OXFORD

Briefings in Bioinformatics, 2025, **26(6)**, bbaf644
https://doi.org/10.1093/bib/bbaf644
Problem Solving Protocol

STIFT: spatiotemporal transcriptomics integration through spatially informed multi-timepoint bridging

Ji Qi[1], Muyang Ge[1], Jishuai Miao[1], Xiaocheng Zhou[1], Zhixiang Lin[1,2,3,*]

[1]Department of Statistics and Data Science, The Chinese University of Hong Kong, Ma Liu Shui, Sha Tin, New Territories, Hong Kong SAR, China
[2]Shenzhen Loop Area Institute, No. 6 Hongmian Road, Futian Free Trade Zone, Futian District, Shenzhen, Guangdong, 518038, China
[3]CUHK Shenzhen Research Institute, No. 10, 2nd Yuexing Road, Nanshan District, Shenzhen, Guangdong, 518000, China

*Corresponding author. CUHK Shenzhen Research Institute, No. 10, 2nd Yuexing Road, Nanshan District, Shenzhen, Guangdong, 518000, China.
E-mail: zhixianglin@cuhk.edu.hk

Abstract

Recent advances in spatial transcriptomics have highlighted the need for integrating spatiotemporal transcriptomics data, defined as spatially resolved gene expression profiles captured across sequential time points in developmental or regenerative processes. We present STIFT (SpatioTemporal Integration Framework for Transcriptomics) specifically designed for integrating spatiotemporal transcriptomics data. STIFT is a three-component framework combining developmental spatiotemporal optimal transport, spatiotemporal graph construction, and a graph attention autoencoder informed by temporal triplet learning. STIFT integrates large-scale 2D or 3D spatiotemporal transcriptomics data, enabling batch effect removal, spatial domain identification, trajectory inference and exploration of developmental dynamics. Applied to axolotl brain regeneration, mouse embryonic development, and 3D planarian regeneration datasets, STIFT removes batch effects and achieves clear spatial domain identification while preserving temporal developmental patterns and biological variations across hundreds of thousands of spots, demonstrating its effectiveness and specificity in integrating spatiotemporal transcriptomics data.

Keywords: spatial transcriptomics; spatiotemporal data; data integration; graph attention autoencoder

Introduction

Gene expression in tissues is a highly dynamic and heterogeneous process that varies across both space and time. Recent advances in spatial transcriptomics technologies have revolutionized our ability to study these complex patterns, enabling insights into cellular heterogeneity, tissue architecture, and spatial gene regulation during development, and disease progression [1].

A fundamental challenge in developmental biology is deciphering how gene expression evolves in both space and time. Spatiotemporal transcriptomics data consist of spatially resolved transcriptomic measurements collected from tissue sections across multiple developmental time points. While spatiotemporal transcriptomics data provide valuable insights into tissue development, integrating these datasets presents unique challenges beyond traditional batch correction. During development, tissues undergo complex morphological changes as cells proliferate and undergo apoptosis [2]. This creates intricate many-to-one relationships between spatial locations across time points, while cells simultaneously differentiate and alter their transcriptional states. Technical batch effects between time points add another layer of complexity to the integration task. Hence, the integration tool tailored for spatiotemporal transcriptomics data becomes especially demanding when processing large-scale developmental datasets. These datasets require sophisticated approaches that can simultaneously model spatial organization, temporal progression, and underlying biological processes of tissue morphogenesis.

Several computational methods have been developed to address the challenge of integrating multiple ST slices. STAGATE [3] uses a graph attention autoencoder (GATE) framework designed to identify spatial domains. STAligner [4] also employs graph attention networks for batch correction and spatial domain identification. SPIRAL [5] integrates data using graph domain adaptation and cluster-aware optimal transport for coordinate alignment. PRECAST [6] unifies spatial factor analysis with spatial clustering and embedding alignment while handling partially shared domain clusters across datasets. Graspot [7] combines graph attention networks with unbalanced optimal transport to integrate multiple ST datasets while preserving biological structure variations. BASS [8] employs a Bayesian hierarchical method for joint clustering and spatial domain detection across multiple samples. GraphST [9] combines graph neural networks with self-supervised contrastive learning to learn informative and discriminative spot representations. DeepST [10] first uses a graph neural network for representation learning and then domain adversarial neural networks for batch effect removal. Existing methods all face significant challenges when integrating spatiotemporal transcriptomics data. While multi-slice integration methods like Graspot and STAligner can effectively reduce batch effects, they treat temporal information

Received: April 11, 2025. **Revised:** November 4, 2025. **Accepted:** November 11, 2025

as another batch effect, overlooking the continuous nature of developmental processes. Existing methods typically process time points as discrete, independent batches, losing crucial trajectory information about cell differentiation, and tissue morphogenesis. The issue becomes even more pronounced when integrating 3D spatiotemporal transcriptomics data, where gene expression is captured across the x, y, and z coordinates at each time point. This is because 3D spatiotemporal transcriptomics data necessitate a holistic modeling of the continuous nature of developmental processes to accurately identify spatial domains, making the incorporation of temporal information even more essential.

Hence, we present STIFT (SpatioTemporal Integration Framework for Transcriptomics), a novel computational framework designed for integrating spatiotemporal transcriptomics data during development and regeneration. STIFT employs a three-component architecture: first, it utilizes developmental spatiotemporal optimal transport to establish probabilistic mappings between spots across consecutive time points, identifying potential parent–child relationships and uncovering critical temporal connections often overlooked by existing multi-slice integration methods. Second, it constructs a spatiotemporal graph by integrating both spatial neighbor relationships within slices and temporal connections between slices collected at adjacent time points, leveraging these dual connections to effectively preserve spatial organization and developmental progression. This makes STIFT suitable for developmental and regenerative processes characterized by complex tissue dynamics, such as cell proliferation and differentiation. Finally, it combines a GATE with triplet learning informed by temporal mappings to generate integrated embeddings that maintain both spatial organization and developmental trajectories. STIFT demonstrates superior batch effect removal and preservation of biological variations compared with existing methods. Its performance advantages are particularly evident in 3D spatiotemporal transcriptomics dataset integration, as it accounts for the continuous nature of developmental processes. By leveraging spatiotemporal information, STIFT establishes itself as a powerful tool for integrating large-scale ($> 10^5$ spots) spatiotemporal transcriptomics data, enabling downstream analyses including spatial domain identification, trajectory inference and exploration of developmental dynamics in developmental and regenerative biology.

Results

Method overview

STIFT is a computational framework designed to integrate spatiotemporal transcriptomics data, enabling the study of continuous developmental and regenerative processes. It addresses batch effect removal while preserving spatial organization and developmental trajectories through three interconnected components that effectively utilize spatiotemporal information (Fig. 1).

The first component, developmental spatiotemporal optimal transport, establishes probabilistic mappings between spots across consecutive time points, capturing progenitor-descendant relationships by analyzing gene expression and spatial coordinates. The second component, spatiotemporal graph construction, integrates intra-slice spatial neighborhoods with inter-slice temporal connections derived from optimal transport mappings to form a unified graph representation. This graph links each spot to both its spatial neighbors within the same slice and its probable progenitors in the previous time point or descendants in the next time point, facilitating integrated modeling of spatial coherence and developmental continuity. The third component, GATE, combines graph neural networks with temporal triplet learning to generate low-dimensional embeddings that preserve both spatial domains and developmental dynamics. The GATE architecture disentangles biological variations from noise by reconstructing gene expressions while enforcing temporal consistency through triplet constraints. Incorporating temporal triplet learning prevents over-correction of biological variations while aligning temporal progression.

After generating the unified low-dimensional embeddings, STIFT can facilitate batch effect removal, spatial domain identification, and exploration of developmental dynamics during developmental and regenerative processes.

Brain regeneration in axolotls: STIFT demonstrates improved batch effect removal while preserving biological variation

To evaluate STIFT's integration performance on spatiotemporal transcriptomics data, we applied it to an axolotl brain regeneration dataset [11] profiled by Stereo-seq [12]. This dataset comprises spatial transcriptome analyses of serial sections along the rostral-caudal axis at seven time points, ranging from 2 to 60 days post-injury (DPI). STIFT effectively integrates the seven slices into a common embedding space (Fig. 2b), significantly reducing the batch effect present in the original dataset while preserving biological variation patterns. The spatial domains identified by STIFT exhibit consistent cell type distributions and localization patterns across multiple time points (Fig. 2a and Supplementary Fig. S6). STAGATE, Graspot, and SPIRAL either over-correct or under-correct for these technical artifacts. STIFT and STAligner successfully reduce batch effects by evenly mixing spots from the same cell types across different slices. However, the embeddings produced by STAligner form a large, homogeneous cluster where the boundaries between distinct cell types are obscure, indicating the loss of biological variation (Fig. 2b). For example, STAligner fails to distinguish between four EGC subtypes and cannot effectively separate nptxEX cells from other cell populations, whereas STIFT achieves clear delineation of these cell types.

