communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-08015-w

Spatially informed graph transformers for spatially resolved transcriptomics

Xinyu Bao[1], Xiaosheng Bai[1], Xiaoping Liu[1] ✉, Qianqian Shi[2,3] ✉ & Chuanchao Zhang[1] ✉

Spatially resolved transcriptomics (SRT) has emerged as a powerful technique for mapping gene expression landscapes within spatial contexts. However, significant challenges persist in effectively integrating gene expression with spatial information to elucidate the heterogeneity of biological tissues. Here, we present a **Spa**tially informed **G**raph **T**ransformers framework, SpaGT, which leverages both node and edge channels to model spatially aware graph representation for denoising gene expression and identifying spatial domains. Unlike conventional graph neural networks, which rely on static, localized convolutional aggregation, SpaGT employs a structure-reinforced self-attention mechanism that iteratively evolves topological structural information and transcriptional signal representation. By replacing graph convolution with global self-attention, SpaGT enables the integration of both global and spatially localized information, thereby improving the detection of fine-grained spatial domains. We demonstrate that SpaGT achieves superior performance in identifying spatial domains and denoising gene expression data across diverse platforms and species. Furthermore, SpaGT facilitates the discovery of spatially variable genes with significant prognostic potential in cancer tissues. These findings establish SpaGT as a powerful tool for unraveling the complexities of biological tissues.

Spatially resolved transcriptomics (SRT) techniques have revolutionized the ability to perform high-throughput sequencing of mRNA while preserving spatial information in multicellular organisms[1–3]. Techniques such as 10x Visium, Slide-seqV2[4], and Stereo-seq[5] have unlocked the potential to explore cellular heterogeneity within tissues, facilitating the discovery of novel cell types and spatially functional regions that remain undetectable through bulk RNA sequencing. However, SRT faces challenges such as data sparsity and noise, which result from various technical limitations, including transcript capture rates and spatial resolution. These limitations pose significant obstacles to the effective representation of SRT data[6–11]. Notably, neighboring locations on tissue sections often share cellular microenvironments and exhibit similar gene expression profiles. Therefore, the integration of locational information with gene expression data can help mitigate noise and bias, enhancing pattern recognition in SRT studies.

In recent years, several computational methods have been developed to integrate gene expression data with spatial information, aiming to cluster spots or nodes with similar gene expression patterns and spatial continuity to detect spatial domains. These spatial domains represent regions with shared biological functions, making them critical for subsequent analyses, such as the identification of spatially variable genes (SVGs) and the discovery of prognostic genes in cancer tissues. Existing computational approaches for detecting spatial domains in SRT data can be broadly categorized into probabilistic models and deep learning models. Probabilistic approaches, such as Hidden Markov Random Fields (HMRF)[12], BayesSpace[13], and SpatialPCA[14], employ specialized probabilistic graphical models designed to integrate spatial information, effectively. In contrast, deep learning approaches, including SpaGCN[15], STAGATE[16], DeepST[17] and GraphST[18], leverage graph convolutional networks and their derivatives to capture dependencies informed by spatial data. These two classes of models address the challenge of modeling spatial domains through distinct methodologies, emphasizing either statistical probabilities or the learning capabilities of neural networks to understand the spatial context embedded in the data. However, existing models may not fully capture the complex and non-linear relationships between gene expressions and their spatial contexts, as they often rely on predefined graph structures that may inadequately represent actual biological interactions[19,20]. Therefore, effectively integrating gene expression data with spatial information to explore expression landscape within tissues remains a significant challenge.

[1]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. [2]Hubei Engineering Technology Research Center of Agricultural Big Data, Huazhong Agricultural University, Wuhan, China. [3]Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan, China. ✉e-mail: xpliu@ucas.ac.cn; qqshi@mail.hzau.edu.cn; chuanchaozhang@ucas.ac.cn

To address these challenges, we propose SpaGT, a spatially informed graph transformer designed to model spatially aware graph representation, denoising gene expression and revealing spatially functional regions by utilizing both expression channels and edge channels. Specifically, SpaGT first converts spatial multimodal data into graph-structure data and subsequently generates expression embeddings and edge embeddings that reflect pairwise structural information between spots. These embeddings serve as input data for transformers, which construct the expression and edge channels to model spatially aware graph representation. SpaGT's structure-reinforced self-attention mechanism iteratively evolves the topological structure and transcriptional signal representation of the spatial graph, enabling the effective integration of both global and spatially localized information, thus serving as an alternative to graph convolution in graph neural networks. By integrating global and localized information from expression data, SpaGT corrects low-quality gene expression and improves the detection of fine-grained spatial domains.

Our comprehensive evaluation demonstrates the superiority of SpaGT by benchmarking it against seven state-of-the-art methods using 17 SRT datasets generated by 10x Visium, Slide-seqV2, Stereo-seq, osmFISH, Seq-Scope and ST technologies. For 12 Dorsolateral Prefrontal Cortex (DLPFC) datasets from 10x Visium, as well as single-cell resolution SRT data from osmFISH and Seq-Scope platforms, SpaGT consistently outperforms other methods in identifying spatial domains, achieving higher accuracy compared to known ground-truth annotations. When applied to triple-negative breast cancer SRT data, SpaGT excels in deciphering tumor heterogeneity and provides deeper biological insights into genes closely associated with cancer. Its robustness is further validated through survival analysis using independent clinical data. Moreover, when applied to mouse hippocampus data from Slide-seqV2 and mouse embryo data from Stereo-seq, SpaGT reveals finer-grained anatomical regions, offering more detailed interpretations of tissue function. Collectively, these results establish SpaGT as a powerful and versatile tool for analyzing SRT data.

Results

Overview of SpaGT

We developed SpaGT, a spatially informed graph transformer framework designed specifically for spatially resolved transcriptomics (SRT) analysis. SpaGT enables the learning of spatially aware graph representation for denoising gene expression data and identifying spatially functional regions. It overcomes the limitations of traditional transformers, which are typically restricted to feature representation training, by simultaneously evolving both transcriptional signal representations and relationship similarities between spots using a deep learning approach (Fig. 1).

Specifically, SpaGT takes graph-structure data, denoted as $G(X_1, A)$, where $X_1 \in R^{M \times N}$ represents M genes across N spots/cells, and $A \in R^{N \times N}$ represents the adjacency matrix derived from spatial multimodal data, as input (Fig. 1a). SpaGT then generates the node and edge channels of transformers by employing expression embeddings H^i and edge embeddings E^i, respectively (Fig. 1b). Through the use of a structure-reinforced self-attention module, SpaGT applies global self-attention as an aggregation mechanism to update the expression embeddings H^i in the graph representation, while also incorporating edge channels to capture spatially local information.

