Yu and Xie *Genome Biology* (2025) 26:427
https://doi.org/10.1186/s13059-025-03895-x

Genome Biology

METHODOLOGY Open Access

Spatial transcriptomic alignment, integration, and 3D reconstruction by STAIR

Yuanyuan Yu[1,2] and Zhi Xie[1*]

*Correspondence:
Zhi Xie
xiezhi@gmail.com
[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangzhou, China
[2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Abstract
Merging multiple slices into a unified 3D atlas is a significant challenge in spatial transcriptomics. Here, we introduce STAIR, an end-to-end solution for alignment, integration, and 3D reconstruction. STAIR employs a heterogeneous graph attention network with spot-level and slice-level attention mechanisms to achieve a unified embedding space and guide unsupervised 3D reconstruction. We demonstrate STAIR's marked improvements in feature integration and 2D alignment across samples and platforms compared to previous methods. Furthermore, STAIR shows first-of-its-kind performance in z-axis reconstruction of parallel slices and seamlessly integrates new slices into existing 3D atlases, providing novel biological insights from a 3D perspective.

Keywords Spatial transcriptome, 3D reconstruction, Alignment, Integration, Heterogeneous graph attention network

Background

Recent advances in spatial transcriptomics (ST) have enabled the measurement of gene expression while preserving the spatial organization of tissues. Various ST techniques [1–7] facilitate exploring molecular programs in their native spatial context. To study organ- or tissue-level architecture in three-dimensional (3D) space, researchers sampled parallel slices of specific samples at certain distance intervals, where each slice captures spatially resolved molecular features in a single plane (x-axis and y-axis) [8–11]. At present, construction of ST-based molecular atlases is ongoing, such as mouse brain [8, 11], macaque brain [10], and drosophila embryo [9]. By revealing topological structures organ-wide, these studies advance our comprehension of molecular drivers of tissue organization.

However, constructing a 3D atlas from these serial slices remains challenging. Because slices are sequenced independently after sectioning, they often exhibit batch effects in gene expression space and lack shared spatial coordinates in physical space. These difficulties are further compounded when integrating slices across different samples, where the relative spatial distances between slices (z-axis) are generally unknown. As a result,

direct 3D atlas reconstruction is hindered, underscoring the need for computational strategies that can jointly integrate, align, and position slices in a coherent 3D space.

Several computational methods have been proposed to analyze multi-slice ST data. PRECAST [12] integrates spatial embeddings across slices through a probabilistic framework. PASTE [13] aligns 2D coordinates by balancing gene expression similarity and spatial proximity using optimal transport. STAligner [14] unifies embedding integration and 2D alignment via graph attention networks and triplet loss, followed by landmark-based 2D registration based on manually selected landmark domains. While these approaches improve cross-slice correspondence in either feature space or physical space, they operate mainly within the 2D plane. More recently, STitch3D [15] extended these efforts toward 3D analysis by jointly modeling multiple slices to reconstruct tissue structures, cell distributions, and developmental trajectories.

Despite substantial progress, several limitations remain. First, current spatial feature integration strategies treat slices uniformly, ignoring slice-level similarity and structural relevance, which restricts the ability to leverage global organ-level information and limits overall integration performance. Second, gene expression batch effects are often entangled with spatial structure. Spatial features constructed directly on batch-affected expression may propagate noise during neighborhood aggregation, while subsequent batch correction may distort true biological variation. Third, existing 2D alignment methods generally neglect the quantitative use of local region contours, which is the critical information for ensuring continuity between adjacent slices. Fourth, existing 3D reconstruction methods (e.g., STAligner, STitch3D) either assume known inter-slice distances or require full spot-level 3D coordinates. However, such information is often unavailable or unreliable, when dealing with slices across different samples lacking standardized anatomical measurements. Consequently, computational inference of relative z-axis positioning is essential for truly assembling multi-sample 3D atlases, which is a capability that current solutions do not provide. Finally, existing 3D atlases offer limited capacity for quantitative annotation, transfer, or integration of new slices, which hinders both the utility of atlas information and the accumulation of knowledge within a fixed spatial reference framework.

To address these challenges, we developed STAIR, an end-to-end framework for spatial feature integration, 2D physical coordinate alignment and 3D construction. STAIR first mitigates batch effects at the gene expression level using a batch-aware nonlinear embedding to ensure that the expression features are harmonized across slices before spatial feature learning and alignment. It then constructs a large heterogeneous graph [16] with the spots from all slices as nodes and assigns node attributes based on the slice each spot belongs to. To achieve accurate spatial feature integration and consistent spatial region partitioning across slices, STAIR employs an attention mechanism for adaptive feature aggregation both within and across slices, with cross-slice aggregation leveraging spot-level and slice-level attention. These results are then used to guide 2D alignment. Crucially, STAIR infers the relative z-axis positioning of parallel slices in a fully unsupervised manner, requiring only standard ST data as input. Moreover, STAIR enables seamless incorporation of new slices into an existing 3D atlas, effectively expanding the reference 3D atlas. We systematically benchmarked STAIR across diverse datasets and demonstrated its superior performance in spatial feature integration and 2D alignment. Notably, STAIR is the first method capable of unsupervised z-axis

reconstruction of parallel slices, as shown in applications to the mouse hypothalamic preoptic area, whole mouse brain, and tumor tissues. STAIR further enables cross-platform atlas expansion, accurately integrating new slices from a different ST platform and transferring atlas annotations.

Results

Overview of the STAIR framework

STAIR achieves integration and alignment of molecular features and physical coordinates for ST data, enabling 3D reconstruction and assimilation of new slices into the reference atlases (Fig. 1). The framework requires only ST data as input and outputs aligned spatial embeddings and reconstructed 3D coordinates. STAIR consists of two core modules (Fig. 1A): STAIR-Emb, which aligns molecular embeddings, and STAIR-Loc, which performs 2D coordinate registration along the x–y plane. The reconstruction and prediction of z-axis positions are based on the high-order semantic information captured by STAIR-Emb. Based on the aligned embeddings, spatial domains and developmental trajectories can be identified through standard clustering and trajectory inference. Thus, using multi-slice ST data alone, STAIR reconstructs 3D tissue organization while preserving both discrete anatomical structures and continuous spatial gradients.

For spatial feature integration (Fig. 1A, left), STAIR first applies an autoencoder with batch factors [17, 18] for nonlinear dimensionality reduction, mitigating batch effects in gene expression. STAIR-Emb then builds a heterogeneous graph in which all spots serve as nodes and their slice identities serve as node attributes. Intra-slice edges are defined by spatial neighborhoods, whereas inter-slice edges are weighted by gene expression similarity because the relative slice positions are unknown. Subsequently, STAIR-Emb aggregates intra-slice and inter-slice neighbor information in sequence, where the inter-slice aggregation employs an attention mechanism [16] composed of spot-level and slice-level attention. Importantly, slice-level attention scores capture higher-order correspondences across slices.

For 2D physical coordinate alignment (Fig. 1A, right), STAIR-Loc conducts a two-step registration between slice pairs. The initial alignment uses rotation and translation guided by spatial features to produce an approximate overlay. The fine alignment then employs the iterative closest point (ICP) algorithm [19] based on boundary spots of the slices and the most aggregated domain. Sequential application of STAIR-Loc across ordered slices reconstructs a stacked 3D space.

To enable fully unsupervised 3D reconstruction from an arbitrary set of parallel slices (Fig. 1B), we define inter-slice semantic distances based on the slice-level attention scores and construct a minimum spanning tree (MST) [20] to infer the relative z-axis coordinates. These inferred z-coordinates then guide the final x–y alignment in STAIR-Loc, producing a continuous 3D atlas. STAIR also supports assimilation of new parallel slices into an existing atlas (Fig. 1C). We identify atlas slices close to the new slice based on the semantic distances, followed by attention-weighted averaging to predict the z-coordinate of the new one. The atlas slice with the closest z-coordinate is then selected as a spatial template to scale and align the x- and y-axes. At this stage, the 3D coordinates for every spot in the new slice have been obtained, allowing annotated atlas information (e.g., standard anatomical region labels) to be accurately propagated to the

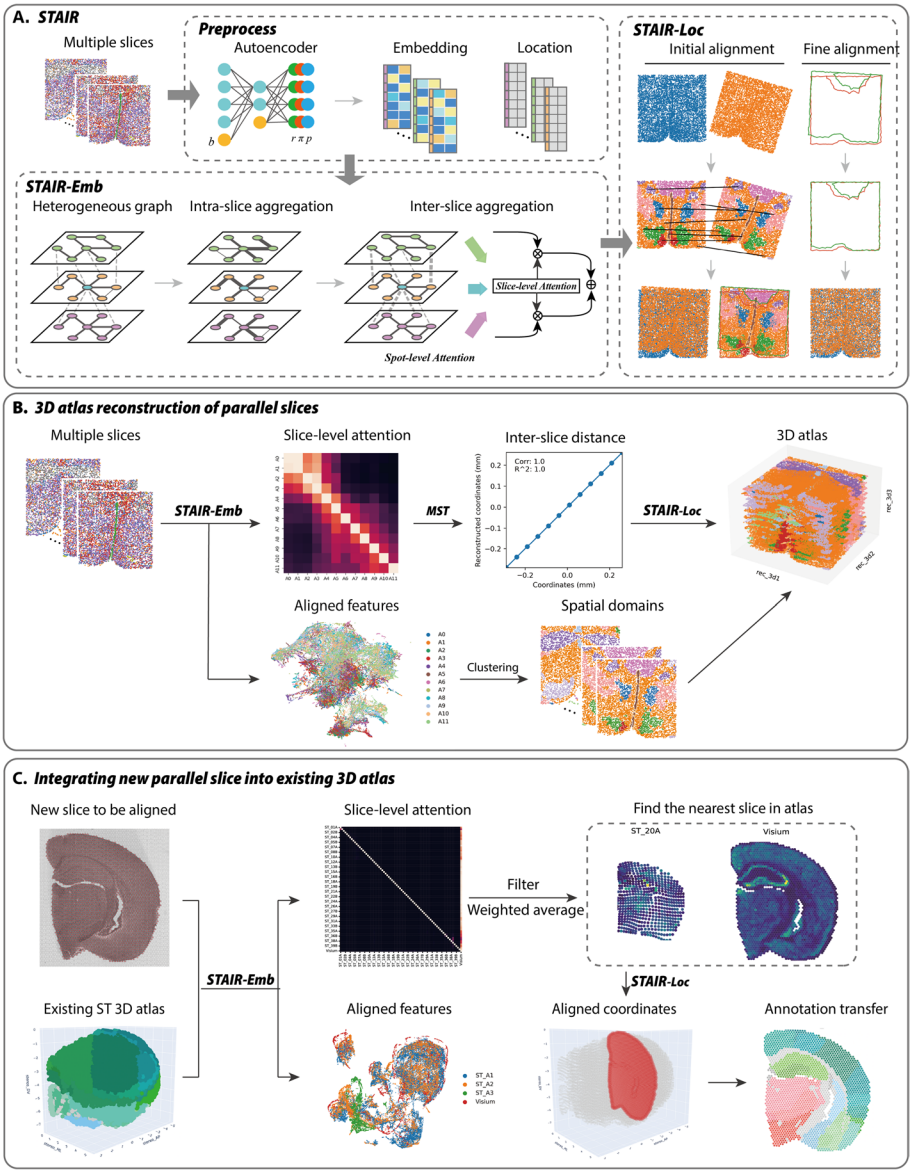

Fig. 1 Overview of STAIR framework. **A** STAIR processes multiple ST slices by utilizing an autoencoder to compress the expression matrices of each slice, resulting in integrated expression features. Subsequently, STAIR-Emb takes the expression features and the 2D coordinates of each slice as input, employing a heterogeneous graph attention network to learn integrated spatial features. Finally, STAIR-Loc utilizes these spatial features to establish the initial alignment of spatial coordinates, followed by refining the alignment further by incorporating boundary points of slices and their respective domains. **B** 3D atlas reconstruction of parallel slices. Without prior knowledge of inter-slice distance, STAIR-Emb derives inter-slice distance matrices from semantic relationships, and a minimum spanning tree (MST) infers relative z-axis positions to guide sequential 2D alignment. **C** Integration of new parallel slices into an existing 3D atlas. STAIR-Emb integrates spatial features of the new slice with the 3D atlas, followed by predicting the new slice's z-axis location and aligning its 2D coordinates (x- and y-axis) with the 3D atlas using STAIR-Loc

newly integrated slice. Such expansion enhances the utility and applicability of the established 3D atlas.

Accurate integration of spatial embeddings across slices by STAIR

We first quantitatively evaluated STAIR's efficacy in integrating spatial embeddings and identifying unified spatial domains across diverse tissue sections, a prerequisite for

downstream analysis. Our initial testing utilized 10X Visium dataset derived from the human dorsolateral prefrontal cortex (DLPFC) [21], spanning three samples with four sequential sections per sample. The original study [21] performed precise manual annotation, delineating white matter (WM) and six gray matter layers ranging from layer 1 to layer 6 to provide ground truth labels (Fig. 2A). First, we conducted separate spatial domain identification from slices of each sample, which had close locations with some differences. STAIR achieved the most accurate domain division results on all three samples (Fig. 2B, Additional file 1: Fig. S1), with respective median adjusted rand index (ARI) values of 0.60, 0.53, and 0.62 (Fig. 2C).