We further compared STIFT with other methods using both batch integration and biological structure preservation metrics. For batch integration evaluation, we employed the batch entropy mixing score, batch average silhouette width (ASW), and integration local inverse simpson index (iLISI). To assess biological structure preservation, we utilized the adjusted rand index (ARI), cell type local inverse simpson index (cLISI, a lower cLISI indicates better mixing of cell types), and cell type ASW (Fig. 2c). STIFT achieved the highest iLISI and batch ASW, as well as a second-place ranking in batch entropy mixing score. In contrast, while Graspot obtained the highest batch entropy mixing score, it performed poorly in ARI, cLISI, and cell type ASW. Specifically, STIFT excelled with the highest ARI and cell-type ASW, alongside the second lowest cLISI. Consequently, STIFT accurately maps consensus cell types across multiple time points compared with existing methods, thereby revealing regeneration dynamics in the axolotl brain.

To illustrate STIFT's ability to distinguish subtle differences among cell subtypes, we focus on the classification of ependymoglial cells (EGCs) as a case study. Previous research in various regenerative species has established that EGCs, which are similar to neural stem cells in mammals, play a critical role in neurogenesis during axolotl brain regeneration [13]. Understanding the dynamics of EGCs is essential for studying the axolotl brain regeneration process, which serves as a valuable model for mammalian brain regeneration [14]. However, differentiating between

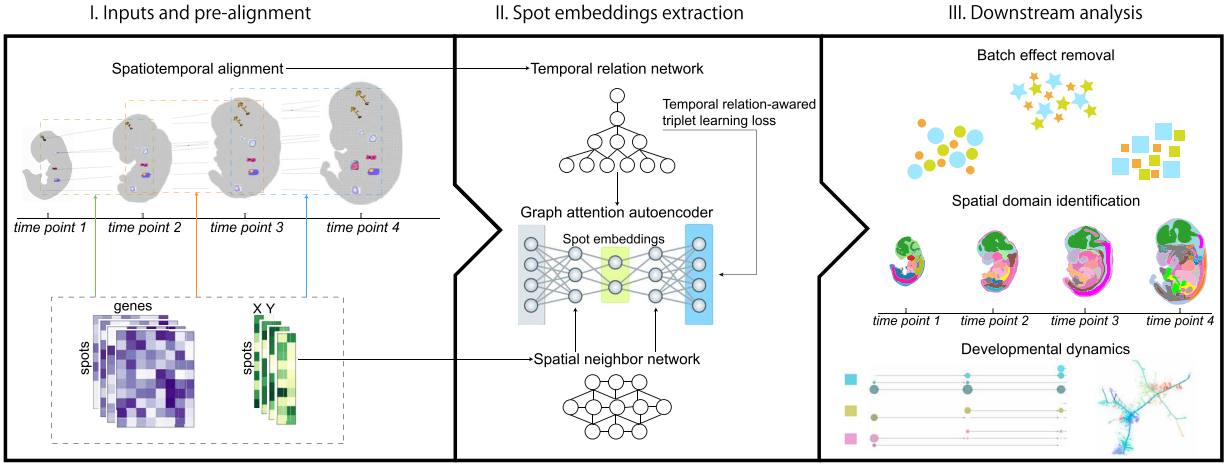

Figure 1. **Overview of STIFT:** STIFT first uses developmental spatiotemporal optimal transport to establish probabilistic mappings between spots across consecutive time points using the gene expression information and spatial coordinates information of all slices; second, it simultaneously constructs a spatial neighbor network using the spatial coordinates information within each slice and a temporal relation network from the probabilistic mappings; third, it integrates these two networks to construct a spatiotemporal graph; finally, it takes the spatiotemporal graph and gene expression information to GATE with triplet learning informed by temporal relations to generate integrated embeddings that preserve both spatial organization and developmental trajectories.

EGC subtypes presents significant challenges for clustering and classification. Using STIFT, we accurately identified four distinct EGC clusters located in the Ventricular Zone and wound site. Each EGC cluster expressed distinct molecular markers [11], corresponding to the subtypes reaEGC, sfrpEGC, wntEGC, and ribEGC. By comparing the expression patterns of these specific marker genes with our STIFT annotations, we found that all gene marker expression patterns aligned with the subtypes identified by STIFT (Fig. 2d and g). The dynamic shifts in EGC subpopulation proportions further validated STIFT's cell type annotation accuracy (Fig. 2e): SfrpEGC exhibited an immediate surge at 2DPI and maintained elevated levels throughout the recovery period, indicating its early response role; ReaEGC demonstrated a biphasic pattern, increasing from 2DPI to 10DPI followed by a gradual decline, while ribEGC and wntEGC showed inverse trajectories with an initial sharp decrease post-injury followed by slow recovery. These shifts in the subtype populations based on STIFT's annotations reflect the injury-induced conversion of local wntEGCs and ribEGCs into reaEGCs, which appear to play a crucial role in supporting tissue regeneration. The eventual restoration of wntEGC and ribEGC populations suggests a return to homeostatic conditions, highlighting the orchestrated nature of the regenerative response [11]. The precise identification of EGC subtypes is also shown in UMAP plots, where the four subtypes are well separated within the STIFT embedding space, with spots from the same subtype clustering closely together. In contrast, competing methods resulted in a mixing of the four EGC subtypes with other cell types and among themselves.

After integrating slices across multiple time points into the common embedding space, trajectory analysis can be conducted based on this embedding space to investigate how EGC cells proliferate and differentiate into nptxEX cells to cover the wound (Fig. 2f). Presumed nptxEX regeneration-related cells from 2 days post-injury (DPI) to 60 DPI were selected to construct a trajectory using StaVia [15]. The atlas view of the trajectory illustrates the lineage of wntEGC/sfrpEGC-reaEGC-WSN-IMN-nptxEX, corroborated by previous RNA velocity analyses [11]. The result shows that reaEGCs proliferate to heal the wound while simultaneously transforming or differentiating into immature neurons (IMN) and mature neurons (nptxEX) for tissue regeneration. Thus, STIFT significantly enhances the integration of multiple slices across developmental and regenerative processes, facilitating spatial domain identification and enabling trajectory inference across multiple time points that reveal complex cellular transitions and differentiation.

Mouse embryonic development: STIFT effectively integrates large-scale datasets

Advanced spatial transcriptomics technologies now generate much larger datasets [16]. To demonstrate that STIFT can integrate large-scale developmental data (usually include $10^5 \sim 10^6$ spots across multiple time points), we utilize a mouse embryo development dataset [12] as an example. This dataset comprises five mouse embryo developmental stages (E11.5 to E15.5) totaling 372 747 spots, with spot numbers increasing across stages. Each spot contains gene expression and spatial coordinates data, requiring efficient computational methods. Only STIFT, STAligner, and STAGATE successfully integrated the dataset, while other methods (SPIRAL and Graspot) failed due to memory limitations on high-end hardware.

Among the three methods capable of processing mouse embryo development data, STIFT effectively reduces batch effects by mixing spots from the same organs across different slices. Most organs, including the brain, liver, heart, spinal cord, and kidney, are distinctly separated in the UMAP visualization of STIFT embeddings, with the boundaries becoming more pronounced when visualized by organ systems. (Fig. 3a). In contrast, STAGATE fails to remove batch effect, and STAligner cannot distinguish organs like spinal cord, lung, and kidney. Evaluation metrics indicate that STIFT performs comparably to STAligner and outperforms STAGATE (Fig. 3b). STIFT successfully identified all major tissues and preserved fine anatomical structures throughout development (Fig. 3d). Many details in spatial domains were well preserved by STIFT, such as the boundary between the forebrain and hindbrain throughout the developmental process, the boundary between the spinal cord and meninges, and smaller structures like the olfactory bulb observed at E12.5.

