools connecting the outside and inside of cells, such as NicheNet<sup>60</sup>, Stereopy-NicheReg3D provides a more definitive and comprehensive network for inferring how cell-niche-specific L-R pairs regulate intracellular regulon activities related to specific cellular functions.

Through our 3D joint analysis pipeline, we have demonstrated how spatially informed extracellular signaling at the niche influences intracellular gene regulation in the cell of interest, surpassing the limitations of 2D data analysis (Supplementary Figs. 30, 31). The integration of CCC and GRN could improve the accuracy of context-specific L-R-TF-TG predictions concerning morphological phenotypical changes. As such, we have derived an improved model of 3D regulation implicating VCM development in cardiac maturation and physiology (Fig. 5l). During heart development, VCMs constitute a fundamental element of heart function, while EC, EP, FM, and blood cells are key components of the micro-environment promoting CM maturation. The niche components collectively or specifically transmit signals through shared or distinct L-R pairs, which further promote or inhibit specific TFs inside VCM cells through specific signaling pathways. These TFs ultimately influence the expression of downstream TFs and TGs, jointly demonstrating the cellular functional state and subtype.

Therefore, Stereopy-NicheReg3D is expected to be a valuable tool with an interactive visualization browser in the 3D space (Supplementary Fig. 32 and Supplementary Movie 1) for better dissecting the functional consequences of spatially informed inter-intracellular regulation networks, thereby facilitating the prediction of cellular function, state, and corresponding phenotype.

## Discussion

Interpreting similarities, differences, and developmental changes across multiple samples is non-trivial to unravel complex biological regulatory mechanisms in multi-sample spatial omics datasets. In this study, we introduce Stereopy, a comprehensive toolkit for managing, analyzing, and visualizing multi-sample data to address this problem. It features essential components such as the MsData container, MSS controller, and a multi-sample transformer, effectively mitigating the challenges encountered in jointly analyzing multi-sample data. Stereopy also provides a wide array of analysis solutions and algorithms tailored specifically for comparative, temporal, and 3D integrative analysis in multi-sample endeavors.

Firstly, we employed Stereopy to validate the co-occurrence of Podocytes with GCs in comparative kidney datasets, identifying *Spp1* as a potentially significant UMOD KI conditional marker. The application of the Stereopy-CCD algorithm proved its efficacy in detecting important cell communities across multiple samples, thereby expanding the scope of diversity analysis in comparative studies. Subsequently, we harnessed the capabilities of Stereopy to delve into temporal mouse embryonic brain datasets, highlighting the function of *Tead1* and the cortical hem in forebrain cortical development. This investigation provided valuable insights into the dynamics of mouse

forebrain development. The Stereopy-TGPI algorithm demonstrated its ability to accurately infer temporal gene patterns by integrating spatial information, thereby revealing potential gene patterns and key TF genes related to forebrain development. Finally, we leveraged Stereopy to explore 3D multi-sample datasets, specifically investigating the developing VCM in a mouse embryonic cardiac dataset. Through this analysis, we identified an *Itgb1*-stimulated co-regulation network, illuminating the complicated inter- and intracellular regulatory mechanisms in the 3D niche-based microenvironment. The Stereopy-NicheReg3D pipeline demonstrated its superiority in identifying more comprehensive cell-type-specific LR pairs and signaling paths compared to existing tools.

Stereopy represents a comprehensive and robust solution that surpasses simply providing functionalities and algorithms for analyzing complex spatial omics datasets. Its advanced features, including batch effect evaluation and removal for multiple samples, as well as multi-sample joint analysis functions such as 3D slice registration, 3D data trajectory inference and visualization, amplify the utility of Stereopy in the field. Moreover, Stereopy incorporates numerous data analysis functions, including well-known functions adapted from R code such as *scTransform* and *SingleR*. It also supports diverse data types, including GEF and GEM files generated by Stereo-seq, as well as the commonly used *h5ad* file format, enabling the analysis of data from different platforms. It is worth noting that Stereopy can analyze SRT datasets as long as they provide both spatial information and gene expression at the same resolution. However, some algorithms bundled within Stereopy perform optimally with high-resolution datasets, rendering them more suitable for high-resolution SRT platforms.

Stereopy has effectively tackled key challenges in multi-sample spatial omics analysis, including data management, analysis module planning, algorithm development, and interactive visualization of 2D/3D data. Nonetheless, there are still opportunities for further improvement to enhance and enrich multi-sample data analysis by accommodating emerging modalities, addressing evolving analysis demands, and integrating additional omics to support scientific research. It is imperative to leverage spatial and feature information, particularly in spatiotemporal datasets (referred to as 4D datasets), to unlock insightful biological discoveries. Stereopy is committed to expanding its repertoire of analysis functions and extending its application domains to diverse areas, including clinical and immune research. The support for multimodal analysis and multi-omics datasets should be prioritized as they provide a wealth of biological information and represent the future of spatial omics technologies.

Despite the prevalence of research involving multiple samples, the research community dedicated to multi-sample analysis remains relatively underdeveloped. This deficiency can be attributed to the absence of a standardized framework that integrates various analysis tools and elucidates the canonical forms of multi-sample multi-omics analysis. The integration of algorithms and tools into a unified framework poses formidable obstacles for the joint analysis of multiple samples, compelling researchers to either forego the valuable insights or invest substantial time in searching for appropriate tools. Stereopy emerges as a foundation for building a vibrant multi-sample omics community and promoting the establishment of canonical forms for data analysis. By introducing the developer mode, Stereopy further encourages contributions from the bioinformatics community, fostering collaborative efforts. With unwavering dedication, Stereopy strives to furnish researchers with a user-friendly toolkit and robust analysis modules.

## Methods

### Comparison of Stereopy, Scanpy, Seurat, and Giotto toolkits for general single-cell analyses

Various toolkits have provided multiple functions for single-cell or spatial transcriptomic analyses. Among these, Scanpy, a Python package, and Seurat<sup>4</sup>, an R package, have gained widespread popularity. Giotto<sup>2</sup>,

another R package, is specifically designed for ST. This comparison aims to comprehensively evaluate the time consumption performance of Stereopy, Scanpy, Seurat, and Giotto toolkits, focusing on general single-cell analysis and spatial transcriptomic analysis, which includes essential steps such as pre-processing, principal components analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), cell neighbor finding, Louvain clustering, Leiden clustering, and gene marker identification. To assess performance on multiple samples, Stereo-seq mouse embryo datasets ranging from E9.5 to E14.5 were utilized. Since Stereopy is the only toolkit capable of analyzing multi-sample data, the remaining toolkits were evaluated by merging the multi-sample data into a single dataset. To ensure fairness in the comparison, all hyperparameters were held constant. Pre-processing involved testing three steps: normalization, *log1p* transformation, and scaling. PCA was performed by retaining the top 30 principal components (PCs) without considering highly variable genes. Cell neighbor analysis was conducted based on PCA results using the top 20 PCs and identifying the 10 nearest neighbors. Louvain and Leiden clustering algorithms were employed, with the default resolution set to 1. Gene marker identification involved testing on pre-annotation clustering results and utilizing a t-test based on an all-versus-rest approach. All toolkits were tested on a Linux machine with a 64-core CPU and 512 GiB of RAM.