Then, we processed twelve slices from three samples simultaneously (Fig. 2D; Additional file 1: Fig. S2A). STitch3D was excluded from this test due to its requirement for 3D coordinates to handle multiple slices, which were not available. Despite the challenge, STAIR maintained the highest consistency with annotations, achieving a median ARI value of 0.65, far exceeding the second-ranked STAligner with a median value of 0.46 (Fig. 2C, right). Notably, simultaneous consideration of three samples proved superior to testing each sample individually, resulting in higher ARI and a clearer demarcation between layer 4, layer 5, and layer 6 (Fig. 2B to D). In Sample 2, STAIR uncovered a small set of spots corresponding to layer 1 (cluster 7) and layer 2 (cluster 5), which had been misclassified as layer 3 (cluster 3) in the original annotation. Validation using layer-specific marker genes (*GFAP* for layer 1, *C1QL2* and *HPCAL1* for layer 2, and *ADCYAP1* and *FREM3* for layer 3) confirmed the presence of these layers (Additional file 1: Fig. S3), demonstrating that STAIR not only ensures robust cross-sample alignment but also refines existing annotations by uncovering biologically meaningful structures previously

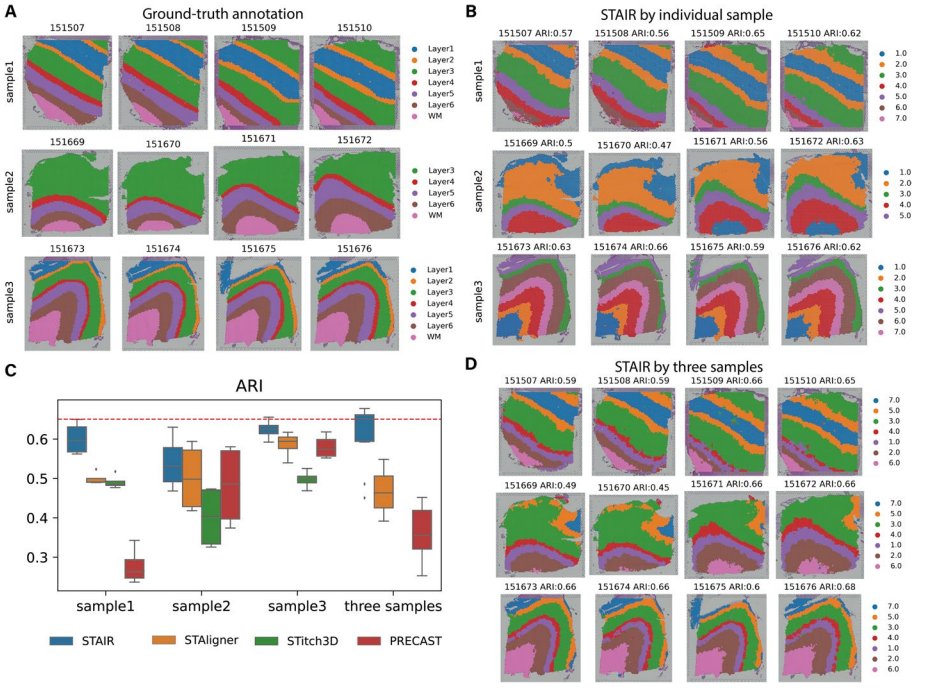

Fig. 2 STAIR effectively integrates heterogeneous spatial transcriptomics data. **A** Ground-truth segmentation of manually annotated regions in 12 DLPFC sections. **B** STAIR's spatial domain identification based on the 4 DLPFC slices for each sample. **C** Boxplots of adjusted rand index (ARI) scores of the four methods applied to the 4 DLPFC slices of each sample and to the total 12 slices of the three samples. In the boxplot, the center line denotes the median, box limits denote the upper and lower quartiles, and whiskers denote the 1.5 × interquartile range. **D** STAIR's spatial domain identification based on the 12 DLPFC slices

overlooked. Subsequently, we performed low-dimensional visualization using uniform manifold approximation and projection (UMAP) [22] based on the spatial embeddings derived by these methods (Additional file 1: Fig. S2B). In the STAIR-based visualization, all spots exhibited an arrangement according to the layers, with thorough mixing between the different samples. In contrast, both STAligner and PRECAST lacked clarity in arranging and distinguishing these known layers. Specifically, STAligner mixed layer 2 and layer 3, as well as layer 4 to layer 6, while PRECAST only distinguished WM. We also quantified the effects of spatial embedding learning and integration using average silhouette width (ASW). We calculated ASW for spatial embeddings with respect to spatial domains (ASW_{domain}) and samples (ASW_{batch}), as well as ASW_{F1} to evaluate overall capability (Methods). STAIR achieved best spatial embedding learning and integrating with the highest ASW_{F1} (Additional file 1: Fig. S2C). STAligner also adequately integrated samples, as indicated by an ASW_{batch} value similar to STAIR. However, it had weaker feature learning capabilities with a much lower ASW_{domain}, which was consistent with the unclear UMAP pattern we observed previously.

Furthermore, we conducted comprehensive ablation studies to evaluate the contribution of each component within STAIR-Emb, as well as the number of slices for each sample to the overall performance improvement (Additional file 2: Note S1, Fig. S4). In summary, STAIR successfully learned and integrated the spatial embeddings, ensuring a consistent spatial region division.

Precise alignment of 2D coordinates across slices by STAIR

A single ST experiment only acquires data from one slice, resulting in the loss of the unified physical space across multiple slices. In this section, we evaluate STAIR's ability in 2D coordinates alignment. We utilized 12 MERFISH slices of the mouse hypothalamus preoptic area [2] with known 3D coordinates. To perform testing, we kept the first slice fixed while introducing random rotations and translations in x- and y-axes to the remain 11 slices (Fig. 3A). Subsequently, we employed STAIR, STAIR-init (initial alignment of STAIR-Loc), PASTE, STitch3D, and STAligner on the 2D shuffled slices and assessed their effectiveness. For the spatial features based 2D alignment methods, including STAIR and STAligner, we first learned and aligned the spatial features of all 12 slices simultaneously, and then performed 11 consecutive pairwise positional alignments while keeping the reference slice fixed.

STAIR achieved the most accurate 2D alignment, aligning each slice to a position very close to the real coordinates (Fig. 3B), with the median rotation and translation errors being 0.03 and 0.04mm respectively (Fig. 3C). STAIR's accurate alignment results partly benefited from the good initial values provided by the initial alignment based on spatial features. The alignment of STAIR-init was already quite impressive, with the medians of the two errors being 0.1 and 0.1mm, ranking second after STAIR. Fine alignment based on the continued spatial domain further improved STAIR's accuracy. Repeating 500 random rotations and translations, both STAIR and STAIR-init achieved similar results (Additional file 1: Fig. S5A). STAligner performs 2D alignment based on spatial feature of manually selected spatial domains. Since anchor domain selection might affect the results, we selected each domain as an anchor in turn. Among them, four domains could produce alignment result, which we named STAligner-0, STAligner-1, STAligner-2, and STAligner-4, respectively. Different anchor domains performed differently, with

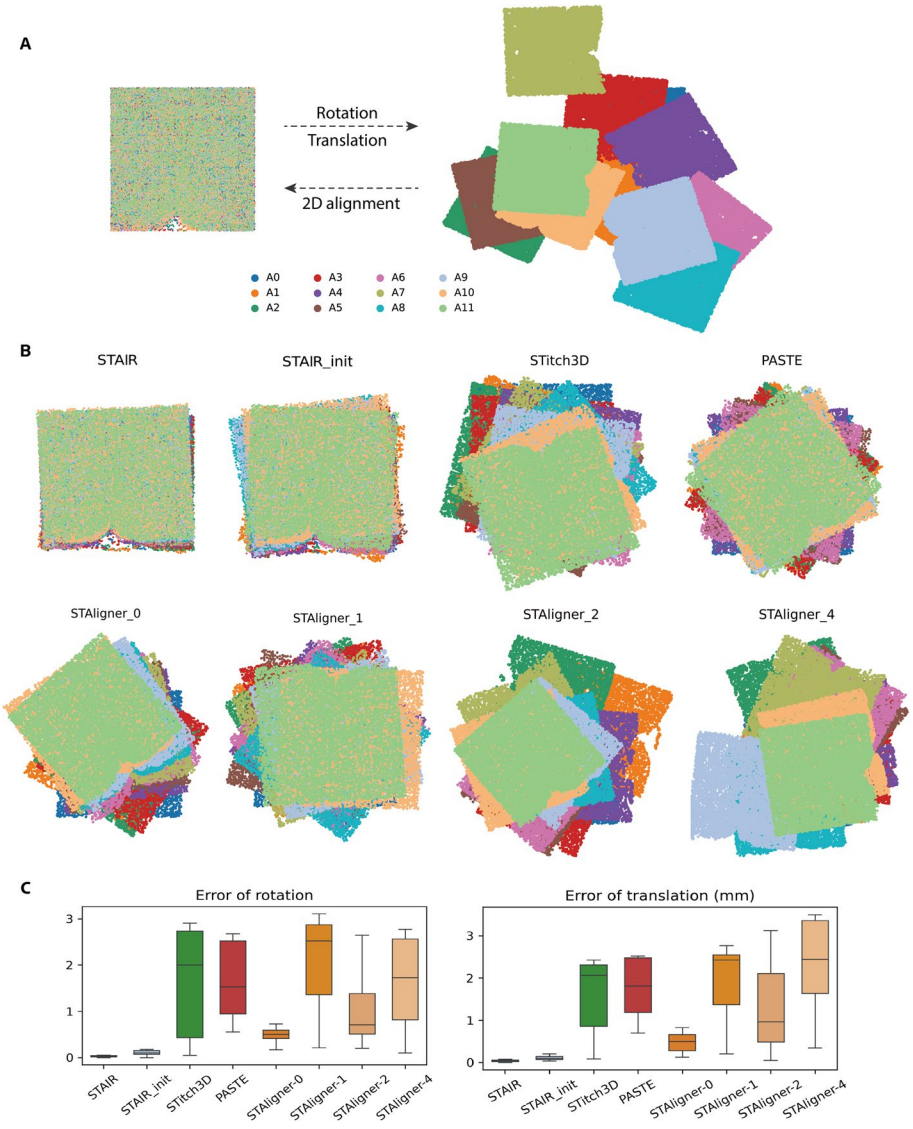

Fig. 3 Precise alignment of 2D coordinates by STAIR. **A** Schematic diagram of spatial position alignment. 2D coordinates in the first slice were fixed, and the remaining 11 slices were randomly rotated and translated. STAIR, STAIR-init, PASTE, STitch3D, and STAligner were employed to align the spatial coordinates of the rotated data. **B** Results of 2D spatial alignment using STAIR, STAIR-init, PASTE, STitch3D, and STAligner. **C** Boxplots showing the rotation (left) and translation (right) errors of each method. In the boxplot, the center line denotes the median, box limits denote the upper and lower quartiles, and whiskers denote the 1.5 × interquartile range

STAligner-0 achieving the best results with median errors of 0.5 and 0.5mm. In contrast, the alignment effects of STitch3D and PASTE were limited, exhibiting rotation and translation errors over 1.5 and 1.8mm, respectively.

We further assessed STAIR's robustness in aligning 2D coordinates. First, we examined the impact of resolution by aggregating neighboring cells into virtual spots with lower resolution. With 2–5 aggregated cells per spot, STAIR maintained the lowest median rotation errors of 0.02–0.07 and translation errors of 0.04–0.10mm (Additional file 1: Fig. S5B). The two errors of STAIR-init were slightly larger, with values ranging from 0.08 to 0.17 and 0.13 to 0.17mm, respectively. Furthermore, assuming that STAligner could make the most accurate manual domain selection for 2D alignment, we reported only the optimal result as STAligner-best. At different resolutions, the rotation

and translation error ranges of STAligner-best were 0.30 to 1.01 and 0.46 to 0.98, respectively. Additionally, given that fine alignment of STAIR relied on spatial domain information, we examined performance across 8–15 domains (Additional file 1: Fig. S5C). STAIR demonstrated stable alignment, with median rotation errors of 0.02 to 0.08 and translation errors of 0.03 to 0.09.

In short, dataset with ground truth coordinates confirmed STAIR's capabilities for precise 2D slice alignment and its robustness across varied resolutions and numbers of spatial domains.

STAIR effectively integrates heterogeneous ST data

With the advancement of ST technology, integrating datasets from different samples and platforms became increasingly important. Effective integration should identify shared signals across samples while retaining biologically unique variations within each individual sample.

We applied STAIR to integrate mouse olfactory bulb data from Stereo-seq [23] and Slide-seqV2 [4] platforms, which differed in spatial resolution and sampling location. Stereo-seq captured the main olfactory bulb (MOB) at sub-single-cell resolution, providing highly detailed spatial information. In contrast, Slide-seqV2 encompassed both MOB and the accessory olfactory bulb (AOB) with a resolution of 10 μm, which approximates but does not precisely achieve single-cell resolution. STAIR accurately delineated the MOB region common to both datasets and the AOB region unique to Slide-seqV2 (Fig. 4A, top), highly consistent with standard allen brain atlas (ABA) partitioning (Additional file 1: Fig. S6A). The AOB region, located in the middle-upper part of the olfactory bulb, comprised two sub-regions: AOBmi and AOBgr. The MOB region encompassed seven sub-regions arranged in a concentric ring: rostral migratory stream (RMS), granule cell layer (GCL), two mitral cell layers (MCLs), external plexiform layer (EPL), glomerular layer (GL), and olfactory nerve layer (ONL). Each region exhibited high expression of its corresponding marker genes (Additional file 1: Fig. S6B). In comparison, STAligner struggled to identify the RMS layer and confused the EPL and GL layers of Stereo-seq (Fig. 4A, middle). PRECAST failed to identify coherent spatial patterns (Fig. 4A, bottom). UMAP visualization demonstrated that STAIR effectively distinguished between AOB and MOB in low-dimensional space, preserving the AOBmi and AOBgr sub-regions in Slide-seqV2 data (Fig. 4B, C). Simultaneously, STAIR seamlessly integrated the shared MOB region, arranging sub-layers consistent with their physical locations. In contrast, while STAligner achieved integration to a certain extent, it failed to discern the distinct difference between AOB and MOB. PRECAST completely mixed the datasets, losing Slide-seqV2 specificity.