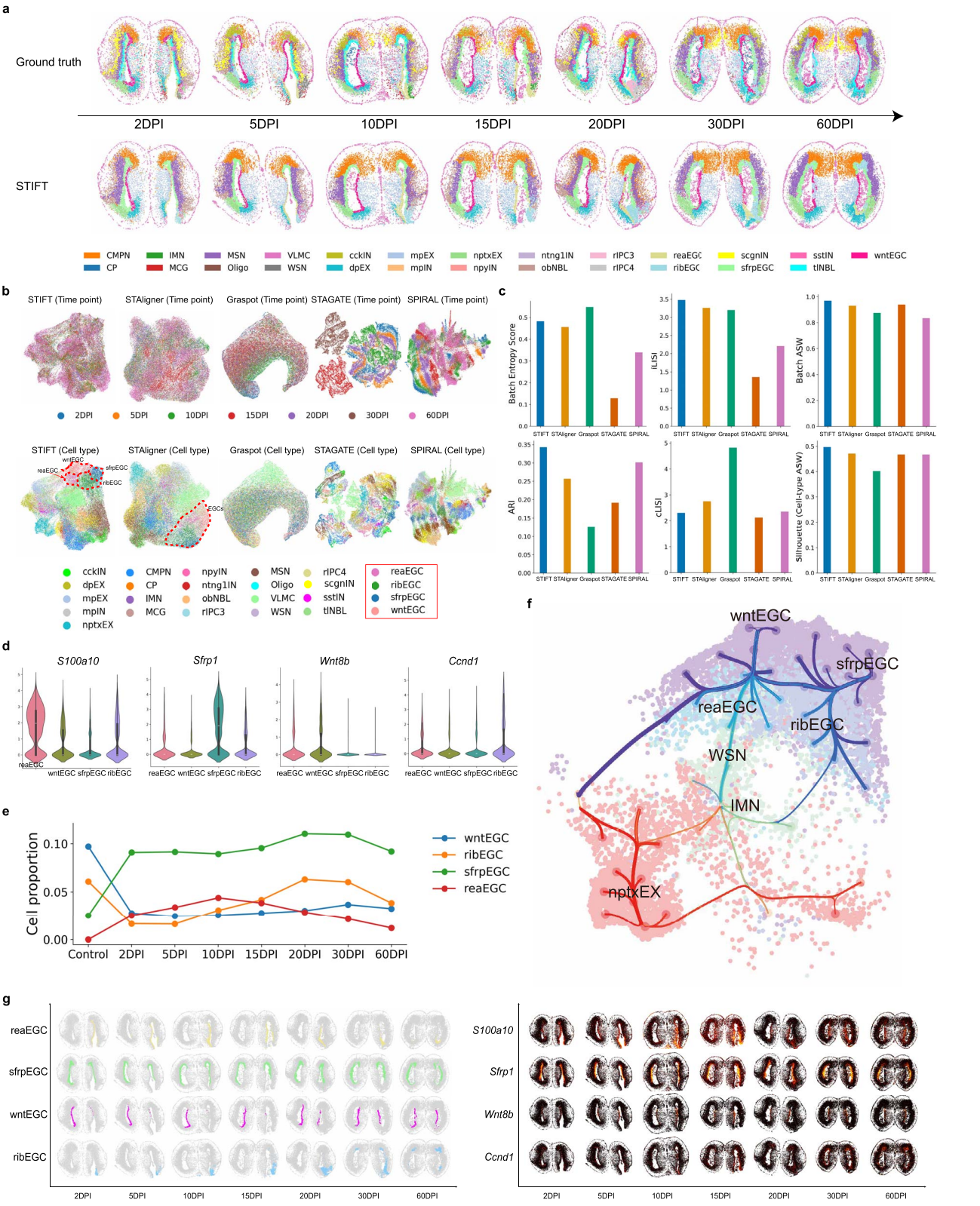

Figure 2. **Performance evaluation of STIFT on the axolotl brain regeneration dataset:** (a) manual annotations of axolotl brains (top) and aligned spatial domains identified by STIFT (bottom) across slices at seven time points; (b) UMAP plots of embeddings colored by slices (top) and manual annotations of cell types (bottom), integrating slices from seven post-injury time points using STIFT, STAligner, STAGATE, Graspot, and SPIRAL; (c) bar plots displaying batch integration metrics (top), including batch entropy mixing score, integration local inverse Simpson index (iLISI), and batch ASW, while biological structure preservation metrics are shown below, featuring ARI, cLISI (where a lower cLISI indicates better biological structure preservation performance), and cell type ASW; (d) violin plot illustrating domain-marker gene expressions in derived EGC clusters as identified by STIFT; (e) the proportion change of ECG cell type populations across regenerative time points; (f) trajectory visualization of regeneration-related cell types on the STIFT UMAP, in which the clear lineage of wntEGC/sfrpEGC-reaEGC-WSN-IMN-nptxEX illustrates how EGC cells proliferate and differentiate to cover the wound during regeneration; (g) spatial domains identified by STIFT (left) as EGCs (reaEGC, sfrpEGC, wntEGC, and ribEGC) across seven time points, with corresponding marker genes displayed on the right.

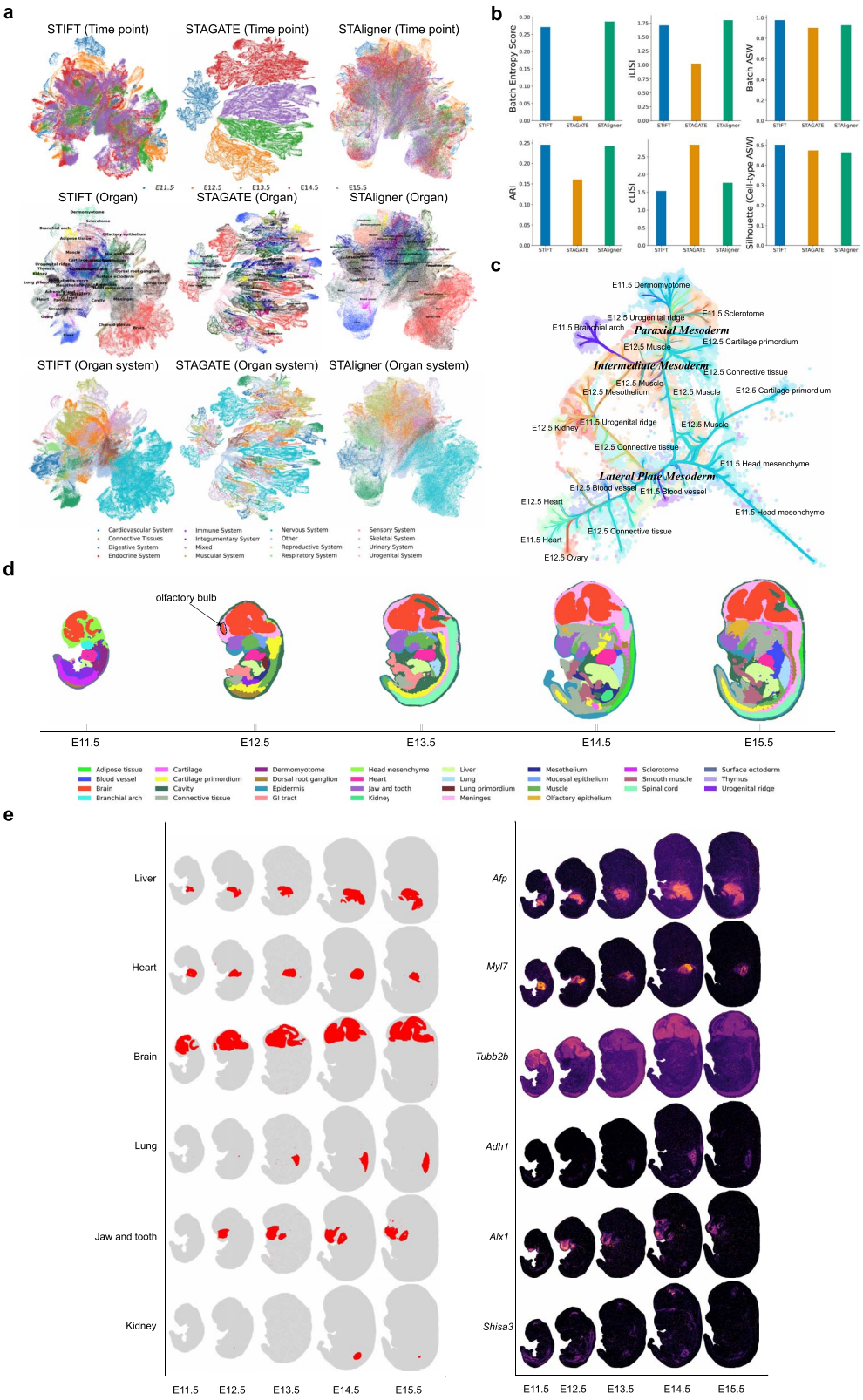

Figure 3. **Performance evaluation of STIFT on the mouse embryonic development dataset:** (a) UMAP plots of embeddings colored by slices using STIFT, STAGATE, and STAligner (top), along with manual annotations of organs (middle) and organ systems (bottom); (b) bar plots displaying batch integration metrics (left) and biological structure preservation metrics (right); (c) trajectory visualization of mesoderm-derived organs at E11.5 and E12.5 using the STIFT UMAP; (d) aligned spatial domains identified by STIFT across slices at five developmental stages; (e) organs identified by STIFT (left), including liver, heart, brain, lung, jaw and tooth, and kidney, with corresponding marker genes shown on the right.

STIFT enabled visualization of organ and tissue developmental trajectories through treemap (Fig. 4) and spatial domain identification (Fig. 3e) across embryonic stages. The analysis revealed distinct developmental patterns supported by existing literature: the heart maintained a consistent relative size throughout development [17], while the liver showed significant growth from E11.5 to E14.5 [18]; nervous system tissues, including the brain and associated structures, displayed remarkable stability in their cell populations [19]; lung development initiated at E13.5 with continued expansion in subsequent stages [20]; the jaw and teeth emerge from E12.5, developing from the branchial arch at stage E11.5, with separation occurring by E13.5 [21]; the kidney begins to develop at E14.5 with a relatively small cell population [22]. These patterns matched expected gene expression distributions (Fig. 3e) and the spatiotemporal plots (Supplementary Fig. S4) and treemap (Supplementary Fig. S5) generated from ground truth annotations.