The spatially aware attention weights S^i, generated by the self-attention module, are used to iteratively update the topological structure information E^i of the graph representation across layers. The core innovation of SpaGT lies in its structure-reinforced self-attention module, which effectively learns and updates the graph representation $G(H^i, E^i)$ throughout the model. Additionally, SpaGT incorporates a clustering-augmented contrastive

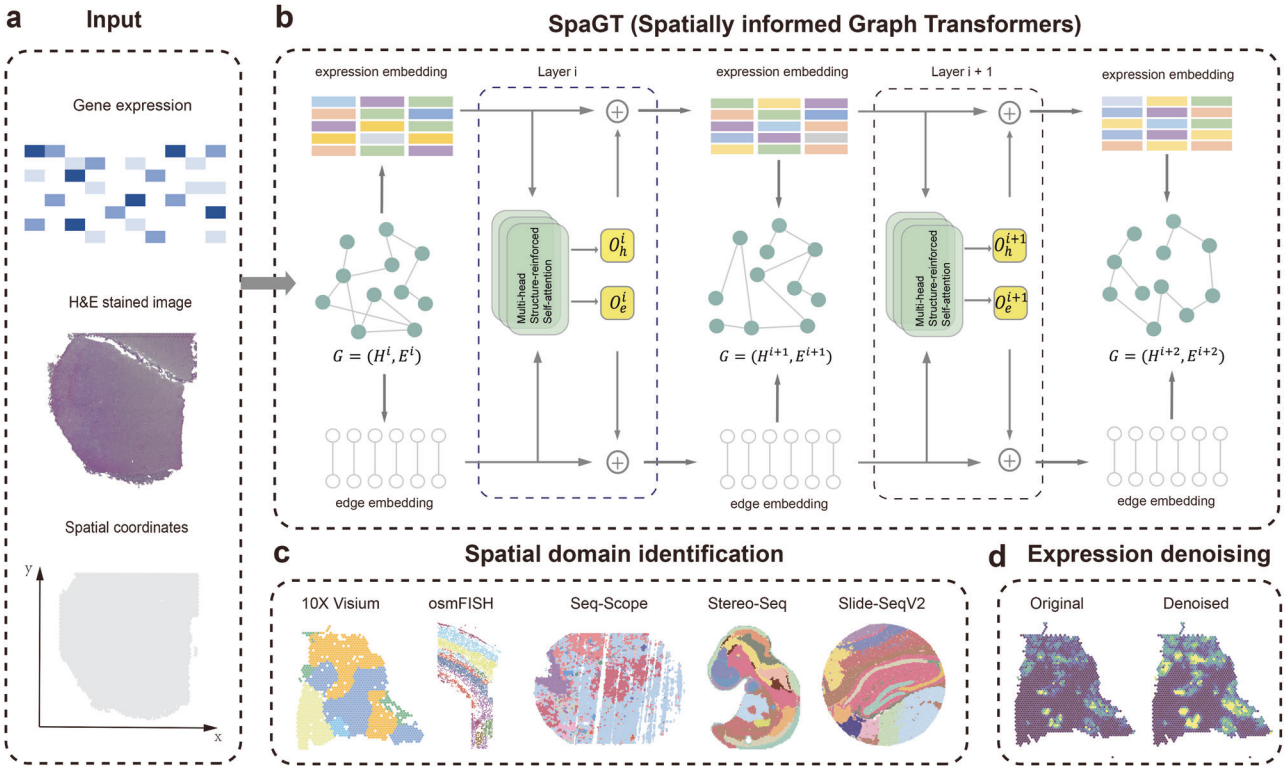


Fig. 1 | The schematic overview of SpaGT. a The three input data types for SpaGT are the expression matrix of SRT data, spatial coordinates, and Hematoxylin and Eosin (H&E)-stained images. **b** SpaGT transforms the input data sources into graph-structured data (i.e., G) from spatial multimodal data and generates expression embedding H^i and edge embedding E^i, respectively. The node and edge information are updated through the structure-reinforced self-attention module and combined with the input data via residual connections to obtain a new graph representation G. In all boxplots, the center line indicates the median; box limits denote the upper and lower quartiles; whiskers extend to 1.5× the interquartile range. Outliers were removed prior to plotting. **c** Demonstrating the application of SpaGT in spatial domain identification for multimodal SRT data. **d** The superiority of SpaGT in data expression denoising.