Furthermore, we integrated mouse embryo data from three platforms: Stereo-seq, Slide-seqV2, and seqFISH. These datasets exhibit significant differences in biological characteristics, experimental properties, spatial resolution, and gene throughput. The mouse embryos were sampled at E9.5 for Stereo-seq [7] and Slide-seqV2 [24], and E8.75 for seqFISH [25]. Both Stereo-seq and Slide-seqV2 performed genome-wide measurements, with Stereo-seq data at bin50 (35.71μm) and Slide-seqV2 at 10 μm spatial resolution. In contrast, seqFISH, an imaging-based ST method, measured 342 genes with single-molecule resolution. Despite these substantial differences, STAIR successfully achieved spatial feature integration (Additional file 1: Fig. S7A) and consistent spatial

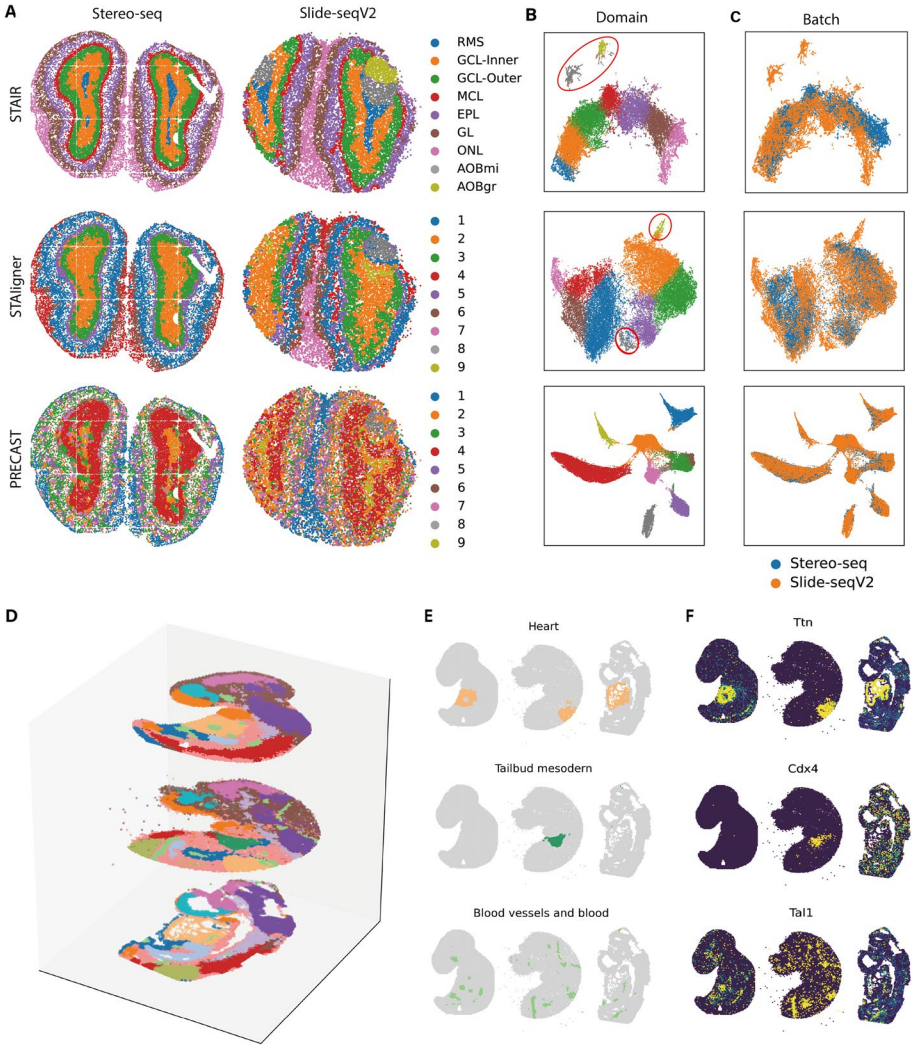

Fig.4 Cross-platform integration of mouse olfactory bulb and embryo datasets. **A** Integrative spatial domain identification of mouse olfactory bulb datasets: Stereo-seq (left) and Slide-seqV2 (right), using STAIR (top), STAligner (middle), and PRECAST (bottom). **B** UMAP visualizations derived from spatial embeddings of STAIR (top), STAligner (middle), and PRECAST (bottom), colored by identified spatial domains. **C** UMAP visualizations derived from spatial embeddings of STAIR (top), STAligner (middle), and PRECAST (bottom), colored by dataset of origin. **D** Spatial visualization of domain identification and 2D alignment results for three mouse embryo datasets: Stereo-seq (top), Slide-seqV2 (middle), and seqFISH (bottom). Colors represent spatial domains identified by STAIR. **E** Heart, somite, and tailbud mesoderm domains identified by STAIR for three mouse embryo datasets: Stereo-seq (left), Slide-seqV2 (middle), and seqFISH (right). **F** Scaled expression of marker genes *Ttn*, *Cdx4*, and *Tal1*, corresponding to heart, tailbud mesoderm, and blood vessels and blood, respectively, in the three mouse embryo datasets

domain identification (Fig. 4D, E, Additional file 1: Fig. S7B, C), while also scaling and aligning the 2D coordinates of these slices (Fig. 4D). STAIR identified domains including forebrain, midbrain, hindbrain, spinal cord, neural crest, somite, heart, blood vessels and blood, liver-gut tube, meninges, connective tissue, tailbud mesoderm, lateral plate mesoderm, and cranial mesoderm. These domains were validated through known marker genes: *Lhx2* [26], *Otx2* [27], *Gbx2* [25], *Hoxb9* [28], *Prrx1* [29], *Meox1* [24], *Ttn* [24], *Tal1* [30], *Afp* [24], *Col4a1* [24], *Postn* [7], *Cdx4* [31], *Foxf1* [32], and *Tbx1* [33], respectively (Fig. 4F, Additional file 1: Fig. S8). Notably, despite significant positional differences in domains such as the heart and somite across the datasets, particularly in Slide-seqV2, they all highly correspond with areas of high *Ttn* gene expression (Fig. 4E, F, top). The

Slide-seqV2 dataset also exhibited a unique tailbud mesoderm region (Fig. 4E, F, middle). Importantly, STAIR not only identified these highly spatially clustered domains but also accurately detected dispersed structures such as blood vessels and blood, and connective tissue (Fig. 4E, F, bottom).

In conclusion, despite the diverse origins of these datasets from distinct platforms with significant biological and technological disparities, STAIR successfully achieved precise feature integration and 2D spatial alignment.

Construction of 3D atlas by STAIR

Although spatial embeddings and 2D coordinates can be aligned across multiple heterogeneous slices, constructing a coherent 3D atlas from multi-sample ST data remains challenging. When slices originate from different samples, their relative positions along the z-axis are typically unknown. To address this challenge, STAIR computationally infers the relative z-axis positions of slices, enabling 3D reconstruction without any pre-existing slice-level z-coordinate information.

We first evaluated z-axis reconstruction on a single-sample dataset to validate the reliability of STAIR's 3D positioning. Specifically, we used 12 MERFISH-generated mouse hypothalamic preoptic area slices (Fig. 5A, left), for which the true z-axis positions were available [2] (Additional file 1: Fig. S9A). STAIR-Emb was applied to integrate the slices (Additional file 1: Fig. S9B, C), and we observed that inter-slice attention scores were strongly negatively correlated with physical distances (Fig. 5A middle, B), highlighting their potential for z-axis inference. Using MST constructed from the attention-derived semantic distance matrix (Methods), STAIR accurately reconstructed the z-axis (Fig. 5A right). We then sequentially aligned the x- and y-axes of slices according to the reconstructed z-axis, yielding a 3D atlas that closely matched the true coordinates (Fig. 5A right). We benchmarked z-axis reconstruction against simpler slice–slice similarity measures, including Pearson correlation of total gene expression and cell-type composition between slices. Pairwise similarity visualization and spearman's ρ correlations with physical distances showed that attention scores consistently provided the strongest negative correlations (Fig. 5B, C, Additional file 1: Fig. S9D). In slices A5–A11, both gene expression and cell-type composition were highly correlated (Fig. 5C, green box), while attention scores accurately reflected their true physical relationship. We quantitatively evaluated z-axis reconstruction using three metrics: Pearson correlation coefficient (PCC) between reconstructed and true coordinates, the normalized mean squared error (NMSE), and the misordering rate (MR) (Methods). Attention-based reconstruction was nearly perfect, achieving PCC = 1 with NMSE = 0 and MR = 0 (Fig. 5D, E). By contrast, simpler similarity metrics yielded much higher NMSE, particularly for cell-type composition, suggesting reduced precision in localization. To further challenge the models, we randomly cropped slices at different angles, removing 20%, 40%, or 60% of spots (Additional file 1: Fig. S10). As crop ratios increased, the performance of simple similarity metrics degraded substantially, whereas attention-based reconstruction remained robust (Fig. 5F). Even under 60% cropping, the median NMSE remained as low as 0.08 (Fig. 5G). This robustness stems from the ability of attention to integrate higher-order semantic associations across locally matched spots rather than relying solely on global slice-level similarity. Finally, when applied to 24 slices from two animals of opposite

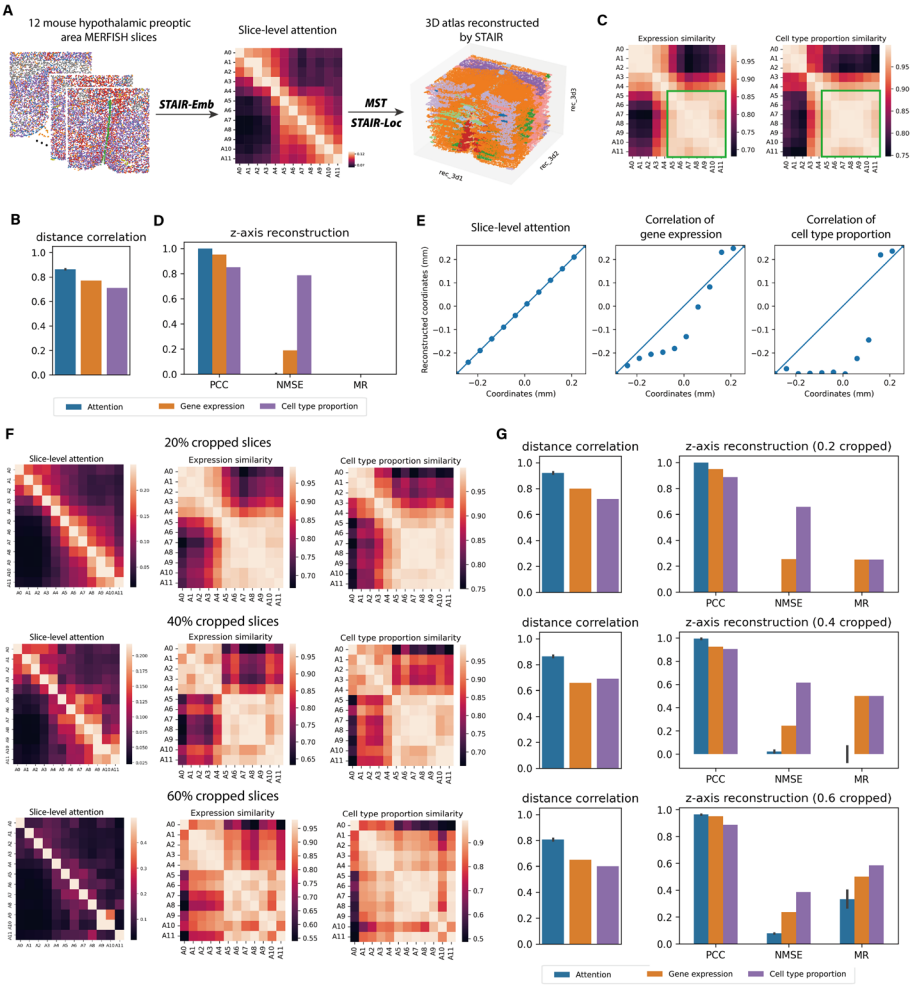

Fig. 5 STAIR reconstructs 3D atlases for the hypothalamic preoptic region. **A** Schematic of 3D reconstruction for 12 parallel MERFISH slices of the hypothalamic preoptic region. STAIR-Emb computes slice-level attention scores, which are then used to infer relative distances along the z-axis. STAIR-Loc aligns the 2D coordinates of each slice guided by these inferred inter-slice distances. **B** Bar plot showing the negative Spearman correlation between pairwise attention scores or simpler similarity metrics and the physical distances between slices. The height of bar for attention-based scores represents the median correlation across 50 repeated integrations, and the error bars indicate the standard deviation. **C** Heatmaps showing slice–slice similarity computed from total gene expression (left) and cell type composition (right). **D** Bar plot of z-axis reconstruction performance, including Pearson correlation coefficient (PCC) between reconstructed and true coordinates, normalized mean squared error (NMSE), and misordering rate (MR). For attention-based scores, bar heights represent the median values across 50 repeated integrations, and error bars show the standard deviation. **E** Scatterplots of reconstructed z-coordinates (y-axis) versus true z-coordinates (x-axis) for reconstructions based on attention scores (left), total gene expression (middle), and cell type composition (right). **F** Heatmaps of slice–slice similarity under 20% (top), 40% (middle), and 60% (bottom) random cropping, computed using attention scores (left), total gene expression (middle), and cell type composition (right). **G** Left: Bar plots showing the negative Spearman correlation between pairwise attention scores or simpler similarity metrics and physical distances under 20% (top), 40% (middle), and 60% (bottom) cropping. Right: Bar plots of z-axis reconstruction performance (PCC, NMSE, and MR) under the same cropping conditions

sexes, STAIR maintained high accuracy, demonstrating robustness to biological variability (Additional file 1: Fig. S11).