The common embedding space of STIFT enables trajectory inference [15] across multiple time points and reveals the path of differentiation for various organs across developmental stages. Taking the progression of mesoderm-derived cells between stages E11.5 and E12.5 as an example (Fig. 3c), we observe that cells from the intermediate mesoderm develop into the urogenital ridge (E11.5), which ultimately gives rise to the kidney (E12.5). Meanwhile, lateral plate mesoderm contributes to the formation of the heart, ovary, and blood vessels, while paraxial mesoderm differentiates into dermomyotome (E11.5), sclerotome (E11.5), and cartilage primordium (E12.5). These lineages are further supported by the existing literature [23, 24]. This trajectory analysis powered by STIFT provides valuable insights into how these various lineages progress over time, highlighting the intricate dynamics of organ development during embryogenesis.

In summary, by integrating large-scale Stereo-seq data, STIFT effectively mapped high-resolution spatial domains across extensive tissue regions and uncovered tissue developmental dynamics. This integration facilitates trajectory inference of organ development from the common embedding space, offering detailed insights into the complexities of organ formation.

Planarian regeneration process: STIFT effectively integrates 3D spatiotemporal transcriptomics data

Recent advancements in spatial transcriptomics have led to the emergence of numerous 3D spatiotemporal datasets, capturing gene expression across x, y, and z coordinates at various time points through multiple consecutive 2D tissue slices. For instance, a dataset profiling the planarian regeneration process [25] spans six time points from whole-worm coronal cryosections. Conventional integration methods often rely on mutual nearest neighbors (MNN) matching based solely on gene expression data, which can result in incorrect pairing of unrelated spots. This challenge is particularly pronounced in 3D spatiotemporal transcriptomics data integration, where complex spatial relationships increase the risk of mismatches. In planarian regeneration data, MNN matching based on gene expression alone creates numerous incorrect pairs, such as head-to-tail pairs or muscle-to-neuronal pairs (Fig. 5d, left). STIFT addresses this limitation by incorporating spatiotemporal relationships to identify likely ancestors and descendants during triplet learning, thereby reducing errors such as head-to-tail mismatches (Fig. 5d, right). This strategy enhances spot matching accuracy and improves 3D spatiotemporal transcriptomics data integration compared with conventional MNN-based methods.

STIFT accurately tracks how cells change during regeneration in 3D by identifying ancestors and their descendants (Fig. 5a). The UMAP visualization of STIFT embeddings demonstrates that spots across slices are well mixed while spatial domains remain distinct with clear expression patterns originated from the variation of different cell types (Fig. 5b). STIFT outperformed other methods in preserving biological variations while removing batch effects (Fig. 5c). These results underscore the importance of incorporating spatiotemporal information in the integration of 3D spatiotemporal transcriptomics datasets. The tissue structures are further validated by higher expressions of marker genes in the corresponding identified spatial domains as observed in top-down views of the 3D planarian model (Fig. 5e). Each domain exhibits clear tissue architecture and can be annotated based on established ground truth patterns: for instance, neuronal and muscle cells are predominantly located near the central pharynx; the neoblast population is widely distributed throughout the body, showing a sharp increase at 6 hours post-amputation (hpa), peaking at 12 hpa during early regeneration before gradually decreasing; epidermal cells, located at the planarian edge, display distinct boundaries with other cell types. By leveraging spatiotemporal relationships in the integration of 3D spatiotemporal transcriptomics datasets, STIFT significantly enhances our ability to characterize complex tissue structures and organ development processes in 3D space.

Discussion

STIFT is designed specifically for integrating large-scale spatiotemporal transcriptomics data during development and regeneration processes. STIFT's three-component architecture effectively preserves both spatial organization and developmental trajectories while removing batch effects. STIFT demonstrates superior performance in several key aspects: it demonstrates superior batch effect removal while maintaining biological variations, effectively preserves spatiotemporal information through developmental spatiotemporal optimal transport, enables detailed mapping of consensus spatial domains across extensive tissue regions, and provides trajectory inference capabilities across multiple time points to reveal differentiation paths for various cell types. Through these features, STIFT outperforms existing methods in both batch integration metrics and biological structure preservation metrics, making it a powerful tool for understanding organ development and tissue organization. A brief guide to interpreting STIFT outputs and biological insights it can reveal is provided in Supplementary Note S3.

To demonstrate generalization across tissues and modalities, we added a spatiotemporal cardiac infarction dataset [26] (see Supplementary Note S1 and Fig. S1) profiled by 10× Visium and a human cerebral cortex development dataset [27] (see Supplementary Note S2 and Fig. S2) profiled by MERFISH. The results consistently demonstrate STIFT's robustness. However, STIFT can still be improved in two aspects: first, as spatial multiomics data become increasingly available, STIFT could expand its capabilities to integrate spatiotemporal data in various modalities, such as spatial proteomics data. In addition, combining gene expression data with spatial locations and histological images could enhance the delineation of morphological structures across multiple samples.

In conclusion, STIFT offers researchers a powerful integration tool for exploring complex developmental and regenerative processes. As the field continues to evolve, STIFT will play a crucial

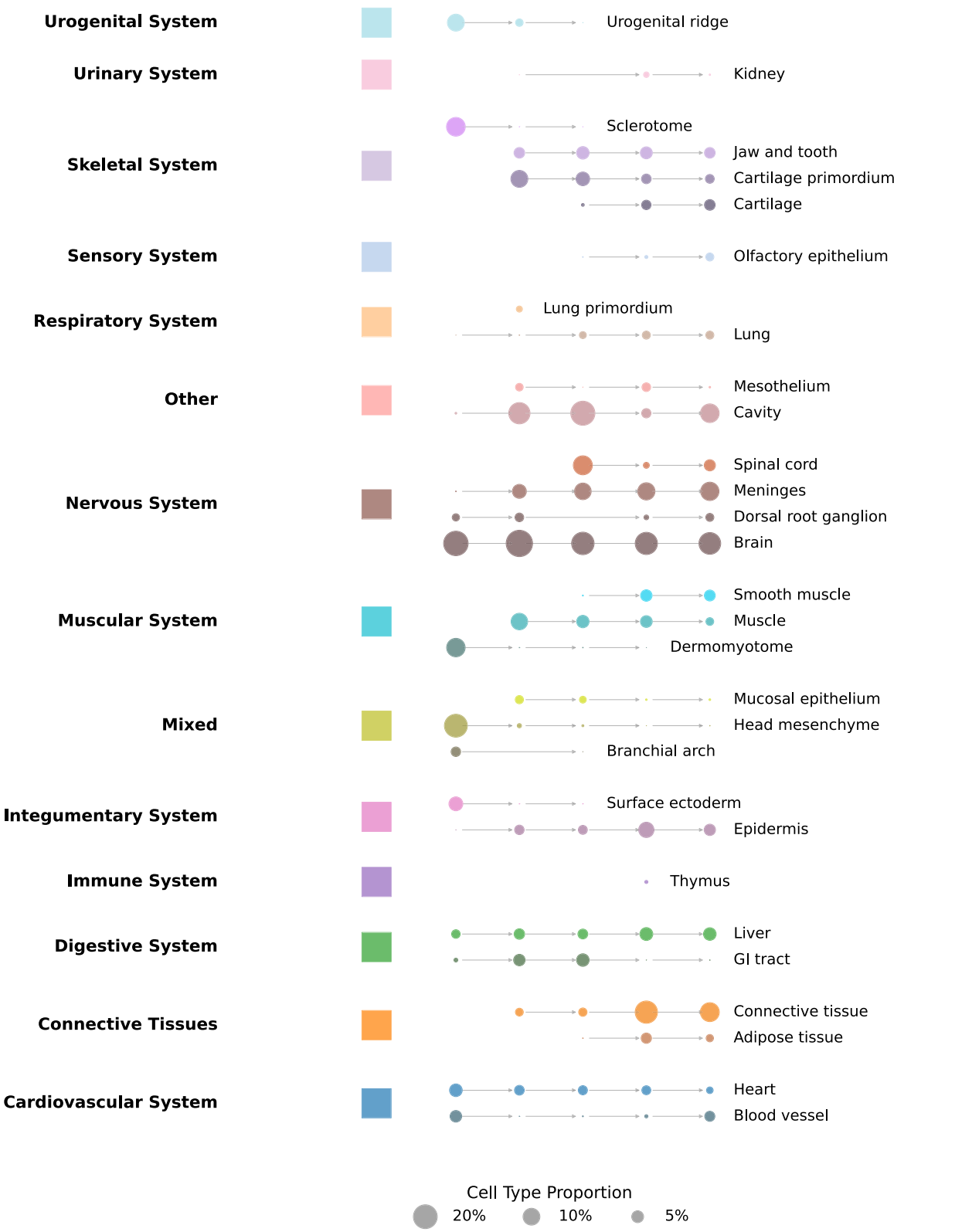

Figure 4. Treemap visualization of tissue developmental dynamics across stages, where each column represents a developmental stage, each row represents a cell type rooted by organ systems, and node sizes indicate relative abundance.