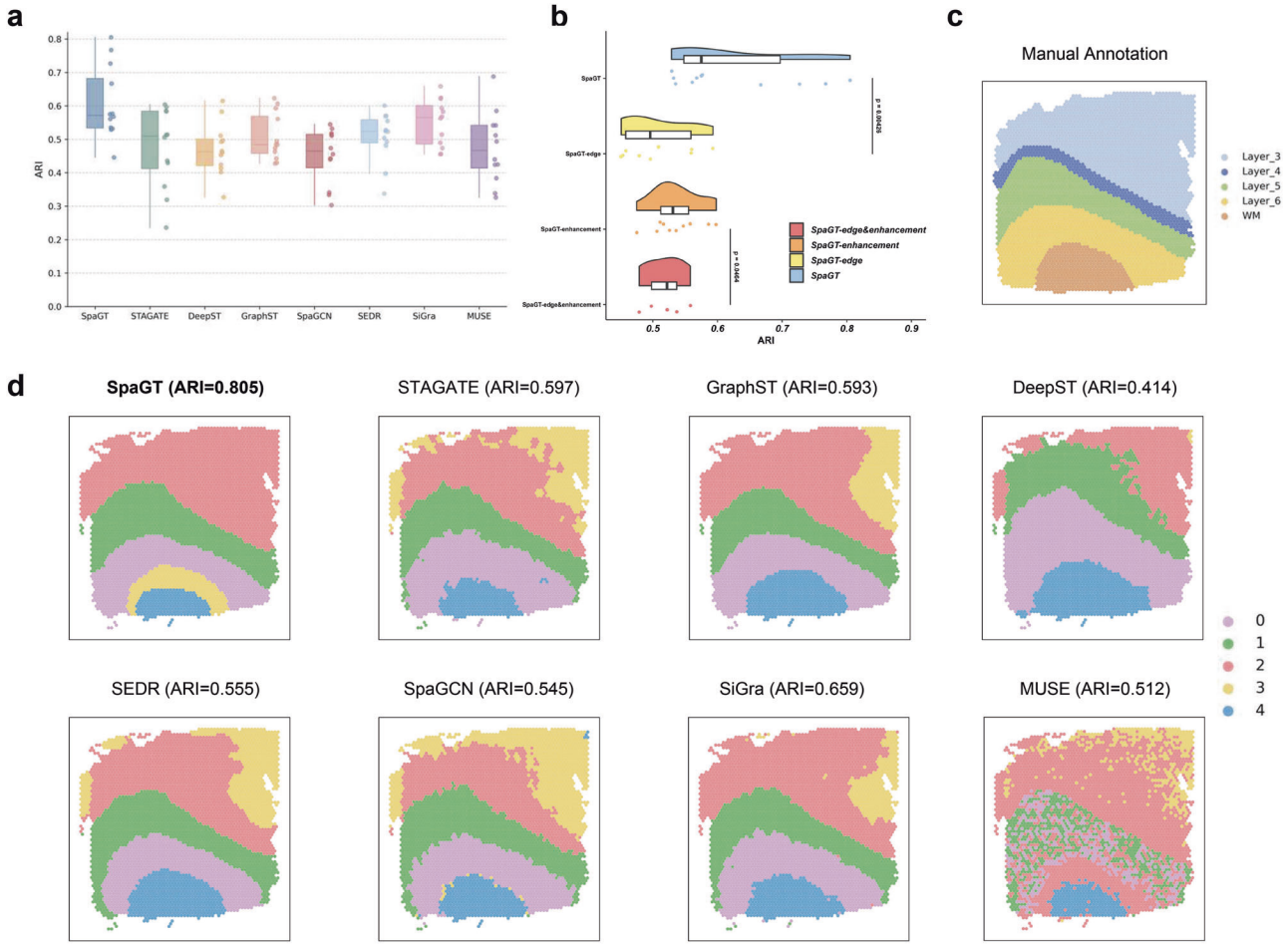


Fig. 2 | Performance comparison of SpaGT with various existing methods in spatial domain identification. a The figure shows a box plot comparison of SpaGT and seven other algorithms on all dorsolateral prefrontal cortex (DLPFC) slices. The *y*-axis represents the Adjusted Rand Index (ARI), which measures the similarity between the spatial domains predicted by each method and the manually annotated. **b** Boxplots of Adjusted Rand Index (ARI) assess the contributions of SpaGT's components to overall performance. **c** The spatial locations of the six cortical layers (L3 to L6) and white matter (WM) in slice 151672 from the DLPFC region are shown, based on manual annotations by Maynard et al. **d** Furthermore, a comparative performance of SpaGT and other advanced methods (such as STAGATE, GraphST, DeepST, SEDR, SpaGCN, SiGra and MUSE) in spatial domain identification on slice 151672 is displayed.

module to ensure that the learned graph representations are suitable for spatial clustering tasks.

Upon completing the training process, SpaGT produces an optimal graph representation $G(H^L, E^L)$ for downstream analysis. The optimal transcriptional signal representation H^L is then utilized to construct a nearest-neighbor network, which is subsequently integrated with the Leiden algorithm to identify spatial domains (Fig. 1c)[21]. Simultaneously, the topological structural information E^L is employed to denoise expression profile $\widetilde{X}_1 = X_1 E^L$, which enhances spatial expression patterns and domain specificity. This denoised data is further used to identify domain-specific differential genes of biological interest (Fig. 1d).

SpaGT's superior efficacy in spatial domain identification

To validate SpaGT's superior accuracy in spatial domain detection, we evaluated its performance on 12 human dorsolateral prefrontal cortex (DLPFC) sections from 10x Visium and compared it to several current state-of-the-art methods. Maynard et al.[22] carefully delineated the DLPFC datasets based on cytological structure and genetic markers, categorizing them into six cortical layers (L1-L6) and the white matter (WM). Our evaluation is based on these manual annotations, used as ground truth, and the Adjusted Rand Index (ARI). We compared SpaGT's performance with STAGATE[16], GraphST[18], DeepST[17], SEDR[23], SpaGCN[15], SiGra[24], and MUSE[25].

In our stratification results, SpaGT achieved the highest median ARI of 0.572, indicating closer alignment with the manual annotations than the other methods (as confirmed by the Wilcoxon signed-rank test, $P < 1e\text{-}10$, as depicted in Fig. 2a and Supplementary Fig. 1). The other methods exhibited varying degrees of clustering accuracy, with STAGATE obtaining a median ARI of 0.510. SEDR, GraphST, DeepST, SpaGCN, SiGra and MUSE achieved median ARIs of 0.524, 0.485, 0.463, 0.465, 0.566 and 0.467, respectively (Supplementary Data 1).

We also systematically assessed the contribution of various SpaGT's components to its overall performance. First, we removed the edge channels and calculated ARI values for the 12 DLPFC sections (Fig. 2b). SpaGT without edge channels (SpaGT$^{-\text{edge}}$) performed significantly worse, with an average ARI of 0.486, compared to the complete SpaGT model, which had an average ARI of 0.607. In contrast, removing only the enhancement matrix X_1 (SpaGT$^{-\text{enhancement}}$) resulted in a better outcome, with an average ARI of 0.517. These findings highlight the critical importance of edge embeddings in capturing spatial topological structure information. When compared to matrix X_0 (SpaGT$^{-\text{edge\&enhancement}}$, average ARI = 0.467), the spatial expression-augmented matrix X_1 indeed enhanced clustering performance, although still not to the level of SpaGT's full graph representation. Thus, spatial expression augmentation effectively reduces noise interference and clarifies spatial patterns, leading to enhanced clustering performance.