### Cell co-occurrence detection algorithm

To explore changes in cell neighborhood, we developed a global co-occurrence method (Supplementary Fig. 8) to reflect spatial distribution relationships between cell types or clusters. The presented co-occurrence method is composed of three essential steps: 1) Calculation of cell-to-cell spatial distance, 2) Construction of a spatial graph, and 3) Counting of cell-type contacts. In the first step, we calculate a cell-cell pairwise spatial distance matrix based on the Euclidean distance metric. Secondly, we construct a cell neighborhood graph using the distance matrix as the adjacency matrix. To ensure the graph captures meaningful spatial relationships, we only retain edges within a defined distance range bounded by a minimal distance threshold and maximal distance threshold. These thresholds can be manually selected based on the specific context of the analysis. After constructing the cell neighborhood graph, we calculate the probability that cell type A has an edge with cell type B. This probability represents the co-occurrence probability of cell type A with cell type B. The following equations explain this process in more detail. We mark cells belonging to the cell type A  $CA[1, N]$  and M cells belonging to  $CB[1, M]$  as follows, of which  $M, N$  refer to:

$$CA_i \in \{CA_1, CA_2, \dots, CA_N\}$$

$$CB_i \in \{CB_1, CB_2, \dots, CB_M\} \quad (1)$$

Cell counts of cell type A are given by the number of A cells that are located around cell type B from the minimum distance to the maximum distance:

$$\text{count}(A, B, \text{min\_distance}, \text{max\_distance}) = C \quad (2)$$

The co-occurrence of A with B is defined as the probability of the cell type B appearing around each cell of the cell type A,  $P(A|B)$  at distance from *min\_distance* to *max\_distance*:

$$\text{Co-occurrence}(A, B) = P(B, |, A) = C/N \quad (3)$$

Notably, the co-occurrence of A with B may not be equal to the co-occurrence of B with A, which equals  $e/n$  and  $e/m$  respectively in our method. This asymmetry in our co-occurrence arises from the fact that the spatial distribution of a cell type includes another cell type and is

more universally distributed. For example, ECs are universally distributed in mouse kidney which means most other cells are connected to ECs while not every EC is connected to other cell types. Owing to the asymmetry of our co-occurrence method, we can also detect the wideness of spatial distribution for cell types. Additionally, we created the co-occurrence result integration method for multiple samples based on the weighted mean of the group, where weights equal to the 17 cell counts ratio in multi-sample data

$$\begin{aligned} \text{Co-occurrence}(A, B)_{\text{group}} \\ = \sum_{i=1}^S \left( \frac{N_i}{\sum_{i=1}^S N_i} * \text{Co-occurrence}(A, B)_i \right) i \in (1, 2, \dots, S) \end{aligned} \quad (4)$$

The grouped co-occurrence is equal to the result of merged multiple samples calculated per sample since all samples originate from the same tissue.

On the other hand, we use the difference to indicate the co-occurrence between two samples or two groups.

$$\begin{aligned} \text{Co-occurrence}(A, B)_{\text{differential}} = \text{Co-occurrence}(A, B)_1 \\ - \text{Co-occurrence}(A, B)_2 \end{aligned} \quad (5)$$

The differential co-occurrence value ranges from -1 to 1, where the positive value represents improvement of co-occurrence, and vice versa. To be noticed, the differential co-occurrence can also be calculated between two groups.

### Benchmarking of co-occurrence algorithm

To compare the performance of cell type co-occurrence of Stereopy with Squidpy, we tested on mouse kidney WT and BTBR samples<sup>30</sup>. Since there is no ground truth for cell co-occurrence. We compared the results with previously reported findings. The co-occurrence is calculated based on the cell spatial neighborhood and the distance traverse from 0 to 180 in steps size of 30, unit same as the resolution of slide-seq V2 technology which is 10  $\mu\text{m}$  (Supplementary Fig. 9a). For Stereopy, we use co-occurrence function with default parameters while for Squidpy we use co\_occurrence with parameters spatial\_key = spatial, interval = np.array([0,30,60,90,120,150,180]) and n\_splits = 1. As a result, Stereopy shows a more obvious co-occurrence of podocytes and GC cells than Squidpy, which is consistent with Marshall's findings<sup>30</sup>. In addition, with the help of grouped and differential co-occurrence among multi-sample analysis, Stereopy is capable of finding the similarities and diversities of cell type co-occurrence among multiple samples. Compared to the significant decrease in co-occurrence of GC, MC with itself in Squidpy, Stereopy can exhibit more significant changes between multiple articles, such as the reduction of co-occurrence between PCT\_1 and PCT\_2 (Supplementary Figs. 9c, 33).

### Cell community detection (CCD) algorithm

The functionality of a tissue is tightly coupled with the cell populations inhabiting it. The neighborhood of each cell impacts its gene expression patterns<sup>5</sup>. Here, we define a cell community as a spatial tissue area that exhibits a consistent distribution of cell types across all of its parts or regions. Therefore, detecting cell communities is of paramount importance to understanding the structure and function of the tissue. The main idea behind defining functional tissue domains (communities) can be narrowed to detecting tissue areas with the same cell mixture (percentages of cell types).

To address this, we developed the Cell Community Detection (CCD) algorithm. For each cell-spot  $cs_{u,v,k}$ , with spatial coordinates  $(u, v)$ , in slice  $k$ ,  $k = 1..K$ , CCD obtains a cell community label  $l_{u,v,k}$ ,  $l_{u,v,k} \in l_1, l_2, \dots, l_M$  (Supplementary Fig. 4). Then, it performs spatial convolution of cells using cell-type aggregating kernels  $c_{d,s}$  with a specific size  $d$

and sliding step  $s$ ,  $\{d, s\} \in [\{d_0, s_0\}, \{d_1, s_1\}, \dots, \{d_W, s_W\}]$ . The inclusion of multiple kernel support empowers CCD to extract both larger uniform cell communities and narrower border communities, which hold significant value in biological comparative studies. The kernels traverse the tissue, systematically scanning each window  $w_{i,d,s,k}$  and calculating the cumulative presence of  $N$  distinct cell types within its receptive field, creating feature vectors  $f_{i,d,s,k}$  of cell-type percentages  $([p_1, p_2, \dots, p_N])$ :

$$f_{i,d,s,k} = c_{d,s} * w_{i,d,s,k} = [p_1, p_2, \dots, p_N] \quad (6)$$

Where  $i$  represent the number of each sliding window when the kernels traverse the tissue.

Feature vectors obtained from all kernel windows across all  $K$  slices are then fed to a clustering algorithm  $CA$  such as Leiden<sup>25</sup>, Spectral<sup>61</sup>, or Hierarchical<sup>62</sup> to obtain community labels  $l_{i,d,s,k}$ :

$$CA(f_{i,d,s,k}) \rightarrow l_{i,d,s,k}, l_{i,d,s,k} \in l_1, l_2, \dots, l_M \quad (7)$$

The number of desired communities  $M$  can be predefined as a parameter (Spectral or Hierarchical clustering) or by setting the resolution of clustering (Leiden). It is worth emphasizing that in the case of multiple slices, the CCD analyzes all window data of all slices together, making the community results more robust, eliminating slice constraints, and facilitating meaningful slice comparisons. Furthermore, it enables the detection of distinct communities between slices, contributing a valuable feature to the analysis.

After obtaining window community labels  $l_{i,d,s,k}$ , the community label  $l_{u,v,k}$  of cell-spot  $cs_{u,v,k}$  is derived by majority voting (MV), using community labels of all windows covering the cell-spot:

$$l_{u,v,k} = \text{MV}(\forall l_{i,d,s,k}, \text{where } cs_{u,v,k} \in w_{i,d,s,k}), l_{u,v,k} \in l_1, l_2, \dots, l_M \quad (8)$$

While cell co-occurrence calculates the probability of cell type  $B$  appearing around each cell of cell type  $A$  in the entire sample, cell community calculates the distribution of cell types within a sliding window.