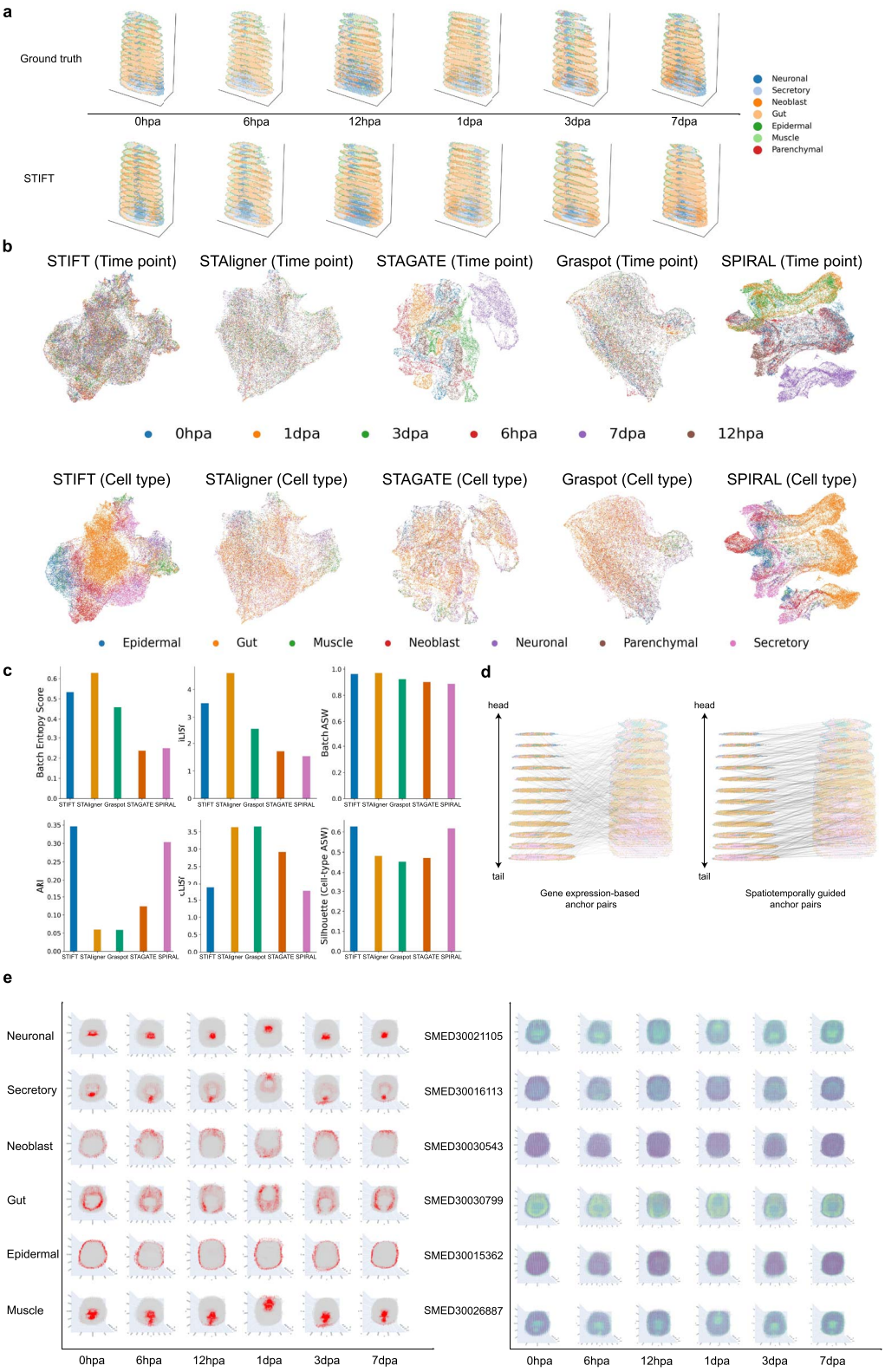

Figure 5. **Performance evaluation of STIFT on the planarian regeneration dataset:** (a) manual annotations of planarians (top) alongside aligned spatial domains identified by STIFT (bottom) across slices at six time points, in which the three-dimensional reconstructed planarian slices were aligned based on their shapes and the centrally located pharynx, with the Z-axis shown in a stretched view; (b) UMAP plots of embeddings colored by slices (top) and manual annotations of cell types (bottom), integrating slices from six time points using STIFT, STAligner, STAGATE, Graspot, and SPIRAL; (c) bar plots displaying batch integration metrics (top) and biological structure preservation metrics (bottom); (d) gene expression-based anchor pairs (left) and spatiotemporally guided anchor pairs (right) between 6 hpa and 12 hpa during planarian regeneration, where within each figure, the left sample represents 6 hpa and the right sample shows 12 hpa, and the anterior end (top) represents the head region, while the posterior end (bottom) corresponds to the tail region of the planarian, and gray lines connect corresponding spots in the same anchor pairs; (e) top-down views of spatial domains identified by STIFT (left) across time points, alongside corresponding marker gene expressions (right), in which spots are rendered transparent to enhance visualization of their distribution within the spatial domains and gene expressions.

role in unlocking the full potential of spatiotemporal transcriptomics data, leading to deeper insights into developmental and regenerative processes and mechanisms.

Methods

Data preprocessing

STIFT processes gene expression and spatial coordinate data from multiple slices across different time points. For each slice, raw gene expressions are normalized to a total count of 10 000 per cell, followed by log-transformation using the SCANPY package [28]. Highly variable genes are identified, selecting the top 10 000 for downstream analysis. To ensure consistency across slices, only the intersection of highly variable genes from all slices is retained, allowing for shared spot features. The spatial maps of slices are proportionally normalized to a 100×100 square. Suppose we have K slices from K different time points. After preprocessing, each slice $\mathcal{S}_k$ at time point t_k is represented by matrices ($X_k \in \mathbb{N}^{n_k \times g}, S_k \in \mathbb{R}^{n_k \times \dim}$), where X_k is the gene expression matrix, and S_k is the spatial coordinate matrix, $k = 1, 2, \cdots, K$. Here, n_k denotes the number of spots, g is the number of genes shared by all slices, and dim specifies the spatial coordinate dimension, which can be either 2 or 3.

Temporal alignment

STIFT employs the differentiable entropy-regularized spatial optimal transport (DEST-OT) algorithm [29] to align spots across multiple time points. This alignment facilitates the study of developmental processes by establishing probabilistic mappings between spots in slices from adjacent time points, t_k and t_{k+1}. Given two slices, $\mathcal{S}_k = (X_k, S_k)$ and $\mathcal{S}_{k+1} = (X_{k+1}, S_{k+1})$, DEST-OT computes an alignment matrix $\mathbf{\Pi}_{k,k+1} \in \mathbb{R}_+^{n_k \times n_{k+1}}$. Each entry $\mathbf{\Pi}_{ij}$ represents the probability that cells in spot i of slice $\mathcal{S}_k$ are progenitors of cells in spot j of slice $\mathcal{S}_{k+1}$.

This probabilistic mapping identifies potential parent–child relationships across time points, facilitating the construction of spatiotemporal graphs that integrate temporal connection information from different time points into the spatial neighbor information of individual slices.

For cases where the number of spots in two slices is computationally prohibitive for optimal transport calculation, a downsampling strategy is employed. Specifically, β (100% $\geq \beta \geq$ 30%) of spots are randomly selected from each slice to perform the alignment. For the remaining unaligned spots, alignment relationships are assigned based on their nearest neighbors that have undergone the alignment process. This approach maintains computational efficiency while preserving the quality of temporal mapping across large-scale datasets.

Spatiotemporal graph construction

In this section, we construct a comprehensive spatiotemporal adjacency matrix by integrating spatial relationships captured through k-nearest neighbors [30] and temporal connections derived from alignment matrices. Initially, for each tissue slice $\mathcal{S}_k$, an undirected neighbor graph is constructed using a k-nearest neighbors approach with a cutoff of r_s neighbors under Euclidean distance. This method captures local spatial relationships by connecting each spot to its nearest neighbors based on spatial coordinates. The graph is denoted as the adjacency matrix A_k, where $A_k[i, j] = 1$ if and only if an edge exists between spot i and spot j. A_k is a symmetric matrix with self-loops added for each spot.