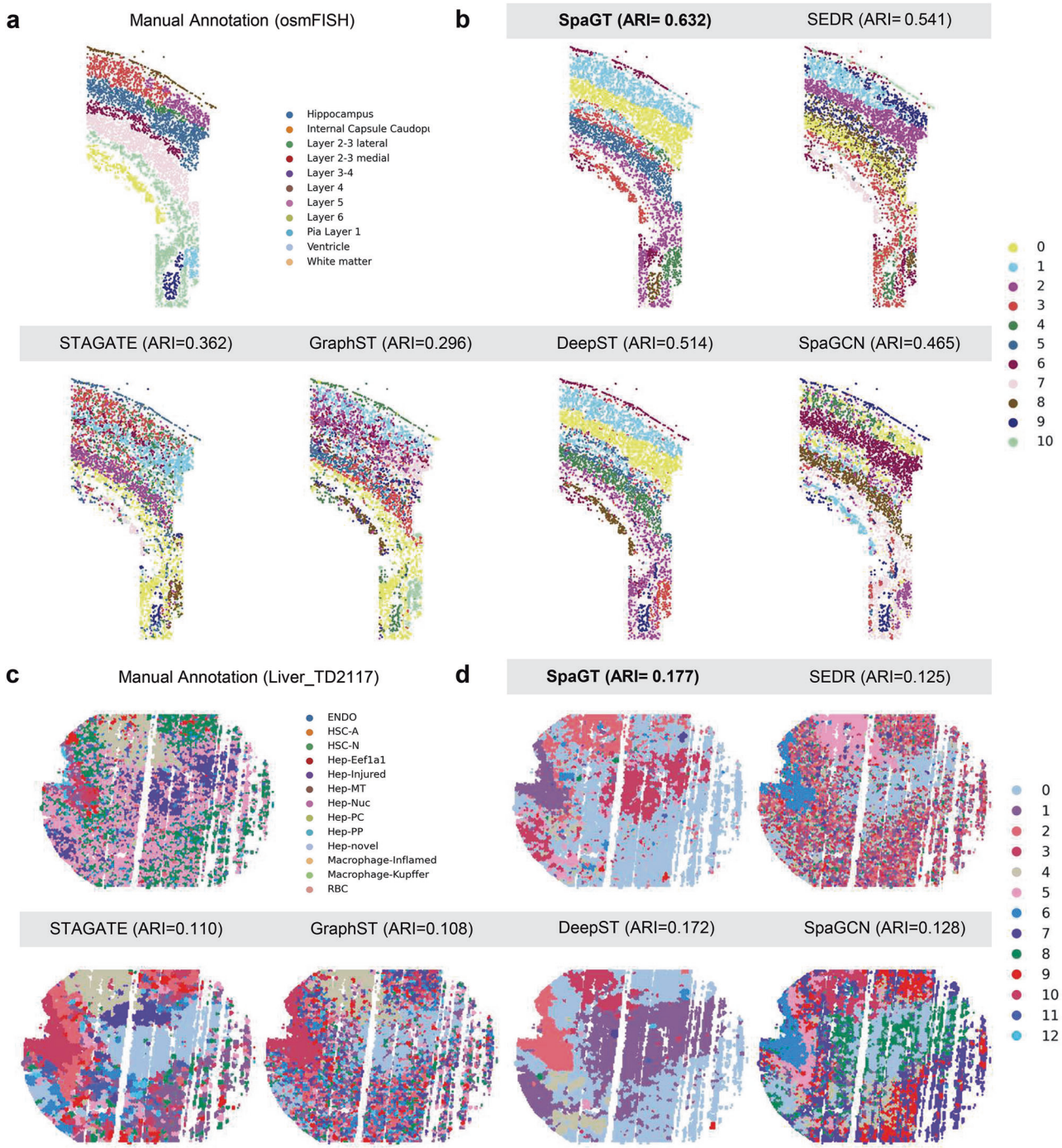


Fig. 3 | SpaGT enhances spatial domain identification in high-resolution SRT data. a Somatic cortex osmFISH data from mice ($n = 4839$ points) was manually annotated into 11 categories. **b** Clustering assignments for the somatic cortex layers generated by SpaGT and five spatial algorithms (DeepST, SpaGCN, GraphST, SEDR, and STAGATE). **c** Ground truth based on Seq-Scope processed datasets for the liver ($n = 18,808$ points). **d** Comparison of spatial subdomains detected by the aforementioned six methods and their corresponding Adjusted Rand Index (ARI).

To further highlight SpaGT's superiority, we specifically examined slice 151672 (Fig. 2c), showcasing the spatial domain identification from each method against the ground truth (Manual Annotation). SpaGT's predictions exhibited a high degree of congruence with the manually annotated domains, achieving an ARI of 0.805 (Fig. 2d). Notably, compared to other methods, SpaGT demonstrated more distinct boundaries and fewer noise points. These results underscore the crucial role that these components play in accurately identifying the spatial structure of tissues.

SpaGT unveils anatomical structures in single-cell resolution SRT data

In addition to its application on 10x Visium data, SpaGT has been extended to analyze single-cell resolution SRT data derived from diverse technological platforms. First, we applied SpaGT to the somatosensory cortex osmFISH dataset[26] (Fig. 3a). Notably, SpaGT achieved a substantial Adjusted Rand Index (ARI) of 0.632, outperforming other methods (as confirmed by the Wilcoxon signed-rank test, $P < 1e\text{-}6$, SEDR: ARI = 0.541; STAGATE: ARI = 0.362; GraphST: ARI = 0.296; DeepST: ARI = 0.514; SpaGCN:

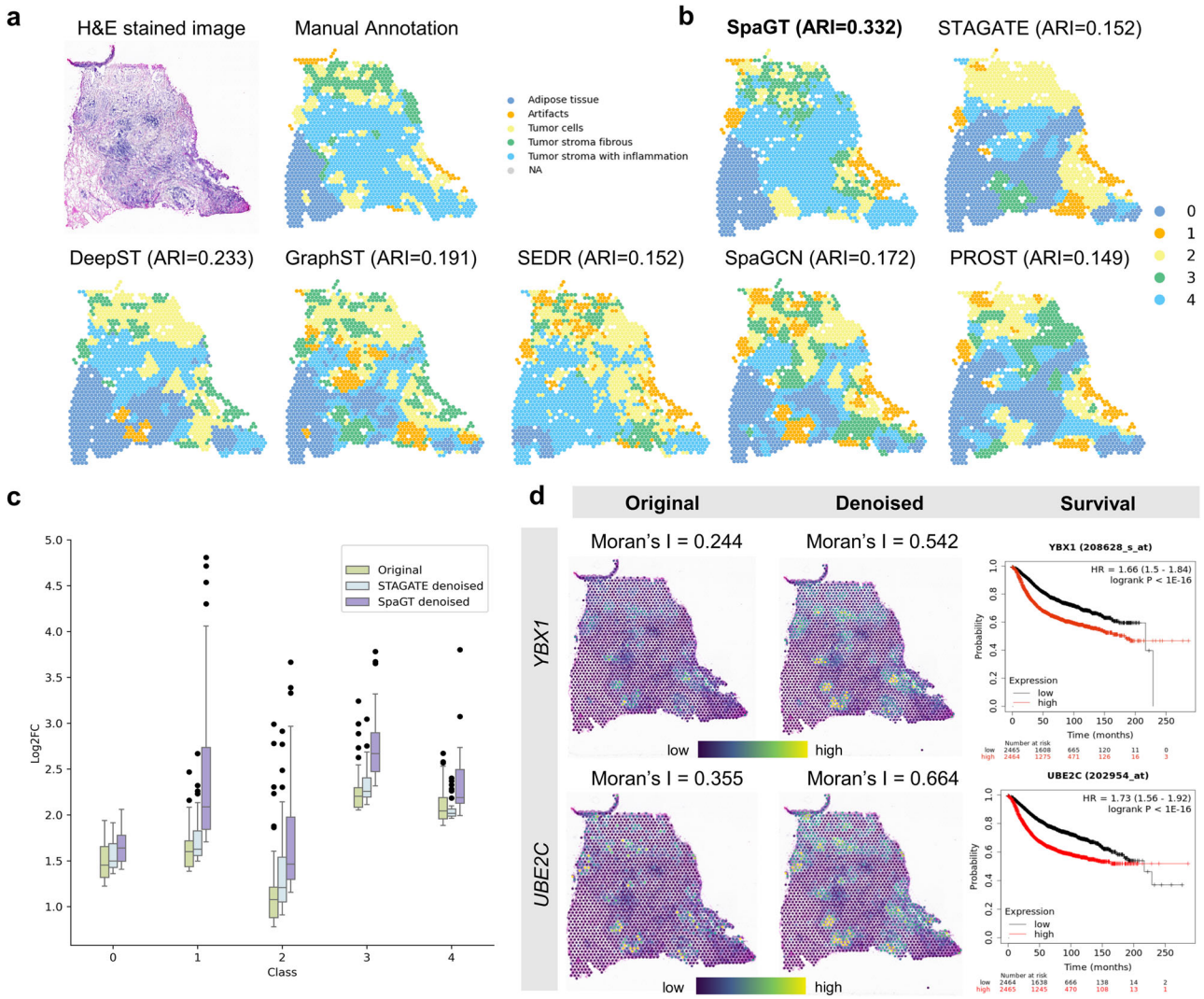


Fig. 4 | SpaGT reveals biological insights into tumor heterogeneity in human triple-negative breast cancer (TNBC) ST data ($n = 1848$ points). a The H&E-stained image (left) depicts different tissue regions annotated by pathologists in the original study. The annotated regions include invasive cancer (Tumor stroma with inflammation), adipose tissue, fibrous tissue, in situ carcinoma (Tumor cells) and tissue artifacts. **b** Comparison of spatial domain identification on breast cancer sections, where cluster ARI is used to compare the similarity between identified spatial domains and ground truth annotations. **c** Changes in gene differential expression and spatial autocorrelation patterns before and after data denoising (FC: fold change in gene expression). In the boxplots, the center line indicates the median; box limits denote the upper and lower quartiles; whiskers extend to 1.5× the interquartile range. Outliers were removed prior to plotting. **d** Spatial expression of selected domain marker genes before (left) and after (middle) data denoising, along with changes in Moran's I. Survival analysis of breast cancer-related genes (*YBX1* and *UBE2C*) (right).

ARI = 0.465). The results demonstrate that SpaGT aligns closely with the anatomical structures, revealing distinct stratification in spatial domains (Fig. 3b and Supplementary Fig. 2).

We also applied SpaGT to a liver slice processed using the Seq-Scope dataset (Fig. 3c), where SpaGT achieved the highest ARI compared to other methods (as confirmed by the Wilcoxon signed-rank test, $P < 1e\text{-}4$, SpaGT: ARI = 0.177; SEDR: ARI = 0.125; STAGATE: ARI = 0.110; GraphST: ARI = 0.108; DeepST: ARI = 0.172; SpaGCN: ARI = 0.128). These results underscore SpaGT's effectiveness in accurately identifying and stratifying spatial domains, confirming its superior performance in handling complex biological datasets (Fig. 3d and Supplementary Fig. 3).

SpaGT reveals biological insights into heterogeneity in triple-negative breast cancer

We applied SpaGT to a highly heterogeneous triple-negative breast cancer (TNBC) tissue section, specifically from a rare subtype of low albumin (CL) tumors[27], to validate the model's ability to differentiate spatial domains within heterogeneous cancer tissues. Using current histopathological images and annotations provided by pathologists (Fig. 4a), we compared SpaGT with other spatial domain identification methods using Adjusted Rand Index (ARI) as a performance metric (Fig. 4b and Supplementary Fig. 4).

SpaGT successfully clustered the tissue into five regions, achieving ARI of 0.332 (Fig. 4a, b), which surpassed the performance of other methods (as confirmed by the Wilcoxon signed-rank test, $P < 1e\text{-}8$, STAGATE: ARI = 0.152; DeepST: ARI = 0.233; GraphST: ARI = 0.191; SEDR: ARI = 0.152; SpaGCN: ARI = 0.172; PROST[28]: ARI = 0.149). Notably, SpaGT significantly enhanced the segmentation of tumor and stromal regions, further emphasizing its proficiency in accurately identifying spatial domains within heterogeneous tissues.

Cancer tissues are often characterized by considerable cellular heterogeneity, and the original expression data is typically sparse and noisy due to technical limitations. To demonstrate SpaGT's efficacy in restoring spatial expression patterns, we compared the expression specificity of spatial domains using log2 fold change (log2FC) across the denoised expression

Table 1 | Comparison of LogFC and gene descriptions before and after denoising in SpaGT analysis

Gene	Original	Denoised	*P*-value	Brief description
IGHG1	2.011	4.537	1.00E-04	An immunoglobulin heavy chain γ1; overexpressed in some B-cell malignancies.
S100A8	1.875	3.801	1.00E-11	A calcium-binding protein that drives inflammation, associated with cancer progression.
YBX1	2.084	3.782	8.10E-09	Y-box binding protein-1, a transcription/translation regulator driving tumor proliferation and drug resistance.
UBE2C	2.546	3.671	0.0019	A ubiquitin-conjugating enzyme promoting cell cycle progression; frequently upregulated in various cancers.
CD24	1.858	3.651	3.50E-10	An adhesion molecule; high expression correlates with increased tumor invasiveness and poor prognosis.
SULT1C2	3.241	3.319	0.00026	A sulfotransferase that may be involved in detoxification or metabolic processes in cancer cells.
DBI	1.735	3.315	1.30E-07	Diazepam-binding inhibitor regulating lipid metabolism; upregulation can fuel tumor growth.
SCD	2.020	3.274	4.00E-04	Stearoyl-CoA desaturase promoting synthesis of unsaturated fatty acids for cancer cell proliferation.
IGKC	1.306	3.235	0.023	Immunoglobulin κ light chain; associated with B-cell involvement in certain tumors.
ALOX15B	2.311	3.187	7.20E-03	A lipoxygenase modulating arachidonic acid metabolism; can support tumor cell growth.

data from SpaGT, the original matrix, and the denoised expression obtained using STAGATE (Fig. 4c). This quantitative comparison highlights SpaGT's effectiveness in improving spatial gene expression patterns. Additionally, we evaluated Moran's I statistics for both the original and denoised expression data (Fig. 4d and Supplementary Fig. 5), which demonstrated clear improvements in expression clarity and substantial noise reduction following the denoising process.