**Determining optimal window size.** The kernel  $c_{d,s}$  parameters, window size  $d$  and sliding step  $s$  are optional. In cases where no specific pair of values is provided, an iterative process is employed to determine the optimal window size. During the first iteration, the initial window size is derived by dividing the smaller spatial range of  $u$  and  $v$  by 100 and rounding the result to the nearest even number. Subsequently, the average number of cells covered by windows of this size is computed. If the calculated average falls below 30, the window size is increased by 10%; conversely, if it exceeds 50, the window size is decreased by 10%. The step is repeated until the average number of cell spots of all windows falls within the desired range<sup>30,48</sup>. The sliding step is then set to half of the determined window size.

**Spatially aware cell type filtering.** Some cell types widely present in all spatial areas of the tissue can form irrelevant and false cell communities with any other cell type, leading to potentially misleading findings. Additionally, these false communities have the potential to substitute more significant and relevant communities. To mitigate these false detections, we incorporated an automatic cell type removal mechanism into the CCD algorithm. This addition enables the identification and elimination of cell types that are uniformly or randomly present in all parts of the tissue (for example, Erythrocytes in mouse embryo brain). The spatial distribution of each cell type in the tissue can be evaluated by measuring the level of cell dispersion, connectedness and grouping. Cell-spots in ST tissue slices are spatially sparse and require special preprocessing for assessing their spatial organization. For each cell type  $n$ ,  $n \in [1..N]$  we create a binary image

$B_n$  with values 1 in all  $(u, v)$  positions of cells of type  $n$ , otherwise 0. The obtained binary image is then downsampled with the rate  $r$ , while all cell pixels are kept in their rounded positions. The rate  $r$  is provided by the user or calculated as a half of the minimal kernel size used. Two spatial metrics are calculated: entropy and scatteredness<sup>60,61</sup>. Entropy provides insight into the frequency of cell appearance in the tissue  $p_{A'}$ ,

$$\text{entropy}_n = -\sum(p_{A'} \cdot \log(p_{A'})) \quad (9)$$

with higher values implying the ratio of cell and non-cell pixels closer to 1. The scatteredness measure quantifies the degree of scattered or isolated regions in a binary image, which provides insight into the spatial dispersion and grouping of cells. It is calculated as the ratio of the number of connected components (objects) in the image to the total number of non-zero pixels. CCD supports setting the threshold values for these metrics to exclude from processing the cell types that are randomly or evenly spread throughout the tissue. Removing cell types with high entropy and scatteredness improves clustering and provides more robust cell communities.

**Information-aware window filtering.** The robustness and quality of CCD depend on the quality of clustering. To ensure stable clustering, feature vectors should contain a significant amount of information, that is, each evaluated window should consist of an adequate number of cell-spots. CCD gathers data on total cell numbers per window, for each kernel size  $d$ , and supports setting a threshold value for the minimum cell-spot number for the window's feature vector to be included in the clustering process. By default, the threshold is set to the value of  $-2\sigma$  of the window cell number distribution. Cell-spots are marked with the label of unknown if there are no cell community-labeled windows that overlap them

$$\text{if } \forall l_{i,d,s,k} = \text{unknown}, \text{ where } c_{u,v,k} \in w_{i,d,s,k} \implies l_{u,v,k} = \text{unknown} \quad (10)$$

Selective exclusion of tissue areas with sparse cell populations prior to clustering mitigates the formation of numerous small clusters, thereby enhancing the robustness of the clustering process. Moreover, if pertinent information is deemed essential from these excluded regions, they can be subject to processing using larger kernels, enabling comprehensive analysis when warranted.

### Benchmarking of cell community detection algorithm

To assess the stability and reliability of Stereopy's CCD, we conducted a comparison with existing algorithms for domain detection on three samples: single sample Stereo-seq mouse embryo whole brain<sup>3</sup>, Slide-seq V2 UMOD KI kidney comparative samples<sup>30</sup>, and Stereo-seq multi-sample adult mouse brain<sup>63</sup> (Supplementary Note 2.1). For a single sample, we included Giotto's Spatial Domain Identification (GSDI)<sup>15</sup>, SpaGCN<sup>28</sup> and GraphST<sup>23</sup> for comparison. In addition, for the multi-sample analysis, we included PRECAST<sup>21</sup>, BASS<sup>22</sup> and BANKSY<sup>29</sup> (Supplementary Note 2.2). Giotto and SpaGCN only support single-sample processing, creating results that require cluster matching to support further analysis. GraphST, BASS, PRECAST and BANKSY are all able to process multiple slices simultaneously. CCD can process a single sample as well as multiple samples simultaneously. SpaGCN was run with the default parameters (resolution = 1.5). Giotto's SDI required adjustment of gene expression and cell location data to a defined input format. Data was normalized with normalizeGiotto using scale-factor = 6000. Then, the functions createSpatialNetwork, binSpect, and initHMRF\_V2 were processed with  $k = 16$  for the brain sample, and  $k = 7$  for the kidney sample. Annotation was extracted with the doHMRF\_V2 function and visualized independently. GraphST is run with default parameters to obtain a 64-dimensional representation of cells. Then, Louvain is applied to cluster each sample by adjusting the

resolution until a similar number of clusters as CCD is achieved. Seurat objects for each slice were created for both BASS and PRECAST, and default values were used for all parameters, together with the desired number of clusters. To enhance the relevance of the results for domain detection, the BANKSY algorithm was executed with a key parameter, lambda, set to 0.8. This parameter value was chosen to optimize the performance of the algorithm in identifying relevant domains. Additionally, the parameter num\_clusters was set to the desired number of clusters, allowing for the specific identification of distinct domains within the dataset. CCD for mouse embryo whole brain sample was ran with win\_sizes = 150, sliding\_steps = 50, cluster\_algo = spectral and n\_clusters = 16, while for multi-sample adult dataset parameters were winsizes = 200, sliding\_steps = 50, cluster\_algo = agglomerative (Hierarchical) and n\_clusters = 16. All parameters were chosen to provide, on average, 30–40 cells per window, while keeping the communities smooth and coherent.

### Evaluation metrics

We utilized two metrics to evaluate the performance of various algorithms in generating results: Scatter and Density BetWeen clusters (S-Dbw)<sup>64</sup> and SD validity index<sup>56</sup>. S-Dbw considers both cluster separation and cluster cohesion. It measures how well-separated clusters are from each other (good separation) while also considering how tightly the data points are grouped within each cluster (good cohesion). SD validity index combines the measures of average cluster scattering and total separation between clusters. These dual considerations make S-Dbw and SD more comprehensive metrics for this purpose than the silhouette score that measures how similar each data point in one cluster is to the data points in the neighboring clusters. Additionally, clear boundaries for these spatial domains are crucial to analyze cell type proportions and spatial distributions within each domain. Thus, we used Moran's I to evaluate the spatial auto-correlation of all spatial domain methods. The total benchmark result can be found in Supplementary Table 1, CCD provides lower S-Dbw and SD scores as well as higher Moran's I than other algorithms, confirming better cluster cohesion and grouping. Meanwhile, we compared the execution time and memory consumption of GSDI, SpaGCN, GraphST, PRECAST, BASS, and CCD (Supplementary Table 2). The execution time of the CCD is faster compared to the GSDI and SpaGCN, demonstrating a speedup of at least 90 and 35 times, respectively. The peak memory consumption is affected by the dimensions of the input file, rendering CCD more efficient due to its independence from gene expression matrices.