Subsequently, the temporal relation matrices $T_{k,k+1}$ and $T_{k+1,k}$ are constructed to capture the most significant temporal connections between consecutive tissue slices at time points t_k and t_{k+1}. These matrices are derived from the alignment matrix $\mathbf{\Pi}_{k,k+1}$, which provides probabilistic mappings between spots in these slices. Given the alignment matrix $\mathbf{\Pi}_{k,k+1} \in \mathbb{R}_+^{n_k \times n_{k+1}}$, where each entry $\mathbf{\Pi}_{ij}$ represents the probability that spot i in slice k is a progenitor of spot j in slice $k+1$, we proceed as follows:

For each spot i in slice $\mathcal{S}_k$, we identify its r_t most likely descendants in slice $\mathcal{S}_{k+1}$ using the alignment probabilities from matrix $\mathbf{\Pi}_{k,k+1}$. The temporal relation matrix $T_{k,k+1}$ is defined as:

$$T_{k,k+1}[i,j] = \begin{cases} 1, & \text{if } j \text{ is among the top} r_t \text{ entries in row } i \text{ of } \mathbf{\Pi}_{k,k+1} \\ 0, & \text{otherwise} \end{cases}$$

Similarly, for each spot j in slice $\mathcal{S}_{k+1}$, we identify its r_t most likely ancestors in slice $\mathcal{S}_k$ and then define $T_{k+1,k}$ as:

$$T_{k+1,k}[j,i] = \begin{cases} 1, & \text{if } i \text{ is among the top} r_t \text{ entries in column } j \text{ of } \mathbf{\Pi}_{k,k+1} \\ 0, & \text{otherwise} \end{cases}$$

This approach ensures that only the most significant temporal connections are retained in the temporal mapping matrices $T_{k,k+1}$ and $T_{k+1,k}$. By focusing on these top-r_t parent–child relationships, we effectively capture dynamic lineage information necessary for understanding developmental processes across time points.

The final step is to construct an integrated spatiotemporal adjacency matrix by concatenating spatial adjacency matrices A_k and temporal relation matrices $T_{k,k+1}$ for all slices. The resulting spatiotemporal adjacency matrix A is represented as:

$$A = \begin{bmatrix} A_1 & T_{1,2} & 0 & \cdots & 0 \\ T_{2,1} & A_2 & T_{2,3} & \cdots & 0 \\ 0 & T_{3,2} & A_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_n \end{bmatrix}$$

In integrated spatiotemporal adjacency matrix A, A_k captures spatial relationships within sample k, and $T_{k,k+1}$ and $T_{k+1,k}$ model directed temporal connections between sample k and sample $k+1$. The integrated spatiotemporal graph A enables continuous developmental trajectory learning instead of treating time points as batches.

Graph attention autoencoder

The GATE [31] in STIFT is designed to learn integrated embeddings that preserve both spatial organization and developmental trajectories. It processes the spatiotemporal graph structure along with normalized gene expression data as node attributes to generate spot-specific embeddings using the attention mechanism.

The GATE consists of two main components: an encoder and a decoder. The encoder processes concatenated normalized gene expression data X and the corresponding graph structure to generate spot representations $\mathbf{z}$. Let L denotes the number of layers in the encoder, and n represents the number of spots in all slices. The output representation of the lth encoder layer for spot i is

formulated as:

$$\mathbf{z}_i^{(l)} = \sigma\left(\sum_{j\in\mathcal{N}_i}\alpha_{ij}^{(l)}\left(\mathbf{W}^{(l)}\mathbf{z}_j^{(l-1)}\right)\right)$$

where $\mathcal{N}_i$ represents all spots connecting to spot i in spatiotemporal graph A, σ is the Sigmoid activation function, and $\mathbf{W}^{(l)}$ is a trainable weight matrix. The initial representation $\mathbf{z}_i^{(0)}$ is the gene expression vector $\mathbf{x}_i$.

The attention coefficient $\alpha_{ij}^{(l)}$ adaptively learns the similarity between neighboring spots:

$$\alpha_{ij}^{(l)} = \frac{\exp(e_{ij}^{(l)})}{\sum_{m\in\mathcal{N}_i}\exp(e_{im}^{(l)})}$$

The relevance score $e_{ij}^{(l)}$ between spots is calculated as:

$$e_{ij}^{(l)} = \sigma\left(\mathbf{v}^{s(l)^T}\left(\mathbf{W}^{(l)}\mathbf{z}_i^{(l-1)}\right) + \mathbf{v}^{r(l)^T}\left(\mathbf{W}^{(l)}\mathbf{z}_j^{(l-1)}\right)\right)$$

with $\mathbf{v}^{s(l)}$ and $\mathbf{v}^{r(l)}$ being trainable weight vectors, and σ is the Sigmoid activation function.

The final encoder layer (spot embedding layer) is directly calculated by:

$$\mathbf{z}_i^{(L)} = \sigma(\mathbf{W}^{(L)}\mathbf{z}_i^{(L-1)})$$

The decoder reconstructs normalized gene expression information by reversing the encoding process. Using the embeddings $\mathbf{z}_i^{(L)}$ as the input of decoder $\hat{\mathbf{z}}_j^{(L)}$, it reconstructs representations in previous layers:

$$\hat{\mathbf{z}}_i^{(l-1)} = \sum_{j\in\mathcal{N}_i}\widehat{\alpha}_{ij}^{(l-1)}\sigma(\widehat{\mathbf{W}}^{(l)}\hat{\mathbf{z}}_j^{(l)})$$

The reconstructed expressions are denoted by $\hat{\mathbf{z}}_i^{(0)}$, which is calculated by:

$$\hat{\mathbf{z}}_i^{(0)} = \sigma(\widehat{\mathbf{W}}^{(1)}\hat{\mathbf{z}}_i^{(1)})$$

To prevent overfitting, it is set that $\widehat{\mathbf{W}}^{(l)} = \mathbf{W}^{(l)}$ and $\hat{\alpha}_{ij}^{(l)} = \alpha_{ij}^{(l)}$.

The objective of GATE is to maximize similarity between input and reconstructed expressions, achieved by minimizing mean square error loss:

$$L_{\text{reconstruction}} = \frac{1}{n}\sum_{i=1}^{n}||\hat{\mathbf{z}}_i^{(0)} - \mathbf{x}_i||_2$$

Spot triplet learning

STIFT employs spot triplet learning [32] to enhance batch effect correction in spatiotemporal transcriptomics data through a triplet loss framework. This technique strategically selects triplets based on temporal mappings $\mathbf{\Pi}_{k,k+1}$ derived from the alignment process between consecutive slices.

For each anchor spot i in slice $\mathcal{S}_k$, positive examples are selected from its progenitors and descendants in the adjacent slices $\mathcal{S}_{k-1}$ and $\mathcal{S}_{k+1}$ based on the largest mapping probabilities. The selection is determined as follows:

$$i_1^+ = \mathbf{argmax}_j(\mathbf{\Pi}_{k-1,k}[j,i]) \quad \text{(most probable progenitor)}$$

$$i_2^+ = \mathbf{argmax}_j(\mathbf{\Pi}_{k,k+1}[i,j]) \quad \text{(most probable descendant)}$$

For edge cases: (i) when $k = 1$, let $i_1^+ := i_2^+$ (most probable descendant); (ii) when $k = K$, let $i_2^+ := i_1^+$ (most probable progenitor).

A negative example i^- is randomly chosen from spots within the same slice $\mathcal{S}_k$. As the total number of spots is usually much larger than the k-nearest neighbors of the anchor spot, the probability of the sampled negative spot being located in the k-nearest neighbors of the anchor is very low. This negative sampling strategy can accelerate model training, especially in large-scale datasets.

The triplet set ($\mathbf{z}_i$, $\mathbf{z}_{i_1}^+$, $\mathbf{z}_{i_2}^+$, and $\mathbf{z}_i^-$) is constructed where $\mathbf{z}_i$ is the embedding of the anchor spot, $\mathbf{z}_{i_1}^+$ and $\mathbf{z}_{i_2}^+$ are the embeddings of positive spots, and $\mathbf{z}_i^-$ is the embedding of the negative spot.

The triplet loss is used to minimize the distance between the anchor-positive pair and maximize the distance between the anchor-negative pair in the latent space. It is formulated as follows:

$$\mathcal{L}_{\text{triplet}} = \frac{1}{n_{\text{tri}}}\sum_{(i,i_1{}^+,i_2{}^+,i^-)}\max\left(0, \frac{1}{2}\left(||\mathbf{z}_i - \mathbf{z}_{i_1^+}||_2^2 + ||\mathbf{z}_i - \mathbf{z}_{i_2^+}||_2^2\right) - ||\mathbf{z}_i - \mathbf{z}_i^-||_2^2 + m\right) \quad (1)$$

where $||\cdot||_2^2$ denotes the squared Euclidean distance, n_{tri} represents the total number of triplets, $\mathcal{C}_{\text{tri}}$ is the set of spot triplets, and m is the margin used to enforce the distance between positive and negative pairs.