The denoised data also uncovered more potential biological insights. Focusing on critical tumor areas (domain 2), we identified genes with a log2 fold change (log2FC) greater than 1 as significantly upregulated, classifying them as differentially expressed genes (DEGs). We presented the log2FC comparison of DEGs before and after denoising, along with a brief description of their functions (Table 1 and Supplementary Table 1). Among the identified DEGs, genes such as *YBX1* and *UBE2C* (Fig. 4d) demonstrated enhanced specificity in cancerous regions, underscoring their potential as biomarkers for pathological evaluation. *YBX1* (Y-box binding protein 1) is intimately linked with glycolytic gene features and is predominantly overexpressed in TNBC. In breast cancer patient samples, *YBX1* expression is positively correlated with genes associated with epithelial-mesenchymal transition (EMT). Suppression of *YBX1* has been shown to downregulate EMT-related genes, consequently reducing tumor migration and invasion in TNBC cell lines, such as *MDA-MB-231* and *BT549*[29]. Similarly, *UBE2C* (ubiquitin-conjugating enzyme 2C) plays a crucial role in cell cycle progression, mitosis regulation, and the targeted degradation of short-lived proteins. The expression of *UBE2C* is notably elevated in cancers compared to normal tissues[30]. Collectively, these results emphasize the robust performance of SpaGT in handling heterogeneous cancer SRT data, effectively balancing biological relevance with spatial smoothness. The denoising analysis further provided novel insights into the underlying biology of tumors, highlighting SpaGT's potential to uncover critical molecular markers and spatial heterogeneity in complex cancer tissues.

SpaGT accurately discerns anatomical structures across high-resolution SRT platforms

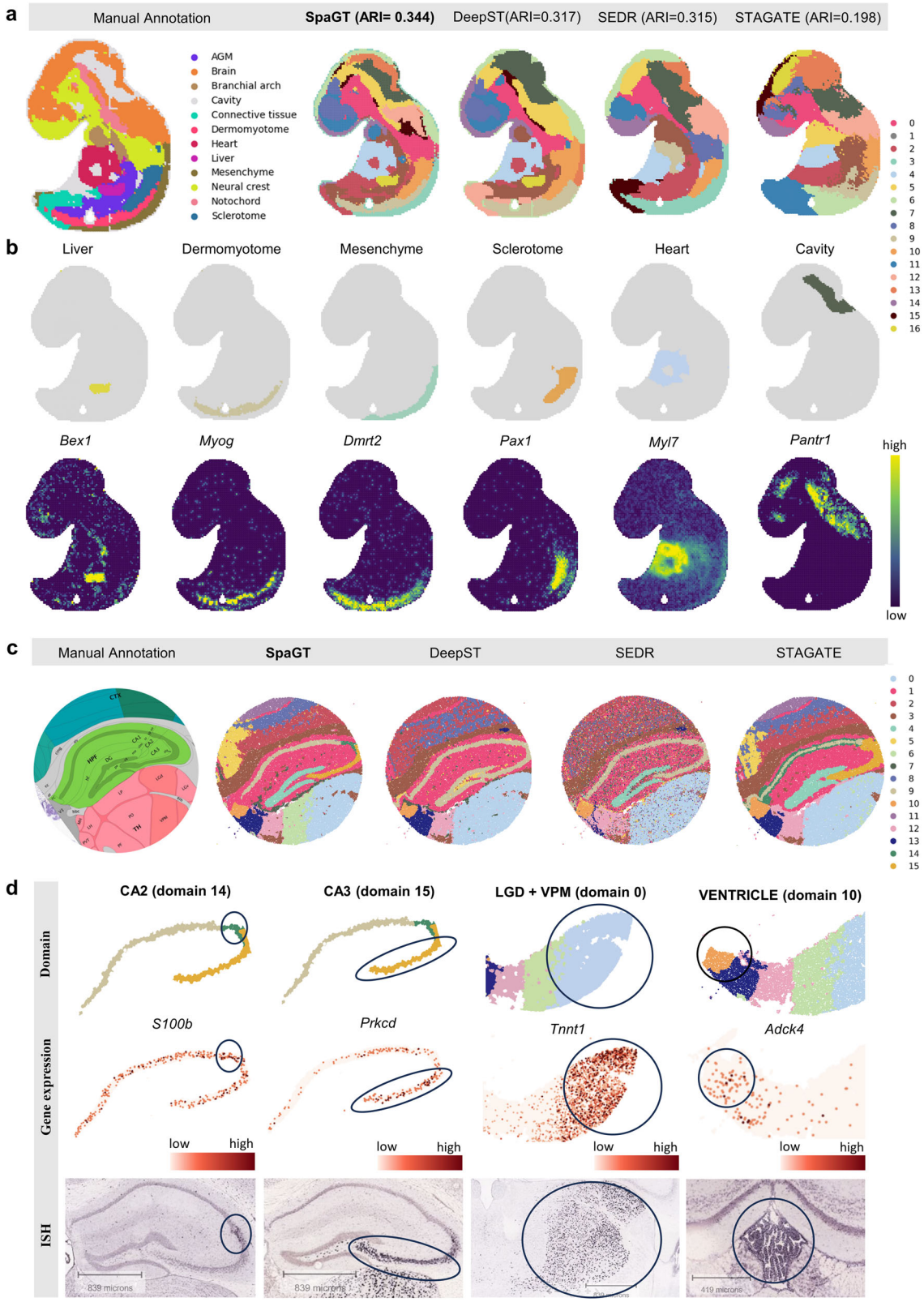
To further demonstrate SpaGT's superior performance across various platforms, we applied spatial domain analysis to high-resolution SRT data. High-resolution datasets often present additional challenges, including a larger number of data points per unit area, which increases the computational load, as well as more intricate biological details that require sophisticated algorithms for accurate pattern detection. Moreover, the higher noise levels typical of these datasets can obscure meaningful biological signals.

Initially, we Initially, we performed spatial domain detection on mouse embryo sections obtained using Stereo-seq (Fig. 5a). Compared to existing manual annotations, SpaGT demonstrated excellent consistency (as confirmed by the Wilcoxon signed-rank test, $P < 1e\text{-}4$, SpaGT: ARI = 0.344; DeepST: ARI = 0.317; SEDR: ARI = 0.315; STAGATE: ARI = 0.198), accurately identifying the spatial domains of distinct tissues. The precision of SpaGT's spatial domains was validated through the expression profiles of corresponding tissue marker genes (Fig. 5b, Supplementary Figs. 6 and 7). Notably, SpaGT effectively identified the shapes of liver and heart sections, which were characterized by the marker genes *Bex1* and *Myl7*[31], respectively.