### Comparison on mouse embryo brain sample, and region analysis.

The mouse embryo brain (Supplementary Fig. 19a), a structurally well-explored sample, was used for comparison of spatial domain detection methods, and further analysis of the biological significance of CCD communities. Supplementary Fig. 5b provides a comparison of spatial regions obtained by Stereopy's CCD, Giotto's SDI, SpaGCN, and GraphST, with the numbers of domains fixed. SpaGCN fails to provide domain integrity. Both GSDI and CCD detect layers in the dorsal pallium, as well as the thalamus. However, CCD provides smoother and more coherent regions (Supplementary Table 1), with the detection of several more separate communities. The cell communities detected by Stereopy's CCD are composed of multiple neighboring cell types and correspond to functional tissue domains. To evaluate the CCD's ability to infer biological function and structure, we analyzed separate regions and their correspondence with known functional and anatomical regions. Supplementary Fig. 5c, d displays the region and composition of two communities which show significant spatial matching with Hotspot<sup>9</sup> gene modules, and anatomical regions from Allen brain map<sup>10</sup>. The orange community represents a cell type-homogenous region, with 70% of dopaminergic neurons (Die GNeu) and 23% of midbrain glutamatergic neuroblasts (Mb Glu Neu) as main

components, where other cell types appear in abundancies less than 4%. Although these cells can be found in other areas of the tissue (Supplementary Fig. 5c, second column), this region is defined by the specific mixture of cell types, that is, a specific tissue domain. This community is spatially matched with the Hotspot gene module, as well as with the anatomical region of dorsal tier of thalamus (Supplementary Fig. 5c, columns three and four). The brown community is heterogeneous and contains, on average, 30% forebrain GABAergic neuron cells (Fb Glu NeuB), 29% cortical intermediate progenitor cells (Corti prog), 13% of cortical or hippocampal glutamatergic neuron cells (CortiHippo Glu Neu) and 10% of cortical glutamatergic neuron cells (Corti Glu Neu) (Supplementary Fig. 5d, first and second column). This region is shown to coincide with the gene module obtained by Hotspot, and when comparing with Allen brain atlas annotation, it corresponds to the mantle zone of dorsal pallium (Supplementary Fig. 5d, third and fourth column). These results confirm the ability of CCD to extract biological information.

**Comparison on multi-sample adult mouse brain sample.** Three samples were processed separately by GSDI, GraphST, and SpaGCN, while BASS, PRECAST, and CCD employed their multi-sample approach (Supplementary Fig. 6b). SpaGCN manages to obtain anatomical regions with clear borders but provides an unstable number of domains for consecutive samples while using the same parameters (Supplementary Fig. 6b). When comparing per sample, domains obtained by GSDI, PRECAST, BASS, GraphST, and CCD are similar by constitution. However, multi-sample processing provides more coherent (Supplementary Table 1) and anatomically matching results with higher reliability of inter-sample domain matching. Selected samples have similar cell type shares (Supplementary Fig. 6c). Thus, the consistency of CCD's communities throughout samples is confirmed with stable tissue-share communities in all samples (Supplementary Fig. 6d) together with lowest S-Dbw and SD scores (Supplementary Table 1). Execution of 3 slices of adult mouse brain CCD finishes in 214 s while consuming 25,716 MB. It costs less in terms of execution time and memory compared to other tools (Supplementary Table 2).

**Comparison on UMOD KI/WT sample.** CCD, BASS, and PRECAST provide joint analysis of both UMOD KI and WT samples, while Giotto's SDI, GraphST, and SpaGCN perform on each sample separately (Supplementary Fig. 7c). We compared the results generated by these algorithms, especially the medulla region according to the annotation obtained from the Marshall et al. paper. CCD provides domains of higher integrity and robustness compared to BASS, PRECAST, GSDI, and SpaGCN, especially in the medulla region on which CCD identified almost the same region with the annotation from the original paper. GSDI, BASS, PRECAST, and SpaGCN detected more than one region and even mixed regions in the medulla area, while CCD and GraphST detected regions consistent with Marshall et al. paper. To further demonstrate the consistency of CCD regions, we calculated the marker genes for each of them. Marker genes show consistency of the gene expression and the cell community region in both UMOD KI and WT samples (Supplementary Fig. 11). CCD manages to process these two kidney samples in 31 seconds while consuming only 684 MB. It costs less in terms of execution time and memory compared to other tools (Supplementary Table 2).

### Temporal gene pattern identification (TGPI) algorithm

It is of interest that the gene expression exhibits a certain pattern during distinct biological processes. Among various kinds of gene patterns, up- or down-regulation is a prevalent phenomenon. Here we devised a approach for identifying genes exhibiting up- or down-regulated expression, utilizing the serial test along time series and cell type trajectory. Note that the t-test statistics might lead to false

positive discovery as described in Lamian<sup>65</sup>. Thus, we have implemented the permutation test and Wilcoxon test in TGPI. By employing a one-tailed test, we obtain statistical scores and *p*-values between consecutive time points. To capture both ends of the spectrum, we calculate *p*-values for both higher and lower expression, thereby characterizing up- and down-regulated genes, respectively.

$$Pl_t = \text{permutation\_test}(\mathbf{exp}_{t+1}, \mathbf{exp}_t, \text{alternative} = 'less')$$

$$t \in \{1, 2, \dots, T-1\}$$

$$Pg_t = \text{permutation\_test}(\mathbf{exp}_{t+1}, \mathbf{exp}_t, \text{alternative} = 'greater')$$

$$t \in \{1, 2, \dots, T-1\} \quad (11)$$

Then we proposed the False Positive Risk Score (FPR Score) to combine the *p*-value so that we can sort out the most up- or down-regulated genes. We proposed the FPR score and integrated its metrics into the TGPI framework. Our hypothesis is rooted in the notion that the alternative probability signifies an increase in gene expression between adjacent time points, with each time point being independent. Consequently, the FPR score serves to highlight the significance of serial up or down-regulation. Its calculation adheres to the following formula:

$$FPR \text{ score} = 1 - \prod_{i=1}^k (1 - P_i) \quad (12)$$

While some genes exhibit straightforward serial up or down-regulation, others manifest more complex patterns during certain biological processes. To automatically capture all types of patterns, we draw inspiration from Mfuzz<sup>35</sup> and employ fuzzy C means clustering of genes. Our methodology, Stereopy, takes into account both spatial and temporal expression features, yielding more biologically significant results.

**Definition 1 (Spatial features).** Stereopy calculates PCA based on rasterized expression on a certain bin size, and uses the first several principal components as spatial features. We use rasterized expression by sum the expression of a gene within a bin size square, which can eliminate the impact of cells with relatively high expression.

$$\mathbf{exp}_{i,j} = \text{sum}(\mathbf{exp}_{x,y}), x = \text{int}\left(\frac{i}{\text{bin size}}\right), y = \text{int}\left(\frac{j}{\text{bin size}}\right) \quad (13)$$

$$f_{\text{spatial}} = \text{PCA}\left(\left\{\mathbf{exp}_{1,1}, \mathbf{exp}_{2,1}, \dots, \mathbf{exp}_{i,j}, \dots, \mathbf{exp}_{I,J}\right\}\right)_{i \in [1, I], j \in [1, J]} \quad (14)$$

$\mathbf{exp}$  represents the expression in coordinate(*i*, *j*).