The final total loss of STIFT is:

$$\mathcal{L} = L_{\text{reconstruction}} + \lambda\mathcal{L}_{\text{triplet}}$$

Here, λ serves as a weighting factor balancing the contributions of reconstruction loss and triplet loss. The optimization process employs stochastic gradient descent with gradient clipping to maintain stability during training.

Training procedure

STIFT employs an iterative training process that integrates GATE training with spot triplet learning. The training procedure consists of two main phases: pretraining and fine-tuning.

Pretraining phase

The GATE is structured with an encoder consisting of two layers, with node dimensions set to 512 and 30, respectively. The architecture is based on STAGATE [3]. Initially, the model is pretrained using only the reconstruction loss over ~500 epochs. This phase focuses on learning the embeddings through the GATE architecture.

Fine-tuning phase

Following pretraining, the model transitions to the fine-tuning phase, which typically consists of 1000 epochs. During this phase,

the training process is iterative, alternating between GATE training and spot triplet learning to refine the learning of spot embeddings.

Optimization details

The model is optimized using stochastic gradient descent with the Adam optimizer [33]. To maintain stability during training, gradient clipping is applied with a threshold of 5.0. The learning rate is set to 0.001 with a weight decay of 0.0001.

Loss function

During fine-tuning, the model is trained using a total loss that incorporates both triplet loss and GATE reconstruction loss:

$$\mathcal{L} = L_{\text{reconstruction}} + \lambda \mathcal{L}_{\text{triplet}}$$

Here, λ serves as a weighting factor balancing the contributions of reconstruction loss and triplet loss (default is 0.1).

Clustering

For identifying distinct spatial domains or cell types within the integrated spatiotemporal transcriptomics data, STIFT employs two widely used community detection algorithms: Louvain [34] and Leiden [35]. These methods are applied to the low-dimensional embeddings obtained from the GATE using the SCANPY package [28].

Evaluation metrics

To assess the performance of STIFT, we employ a comprehensive set of evaluation metrics that measure both the quality of batch integration and the preservation of biological structures. These metrics are computed on the low-dimensional embeddings generated by the GATE.

Batch integration metrics

Integration Local Inverse Simpson Index (iLISI)

iLISI measures batch integration by quantifying the diversity of batch labels in the local neighborhood of each cell. For a given cell i, it is calculated as:

$$\text{iLISI}_i = \frac{1}{\sum_{b \in B} p_{i,b}^2}$$

where B is the set of all batch labels, and $p_{i,b}$ is the proportion of cells with batch label b in the local neighborhood of cell i. The neighborhood is typically defined using k-nearest neighbors in the embedding space. Higher iLISI scores are desirable, indicating good batch integration [36].

Batch Entropy Mixing Score

This metric quantifies the degree of mixing between different batches in the integrated embedding space. A higher score indicates better batch integration. We calculate this using a k-nearest neighbor approach with 100 neighbors, 100 pools, and 100 samples per pool [36].

Batch Average Silhouette Width

The batch ASW measures the separation between batches in the embedding space. A lower ASW indicates better integration, as it suggests less batch-specific clustering. We compute this as:

$$\text{Batch ASW} = 1 - \frac{1}{n}\sum_{i=1}^{n} |s_i|$$

where s_i is the silhouette score for the ith spot, and n is the total number of spots [37].

Biological structure preservation metrics

Adjusted Rand Index

The ARI measures the agreement between the clustering results and the known cell-type labels. It is computed as:

$$\text{ARI} = \frac{\text{RI} - \mathbb{E}[\text{RI}]}{\max(\text{RI}) - \mathbb{E}[\text{RI}]}$$

where RI is the Rand Index. A higher ARI indicates better correspondence between clusters and true cell types [38].

Cell Type Local Inverse Simpson Index

cLISI measures the preservation of cell-type separation. It is calculated similarly to iLISI but uses cell type labels instead of batch labels:

$$\text{cLISI}_i = \frac{1}{\sum_{c \in C} p_{i,c}^2}$$

where C is the set of all cell type labels, and $p_{i,c}$ is the proportion of cells with cell type label c in the local neighborhood of cell i. Lower cLISI scores are desirable, indicating maintained biological distinctions [36].

Cell type Average Silhouette Width

This metric assesses how well cell types are separated in the embedding space. It is calculated as:

$$\text{Cell type ASW} = 0.5 \cdot \left(\frac{1}{n}\sum_{i=1}^{n} s_i + 1\right)$$

where s_i is the silhouette score for the ith spot based on cell-type labels. A higher score indicates better preservation of cell-type structure [37].

These metrics collectively provide a comprehensive evaluation of STIFT's performance in integrating spatiotemporal transcriptomics data while preserving important biological information.

Benchmarking methods

Since there are no integration methods specifically designed for spatiotemporal transcriptomics data integration, we compared STIFT with four following methods for spatial transcriptomics data integration and alignment to evaluate its performance.

STAligner [4] is a method for aligning spatial transcriptomics datasets that accounts for partially matched tissue sections and local nonlinear distortions. We applied STAligner with hidden dimensions [512, 30], 2000 epochs, a learning rate of 0.001, gradient clipping threshold of 5.0, margin of 1.0, and weight decay of 0.0001.

STAGATE (Spatial and Temporal Adaptive Graph ATtention autoEncoder) [3] is a graph attention auto-encoder framework designed to identify spatial domains. We used STAGATE with hidden dimensions [512, 30], 2000 epochs, a learning rate of 0.001, gradient clipping threshold of 5.0, margin of 1.0, and weight decay of 0.0001.

Graspot [7] is a graph attention network for spatial transcriptomics data integration with unbalanced optimal transport. We implemented Graspot using hidden dimensions [512, 30], 1000 epochs, a learning rate of 0.001, gradient clipping threshold of 5.0, and weight decay of 0.0001.

SPIRAL [5] is graph domain adaptation-based data integration method. We implement SPIRAL using autoencoder dimension [512, 32], GraphSAGE dimension [512, 32], discriminator [28, [32, 16]], 100 epochs, batch size of 1024, learning rate of 0.001, weight decay of 0.0005, beta of 1.0, and KNN parameter of 6.

For all methods, we followed these general steps: we preprocessed the data using the same pipeline as STIFT. We applied the respective integration method with the specified parameters. We performed clustering on the integrated data using Louvain and Leiden algorithms under the same resolution as STIFT. We visualized results using UMAP [39]. All methods were implemented using PyTorch and run on a NVIDIA A100 80 GB GPU for consistent comparison. Runtime and memory benchmarks are reported in Supplementary Note S4 and Fig S3. Hyperparameter sensitivity analysis results are reported in Supplementary Fig. S7.

Key Points

- STIFT fully utilizes spatiotemporal information through developmental optimal transport, spatiotemporal graph construction, and a graph attention autoencoder informed by temporal triplet learning.
- STIFT demonstrates superior batch effect removal and preservation of biological variations compared to existing methods when assessed across multiple performance metrics.
- STIFT can effectively integrate large-scale ($> 10^5$ spots) 2D or 3D spatiotemporal transcriptomics data.
- STIFT enables downstream analyses including spatial domain identification, trajectory inference and exploration of developmental dynamics in developmental and regenerative biology.

Acknowledgements

The authors would like to express their gratitude to Hongyu Zhao for his insights and feedback, which significantly enriched this work.

Author contributions

Z.L. and J.Q. conceived the project. J.Q. developed the method, implemented the software, and wrote the manuscript. J.Q., M.G., J.M., and X.Z. performed the data analyses. Z.L. revised the manuscript. All authors read and approved the final manuscript.

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Funding

This work is supported in part by the Chinese University of Hong Kong Science Faculty's Collaborative Research Impact Matching Scheme (CRIMS 4620033 to Z.L.), 1+1+1 CUHK-CUHK(SZ)-GDSTC Joint Collaboration Fund (2025A0505000057 to Z.L.), and Research Grants Council, University Grants Committee (GRF 14301120 to Z.L. and GRF 14300923 to Z.L.).

Data availability

Axolotl regeneration data

The axolotl regeneration data [11] provide spatially resolved transcriptomic data visualizing gene expression across regeneration and development stages of axolotl telencephalon at single-cell resolution using Stereo-seq. The dataset includes spatial transcriptome analyses on serial sections along the rostral–caudal axis at various time points after injury. The days post injury (DPI) and their corresponding number of spots are: 2 DPI (7668 spots), 5 DPI (8106 spots), 10 DPI (9436 spots), 15 DPI (8996 spots), 20 DPI (10 462 spots), 30 DPI (9252 spots), and 60 DPI (10 964 spots). The injury model involved removing a reproducible portion of dorsal pallium in the left telencephalic hemisphere of 11 cm length axolotls. The data (CNGB Project ID CNP0002068) can be accessed via https://db.cngb.org/stomics/artista.