We next tested STAIR on a more challenging dataset comprising 40 coronal half-brain slices generated from three different samples using the ST platform [8], spanning the olfactory bulb to the hindbrain along the anterior–posterior (AP) axis. Bregma coordinates from the original study [33] serving as ground truth z-axis positioning (Fig. 6A, left). STAIR-Emb efficiently integrated spatial features and delineated spatial regions

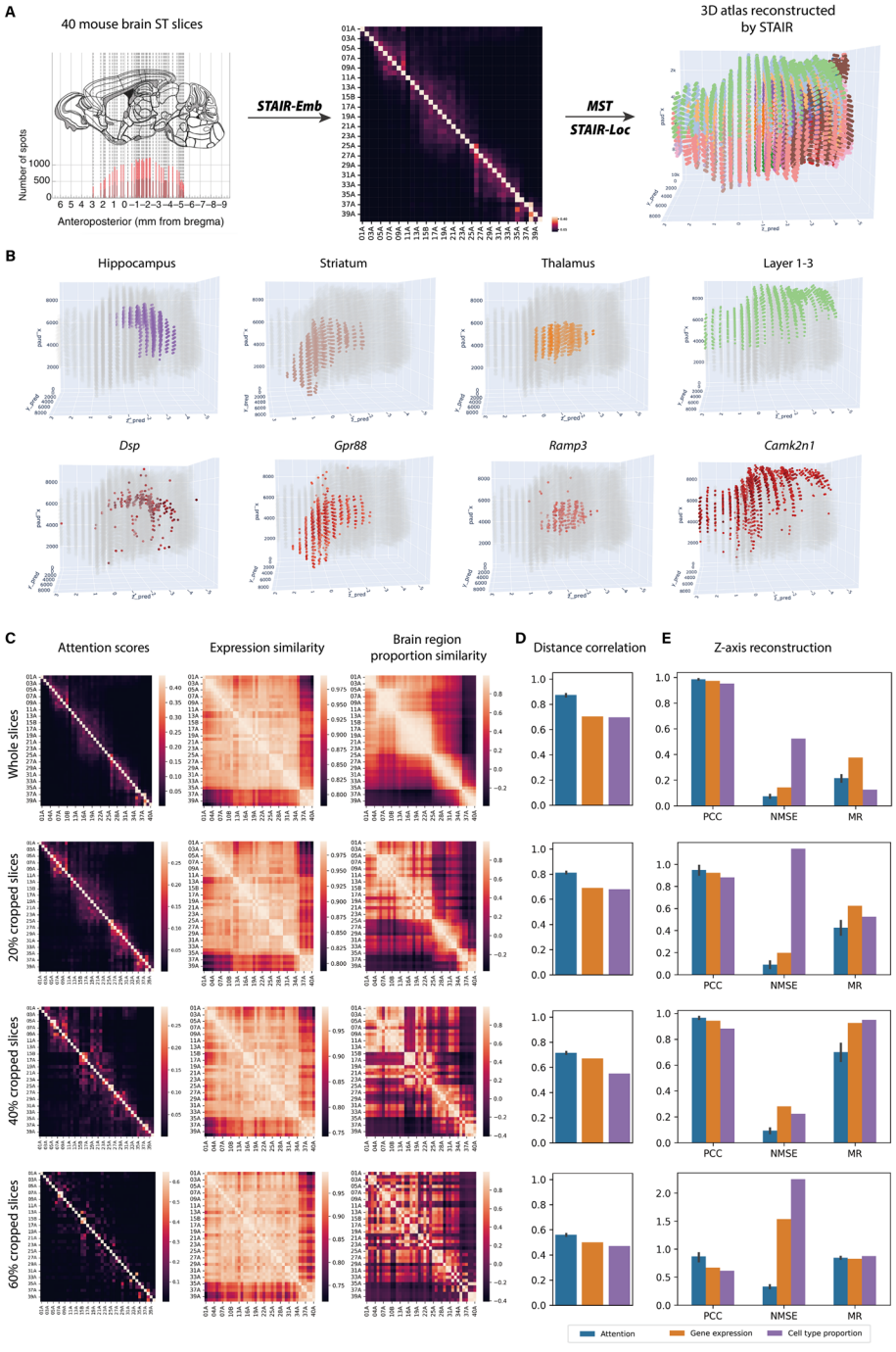

Fig. 6 STAIR reconstructs 3D atlases for the ST mouse brain. **A** Left: Distribution of 40 coronal sections generated by the ST platform used to generate the atlas, adapted from the Ortiz's work [8]. Middle: Heatmap of attention scores across 40 slices in the of ST mouse brain data. Right: Visualization of reconstructed 3D coordinates, colored by spatial domains. **B** Visualization of reconstructed 3D coordinates, colored by spatial domains (top) and their corresponding marker genes (bottom). **C** Heatmaps of slice–slice similarity under 20% (top), 40% (middle), and 60% (bottom) random cropping, computed using attention scores (left), total gene expression (middle), and cell type composition (right). **D** Bar plots showing the negative Spearman correlation between pairwise attention scores or simpler similarity metrics and physical distances under 20% (top), 40% (middle), and 60% (bottom) cropping. **E** Bar plots of z-axis reconstruction performance (PCC, NMSE, and MR) under the same cropping conditions

(Additional file 1: Fig. S12A). Despite the complexity of this dataset, inter-slice attention scores remained strongly negatively correlated with physical distances (Fig. 6A, middle), enabling accurate reconstruction of the AP-axis coordinates. Subsequently, STAIR-Loc aligned dorsoventral (DV) and mediolateral (ML) coordinates following the reconstructed AP order, completing the 3D reconstruction (Fig. 6A, right). The accuracy of spatial localization was further validated by the high expression of region-specific marker genes, including *Dsp* [34] in the hippocampus, *Gpr88* [35, 36] in the striatum, *Ramp3* [37, 38] in the thalamus, and *Camk2n1* [39] in the superficial cortex (Fig. 6B). We further compared attention-based reconstruction with simpler similarity metrics under both full-slice and cropped conditions (20%, 40%, 60%) (Fig. 6C–E, Additional file 1: Fig. S12B, C, S13). Attention consistently outperformed other metrics in PCC and NMSE, confirming its superiority in capturing higher-order semantic relationships. MR was relatively unsatisfactory across all methods. This can be explained by the fact that MR is a highly sensitive metric: even small deviations in reconstructed positions, which may have little impact on PCC or NMSE, can still be counted as ordering errors, thereby inflating the MR values.

Finally, we applied STAIR to two large, orthogonal MERFISH mouse brain datasets comprising 66 coronal and 23 sagittal slices respectively, each containing over one million cells. STAIR accurately reconstructed the z-axis positions and sequentially aligned the x- and y-axes (Fig. 7A, B). STAIR-derived spatial features enabled precise UMAP visualization and brain region delineation (Fig. 7C, D), while seamlessly integrating different samples and slices (Additional file 1: Fig. S14A, S15A). From anterior to posterior, we successfully identified olfactory regions and subregions, superficial (Layers 1–3) and deep (Layers 4–6) layers of the isocortex, striatum, hippocampus (CA and DG), thalamus, hypothalamus, midbrain, hindbrain, and cerebellum (granular, Purkinje–Bergmann, and molecular layers). Additional structures including fiber tracts, the ventricular system (VS), and meninges were also recovered. Each region was faithfully reconstructed in 3D with strong spatial continuity (Additional file 1: Fig. S14B, S15B, S16, S17). We further benchmarked z-axis reconstruction by comparing attention scores with simpler similarity metrics. Attention consistently yielded the strongest correlations with physical distance and the highest accuracy (Fig. 7E–G, Additional file 1: Fig. S18A–C, S19A–C), highlighting its ability to capture higher-order inter-slice relationships beyond intuitive similarity. Moreover, attention remained robust under random cropping (20%, 40%, 60%), again outperforming simpler measures (Additional file 1: Fig. S18D, S19D, S20, S21).

3D modeling of breast tumor microenvironment

To further illustrate biological insights provided by 3D atlas, we analyzed the HER2-positive breast cancer [40] ST data comprising three consecutive slices (H1 to H3) (Additional file 1: Fig. S22A). The pathologists [40] annotated one slice (H1) with six tissue types: invasive cancer, adipose tissue, connective tissue, breast glands, in situ cancer, and immune infiltrates (Fig. 8A).

First, STAIR integrated the spatial embeddings of all the three slices, yielding spatial domains highly consistent with the pathological annotations (Additional file 1: Fig. S22B). On slice H1, STAIR achieved an ARI of 0.36, surpassing that of STAligner, STitch3D, and PRECAST, which ranged from 0.30 to 0.32 (Additional file 1: Fig. S22C).

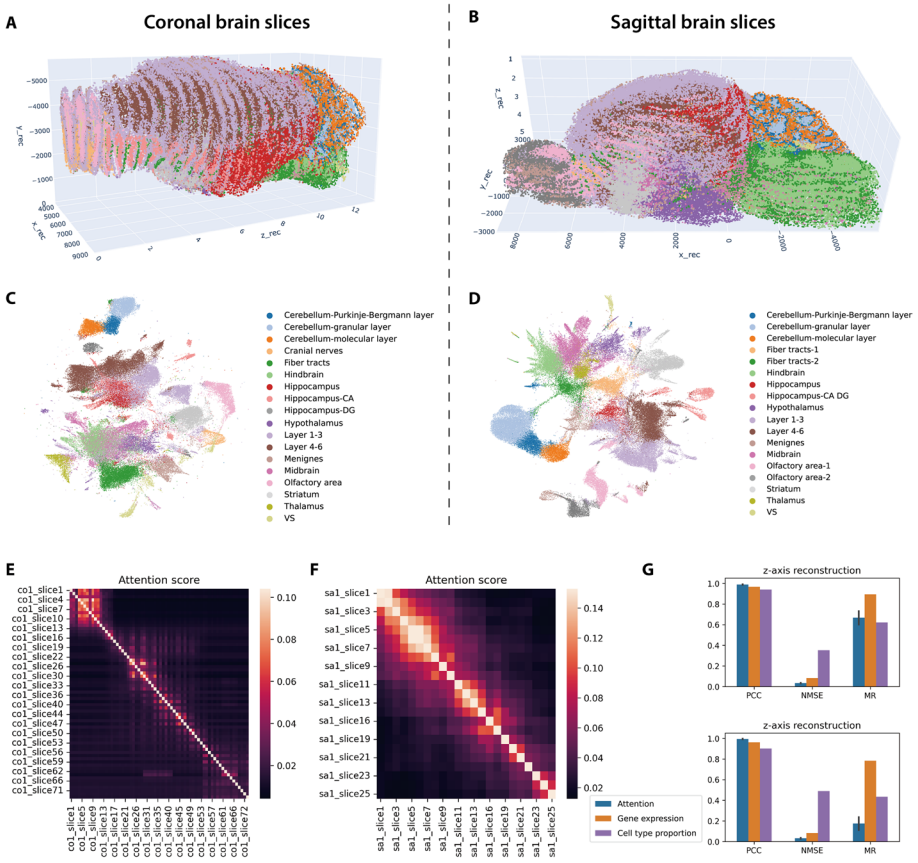

Fig. 7 STAIR reconstructs 3D atlases for coronal and sagittal MERFISH mouse brain. **A**, **B** Visualization of the 3D coordinates reconstructed by STAIR for coronal (**A**) and sagittal (**B**) mouse brains. The z-axis represents slice-wise reconstruction, while the x- and y-axes are aligned according to the reconstructed z-axis order. Each dot represents a cell, with 100,000 cells randomly subsampled for visualization. Points are colored by spatial domains. **C**, **D** UMAP visualization of coronal (**C**) and sagittal (**D**) mouse brains based on STAIR-derived spatial features, colored by spatial domains. Each dot represents a cell, with 100,000 cells randomly subsampled for visualization. **E** Heatmap of attention scores across 66 coronal slices. **F** Heatmap of attention scores across 23 sagittal slices. **G** Bar plots showing z-axis reconstruction performance (PCC, NMSE, and MR) for coronal (top) and sagittal (bottom) mouse brains

Recognizing the inherent distinctions between transcriptome and pathological phenotypes, we deconvoluted [41] each spot with scRNA-seq [42] data to facilitate spatial domain annotation (Additional file 1: Fig. S23). STAIR's spatial regions were annotated as connective tissue, immune cancer, breast glands, adipose tissue, fibrous tissue near the tumor, invasive cancer, and two in situ cancer regions (Fig. 8B). In contrast, STAligner failed to differentiate the in situ cancer area from the invasive cancer area (cluster 0) (Additional file 1: Fig. S22B). STitch3D could not distinguish between the two in situ cancer areas (cluster 4). PRECAST struggled to separate the two spatially distinct in situ cancer areas (cluster 2, 3, and 5). Additionally, none of them could detect a heterogeneous region near tumor with fewer cancer cells and more immune cells (cluster 7).