Additionally, we analyzed mouse hippocampus data using Slide-seqV2, referencing the anatomical diagrams from the Allen Mouse Brain Atlas[32]. Our spatial domain results closely mirrored the established Allen structures (Fig. 5c and Supplementary Fig. 8). Specifically, SpaGT accurately identified regions such as the CA2 and CA3 areas of the hippocampus, the ventricles, the dorsal part of the lateral geniculate complex (LGD), and the ventral posteromedial nucleus (VPM) of the thalamus. We validated the spatial domain results by examining the expression profiles of marker genes, conducting independent ISH (In Situ Hybridization) image and evaluating the average expression levels of region-specific differentially expressed genes (Supplementary Fig. 9). For instance, the marker gene *S100b*[33] and *Prkcd*[34] have the high expression in the CA2 and CA3 regions, which corresponded to ISH image, respectively. Additionally, the LGD and VPM of the thalamus that we accurately identified were consistent with the high expression the marker gene *Tnnt1*[35], while *Adck4*[36] was prominently expressed in the Ventricle region (Fig. 5d). This precision underscores SpaGT's effectiveness in spatially resolving complex tissue structures at a granular level. These findings illustrate that SpaGT is not only well-suited for large-scale, high-resolution datasets but also versatile across a range of platforms, demonstrating robustness in spatial domain analysis.

Discussion

Recent advancements in Spatially Resolved Transcriptomics (SRT) have revolutionized the study of tissue and cancer heterogeneity within spatial environments. However, challenges such as the inherent sparsity and noise of SRT datasets necessitate the development of computational methods that can effectively integrate gene expression with spatially local information. In response to these challenges, we propose SpaGT, a spatially informed graph transformers framework, which simultaneously models the expression embedding and structural information of spatially aware graph representation. SpaGT enhances gene expression analysis and reveals spatial functional regions by employing a structure-reinforced self-attention mechanism that replaces traditional graph convolution with global self-attention, enabling the effective integration of both global and spatially localized information. To assess the performance of the method, SpaGT was applied to spatial resolved transcriptomics (SRT) data at varying resolutions. For low-resolution SRT data from the 10x Visium platform, SpaGT successfully delineates tissue functional regions through spatial clustering, as each spot contains multiple cells (ranging from 1 to 10). For high-resolution SRT data from the Slide-seqV2 and Stereo-seq platforms, SpaGT identifies functional regions within the tissue, with each spot encompassing between 1 and 2 cells. In the case of single-cell resolution SRT data from the Seq-Scope and osmFISH platforms, SpaGT can identify cell types, as each spot contains a single cell. Thus, SpaGT can effectively identify functional regions or cell types in tissues, irrespective of the resolution of the SRT data. Furthermore,



the ability of SpaGT to process SRT data across different resolutions underscores its adaptability to datasets of various scales, thereby enhancing the tool's versatility.

The distinction between SpaGT and graph neural networks (GNNs) that utilize localized convolutional aggregation lies in SpaGT's adoption of a structure-reinforced transformer architecture, which offers significant advantages, particularly in integrating global and spatially localized information. While localized convolutional aggregations are adept at leveraging spatially local relationships, they often fall short in capturing broader global spatial contexts due to their limited receptive fields. Attention-based Graph Neural Networks (GNNs), such as Graph Attention Networks (GAT), leverage the attention mechanism to aggregate neighborhood information

Fig. 5 | SpaGT accurately discerns anatomical structures across high-resolution SRT platforms. a Displays annotated anatomical structures in Stereo-seq data of mouse embryos ($n = 5913$ points) on the left. A comparison of spatial domain detection results between SpaGT and other methods on the right was showcased on spatial coordinates. **b** SpaGT accurately identifies organs such as Liver, Heart, and Sclerotome and displays marker genes corresponding to their spatial domains. **c** Identifies tissue structures and functional regions in mouse hippocampus Slide-seqV2 data ($n = 41,786$ points). The corresponding anatomical definitions obtained from the Allen Mouse Brain Atlas (first image in A) serve as a reference on the left. A comparison of spatial domains determined by different algorithms on the Slide-seqV2 mouse hippocampus dataset is shown with spatial coordinates on the right. **d** Fine anatomical regions such as CA3, thalamus (LGD + VPM), CA2, and ventricular sections detected by SpaGT. In the top image, each spatial domain determined by SpaGT is indicated: the dark green region (domain 14) represents CA2, the yellow region (domain 15) represents CA3, the light blue region (domain 0) represents LGD + VPM, and the orange region (domain 10) represents the ventricle. In the middle image, the expression of domain-specific differential genes for each region is shown, while the corresponding ISH is presented in the bottom image.

for each node, emphasizing the use of local information within the network. In contrast, transformers excel at modeling long-range dependencies within data. The traditional Graph Transformers incorporate positional encoding on top of the Transformer framework, employing the self-attention mechanism to aggregate global information for learning node representations, thus prioritizing global context. SpaGT introduces a structure-reinforced self-attention mechanism, which effectively integrates both global and local information to learn robust spatial graph representations. SpaGT leverages this capability to consider interactions across the entire spatial domain, not just between immediate neighbors. This is particularly beneficial in spatial transcriptomics, where the spatial arrangement of cells or spots influences gene expression patterns over larger areas, beyond local scales. Such a global perspective is essential for accurately identifying spatial domains that are defined by subtle variations in gene expression distributed across an entire tissue section.

Additionally, the transformer architecture in SpaGT offers flexibility and scalability in processing large datasets, which are typical of spatial transcriptomics studies. Unlike graph convolutional neural networks (GCNs) or other localized convolutional aggregation models that require a predefined graph structure—thus limiting adaptability and increasing computational complexity—SpaGT efficiently handles varying data sizes and complexities without the need for extensive network modifications. In summary, SpaGT not only overcomes the limitations of traditional GNNs in processing only local information, but it also enhances the model's ability to integrate and interpret both global and local spatial information. This leads to more accurate and robust identification of spatial domains, making SpaGT a superior choice for advanced spatial transcriptomics analysis.

While SpaGT excels in spatial domain identification and gene expression enhancement, it does not inherently explore the deeper molecular mechanisms that govern tissue architecture and function. This limitation may restrict its utility for comprehensive biological interpretations that require an understanding of underlying molecular interactions. Therefore, the next crucial step in SpaGT's development is the integration of gene regulatory insights into the model, facilitating deeper biological analyses. This enhancement would transform SpaGT into a powerful tool for uncovering the molecular and functional underpinnings of biological tissues and their responses to various physiological and pathological stimuli.