**Definition 2 (Temporal features).** Stereopy utilizes the result from serially up/downregulated genes as input. We use the serial greater *p*-value  $Pg_i$  and serial less *p*-value  $Pl_i$  as features for each gene based on the following formula:

$$f_{\text{temporal}} = \frac{Pl_i - Pg_i}{|Pl_i - Pg_i|} \times \max(1 - Pl_i, 1 - Pg_i), f_{\text{temporal}} \in (-1, 1) \quad (15)$$

The  $Pg_i$  represents the probability of the gene expression not being up-regulated at the *i*+1 timepoint compared with the *i* timepoint. Consequently,  $1 - Pg_i$  can be interpreted as the probability of the gene being up-regulated at the *i*+1 timepoint according to the alternative hypothesis. Similarly,  $1 - Pl_i$  represents the probability of the gene being down-regulated at the *i*+1 timepoint. By combining the probabilities of up and down regulation, with a minus sign to indicate down regulation, we assign lower values of  $f_i$  to represent down-regulated genes and higher values to represent upregulated genes. In

this manner, we consider  $f_i$  as the tendency of a gene between adjacent time points. In contrast to Mfuzz, which utilizes mean expression values as input, Stereopy's temporal feature places greater emphasis on the tendency rather than the original gene expression levels.

**Definition 3** (Spatiotemporal features). To incorporate both temporal and spatial information, we concatenate the scaled spatial features with the first  $N$  spatial features and to temporal features for each gene. A parameter alpha is also introduced to weigh the effect of spatial features.

$$\mathbf{f} = \text{concat}(\mathbf{f}_{\text{temporal}}, \alpha \mathbf{f}_{\text{spatial}}[1 : N_{\text{spatial feature}}]) \quad (16)$$

Finally, we employ the Fuzzy C means algorithm to cluster genes into distinct groups. The main objective of the Fuzzy C means algorithm is to minimize  $J$  according to the following equation:

$$J(u, v) = \sum_{i=1}^N \sum_{j=1}^C u_{i,j}^m |f_i - v_j|^2, i \in [1, N], j \in [1, C] \quad (17)$$

$$u_{i,j} = \frac{1}{\sum_{k=1}^C \left( \frac{|f_i - v_j|}{|f_i - v_k|} \right)^{\frac{2}{m-1}}} \quad (18)$$

$$v_j = \frac{\sum_{i=1}^N u_{i,j}^m f_i}{\sum_{i=1}^N u_{i,j}^m} \quad (19)$$

In this equation,  $f_i$  represents the feature that combines both temporal and spatial, and  $v_j$  belongs to the center of each cluster.  $m$  is the fuzziness and is set to 2 by default.  $u_{i,j}^m$  denotes the membership of the  $i$  gene in the  $j$  cluster:

$$\sum_{j=1}^C u_{ij} = 1, u_{ij} > 0 \quad (20)$$

The termination condition of the iteration is defined as follows, where  $\varepsilon$  is the threshold with default as  $10^{-15}$ :

$$\max \left( |u_{i,j}^{k+1} - u_{i,j}^k| \right) < \varepsilon \quad (21)$$

By incorporating spatial features, Stereopy's TGPI can differentiate gene clusters that exhibit similar temporal expression patterns but varying levels of spatial differential expression. This integration of spatial information enhances the biological significance of the results obtained (Supplementary Figs. 17, 18).

### Benchmarking of temporal gene pattern identification algorithm

The evaluation of Stereopy's TGPI contains two important modules: 1) the false positive risk score FPR score proposed in TGPI to find the up-/down-regulated genes. 2) the whole temporal gene pattern detection algorithm TGPI in spatial-resolved temporal datasets.

**False positive risk score evaluation.** We benchmarked our proposed  $p$ -value combination metric FPR score on the temporal mouse forebrain datasets. We made three comparisons with respect to three cell types including dorsal forebrain, forebrain neuronal intermediate progenitor, and forebrain cortical glutamatergic in these datasets (Supplementary Figs. 14–16). According to the occurrence of three cell types, three comparisons contain 7, 5, and 3 time points respectively. FPR score can find real continuously up-/down-regulated genes that have a stable tendency of rise or fall gene expression along with the time series in all these comparisons.

**Temporal gene pattern detection algorithm evaluation.** We first tested the effect of spatial features.  $N$  top spatial features range from 3 to 6 are tested. Taking *Foxg1*, *Hes5*, and *Mab21l2* as examples, we tested the 4 nearest neighbors (NN) of these genes according to Euclidean distance based on  $N$  top spatial features. (Supplementary Fig. 17). From the result we observed that a similar spatial expression pattern is detected in each 4NN gene. The higher the  $N$  spatial feature is, the more similar spatial expression patterns can be observed (Supplementary Fig. 17). Additionally, as the  $N$  spatial feature reaches 5, the 4NN genes tend to be constant. Since the  $N$  spatial feature can reflect the spatial expression feature, we tested its influence on TGPI (Supplementary Fig. 18a). The result indicated that the increment of the  $N$  spatial feature resulted in higher consistency of genes in a temporal pattern to some extent (blue box). Moreover, with the help of spatial features, TGPI can distinguish genes with similar temporal patterns. For example, Cluster 2 and Cluster 8 of TGPI with  $N$  spatial feature equal to 3 are similar in temporal expression pattern and divergence in spatial expression pattern (Supplementary Fig. 18b, c).

To evaluate TGPI's performance on real datasets, we compared the TGPI algorithm with another time series gene pattern method called Mfuzz<sup>35</sup>. The Stereo-seq mouse embryo brain data from E9.5 to E16.5, which is the subset of the mouse embryo dataset with annotation as Brain, is used to evaluate the performance of TGPI<sup>3</sup>. Genes were clustered into 8 clusters for both Stereopy and Mfuzz. To evaluate the performance of the gene pattern results, we calculate the Pearson's correlation of gene expression with the pseudotime and ANOVA test among time points. The F-score of the ANOVA test is used to reflect the divergence between time points. If a certain gene is more related to time points, the F score will be higher. We calculated the top 1000 genes of each TGPI cluster ordered by weights of clustering results for both Stereopy and Mfuzz. The results show that most TGPI clusters exhibit not only a higher F score in the ANOVA test among time points but also higher Pearson's correlation with pseudotime, which means genes within the gene pattern identified by TGPI are more related to both time point and pseudotime (Supplementary Fig. 19a). Moreover, from the GO enrichment results, we concluded that TGPI is more capable of grouping genes with the same expression pattern and functions (Supplementary Fig. 19b, c). We conducted GO enrichment analysis on the top 20 genes of each cluster for both TGPI and Mfuzz. As a result, 6 gene pattern clusters of TGPI enriched GO terms while only 3 clusters enriched GO terms for Mfuzz with the same  $p$ -value cut off ( $p = 0.05$ ).