Next, we reconstructed the 3D coordinates. The inferred z-axis distance, at 0.52/0.43, closely corresponded the true distance ratio. We then aligned the x-axis and y-axis positions in the order of the z-axis, revealing a continuous structure for each spatial region in the 3D space (Fig. 8C). While STAligner and STitch3D cannot infer z-axis, we compared STAIR with them in 2D coordinates alignment (Additional file 1: Fig. S22D) and employed LISI metric to assess the spatial clustered pattern of domains in the stacked

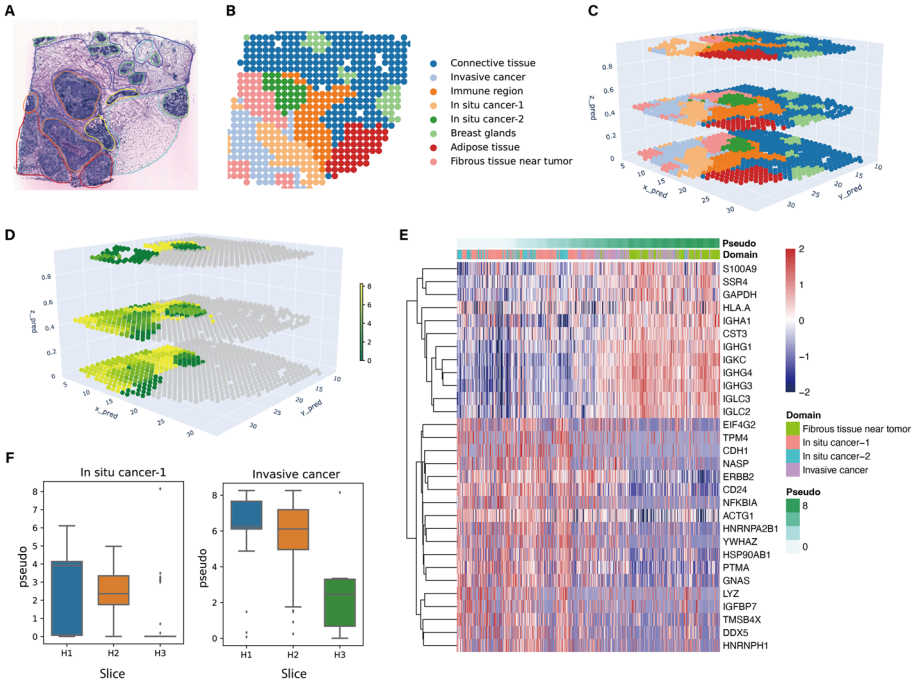

Fig. 8 3D reconstruction and analysis of HER2 + breast cancer slices. **A** Annotations of slice H1 in the original study [40] into six distinct categories: invasive cancer (red), adipose tissue (cyan), connective tissue (blue), breast glands (green), in situ cancer (orange) and immune infiltrates (yellow). **B** 2D spatial visualization shows the domains identified by STAIR in slice H1. That of the other two slices are displayed in Fig. S6B. **C** Visualization of reconstructed 3D coordinates, colored by spatial domains. **D** Pseudo-time of each spot inferred by Monocle3 based on spatial embedding from STAIR. **E** Heatmap displaying genes with expression changes along the Monocle-derived pseudotime, with spots ordered by pseudo-time. **F** Boxplot shows the pseudo-time of the spots in in situ cancer-1 (left) and invasive cancer (right) for each slice. In the boxplot, the center line denotes the median, box limits denote the upper and lower quartiles, and whiskers denote the 1.5 × interquartile range

2D space, which was a comprehensive measure of spatial domain identification and 2D coordinates alignment (Additional file 1: Fig. S22E). STAIR showed best performance with the lowest median LISI value of 1.53.

Furthermore, we examined developmental trajectories [43] and 3D heterogeneity of tumor-associated domains. To select the initial domain of the developmental trajectory, we conducted a differential expression analysis of two in situ cancer regions (Additional file 1: Fig. S24A). In situ cancer-1, marked by high expression of the *ERBB2* [44] gene and S100 family genes [45], showcased stronger malignancy and invasive potential. In contrast, in situ cancer-2, chosen as the starting region for trajectory inference, showed a relatively lower malignancy and a higher immune level, confirmed by overexpressed HLA family [46] genes and enriched immune cells. The trajectory unveiled a progression from in situ cancer-2 to in situ cancer-1, invasive cancer, and fibrous tissue near the tumor (Fig. 8D, Additional file 1: Fig. S24B), with genes exhibiting expression changes along the developmental trajectory (Fig. 8E). Notably, within the same area, tumor invasion exhibited inter-slice heterogeneity on the z-axis. For instance, in situ cancer-1 displayed a trend from H3 to H1, while in situ cancer-2 exhibited a trend from H1 and H3 to H2. Invasive cancer also followed a progression from H3 to H1 (Fig. 8F, Additional file 1: Fig. S24C).

To conclude, STAIR discerned tumor heterogeneity that cannot be identified by the other methods. The 3D reconstruction of tumor tissue enabled analyzing heterogeneity of invasion paths in a 3D view.

Assimilating new sections into a reference atlas with STAIR

Finally, we evaluated STAIR's capability to integrate new slices into an existing 3D atlas, which enhances the utility of reference atlases and allows for their expansion and updating.

We first evaluated STAIR's ability to predict the z-axis positions of new slices using slice-level attention scores (Methods). To this end, STAIR-Emb was first applied to integrate all slices, capturing higher-order attention scores across the dataset. We then conducted a leave-one-out cross-validation for z-axis prediction. Specifically, for each iteration, the true z-axis coordinate of a slice was sequentially masked and predicted using the z-coordinates of the remaining slices along with the corresponding slice–slice similarity measures relative to the masked slice. Attention-based predictions were compared with those obtained from simpler slice–slice similarity measures, including total gene expression correlation and cell type composition similarity. Predictions based on attention consistently achieved higher PCC and lower NMSE and MR than the simpler metrics (Additional file 1: Fig. S25). To further test robustness, we applied random cropping to the slices at 20%, 40%, and 60%, and repeated the z-axis prediction (Additional file 1: Fig. S26), demonstrating that attention-based z-axis prediction remains highly robust even when slices are partially overlapped.

Next, we evaluated the capability to integrate new slices into the existing 3D atlas. We integrated a brain slice from the Visium platform into a 3D ST-platform brain atlas [8], which had been aligned to the Allen mouse common coordinate framework (CCF) [47] 3D space through experimental information and image registration. We input 40 ST slices and the Visium slice into STAIR-Emb for inter-slice attention scoring (Fig. 9A) and seamlessly mixed their spatial embeddings (Fig. 9B). By predicting the AP axis coordinate of the Visium slice, we identified the nearest atlas slice, 20 A, in the 3D atlas. Using the 2D coordinates of 20 A as a reference template, STAIR-Loc performed scaling, rotation, and translation on the new Visium slice in the DV and ML axes, finally obtaining CCF 3D coordinates for each Visium spot (Fig. 9C). Additionally, leveraging detailed spatial anatomical region annotations from the ABA, we secured regional information for each Visium spot (Fig. 9D). The UMAP-based visualization revealed internal clustering of anatomical regions and separation between regions (Fig. 9E). We further demonstrated the accuracy of the regional annotation for the Visium slice based on the expression of specific markers, such as *Cabp7*, *Hpca1*, *Gpr88*, *Rora*, *Mbp* and *Pmch*, which exhibited high expression in hippocampus, hippocampal formation retrohippocampal region, striatum, thalamus, fiber tracts and hypothalamus, respectively. To summarize, STAIR enabled integrating new slices into a 3D atlas which might be generated from a different ST platform.

Discussion

The rapid advances in ST have ushered in new opportunities for exploring tissue architecture with gene expression patterns. However, connecting perspectives across discrete 2D slices to enable unambiguous 3D biological comprehension presented persistent

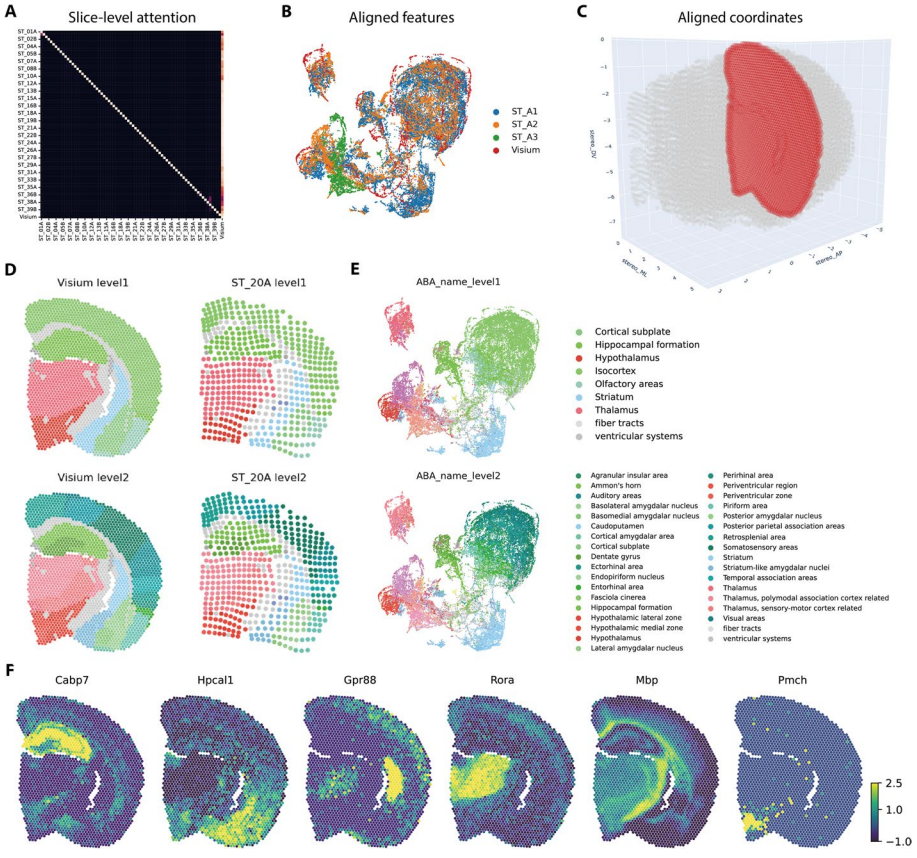

Fig. 9 Assimilating new sections into a reference atlas. **A** Heatmap depicting attention scores among 41 mouse brain slices, including 40 slices of the reference atlas from the ST platform, along with an additional slice generated from the Visium platform. **B** UMAP visualization of spatial embedding generated by STAIR, with colors indicating the respective sample of origin. **C** Visualization of the unified three-dimensional space after aligning the coordinates of the Visium slice with the 3D atlas. Spots from the Visium slice and ST slices are shown in red and gray respectively. **D** Spatial visualization of the Visium slice (left) and the ST slice (right) closest to Visium slice, with colors indicating the anatomical regions of the first (top) and the second (bottom) levels. **E** UMAP visualization of spatial embedding generated by STAIR, with colors indicating the anatomical regions of the first (top) and second (bottom) levels. **F** Spatial visualization of known regional marker genes

computational challenges. Here, we introduce STAIR, a unified framework for accurate spatial embedding and coordinate alignment, enabling pioneering 3D atlas construction from ST slices only.

The superior performance of STAIR derives from the attention mechanism of its heterogeneous graph attention network, which supports dual-level adaptive representation learning. By capturing both spot- and slice-level dependencies, the attention mechanism adaptively aggregates spatial features, allowing each spot to determine its affinity with others based on their spatial context. This results in compact embeddings for shared anatomical structures across slices while preserving distinct representations for slice-specific regions, such as the two AOB regions unique to the Slide-seqV2 mouse olfactory bulb dataset. Consequently, STAIR integrates cross-slice spatial features while retaining the biological heterogeneity of slice-specific regions. This adaptive representation learning further facilitates physical spatial alignment and reconstruction. The embeddings enable the identification of slice anchor pairs for initializing 2D alignments. Moreover, the attention mechanism is highly interpretable and versatile: high-order semantic modeling captures the physical relationships between slices and provides a novel z-axis

positioning strategy, supporting multi-sample 3D reconstruction and seamless integration of new slices.

In addition, STAIR mitigates inter-slice batch effects at the expression level before constructing and aligning spatial features. Expression profiles from different slices often exhibit systematic variations introduced by experimental or technical differences. Directly constructing spatial graphs from such uncorrected data can propagate batch noise during neighbor aggregation and distort downstream spatial alignment. By incorporating a batch-aware nonlinear embedding, STAIR harmonizes expression features across slices. This early correction substantially reduces technical variation, allowing the heterogeneous graph that integrates nodes from multiple slices to capture more reliable neighborhood relationships and produce more accurate spatial features.

We performed several ablation studies to validate the contributions of spot-level and slice-level attention mechanisms in spatial domain identification, spatial feature integration, and 2D alignment (Additional file 1: Fig. S4, Additional file 2: Note S1). Additionally, we quantified running times and GPU memory consumption to demonstrate STAIR's scalability on large datasets and systematically benchmarked the effects of spot number, gene number, and slice number on computational resource usage (Additional file 1: Fig. S17, Additional file 2: Note S2).

While most of the previous ST aligners focus on 2D alignment or integration, STAIR and STitch3D take multiple parallel ST slices as input and aim to create integrated 3D spatial atlases. However, STitch3D requires the z-axis coordinates of these parallel slices as prior information, whereas STAIR does not. This distinction is crucial because shared z-axis coordinates are often unavailable when slices originate from different samples. By inferring relative z-axis positions directly from the data, STAIR enables 3D reconstruction across multi-sample parallel slices without the need for predefined spatial coordinates. We employed three complementary metrics: PCC, RMSE, MR. PCC and RMSE assess the accuracy of continuous z-position estimation, while MR evaluates errors in discrete slice ordering. STAIR achieves extremely high correlation and very low RMSE when compared with ground-truth z-coordinates, even on the most challenging datasets such as the MERFISH coronal mouse brain and locally cropped scenarios. In contrast, MR is inherently more sensitive, particularly in densely sampled or locally cropped datasets, where adjacent slices exhibit high morphological and transcriptional similarity. In such cases, even minimal deviations in continuous space may result in larger apparent disordering. Additionally, STAIR-init and STAligner both selected spots pairs with similar features to guide physical alignment. However, STAligner constrains spot pairs in manually selected regions, which may not coincide with the regions that would yield optimal results. This approach also narrows the range of point pair selection, which may lead to some areas failing to find matching point pairs. For instance, among the 13 regions identified by STAlianger, only four regions could find matching point pairs. We found that even the optimal STAligner alignment result still exhibited weaknesses, which may be related to the limited effectiveness of STAligner's feature alignment.