Mouse embryo development data

The mouse embryo development data [12] comprise Stereo-seq data from sagittal sections across five developmental stages. The stages and their corresponding number of spots are as follows: E11.5 (30 124 spots), E12.5 (51 365 spots), E13.5 (77 369 spots), E14.5 (102 519 spots), and E15.5 (113 350 spots). Each spot contains gene expression data and spatial coordinates, covering major organs and tissues developing during these embryonic stages. The data (CNGB Project ID CNP0001543) can be accessed via https://db.cngb.org/stomics/mosta.

Planarian regeneration data

The planarian regeneration data [25] consist of spatial transcriptomics data from whole-worm coronal cryosections at multiple time points during regeneration, profiled using 10× Visium technology. The time points and their respective number of spots are: 0 hpa with 8949 spots, 6 hpa with 7304 spots, 12 hpa with 10 156 spots, 24 hpa with 8423 spots, 3 days post-amputation (dpa) with 8365 spots, and 7 dpa with 10 210 spots. Three-dimensional reconstructed planarian regeneration models at these six time points were aligned based on their shapes and the centrally located pharynx. Cell2location [40] was used to deconvolute and annotate the 3D cell distribution profiles of planarians at each regenerating time point. The data (GSA ID CRA007941) can be accessed via https://ngdc.cncb.ac.cn/omix/release/OMIX003867.

All ground truth annotation sources used in the manuscript are listed in Supplementary Note S5.

Software availability

The STIFT Python package is freely available as open-source software under the MIT License at GitHub (https://github.com/cuhklinlab/STIFT) and Figshare (https://figshare.com/articles/software/STIFT_Spatiotemporal_Transcriptomics_Integration_Through_Spatially_Informed_Multi-Timepoint_Bridging/28758650). The current stable release (v1.0.0) includes tutorials and documentation for spatiotemporal transcriptomics integration workflows.

References

1. Wang S, Lin S, Yang C. The dawn of spatiotemporal transcriptomics. *Biomed Anal* 2024;**1**:154–61. https://doi.org/10.1016/j.bioana.2024.06.002
2. Voss AK, Strasser A. The essentials of developmental apoptosis. *F1000Research* 2020;**9**:148. https://doi.org/10.12688/f1000research.21571.1
3. Dong K, Zhang S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention

auto-encoder. *Nat Commun* 2022;**13**:1739. https://doi.org/10.1038/s41467-022-29439-6
4. Zhou X, Dong K, Zhang S. Integrating spatial transcriptomics data across different conditions, technologies and developmental stages. *Nat Comput Sci* 2023;**3**:894–906. https://doi.org/10.1038/s43588-023-00528-w
5. Guo T, Yuan Z, Pan Y. *et al.* SPIRAL: integrating and aligning spatially resolved transcriptomics data across different experiments, conditions, and technologies. *Genome Biol* 2023;**24**:241.
6. Liu W, Liao X, Luo Z. *et al.* Probabilistic embedding, clustering, and alignment for integrating spatial transcriptomics data with precast. *Nat Commun* 2023;**14**:296. https://doi.org/10.1038/s41467-023-35947-w
7. Gao Z, Cao K, Wan L. Graspot: a graph attention network for spatial transcriptomics data integration with optimal transport. *Bioinformatics* 2024;**40**:ii137–45.
8. 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* 2022;**23**:168. https://doi.org/10.1186/s13059-022-02734-7
9. Long Y, Ang KS, Li M. *et al.* Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat Commun* 2023;**14**:1155. https://doi.org/10.1038/s41467-023-36796-3
10. Chang X, Jin X, Wei S. *et al.* DeepST: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Res* 2022;**50**:e131–1.
11. Wei X, Sulei F, Li H. *et al.* Single-cell stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science* 2022;**377**:eabp9444. https://doi.org/10.1126/science.abp9444
12. Chen A, Liao S, Cheng M. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 2022;**185**:1777–1792.e21. https://doi.org/10.1016/j.cell.2022.04.003
13. Joven A, Simon A. Homeostatic and regenerative neurogenesis in salamanders. *Prog Neurobiol* 2018;**170**:81–98. https://doi.org/10.1016/j.pneurobio.2018.04.006
14. Kirkham M, Hameed LS, Berg DA. *et al.* Progenitor cell dynamics in the newt telencephalon during homeostasis and neuronal regeneration. *Stem Cell Rep* 2014;**2**:507–19. https://doi.org/10.1016/j.stemcr.2014.01.018
15. Stassen SV, Kobashi M, Lam EY. *et al.* StaVia: spatially and temporally aware cartography with higher-order random walks for cell atlases. *Genome Biol* 2024;**25**:224.
16. Fang S, Chen B, Zhang Y. *et al.* Computational approaches and challenges in spatial transcriptomics. *Genomics Proteomics Bioinform* 2023;**21**:24–47. https://doi.org/10.1016/j.gpb.2022.10.001
17. El Azhar Y, Sonnen KF. Development in a dish—in vitro models of mammalian embryonic development. *Front Cell Dev Biol* 2021;**9**:655993. https://doi.org/10.3389/fcell.2021.655993
18. Drakhlis L, Biswanath S, Farr C-M. *et al.* Human heart-forming organoids recapitulate early heart and foregut development. *Nat Biotechnol* 2021;**39**:737–46. https://doi.org/10.1038/s41587-021-00815-9
19. Müller F, O'Rahilly R. The development of the human brain, including the longitudinal zoning in the diencephalon at stage 15. *Anat Embryol* 1988;**179**:55–71. https://doi.org/10.1007/BF00305100
20. Herriges M, Morrisey EE. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* 2014;**141**:502–13. https://doi.org/10.1242/dev.098186
21. Ye Q, Bhojwani A, Hu JK. Understanding the development of oral epithelial organs through single cell transcriptomic analysis. *Development* 2022;**149**:dev200539. https://doi.org/10.1242/dev.200539
22. Bais AS, Cerqueira DM, Clugston A. *et al.* Single-cell RNA sequencing reveals differential cell cycle activity in key cell populations during nephrogenesis. *Sci Rep* 2021;**11**:22434. https://doi.org/10.1038/s41598-021-01790-6
23. Lua I, Asahina K. The role of mesothelial cells in liver development, injury, and regeneration. *Gut and Liver* 2016;**10**:166–76. https://doi.org/10.5009/gnl15226
24. Tianhao M, Liqin X, Zhong Y. *et al.* Embryonic liver developmental trajectory revealed by single-cell RNA sequencing in the Foxa2eGFP mouse. *Commun Biol* 2020;**3**:642. https://doi.org/10.1038/s42003-020-01364-8
25. Cui G, Dong K, Zhou J-Y. *et al.* Spatiotemporal transcriptomic atlas reveals the dynamic characteristics and key regulators of planarian regeneration. *Nat Commun* 2023;**14**:3205. https://doi.org/10.1038/s41467-023-39016-0
26. Jung S-H, Hwang B-H, Shin S. *et al.* Spatiotemporal dynamics of macrophage heterogeneity and a potential function of Trem2hi macrophages in infarcted hearts. *Nat Commun* 2022;**13**:4580. https://doi.org/10.1038/s41467-022-32284-2
27. Qian X, Coleman K, Jiang S. *et al.* Spatial transcriptomics reveals human cortical layer and area specification. *Nature* 2025;**644**:153–63.
28. Alexander Wolf F, Angerer P, Theis FJ. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;**19**:15.
29. Halmos P, Liu X, Gold J. *et al.* DeST-OT: alignment of spatiotemporal transcriptomics data. *Cell Syst* 2025;**16**:101160.
30. Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. *Am Stat* 1992;**46**:175–85. https://doi.org/10.1080/00031305.1992.10475879
31. Veličković P, Cucurull G, Casanova A. *et al.* Graph Attention Networks. In: *International Conference on Learning Representations (ICLR)*, 2018.
32. Simon LM, Wang Y-Y, Zhao Z. Integration of millions of transcriptomes using batch-aware triplet neural networks. *Nat Mach Intell* 2021;**3**:705–15. https://doi.org/10.1038/s42256-021-00361-8
33. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014.
34. Blondel VD, Guillaume J-L, Lambiotte R. *et al.* Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* 2008;**10**:P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008
35. Traag VA, Waltman L, Van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019;**9**:1–12.
36. Korsunsky I, Millard N, Fan J. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;**16**:1289–96. https://doi.org/10.1038/s41592-019-0619-0
37. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987;**20**:53–65. https://doi.org/10.1016/0377-0427(87)90125-7
38. Hubert L, Arabie P. Comparing partitions. *J Classif* 1985;**2**:193–218. https://doi.org/10.1007/BF01908075
39. McInnes L, Healy J, Melville J. Umap: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426. 2018.
40. Kleshchevnikov V, Shmatko A, Dann E. *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* 2022;**40**:661–71. https://doi.org/10.1038/s41587-021-01139-4

Briefings in Bioinformatics, 2025, **26(6)**, bbaf644
https://doi.org/10.1093/bib/bbaf644
Problem Solving Protocol