### 3D cell-niche regulatory network prediction algorithm

The Stereopy-NicheReg3D algorithm, designed for predicting 3D cell-niche regulatory networks, encompasses an initial step of adaptive 3D niche reconstruction, followed by cell-niche communication prediction at the single-cell resolution. This analysis is based on the high-quality reconstructed 3D architecture of a series of adjacent 2D sections, which necessitates the correlations between spots/genes across different sections and within each section in the preliminary cell clustering and annotation to guide the reconstruction and the process of the 3D integration. In this study, the alignment has been completed by using the automated Fused Gromov-Wasserstein Optimal Transport-based algorithm, ST-GEARS<sup>66</sup>, and additional manual curations. Note that ST-GEARS improves the alignment accuracy by incorporating both gene expression and structural similarity into the optimization process of anchors with cell type-specific distributive constraints. These anchors further connect to the slice-specific groups of spots and genes and transmit the elastic registration across each entire slice, recovering the 3D regionalization of each cell type. We investigated potential batch effects using BatchEval<sup>67</sup>. The results suggested that there were no significant batch effects among the 59 2D samples.

To ensure the accurate modeling of juxtacrine signaling processes, three fundamental principles are proposed to govern the participation of cells:

1. Bordering principle: Cell-cell communications (CCC) between each pair of cell groups should be restricted to their border regions within a reasonable spatial distance. This principle is underpinned by the hypothesis that extracellular CCC predominantly occurs through signaling molecules released by sender cells, which can be sensed by receiver cells in close proximity.
2. Non-confusion principle: CCC analysis should be confined to regions where the cellular composition is relatively uncomplicated. As the number of cell types within a region increases, accurately determining the exact source and target of the signals becomes challenging.
3. Multi-slice principle: The niche should be constructed in a 3D environment rather than a 2D plane. Therefore, employing a 3D multi-slice analysis that more accurately simulates the cellular environment is preferable over performing CCC analysis on individual 2D slices.

In accordance with these principles, the fundamental step in investigating cell-type interactions involve the construction of the 3D niche associated with the central cells. This task is accomplished through an adaptive edge detection approach. The gene expression data, sequenced and annotated in multiple spots that form an organized 3D point cloud denoted as  $\mathcal{C}$ , serve as the basis. Given a central group or type of cells of interest  $T$ , the identity of any spot  $p_i \in \mathcal{C}$  can be represented with an indicator function

$$I_T(p_i) = \begin{cases} 1, & \text{if the spot is assigned to cell type } T \\ 0, & \text{if the spot is assigned to other types} \\ -1, & \text{if the spot captures nothing} \end{cases} \quad (22)$$

Define all the neighboring spots of  $p_i$  of cell type  $T$  within a cube of size  $s$  as  $\mathcal{N}_i(s) = \{n_1, n_2, \dots, n_k\}$  (excluding unassigned spots). The shift of the centroid for  $p_i$  is computed as

$$D_i = 1 - \frac{1}{|\mathcal{N}_i(s)|} \sum_{j=1}^k I_T(n_j) \quad (23)$$

We construct the border region of cell type  $T$  by identifying spots with relatively large shifts. The threshold is chosen in an adaptive and dynamic manner, catering to the local confusion level of cell types. Let the information entropy in the square centered at  $p_i$  be

$$S_i = - \sum_{j=1}^n p(x_j) \log p(x_j) \quad (24)$$

where  $n$  represents the total number of cell types in the data, and  $p(x_j)$  is the frequency of cell type  $j$  in the cube.

**Definition 4** (Border region). Given a cell type  $T$ , the border region of  $T$  is defined as

$$B_T = \{p_i \in T \mid D_i > \theta \cdot S_i\} \quad (25)$$

Here,  $\theta$  is a constant parameter controlling the magnitude of the threshold. If the square area is highly fused, then a stricter threshold is applied. With all spots identified within the border region, we can construct the 3D niche using a variable-sized sliding cube at the cellular resolution.

**Definition 5** (Niche). Given a cell type  $T$ , let  $S_k$  denote the set of all cells of type  $k$ ,  $\mathcal{N}_i(\cdot)$  represent the cube centered at the spot  $p_i$ , and the niche for cell type  $T$  is defined as follows:

$$N_T = \bigcup_{k \neq T} \{p_i \in B_T, p_j \in S_k \mid p_j \in \mathcal{N}_i(s(n))\} \quad (26)$$

where the size  $s$  of each  $\mathcal{N}_i$  is determined such that there are at least  $n$  spots of type  $k$  within the cube.

Next, we perform a label permutation-based statistical CCC analysis to identify significant cell-niche L-R pairs by incorporating both L-R gene co-expression and 3D colocalization of the cells. In brief, we collect potential L-R pairs and construct a customized Liana consensus database<sup>43</sup> ([https://github.com/saezlab/liana-py/tree/main/liana/resource/omni\\_resource.csv](https://github.com/saezlab/liana-py/tree/main/liana/resource/omni_resource.csv))

$$Pair(L, R) = \bigcup_{i \in \{20 \text{ databases}\}} Pair_i(L, R) \quad (27)$$

It is converted to the CellPhoneDB format using the command `cellphonedb database generate`, where input files of protein, gene, complex, and interactions are created according to the Liana consensus lists. We then follow a similar approach reported by CellPhoneDB<sup>45</sup> to compute the average expression level of the ligand in the sender cells and that of the receptor in the receiver cells within the previously calculated niche.

**Definition 6** (Communication Score). Given the average expression level of ligand  $l$  in sender cell type  $i$ ,  $x_i^l$ , and the average expression level of the receptor  $r$  in receiver cell type  $j$ ,  $x_j^r$ , the communication score is defined as the mean value of the average L-R expression within a 3D niche:

$$s_{ij}^{lr} = \frac{1}{2} (x_i^l + x_j^r) \quad (28)$$

The significance of the true communication scores is evaluated through random shuffling of cell-type labels of spots within the niche for  $m$  times. The  $p$ -value is defined as the proportion of random shuffles that reach a score higher than the true score:

$$p = \frac{\sum_{k=1}^m I_{\{x > s_T\}}(s_k)}{m} \quad (29)$$

where  $s_T$  is the true communication score,  $s_k$  is the calculated communication score at the  $k$ th shuffle,  $I_{\{x > s_T\}}(\cdot)$  is the indicator function defined as

$$I_{\{x > s_T\}}(x) = \begin{cases} 1, & \text{if } x > s_T \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

Typically, a  $p$ -value smaller than 0.05 suggests that the corresponding L-R interaction is statistically significant. This suggests that the observed true communication score is unlikely to be solely attributed to chance.

Cells receive signals from spatially proximate cells within the niche through the expression of receptor proteins. These signals further initiate a cascade of intricate intra-cellular regulatory events that culminate in the modulation of gene expression within the recipient cells. This cascade of gene regulation forms a crucial mechanism by which cells respond and adapt to their microenvironment.

To demonstrate the possible regulation mechanism, we eventually connect significant L-R interactions detected in the cell-niche communication analysis with the TF-centered regulons identified by the SpaGRN<sup>68</sup> analysis based on the integrated weighted ligand-signaling network from Nichenet-v2<sup>60</sup> ([https://zenodo.org/record/7074291/files/weighted\\_networks\\_nsga2r\\_final\\_mouse.rds](https://zenodo.org/record/7074291/files/weighted_networks_nsga2r_final_mouse.rds)). This database contains 3,865,137 rows, each of which represents a pair of directed signaling interactions with a specific weight prioritized using 57 data sources. We convert the whole network data into a weighted directed graph  $G = (V, E, W)$ , where  $W: E \rightarrow \mathbb{R}$  is the weight for each edge. For a given receptor,  $v_{source}$  and TF,  $v_{target}$ , we search for the shortest expressed path between the two nodes using the Dijkstra

method and consider it as the potential signaling path between them. The distance of each graph edge is defined as the reciprocal of its weight:

$$d_{ij} = \frac{1}{w_{ij}} \quad (31)$$

where  $w_{ij}$  is the weight of the edge  $e_{ij} = (v_i, v_j)$  connecting node  $i$  and  $j$ . For each possible path  $P$ , its corresponding distance is defined as

$$D(P) = \sum_{e \in E(P)} d_e \quad (32)$$

where  $E(P)$  is all the edges that make up the path  $P$ .