Despite these advances, STAIR has several limitations that warrant further investigation. First, in 2D alignment, STAIR-Loc assumes that global tissue contours and strongly aggregated regions are preserved across consecutive slices. This assumption generally holds for whole organs or tissues prepared for 3D reconstruction (e.g., the mouse brain and embryo datasets described above). However, when slices are sampled from

manually dissected regions or when the tissue area exceeds the field of view, consistent contours may be absent. In such cases, STAIR may face challenges, highlighting the need for future methods capable of handling partial overlap or incomplete structural continuity. Second, in z-axis position reconstruction, STAIR assumes that slices are approximately parallel. While this simplification facilitates computational inference, it fails to capture common scenarios where sections are non-parallel or cut at varying angles during tissue preparation. Extending z-axis inference to accommodate non-parallel slices would therefore be highly valuable. Moreover, when the cutting angle is unknown, its estimation becomes critical for achieving accurate alignment, for example, in integrating sagittal and coronal ST slices to construct a dense 3D molecular atlas of the mouse brain. This estimation is particularly challenging when relying solely on the relatively sparse ST spots. In such cases, incorporating H&E-stained histological images, which preserve continuous tissue morphology, could offer a promising solution. Third, the reconstructed 3D atlases produced by current methods, including those generated by STAIR, may exhibit a slight straightening effect, known as the "banana slicing problem" [48]. This phenomenon arises because, without incorporating external geometric priors, a curved 3D structure cannot be fully recovered from a series of 2D cross-sectional slices. As STAIR currently does not assume any prior knowledge of global curvature, it approximates the overall structure as linear. Future extensions could integrate external or shape-based priors, such as tissue contours or MRI-derived reference geometries, to regularize global structure and recover the natural anatomical curvature of the reconstructed tissue. Finally, regarding the assimilation of new slices, integrating an additional slice currently requires retraining the model jointly with the new slice and the existing atlas, rather than allowing zero-shot generalization. While this design ensures accurate integration by leveraging global slice-to-slice relationships, it increases computational cost and constrains scalability in rapidly growing datasets. Future extensions may focus on enabling efficient zero-shot or incremental assimilation strategies, which would further enhance the practicality of STAIR for large-scale, dynamic 3D atlas construction.

Conclusions

In conclusion, STAIR provides a unified algorithmic approach that advances the analysis of multi-slice ST data, encompassing alignment, integration, and 3D reconstruction. This advancement enables the creation of comprehensive spatial atlases for a diverse array of organs, thereby facilitating the exploration of molecular mechanisms underlying tissue structure phenotypes within a 3D spatial context. Additionally, STAIR integrates additional slices into the existing 3D atlas, facilitating the expansion of the atlas and computationally enhancing its practical value.

Methods

STAIR algorithm

STAIR is designed to align spatial features and 3D coordinates across multiple slices. Assume there are S slices to be aligned, denoted as $O_1, O_2, \cdots, O_s$, where each slice contains $N_1, N_2, \cdots, N_s$ spots, respectively. Among these slices, there are G common genes. Consequently, the gene expression data can be represented as an $N \times G$ dimensional matrix X, where $N = N_1 + N_2 + \cdots + N_s$ is the total number of spots across all slices. The spatial coordinate data is stored in an $N \times 2$ or $N \times 3$ dimensional matrix Y,

depending on whether the z-dimension exists, and an $N \times 2$ coordinate matrix is sufficient as input for STAIR. The respective gene expression matrices and spatial coordinate matrices of $O_1, O_2, \cdots, O_s$ can be represented as $X_1, X_2, \cdots, X_s$ and $Y_1, Y_2, \cdots, Y_s$, respectively.

Data preprocessing

The raw gene expression counts in X are normalized by library size and then log-transformed to obtain the normalized expression matrix $\tilde{X}$. To remove batch effects in expression features across various slices, we employed an autoencoder alongside batch annotation [17] B. If the n th spot is from the slice O_s, we set $b_{ns} = 1$; otherwise, we set $b_{ns} = 0$. Given the assumptions of negative binomial (NB) or zero-inflated negative binomial (ZINB) distributions in gene expression, the preprocessing framework is:

$$Z = f_E\left(\tilde{X}||B\right) \tag{1}$$

$$Z' = f_{D1}\left(Z||B\right) \tag{2}$$

$$(\Pi,\ R,\ P)\ = f_{D2}\left(Z'\right) \tag{3}$$

where $\|$ is the concatenation operation. The encoder f_E, containing two layers, compresses the expression matrix into a m-dimensional low-dimensional feature matrix Z. The decoder consists of f_{D1} and f_{D2}. f_{D1} decodes Z and batch information into a $m\prime$ -dimensional feature matrix Z', followed by f_{D2} which decodes the parameters of the NB distribution (R, P) or the ZINB distribution (Π, R, P) based on Z'. The final expression features Z is obtained by minimizing the reconstruction loss between X and the NB/ZINB distribution parameterized by $(R, P)/(\Pi, \mathrm{R}, \mathrm{P})$. The loss function is the NB/ZINB-based negative log-likelihood.

Spatial embedding integration

The STAIR-Emb module uses a heterogeneous graph attention network [16] to integrate spatial embeddings across slices.

Construction of heterogeneous graph We construct a heterogeneous graph where all spots across slices serve as nodes, whose attributes are determined by their original slices. Different approaches are employed to establish edges between spots within the same slice and spots across different slices.

For spots within the same slice, we construct homogeneous edges based on their physical location. For a spot with index
i in slice O_I, the set of its intra-slice neighbors $\mathcal{N}_i$ is identified by K-Nearest Neighbor (KNN) based on spatial coordinates. The intra-slice adjacency relationships in O_I is stored in an adjacency matrix A_I.

For spots from different slices, since their relative physical coordinates are unknown, we construct heterogeneous edges based on the expression similarity obtained in the preprocessing step. Only nodes with highly consistent expression features are connected. For spots i from slice O_I and j from slice O_j, the edge exists if their expression

similarity $z_i \cdot z_j > \mathrm{t}$. We usually set $\mathrm{t} = 0.9$ and denote the inter-slice neighbors of spot i from slice O_J as $\mathcal{N}_i^{IJ}$.

Intra-slice information aggregation To capture the spatial features of each spot, we first aggregate information from intra-slice neighbors based on homogeneous edges independently for each slice. We employ two graph attention network (GAT) layers [49] to derive the intra-slice representation of spatial features. For slice O_I, its intra-slice representation U_I^{hom} can be expressed as:

$$U_I^{hom} = GAT_2\left(GAT_1\left(Z_I, A_I\right), A_I\right) \tag{4}$$

where Z_I represents the expression feature of slice O_I derived from the preprocessing step, while A_I denotes its corresponding adjacency matrix. Furthermore, the intra-slice representation of spot i is denoted as u_i^{hom}, and the collective intra-slice representation of all slices, consisting of $U_1^{hom}, U_2^{hom}, \cdots, U_S^{hom}$, is denoted as U^{hom}.

In each GAT layer, denoting the input feature as $N \times m_0$ dimensional matrix $H = (h_1, h_2, \ldots, h_N)$ and the output feature as $N \times m_0\prime$ dimensional matrix $H\prime = (h_1\prime, h_2\prime, \ldots, h_N\prime)$, then for spot i with input feature h_i, its output feature $h_i\prime$ can be formulated as:

$$h_i^{'} = ELU\left(\sum_{j \in N_i} a_{ij} W\, h_j\right) \tag{5}$$

where W is the $m_0\prime \times m_0$ dimensional weight matrix, α_{ij} is the normalized attention score using the SoftMax function:

$$\alpha_{ij} = Softmax_j\left(e_{ij}\right) = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(e_{ik})} \tag{6}$$

where $e_{ij} = a^T\left(Wh_i \parallel Wh_j\right)$, a is a learnable $2\, m_0\prime$ dimensional vector, and $\parallel$ denotes the concatenation operation.

Inter-slice information aggregation Further, we employ spot and slice level attention to aggregate inter-slice information adaptively.

First, we learn spot-level attention. For two spots i and j from slices O_I and O_J connected by heterogeneous edge, the corresponding spot-level attention e_{ij}^{IJ} represents the importance of spot j to i:

$$e_{ij}^{IJ} = LeakyReLU(q_{IJ}^T[W_{IJ}^1 u_i^{hom} || W_{IJ}^2 u_j^{hom}]) \tag{7}$$

where W_{IJ}^1 and W_{IJ}^2 are $m \times m$ dimensional weight matrices specific to slice pairs, q_{IJ} is a learnable $2m$ dimensional vector. The final inter-slice spot level attention α_{ij}^{IJ} is obtained by normalizing e_{ij}^{IJ}:

$$u_i^{IJ} = ReLU\left(\textstyle\sum_{j \in \mathcal{N}_i^{IJ}} \alpha_{ij}^{IJ} \cdot u_j^{hom}\right). \tag{8}$$

$$\alpha_{ij}^{IJ} = Softmax_j\left(e_{ij}^{IJ}\right) = \frac{\exp\left(e_{ij}^{IJ}\right)}{\sum_{k \in \mathcal{N}_i^{IJ}} \exp\left(e_{ik}^{IJ}\right)}. \tag{9}$$

For node i in slice O_I, its spot-level inter-slice representation based on neighbors in slice O_J is:

Then, we employ slice-level attention mechanisms to aggregate the spot-level inter-slice representations, yielding the final inter-slice representation. Different slices contribute unequally to target slice, and we adaptively learn how much each slice contributes to others. For the slice O_I to be learned, the slice-level importance of slice O_J to O_I is:

$$w_{IJ} = \frac{1}{N_I}\sum_{i\in O_I} q^T \cdot \tanh(W \cdot u_i^{IJ} + b), \tag{10}$$

where W and b are $m \times m$ dimensional weight matrix and m dimensional bias vector, q is a learnable m dimensional vector. Then, we normalize w_{IJ} to get the final slice level attention score:

$$\beta_{IJ} = \frac{\exp(w_{IJ})}{\sum_{K\neq I}\exp(w_{IK})}. \tag{11}$$

Therefore, the final inter-slice representation of slice O_I based on other slices is:

$$U_I^{het} = \sum_{J\neq I}\beta_{IJ}U^{IJ}. \tag{12}$$

The inter-slice representation of all slices consisting of $U_1^{het}, U_2^{het}, \cdots, U_S^{het}$ is denoted as U^{het}.

Model learning and training The final spatial feature matrix U is obtained by combining the intra-slice representation U^{hom} and inter-slice representation U^{het}:

$$U = \lambda \cdot U^{hom} + (1-\lambda)\cdot U^{het}, \tag{13}$$

where λ weights the homogeneous and heterogeneous components with default value 0.8. The model is trained to ensure U reliably represents the original spatial gene expression features Z. The objective function is defined as the mean squared error between U and Z:

$$Loss = MSE(U, Z) \tag{14}$$

In the training process, m is set to 32, m' is set to 128. The Adaptive Moment Estimation (Adam) optimizer is used for training with a learning rate of 0.001. The default number of iterations is 150. Additionally, the pre-processing module is trained independently, with a learning rate of 0.001 and 100 iterations.

Slice-level attention-based reconstruction and prediction in z-axis

The attention score at slice level, denoted as β_{IJ}, characterizes the collective impact of slice O_J on O_I within the spatial embedding learning process. This encapsulates intricate higher-order spatial semantic insights. Notably, the strong correlation between these scores and the actual inter-slice distances suggests the potential for inferring the physical positions of slices.

De novo z-axis reconstruction for multiple slices When lacking information about the third-dimensional coordinates $(y_{(3)}^1, \cdots, y_{(3)}^I, \cdots y_{(3)}^S)$ of the parallel slices

$O_1, O_2, \cdots, O_s$, we can reconstruct their positional relationships based on the inter-slice attention scores. Specifically, we start by defining the semantic distances between slices O_J and O_I as $d_{IJ} = 1 - \frac{\beta_{IJ}+\beta_{JI}}{2}$, and then construct a MST based on Kruskal algorithm [20] using *network* package implemented in Python. Subsequently, we select the given root slice O_I and set its position as $\widetilde{y}^I_{(3)} = 0$. The position of remaining nodes is then determined by adding the distance between each node and its parent node. Finally, the positions are normalized to the range of 0 to 1, yielding the reconstructed results $(\widetilde{y}^1_{(3)}, \widetilde{y}^2_{(3)} \cdots, \widetilde{y}^S_{(3)})$. Furthermore, for comparison with actual coordinates in the article, we additionally scale the normalized distances $(\widetilde{y}^1_{(3)}, \widetilde{y}^2_{(3)} \cdots, \widetilde{y}^S_{(3)})$ to match the real coordinate range.

Z-axis prediction for new slice When the z-axis coordinates $(y^1_{(3)}, \cdots, y^I_{(3)}, \cdots y^S_{(3)})$ are known, in order to align the new slice O_{new} into the unified 3D space, we predict $y^{new}_{(3)}$ based on the attention vectors between O_{new} and $O_1, O_2, \cdots, O_s$, denoted as $\beta_1 = (\beta_{new,1}, \beta_{new,2}, \cdots, \beta_{new,S})$ and $\beta_2 = (\beta_{1,new}, \beta_{2,new}, \cdots, \beta_{S,new})$.
We first sort the elements of β_1 and β_2 in descending order, and choose the top M elements $\widehat{\beta}_1 = (\beta_{new,R^1_1}, \cdots, \beta_{new,R^1_M})$ and $\widehat{\beta}_2 = (\beta_{R^2_1,new}, \cdots, \beta_{R^2_M,new})$. Following this, we identify the intersection of these element indices $\{R_1, \cdots, R_{M\prime}\}$, where $M\prime \leq M$. Finally, we perform weighted average on the z-axis corresponding to $R_1, \cdots, R_{M\prime}$ to get the prediction of $y^{new}_{(3)}$:

$$\widehat{y}^{new}_{(3)} = \frac{\sum_{K\in\{R_1,\cdots,R_{M\prime}\}} \frac{\beta_{new,K}+\beta_{K,new}}{2} \cdot y^K_{(3)}}{\sum_{K\in\{R_1,\cdots,R_{M\prime}\}} \frac{\beta_{new,K}+\beta_{K,new}}{2}} \tag{15}$$

Spatial location alignment in x-axis and y-axis

The STAIR-Loc module implements a two-stage approach for 2D alignment along the x-axis and y-axis. In the initial stage, spatial features are leveraged to identify precisely matched spot pairs, from which the initial transformation matrix is derived. The fine alignment stage begins by identifying informative spots that effectively capture both global and local information within slices. Subsequently, the ICP algorithm [19] is applied to these spots to obtain the final alignment results.

Initial alignment Consider slice O_I with N_I spots and slice O_J with N_J spots, along with their respective spatial feature sets $\left\{u^I_1, u^I_2, \cdots, u^I_{N_I}\right\}$ and $\left\{u^J_1, u^J_2, \cdots, u^J_{N_J}\right\}$. First, we measure the cosine similarity of spatial embeddings between spots from different slices. The mutual nearest neighbors (MNN) of $k = 1$ located in different slices form pairs for the initial alignment, resulting in curated pairs $\{(i_1, j_1), (i_2, j_2), \cdots, (i_n, j_n)\}$ along with their 2D coordinates $\{\left(y^I_{i_1}, y^J_{j_1}\right), \left(y^I_{i_2}, y^J_{j_2}\right), \cdots, (y^I_{i_n}, y^J_{j_n})\}$. If the slices are from different platforms, scaling may be necessary. We use the distance between $\{i_1, i_2, \cdots, i_n\}$ and that between $\{j_1, j_2, \cdots, j_n\}$ to determine the scaling factor:

$$scale = Median_{k_1\neq k_2} \left\{ \frac{\left| y^J_{j_{k_1}} - y^J_{j_{k_2}} \right|}{\left| y^I_{i_{k_1}} - y^I_{i_{k_2}} \right|} \right\}. \tag{16}$$

Subsequently, we minimize the 2D distance of each pair to determine the optimal initial rotation matrix R^0 and translation vector t^0:

$$Loss_{init} = \sum_{k=1}^{n} \left| y_{i_k}^I - \left(R y_{j_k}^J + t \right) \right|^2. \tag{17}$$

Singular value decomposition [50] solves for R^0 and t^0. All spots are then aligned to their targets accordingly.

Fine alignment The fine alignment stage involves three main steps. First, we select highly concentrated regions that are common to both slices, according to median of local inverse simpson's index (LISI) value which implemented in R package *lisi*. Subsequently, we identify the concave hulls [51] corresponding to the chosen regions and the slices. These concave hulls serve to depict both the local attributes of the region and the broader shape characteristics of the entire slice. Finally, ICP [19] algorithm is employed on these sets of informative spots to achieve precise fine registration, resulting in **o**ur desired alignment outcome.

Aligning new brain slices into existing ABA atlases

We employed a dataset consisting of 40 ST coronal brain slices with CCF 3D coordinate information, derived from a previous study [8], to serve as a foundational reference of whole-brain framework. To obtain the CCF 3D coordinates and anatomical region annotations of each spot in new slice, we incorporated it into the reference dataset through the following steps.

First, we integrate the spatial embeddings of the new slice O_{new} with the reference set $\{O_1, O_2, \cdots, O_{40}\}$, obtaining the attention scores associated with the O_{new} and each reference slice. By applying formula (15), we derive the predicted AP coordinate of O_{new}, denoted as y_{AP}^{new}. Then, we perform two-stage alignment with scaling to align the coordinates in ML and DV plane (see Spatial location alignment), effectively integrating the 3D coordinates of O_{new} into the CCF. Finally, anatomical regions are assigned to each new spot based on location-specific information within the ABA annotation file *annotation_25.nrrd* from https://portal.brain-map.org/, establishing a mapping of the anatomical context for slice O_{new}.

Toolkit for ST analysis

Clustering for spatial domains

Clustering is conducted on integrated spatial embeddings to obtain unified spatial domains across ST slices. We employ mClust [52] clustering method implemented in R package rmclust v5.4.9.

Spatial trajectory inference

Trajectory inference based on spatial embeddings is used to track the development in spatial dimension. We utilize Monocle3 [43] v1.0.0 to perform the pseudo time inference for each spot based on UMAP derived from spatial embedding of STAIR by applying *learn_graph* and *order_cells* in package. Function *graph_test* is employed to find genes that change with pseudo time.

Deconvolution

For HER2 + breast cancer dataset, we perform deconvolution to analyze the cell type composition using software Cell2location [41] v0.1.3 and annotated scRNA-seq dataset [42]. Cell2location is run according to the tutorial and default parameters.

Differential expression analysis

We employ Seurat [53] v4.2.1 to perform differential expression analysis. Differential expression analysis is used to identify cluster-specific marker genes where all the clusters are pairwise compared using the Wilcoxon method. Each identified marker gene was expressed in a minimum of 25% of cells and at a minimum log fold change threshold of 0.25.

Evaluation

Evaluation of spatial embedding alignment

Adjusted rand index (ARI) ARI measures the consistency between spatial domains identified by different algorithms and the known anatomical region labels. Given the contingency table of intersections between the algorithm-generated domains and annotation-based labels, it is calculated as:

$$ARI = \frac{\sum_{ij}\binom{n_{ij}}{2} - [\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}]/\binom{n}{2}}{\frac{1}{2}[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}] - [\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}]/\binom{n}{2}}, \tag{18}$$

where n_{ij}, a_i and b_i are values from the contingency table. ARI values range from −1 to 1. Higher ARI indicates greater agreement with the ground truth annotations.

Average silhouette width (ASW) ASW evaluates how well the features match true clusters in the data. For every spot i, Silhouette width $S(i)$ is calculated as:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{19}$$

where $a(i)$ is the average distance between i and spots in its own cluster, and $b(i)$ is that to adjacent cluster spots. ASW values range from −1 to 1, with greater ASW indicates better match.

In this paper, we employ ASW_{domain} and ASW_{batch} to measure the fitness of spatial features to known spatial regions and batches of slices, respectively. A larger ASW_{domain} and a smaller ASW_{batch} value represent stronger feature learning and integration capabilities. To evaluate the comprehensive performance of spatial features, we calculated the harmonic mean ASW_{F1} of ASW_{domain} and $1 - ASW_{batch}$:

$$ASW_{F1} = \frac{2(1 - ASW_{batch})ASW_{domain}}{1 - ASW_{batch} + ASW_{domain}}. \tag{20}$$

Higher ASW_{F1} indicates stronger comprehensive ability of feature integration and biological specificity retention.

Evaluation of spatial alignment in x-axis and y-axis

To assess STAIR's capability for spatial alignment, we conducted simulations involving multiple slices, each with known 3D coordinates. Specifically, we maintained

the spatial location of the first slice unchanged, while applying random rotations and translations to the remaining slices. Rotation angle $\theta \in (0, \pi)$, translation distance $t_{(1)} \in (-a_{(1)}, a_{(1)}), t_{(2)} \in (-a_{(2)}, a_{(2)})$, where $a_{(1)} = \max(y_{(1)}) - \min(y_{(1)})$, $a_{(2)} = \max(y_{(2)}) - \min(y_{(2)})$. Note that the ground truth rotation angle and translation distance of $S-1$ slices are $\left\{\theta^2, \theta^3, \cdots, \theta^S\right\}$ and $\{(t^2_{(1)}, t^2_{(2)}), (t^3_{(1)}, t^3_{(2)}), \cdots, (t^S_{(1)}, t^S_{(2)})\}$, respectively. The error of rotation and translation are:

$$\Delta_\theta = \frac{1}{S-1} \sum_{I=2}^{S} \left| \theta^I - \widehat{\theta}^I \right| \tag{21}$$

$$\Delta_t = \frac{1}{S-1} \sum_{I=2}^{S} \sqrt{\left(t^I_{(1)} - \widehat{t}^I_{(1)}\right)^2 + \left(t^I_{(2)} - \widehat{t}^I_{(2)}\right)^2} \tag{22}$$

where $\widehat{\theta}^I$ and $(\widehat{t}^I_{(1)}, \widehat{t}^I_{(2)})$ are rotation angle and translation distance of slice A_I obtained by the algorithm to be evaluated, respectively.

Evaluation of spatial domain identification and 2D coordinate alignment in HER2 + breast cancer

LISI measures the degree of local mixing, and we use it to evaluate the spatial aggregation pattern of domains in stacked 2D space. For each spot i with spatial domain label , LISI value of this spot is formulated as:

$$LISI\,(i) = \frac{1}{\sum_{l \in L} p_i(l)}, \tag{23}$$

where $p_i\,(l)$ is the probability that the spatial domain label exists in the local neighborhood of sample , and L is the set of spatial domains. Local neighborhoods are selected by stacked 2D coordinates derived by 2D alignment methods. The value of LISI is in the range of $[1, \infty)$, and smaller LISI indicates better aggregation pattern.

Evaluation of z-axis reconstruction and prediction

We employed three metrics to evaluate the performance of z-axis reconstruction and prediction: the Pearson correlation coefficient (PCC), normalized mean squared error (NMSE), and misordering rate (MR). Denoting the ground truth z-axis coordinate of the S slices as $\{\mathrm{y}^1_{(3)}, \mathrm{y}^2_{(3)}, \cdots, \mathrm{y}^{\mathrm{S}}_{(3)}\}$, and the reconstructed or predicted outcomes are $\{\widehat{\mathrm{y}}^1_{(3)}, \widehat{\mathrm{y}}^2_{(3)}, \cdots, \widehat{\mathrm{y}}^{\mathrm{S}}_{(3)}\}$. The PCC and NMSE values are calculated as:

$$PCC = \frac{\sum_{I=1}^{S} \left(y^I_{(3)} - \overline{y}_{(3)}\right)\left(\widehat{y}^I_{(3)} - \overline{\widehat{y}}_{(3)}\right)}{\sqrt{\sum_{I=1}^{S} \left(y^I_{(3)} - \overline{y}_{(3)}\right)^2} \sqrt{\sum_{I=1}^{S} \left(\widehat{y}^I_{(3)} - \overline{\widehat{y}}_{(3)}\right)^2}} \tag{24}$$

$$NMSE = \frac{\sum_{I=1}^{S} \left(y^I_{(3)} - \widehat{y}^I_{(3)}\right)^2}{\sum_{I=1}^{S} \left(y^I_{(3)} - \overline{y}_{(3)}\right)^2} \tag{25}$$

where $\overline{y}_{(3)} = \frac{1}{S} \sum_{I=1}^{S} y^I_{(3)}$ *and* $\overline{\widehat{y}}_{(3)} = \frac{1}{S} \sum_{I=1}^{S} \widehat{y}^I_{(3)}$.

For the MR, we first sort slices according to their true coordinates $\{y^I_{(3)}\}$ and reconstructed or predicted coordinates $\{\widehat{y}^I_{(3)}\}$. Let $\pi\,(I)$ and $\widehat{\pi}\,(I)$ denote the ranking

positions of slice I in the true and predicted orders, respectively. The MR is then defined as the proportion of slices with mismatched ranks:

$$MR = \frac{1}{S}\sum_{I=1}^{S}\mathbb{I}\left(\pi\left(I\right) \neq \widehat{\pi}\left(I\right)\right) \tag{26}$$

where $\mathbb{I}\left(\cdot\right)$ is the indicator function.

Assessment of alternative methods

We conducted a comparative evaluation of STAIR against other alignment methods, including STAligner, Stitch3D, PRECAST, and PASTE. In our assessment, we employed the default parameters for all methods unless specific parameters were outlined in the original text or tutorial.

STAligner STAligner integrates ST data across different conditions, technologies, and developmental stages. It employs STAGATE and triplet loss to integrate the ST datasets until batch-corrected embeddings are generated. It further considers shared spatial domain and MNNs identified by STAligner as corresponding pairs to guide the 2D alignment. We downloaded the package STAligner v1.0.0 from https://github.com/zhoux85/STAligner, and ran STAligner following its tutorial https://staligner.readthedocs.io/en/latest/index.html.

STitch3D STitch3D first unified 3D spatial coordinates for spots using ICP or PASTE, followed by graph construction based on 3D coordinates. It performed spatial embedding learning and integration by graph attention network and slice- and gene-specific parameters. We downloaded the package Stitch3D v1.0.3 from https://github.com/YangLabHKUST/STitch3D, and ran STitch3D following https://stitch3d-tutorial.readthedocs.io/en/latest/tutorials/index.html. Given the requirement for scRNA-seq datasets from the same tissue, we utilized single cell DLPFC and HER2 + breast cancer data accessed at the Gene Expression Omnibus (GEO) under the accession code GSE144136 [54] and GSE176078 [55], respectively.

PRECAST PRECAST is a probabilistic method for spatial embedding learning, clustering, and alignment. We downloaded the R package PRECAST v1.6.1 from https://github.com/feiyoung/PRECAST/, and ran PRECAST following its tutorial https://feiyoung.github.io/PRECAST/index.html

PASTE PASTE provides the flexibility to align two slices either through pairwise alignment or to simultaneously align multiple slices using center alignment. We opted for pairwise alignment in our testing, as other approaches also employ pairwise processes. In pairwise slice alignment, it aims to find the best possible way to connect spots in one slice with spots in another slice, followed by constructing a stacked 3D alignment of a tissue. The connection, denoted as Π, is chosen to reduce both the differences in gene expression patterns between connected spots from different slices and the differences in physical distances between connected spots within the same slice. Parameter α was used to balance these two differences, and we set its default value of $\alpha = 0.1$ in our test. We download the package PASTE v1.3.0 from https://github.com/raphael-group/paste/tree/

main, and ran PASTE following https://github.com/raphael-group/paste/tree/main/docs/source/notebooks.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03895-x.

Additional file 1.

Additional file 2.

Acknowledgements
We would like to thank all the data and software contributors who make this research possible. We also thank Zhongshan Ophthalmic Center and the Center for Precision Medicine at Sun Yat-sen University for the long-term support.

Peer review information
Andrew Cosgrove and Claudia Feng were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
Z.X. conceived and supervised the study. YY.Y. and Z.X. designed the study. YY.Y. analyzed the data. YY.Y. and Z.X. wrote the manuscript. All authors read and approved the manuscript.