### Benchmarking of cell-niche communication prediction algorithm

To demonstrate the algorithm efficiency, we systematically compared the general features of the Stereopy-NicheReg3D module with two other algorithms, namely CellPhoneDB<sup>45</sup> and NICHES<sup>50</sup>, using the same mouse heart dataset (Supplementary Table 6). We slightly modified two software tools to enable them to analyze the 3D SRT data. For the CellPhoneDB implementation, the spatial relationship of VCM and other cell clusters was initially provided, and the default parameters were used to obtain significantly expressed to- and from-VCM L-R pairs. NICHES was adopted to obtain single-cell-resolution interaction results. We then integrated the expression of L-R pairs coming from each niche component and landing on the VCM cells by summing the L-R expressions, and identified the cell type-specific L-R pairs using the Seurat FindAllMarkers function.

We benchmarked the performance of Stereopy-NicheReg3D and the other two tools on the same Linux system with an Intel Core Processor (Broadwell, IBRS) of 30 threads and 512 GB memory. Both Stereopy and NICHES enable the investigation of sender-receiver interaction at the single-cell level, which is usually computationally prohibitive for CCC analysis thanks to Stereopy's niche extraction and NICHES's subsampling strategies. These strategies also accelerate the computation compared to the whole cluster-based approach employed by CellPhoneDB (Supplementary Fig. 27). While CellPhoneDB does offer the option of subsampling, it is important to note that subsampling might preclude a complete view of CCC structure and risk obscuring significant L-R pairs. As a result, in terms of the number of specific CCC interactions, Stereopy obtains the most specific L-R pairs in all VCM-niche cases except ACM-VCM, almost covering those derived by other tools (Supplementary Fig. 27).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The processed datasets have been deposited in published papers. Slide-seq2 datasets: mouse kidney datasets are downloaded from<sup>30</sup>, of which Puck\_191204\_22.h5ad and Puck\_191204\_15.h5ad are used as BTBR WT and *ob/ob* sample respectively and Puck\_191223\_19.h5ad and Puck\_200104\_07.h5ad are used as WT and UMOD KI sample respectively. Stereo-seq datasets: a sample of 12 weeks adult mouse brain, mouse embryo SRT samples from E9.5 to E16.5, and entire 3D mouse embryonic heart datasets are downloaded from StomicsDB MOSTA<sup>71</sup>. Three adjacent samples of coronal mouse brain are downloaded from Spatial-ID<sup>63</sup>. Source data are provided with this paper.

### Code availability

The code used to develop the model, perform the analyses and generate results in this study is publicly available and has been deposited

in GitHub repository at <https://github.com/STOmics/Stereopy><sup>72</sup>, under MIT license. The specific version of the code associated with this publication is archived in Zenodo and is accessible via 10.5281/zenodo.14722436. The documentation of Stereopy can be found at: <https://stereopy.readthedocs.io/en/latest/>. All the code to reproduce the result of the analysis can be found at the following GitHub repository: <https://github.com/STOmics/Stereopy/tree/main/docs/source/Tutorials>. Users are permitted to reuse, modify, and distribute the code in accordance with the terms of the license. Any modifications to the code should appropriately credit the original authors as outlined by the license terms.

### References

- Velten, B. & Stegle, O. Principles and challenges of modeling temporal and spatial omics data. *Nat. Methods* **20**, 1462–1474 (2023).
- Mayr, U., Serra, D. & Liberali, P. Exploring single cells in space and time during tissue development, homeostasis and regeneration. *Development*, **146**, dev176727 (2019).
- Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022).
- Rodrigues, S. G. et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
- Chen, K. H. et al. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
- Eng, C. L. et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235–239 (2019).
- Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, 6400 (2018).
- Janesick, A. et al. High resolution mapping of the tumor micro-environment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* **14**, 8353 (2023).
- Lin, Y. et al. Atlas-scale single-cell multi-sample multi-condition data integration using scMerge2. *Nat. Commun.* **14**, 4272 (2023).
- Chen, W. T. et al. Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell* **182**, 976–991.e19 (2020).
- Roth, R. et al. Single-cell and spatial transcriptomics approaches of cardiovascular development and disease. *BMB Rep.* **53**, 393–399 (2020).
- Crosse, E. I. et al. Multi-layered Spatial Transcriptomics Identify Secretory Factors Promoting Human Hematopoietic Stem Cell Development. *Cell Stem Cell* **27**, 822–839.e8 (2020).
- Lohoff, T. et al. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat. Biotechnol.* **40**, 74–85 (2022).
- Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods*. **19**, 171–178 (2022).
- Dries, R. et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* **22**, 78 (2021).
- Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
- Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
- Gayoso, A. et al. A Python library for probabilistic analysis of single-cell omics data. *Nat. Biotechnol.* **40**, 163–166 (2022).
- Fang, S., et al. Computational Approaches and Challenges in Spatial Transcriptomics. *Genomics Proteomics Bioinformatics* **21**, 24–47 (2023).
- Bredikhin, D., Kats, I. & Stegle, O. MUON: multimodal omics analysis framework. *Genome Biol.* **23**, 42 (2022).
- Liu, W. et al. Probabilistic embedding, clustering, and alignment for integrating spatial transcriptomics data with PRECAST. *Nat. Commun.* **14**, 296 (2023).

22. Li, Z. & Zhou, X. BASS: multi-scale and multi-sample analysis enables accurate cell type clustering and spatial domain detection in spatial transcriptomic studies. *Genome Biol.* **23**, 168 (2022).
23. Long, Y. et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat. Commun.* **14**, 1155 (2023).
24. Guo, L., et al. VT3D: a visualization toolbox for 3D transcriptomic data. *J. Genet. Genomics* **50**, 713–719 (2023).
25. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
26. Blondel, V. D. et al. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P10008 (2008).
27. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
28. Hu, J. et al. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).
29. Singhal, V. et al. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Genet.* **56**, 431–441 (2024).
30. Marshall, J. L. et al. High-resolution Slide-seqV2 spatial transcriptomics enables discovery of disease-specific cell neighborhoods and pathways. *iScience* **25**, 104097 (2022).
31. Hong, S. Y. et al. Identification of the pivotal role of SPP1 in kidney stone disease based on multiple bioinformatics analysis. *BMC Med. Genomics* **15**, 7 (2022).
32. Patel, Y. P. et al. SPP1 and UMOD gene variants are synergistically associated with risk of renal stone disease. *Gene* **863**, 147264 (2023).
33. Saito, T. et al. Apolipoprotein E-related glomerular disorders. *Kidney Int.* **97**, 279–288 (2020).
34. Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
35. Kumar, L. et al. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* **2**, 5–7 (2007).
36. Tigani, W. et al. Foxg1 Upregulation Enhances Neocortical Activity. *Cereb. Cortex* **30**, 5147–5165 (2020).
37. Kohro, Y. et al. Spinal astrocytes in superficial laminae gate brainstem descending control of mechanosensory hypersensitivity. *Nat. Neurosci.* **23**, 1376–1387 (2020).
38. Zhang, Y. et al. Transcription factor 4 controls positioning of cortical projection neurons through regulation of cell adhesion. *Mol. Psychiatry* **26**, 6562–6577 (2021).
39. Zeira, R. et al. Alignment and integration of spatial transcriptomics data. *Nat. Methods* **19**, 567–575 (2022).
40. Chen, X. et al. Whole-cortex in situ sequencing reveals input-dependent area identity. *Nature* <https://doi.org/10.1038/s41586-024-07221-6> (2024).
41. Scadden, D. T. Nice neighborhood: emerging concepts of the stem cell niche. *Cell* **157**, 41–50 (2014).
42. Jing, X. et al. The Intracellular and Extracellular Microenvironment of Tumor Site: The Trigger of Stimuli-Responsive Drug Delivery Systems. *Small. Methods* **6**, e2101437 (2022).
43. Dimitrov, D. et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13**, 3224 (2022).
44. Litviňuková, M. et al. Cells of the adult human heart. *Nature* **588**, 466–472 (2020).
45. Garcia-Alonso, L. et al. Single-cell roadmap of human gonadal development. *Nature* **607**, 540–547 (2022).
46. Tucker, N. R. et al. Transcriptional and Cellular Diversity of the Human Heart. *Circulation* **142**, 466–482 (2020).
47. Larson, A. et al. Altered intercellular communication and extracellular matrix signaling as a potential disease mechanism in human hypertrophic cardiomyopathy. *Sci. Rep.* **12**, 5211 (2022).
48. Zhang, J. et al. Cardiac differentiation of human pluripotent stem cells using defined extracellular matrix proteins reveals essential role of fibronectin. *eLife* **11**, e69028 (2022).
49. Keefe, J. A. et al. Role of Ca<sup>2+</sup> in healthy and pathologic cardiac function: from normal excitation-contraction coupling to mutations that cause inherited arrhythmia. *Arch. Toxicol.* **97**, 73–92 (2023).
50. Raredon, M. S. B. et al. Comprehensive visualization of cell-cell interactions in single-cell and spatial transcriptomics with NICHES. *Bioinformatics* **39**, btac775 (2022).
51. Liu, J. et al. Wnt/ $\beta$ -catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Sig. Transduct. Target. Ther.* **7**, 3 (2022).
52. Wang, J. et al. The Hippo pathway in the heart: pivotal roles in development, disease, and regeneration. *Nature Reviews. Cardiol.* **15**, 672–684 (2018).
53. Sundaresan, N. R. et al. The sirtuin SIRT6 blocks IGF-Akt signaling and development of cardiac hypertrophy by targeting c-Jun. *Nat. Med.* **18**, 1643–1650 (2012).
54. Torrente, Y. et al. Role of Insulin-Like Growth Factor Receptor 2 across Muscle Homeostasis: Implications for Treating Muscular Dystrophy. *Cells* **9**, 441 (2020).
55. Landstrom, A. P., Dobrev, D. & Wehrens, X. H. T. Calcium Signaling and Cardiac Arrhythmias. *Circ. Res* **120**, 1969–1993 (2017).
56. Zhang, M. et al. Calcium/calmodulin-dependent protein kinase II couples Wnt signaling with histone deacetylase 4 and mediates dishevelled-induced cardiomyopathy. *Hypertension* **65**, 335–344 (2015).
57. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
58. Gan, P. et al. RBPMs is an RNA-binding protein that mediates cardiomyocyte binucleation and cardiovascular development. *Dev. Cell* **57**, 959–973.e7 (2022).
59. Yang, P. et al. The Association of the Copy Number Variation of the MLLT10 Gene with Growth Traits of Chinese Cattle. *Animals* **10**, 250 (2020).
60. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
61. Luxburg, U. et al. A tutorial on spectral clustering. *Statistics and Computing* **17**, 395–416 (2007).
62. Jansegers, M., Vanderschueren, C. & Enghels, R. J. C. L. Hierarchical Group. *Optim. objective Funct.* **58**, 236–244 (2015).
63. Shen, R. et al. Spatial-ID: a cell typing method for spatially resolved transcriptomics via transfer learning and spatial embedding. *Nat. Commun.* **13**, 7640 (2022).
64. Halkidi, M. & Vazirgiannis, M. Clustering validity assessment: finding the optimal partitioning of a data set. In: *IEEE International Conference on Data Mining* (IEEE, 2001).
65. Hou, W. et al. A statistical framework for differential pseudotime analysis with multiple single-cell RNA-seq samples. *Nat. Commun.* **14**, 7286 (2023).
66. Xia, T. et al. ST-GEARS: Advancing 3D downstream research through accurate spatial information recovery. *Nat. Commun.* **15**, 7806 (2024).
67. Zhang, C. et al. BatchEval Pipeline: batch effect evaluation workflow for multiple datasets joint analysis. *GigaByte* **2024**, giga-byte108 (2024).

68. Li, Y. et al. SpaGRN: investigating spatially informed regulatory paths for spatially resolved transcriptomics data. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/2023.11.19.567673v1> (2023).
69. Guo, X. et al. CNSA: a data repository for archiving omics data. Database, 2020. (2020).
70. Chen Fengzhen, Y. L. et al. CNGBdb: China National GeneBank DataBase. *Yi Chuan* **42**, 799–809 (2020).
71. Xu, Z., et al. STOmicsDB: a comprehensive database for spatial transcriptomics data sharing, analysis and visualization. *Nucleic Acids Res.* **52**, D1053–D1061 (2023).
72. Fang, S., et al. Stereopy: modeling comparative and spatiotemporal cellular heterogeneity via multi-sample spatial transcriptomics. *Stereopy*, <https://doi.org/10.5281/zenodo.14722436> (2025).

## Acknowledgements

This work is part of the SpatioTemporal Omics Consortium (STOC) paper package. A list of STOC members is available at: <http://sto-consortium.org>. We acknowledge the Stomics Cloud platform (<https://cloud.stomics.tech/>) to provide convenient ways for analyzing spatial omics datasets. We would like to thank Kai Han, Dantong Wang, Ping Qiu, Yunjia Zhang, Haohao Deng, Chang Shi, and Junfu Guo for their help. We acknowledge the CNGB Nucleotide Sequence Archive (CNSA)<sup>69</sup> of China National GeneBank DataBase (CNGBdb)<sup>70</sup> for maintaining the MOSTA database. This work is supported by National Key R&D Program of China (2022YFC3400400), National Natural Science Foundation of China (32300526, 32100514), General Program (Key Program, Major Research Plan) of National Natural Science Foundation of China (32170439).

## Author contributions

S.S.F., M.Y.X., L.C., X.B.L., M.B., L.W.T., V.K. and L.D.G. were responsible for algorithm development and implementation. Y.L., T.Y.X., B.J. and N.M. analyzed the data. C.L., M.N.C., Q.L. and Y.M.L. contributed to the data preprocessing. L.Y.G., J.H.H., L.A.L., Z.B.W., P.Q., L.Y.W. and F.Z.C. contributed to cleaning up the code and maintaining the website of Stereopy. Z.Y.Y., L.N.H., Z.Q.D., J.H.L., M.L., C.Z., Q.K. and S.K.L. contributed key ideas and advice. X.X., Y.X.L., A.C., Yong Zhang, G.Y.F. and Yi Zhao supervised the study.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-58079-9>.

**Correspondence** and requests for materials should be addressed to Yi Zhao, Guangyi Fan, Yong Zhang, Ao Chen, Yuxiang Li or Xun Xu.

**Peer review information** *Nature Communications* thanks Zilong Zhang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025