Funding
Funding was provided by National Natural Science Foundation of China (32470705), Key-Area Research and Development Program of Guangdong Province (2023B1111020006), Science and Technology Program of Guangzhou, China (2025A03J3990).

Data availability
All datasets analyzed in this study are publicly available. The Visium DLPFC dataset was obtained from https://github.com/LieberInstitute/spatialDLPFC [56], and the MERFISH hypothalamic preoptic region dataset is hosted on Dryad at 10.5061/dryad.8t8s248 [57]. Olfactory bulb datasets generated by Stereo-seq and are available at https://github.com/JinmiaoChenLab/SEDR_analyses [58] and https://singlecell.broadinstitute.org/ [59]. Mouse embryo datasets produced using Stereo-seq, Slide-seqV2, and seqFISH are available fromhttps://db.cngb.org/stomics/ [60], GEO accession GSE197353 [61], and https://crukci.shinyapps.io/SpatialMouseAtlas/ [62]. The ST mouse brain atlas is accessible at http://molecularatlas.org/ [63], and the corresponding aligned Visium dataset is provided by 10 × Genomics at https://www.10xgenomics.com/resources/datasets/mouse-brain-coronal-section-1-ffpe-2-standard [64]. The HER2 + breast cancer ST dataset is available at 10.5281/zenodo.4751624 [65], and the corresponding single-cell dataset can be accessed under GEO accession GSE176078 [55]. MERFISH whole mouse brain datasets are archived in the Brain Image Library at 10.35077/act-bag [66]. STAIR is publicly available as an open-source Python package at https://github.com/yuyuanyuana/STAIR [67] and is released under the MIT License. The source code and datasets are deposited in Zenodo at https://zenodo.org/records/17672066 [68].

Declarations

Ethics approval and consent to participate
No ethnical approval was required for this study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 July 2024 / Accepted: 1 December 2025
Published online: 15 December 2025

References

1. Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353:78–82.
2. Moffitt JR, Bambah-Mukku D, Eichhorn SW, Vaughn E, Shekhar K, Perez JD, et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science. 2018. https://doi.org/10.1126/science.aau5324.
3. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019;363:1463–7.
4. Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, et al. Highly sensitive spatial transcriptomics at near-cellular resolution with slide-seqV2. Nat Biotechnol. 2021;39:313–9.
5. Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature. 2019;568:235–9.

6. Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science. 2018;361:eaat5691.
7. Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell. 2022;185(1777–1792):e1721.
8. Ortiz C, Navarro JF, Jurek A, Martin A, Lundeberg J, Meletis K. Molecular atlas of the adult mouse brain. Sci Adv. 2020;6:eabb3446.
9. Wang M, Hu Q, Lv T, Wang Y, Lan Q, Xiang R, et al. High-resolution 3D spatiotemporal transcriptomic maps of developing Drosophila embryos and larvae. Dev Cell. 2022;57(1271–1283):e1274.
10. Chen A, Sun Y, Lei Y, Li C, Liao S, Meng J, et al. Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. Cell. 2023;186(3726–3743):e3724.
11. Zhang M, Pan X, Jung W, Halpern AR, Eichhorn SW, Lei Z, et al. Molecularly defined and spatially resolved cell atlas of the whole mouse brain. Nature. 2023;624:343–54.
12. Liu W, Liao X, Luo Z, Yang Y, Lau MC, Jiao Y, et al. Probabilistic embedding, clustering, and alignment for integrating spatial transcriptomics data with PRECAST. Nat Commun. 2023;14:296.
13. Zeira R, Land M, Strzalkowski A, Raphael BJ. Alignment and integration of spatial transcriptomics data. Nat Methods. 2022;19:567–75.
14. Zhou X, Dong K, Zhang S. Integrating spatial transcriptomics data across different conditions, technologies and developmental stages. Nat Comput Sci. 2023;3:894–906.
15. Wang G, Zhao J, Yan Y, Wang Y, Wu AR, Yang C. Construction of a 3D whole organism spatial atlas by joint modelling of multiple slices with deep neural networks. Nat Mach Intell. 2023;5:1200–13.
16. Wang X, Ji H, Shi C, Wang B, Ye Y, Cui P, Yu PS: Heterogeneous Graph Attention Network. In *The World Wide Web Conference*. pp. 2022–2032. San Francisco, CA, USA: Association for Computing Machinery; 2019:2022–2032.
17. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. Nat Methods. 2018;15:1053–8.
18. Eraslan G, Simon LM, Mircea M, Mueller NS, Theis FJ. Single-cell RNA-seq denoising using a deep count autoencoder. Nat Commun. 2019;10:390.
19. Besl PJ, McKay ND. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 1992;14:239–56.
20. Kruskal JB. On the shortest spanning subtree of a graph and the traveling salesman problem. Proc Am Math Soc. 1956;7:48–50.
21. Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat Neurosci. 2021;24:425–36.
22. McInnes L, Healy J, Melville J: UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. pp. arXiv:1802.03426; 2018:arXiv:1802.03426.
23. Fu H, Xu H, Chong K, Li M, Ang KS, Lee HK, Ling J, Chen A, Shao L, Liu L, Chen J: Unsupervised Spatially Embedded Deep Representation of Spatial Transcriptomics. bioRxiv 2021:2021.2006.2015.448542.
24. Sampath Kumar A, Tian L, Bolondi A, Hernández AA, Stickels R, Kretzmer H, et al. Spatiotemporal transcriptomic maps of whole mouse embryos at the onset of organogenesis. Nat Genet. 2023;55:1176–85.
25. Lohoff T, Ghazanfar S, Missarova A, Koulena N, Pierson N, Griffiths JA, et al. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. Nat Biotechnol. 2022;40:74–85.
26. Porter FD, Drago J, Xu Y, Cheema SS, Wassif C, Huang SP, et al. Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. Development. 1997;124:2935–44.
27. Kurokawa D, Ohmura T, Sakurai Y, Inoue K, Suda Y, Aizawa S. Otx2 expression in anterior neuroectoderm and forebrain/midbrain is directed by more than six enhancers. Dev Biol. 2014;387:203–13.
28. Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development. 2005;132:2215–23.
29. Fazilaty H, Rago L, Kass Youssef K, Ocana OH, Garcia-Asencio F, Arcas A, et al. A gene regulatory network to control EMT programs in development and disease. Nat Commun. 2019;10:5115.
30. Elefanty AG, Begley CG, Hartley L, Papaevangeliou B, Robb L. Scl expression in the mouse embryo detected with a targeted lacZ reporter gene demonstrates its localization to hematopoietic, vascular, and neural tissues. Blood. 1999;94:3754–63.
31. Mok G, Turner S, Smith E, Mincarelli L, Lister A, Lipscombe J, Uzun V, Haerty W, Macaulay I, Münsterberg A: Single-cell and spatial transcriptomics of the avian embryo tailbud. bioRxiv 2024:2024.2002.2005.578917.
32. Alber AB, Marquez HA, Ma L, Kwong G, Thapa BR, Villacorta-Martin C, et al. Directed differentiation of mouse pluripotent stem cells into functional lung-specific mesenchyme. Nat Commun. 2023;14:3488.
33. Gopalakrishnan S, Comai G, Sambasivan R, Francou A, Kelly RG, Tajbakhsh S. A cranial mesoderm origin for esophagus striated muscles. Dev Cell. 2015;34:694–704.
34. Hagihara H, Fujita M, Umemori J, Hashimoto M, Miyakawa T. Immature-like molecular expression patterns in the hippocampus of a mouse model of dementia with Lewy body-linked mutant beta-synuclein. Mol Brain. 2018;11:38.
35. Spark DL, Vermeulen MH, de la Fuente Gonzalez RA, Hatzipantelis CJ, Rueda P, Sepehrizadeh T, et al. Gpr88 deletion impacts motivational control without overt disruptions to striatal dopamine. Biol Psychiatry Glob Open Sci. 2023;3:1053–61.
36. Laboute T, Gandia J, Pellissier LP, Corde Y, Rebeillard F, Gallo M, et al. The orphan receptor GPR88 blunts the signaling of opioid receptors and multiple striatal GPCRs. Elife. 2020. https://doi.org/10.7554/eLife.50519.
37. Wende B, Beyer AL, Ruhnke N, Kaemmerer D, Sanger J, Schulz S, et al. Expression of the calcitonin receptor-like receptor (CALCRL) in normal and neoplastic tissues. Int J Mol Sci. 2023. https://doi.org/10.3390/ijms24043960.
38. Oliver KR, Kane SA, Salvatore CA, Mallee JJ, Kinsey AM, Koblan KS, et al. Cloning, characterization and central nervous system distribution of receptor activity modifying proteins in the rat. Eur J Neurosci. 2001;14:618–28.
39. Wei JR, Hao ZZ, Xu C, Huang M, Tang L, Xu N, et al. Identification of visual cortex cell types and species differences using single-cell RNA sequencing. Nat Commun. 2022;13:6902.
40. Andersson A, Larsson L, Stenbeck L, Salmen F, Ehinger A, Wu SZ, et al. Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. Nat Commun. 2021;12:6012.

41. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell 2location maps fine-grained cell types in spatial transcriptomics. Nat Biotechnol. 2022;40:661–71.
42. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. Nat Genet. 2021;53:1334–47.
43. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature. 2019;566:496–502.
44. Tan M, Yu D. Molecular mechanisms of erbB2-mediated breast cancer chemoresistance. Adv Exp Med Biol. 2007;608:119–29.
45. Chen H, Xu C, Jin Q, Liu Z. S100 protein family in human cancer. Am J Cancer Res. 2014;4:89–115.
46. Kaneko K, Ishigami S, Kijima Y, Funasako Y, Hirata M, Okumura H, et al. Clinical implication of HLA class I expression in breast cancer. BMC Cancer. 2011;11:454.
47. Wang Q, Ding SL, Li Y, Royall J, Feng D, Lesnar P, et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell. 2020;181(936–953):e920.
48. Malandain G, Bardinet E, Nelissen K, Vanduffel W. Fusion of autoradiographs with an MR volume using 2-D and 3-D linear transformations. Neuroimage. 2004;23:111–27.
49. Yu Y, He Y, Xie Z: SECE: accurate identification of spatial domain by incorporating global spatial proximity and local expression proximity. bioRxiv. 2023:2023.2012.2026.573377.
50. Arun KS, Huang TS, Blostein SD. Least-squares fitting of two 3-D point sets. IEEE Trans Pattern Anal Mach Intell. 1987;PAMI-9:698–700.
51. Edelsbrunner H, Mücke EP. Three-dimensional alpha shapes. ACM Trans Graph. 1994;13:43–72.
52. Scrucca L, Fop M, Murphy TB, Raftery AE. Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R j. 2016;8:289–317.
53. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(3573–3587):e3529.
54. Turecki G: Single-nucleus RNA-seq in the post-mortem brain in major depressive disorder. Datasets. Gene Expression Omnibus. 2020. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144136.
55. Swarbrick A, Wu S, Al-Eryani G, Roden D: A single-cell and spatially resolved atlas of human breast cancers. Datasets. Gene Expression Omnibus. 2021. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176078.
56. Collado-Torres L: LieberInstitute/spatialLIBD: spatialLIBD: initial bioconductor submission. Datasets. Zenodo. 2020. https://zenodo.org/records/4730634.
57. Moffitt JR, Bambah-Mukku D, Eichhorn SW, Vaughn E, Shekhar K, Perez JD, et al. Data from: Molecular, spatial and functional single-cell profiling of the hypothalamic preoptic region. Datasets Dryad. 2018. https://doi.org/10.5061/dryad.8t8s248.
58. H. X: Data for SEDR analysis. Datasets. Github. 2023. https://github.com/JinmiaoChenLab/SEDR_analyses.
59. Stickels M: Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Datasets. Single cell portal. 2020. https://singlecell.broadinstitute.org/single_cell/study/SCP815/highly-sensitive-spatial-transcriptomics-at-near-cellular-resolution-with-slide-seqv2.
60. Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, Qiu X, Yang J, Xu J, Hao S, et al: MOSTA: Mouse Organogenesis Spatiotemporal Transcriptomic Atlas. Datasets. STOmicsDB. 2021. https://db.cngb.org/stomics/datasets/STDS0000058.
61. A SK, L T, A M, F C: Spatial transcriptomic maps of whole mouse embryos. Datasets. Gene Expression Omnibus. 2022. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197353.
62. Lohoff T, Ghazanfar S, Missarova A, Koulena N, Pierson N, Griffiths JA, Bardot ES, Eng CL, Tyser RCV, Argelaguet R, et al: Highly multiplexed spatially resolved gene expression profiling of mouse organogenesis. Datasets. 2022. https://crukci.shinyapps.io/SpatialMouseAtlas/.
63. Ortiz C, Navarro JF, Jurek A, Martin A, Lundeberg J, Meletis K: A Molecular Atlas of the Adult Mouse Brain. Datasets. 2020. https://molecularatlas.org/
64. Mouse Brain Coronal Section "Background" (FFPE). Datasets. *10X Genomics*. 2022. https://www.10xgenomics.com/datasets/mouse-brain-coronal-section-1-ffpe-2-standard.
65. Andersson A: Spatial deconvolution of HER2-positive Breast cancer delineates tumor-associated cell type interactions. Datasets. Zenodo. https://zenodo.org/records/4751624.
66. Zhuang X, Jung W, Zhang M: A molecularly defined and spatially resolved cell atlas of the whole mouse brain. Datasets. Brain Image Library. 2023. https://doi.brainimagelibrary.org/doi/https://doi.org/10.35077/act-bag.
67. Yu Y: Code for STAIR: spatial transcriptomic alignment, integration, and 3D reconstruction. GitHub. 2025. https://github.com/yuyuanyuana/STAIR.
68. Yu Y: Data for STAIR: spatial transcriptomic alignment, integration, and 3D reconstruction. Zenodo. 2024. https://zenodo.org/records/17672066.

Publisher's Note