Methods

Data collection and general preprocessing

We utilize datasets from 10x Visium, Slide-seqV2, Stereo-seq and cancer ST data[27]. Employing the SCANPY[37] Python package, we initially selected the top 3000 highly variable genes (HVGs). Subsequently, we performed log transformation and normalization on the raw gene expression data, which serves as input for SpaGT.

Building spatially informed graph transformers framework

Constructing graph-structure data from SRT data. SpaGT adopts a strategy inspired by SpaGCN to convert spatial multimodal data into a weighted graph $G(X_0, A)$, where matrix A represents the relationships between each spot and its spatial neighbors, thereby capturing the spatially local structure inherent in SRT data.

SpaGT calculates the Euclidean distance between each pair of spots based on their spatial coordinates and corresponding histology image information and selects the k-nearest spatial neighbors[38] for each spot. The cosine distance between these neighboring points, based on the PCA embedding of gene expression (i.e., PCs), is then computed and exponentially transformed into the relationship matrix A, which represents the relationship between spots.

$$A_{ij} = \frac{D_{ij}}{\sum_{i=0}^{N} D_{ij}},\ D = \exp(2 - cosine\_dist(U)), \tag{1}$$

where matrix $U \in R^{15 \times N}$ is a low-dimensional matrix consisting of 15 principal components (PCs). Matrix D represents the cosine distance of these neighboring spots.

Enhancing spatial expression. The gene expression at each spot is augmented by integrating shared information from its surrounding neighbors, resulting in an augmented spatial expression for the corresponding spot. This augmentation is achieved using the relationship matrix A, and the augmented expression matrix X_1 is computed as follows:

$$X_1 = X_0 + \alpha X_0 A, \tag{2}$$

where the tunable parameter α (default to 1.0) can be manually adjusted based on observed outcomes, allowing control over the influence of spatial neighborhood similarity on spatial domains identification (Supplementary Fig. 10).

Spatially informed graph transformers for modeling graph representation. Given the expression embedding H^i and edge embedding E^i as input, SpaGT builds spatially informed graph transformers to learn spatially aware graph representation. This architecture consists of a multi-head structure-reinforced self-attention module, a graph representation update component, and an SVD-based positional encodings component.

- *Multi-head structure-reinforced self-attention module:* In this module, SpaGT introduces edge channels into the self-attention architecture, facilitating the integration of both global and spatially local information. The self-attention process in the i-th layer and for the t-th attention head is expressed as follows:

$$Attentation(Q_t^i, K_t^i, V_t^i) = V_t^i S_t^i,\ i = 1, \ldots, L \tag{3}$$

$$S_t^i = softmax\left(clip\left(\frac{(Q_t^i)^T \cdot K_t^i}{\sqrt{d_t}}\right) + E_S^i\right) \odot \sigma(G_S^i),\ i = 1, \ldots, L \tag{4}$$

where $Q_t^i = P_{Q,t}^i H^i \in R^{d_t \times N}$, $K_t^i = P_{K,t}^i H^i \in R^{d_t \times N}$, $V_t^i = P_{V,t}^i H^i \in R^{d_t \times N}$ represent the query, key and value matrices of the t-th head in i-th layer. The dimensionality of the queries and keys is denoted as d_t. The operator $\odot$ denotes elementwise product. $E_S^i = P_{S,E} E^i \in R^{N \times N}$ is the bias term added to attention weight, allowing spatial information to influence the global aggregation process. $G_S^i = P_{S,G} E^i \in R^{N \times N}$, using the sigmoid function $\sigma(\cdot)$, enables spatial information to gate the values

before aggregation, controlling the flow of information between spots. The scaled dot product is constrained by the $clip(\cdot)$ operation, improving numerical stability. $P^i_{Q,t}$, $P^i_{K,t}$, $P^i_{V,t}$, $P_{S,E}$ and $P_{S,G}$ are learned projection matrices.

- *Updating graph representation*: Using the structure-reinforced self-attention module, the updated term O^i_h for expression embedding and the updated term O^i_e for edge embedding are computed as follows:

$$O^i_h = concact\left(V^i_1 S^i_1, \ldots, V^i_t S^i_t, \ldots, V^i_\Gamma S^i_\Gamma\right),\ i = 0, \ldots, L-1 \quad (5)$$

$$O^i_e = \sum_{t=1}^{\Gamma} S^i_t/\Gamma,\ i = 0, \ldots, L-1 \quad (6)$$

Based on these updated terms (i.e., O^i_h and O^i_e), the expression embedding H^i and edge embedding E^i of graph representation in the i-th layer are updated as follows:

$$\begin{aligned} H^i &= H^{i-1} + O^{i-1}_h = H^{i-1} \\ &\quad + concact\left(V^{i-1}_1 S^{i-1}_1, \ldots, V^{i-1}_t S^{i-1}_t, \ldots, V^{i-1}_\Gamma S^{i-1}_\Gamma\right), \\ &\quad i = 1, \ldots, L \end{aligned} \quad (7)$$

$$E^i = E^{i-1} + O^{i-1}_e = E^{i-1} + \sum_{t=1}^{\Gamma} S^{i-1}_t/\Gamma,\ i = 1, \ldots, L \quad (8)$$

where $concact(*)$ denotes concatenation along the rows. We denote the low-dimensional embedding of gene expression X_1 and the relationships matrix A between each spot and its spatial neighbors as H^0 and E^0, respectively. We apply a feed-forward sublayer (FFN) and Layer Normalization (LN) to obtain the final graph representation as:

$$H^i = LN\left(H^i + FFN^i_H\left(LN\left(H^i\right)\right)\right), i = 1, \ldots, L \quad (9)$$

$$E^i = LN\left(E^i + FFN^i_E\left(LN\left(E^i\right)\right)\right), i = 1, \ldots, L \quad (10)$$

- *SVD-based positional encodings*: We decomposed the relationship matrix A between each spot and its spatial neighbors using Singular Value Decomposition (SVD) to obtain positional encodings.

$$A = U\Sigma V^T = \left(U\sqrt{\Sigma}\right)\left(V\sqrt{\Sigma}\right)^T = \widetilde{U}\widetilde{V}^T,\ P = \left(\widetilde{U}||\widetilde{V}\right)W_{PE} \quad (11)$$

where $U, V \in R^{M\times r}$ are matrices containing the largest r left and right singular vectors as columns, respectively, corresponding to the top r singular values in the diagonal matrix $\Sigma \in R^{r\times r}$. The operator $||$ denotes concatenation along the columns. $W_{PE} \in R^{2r\times N}$ is a learned projection matrix, and the matrix $P \in R^{M\times N}$ serves as positional encodings.

Clustering-augmented contrastive module

Through the aforementioned process, we obtained spatially aware graph representations; however, these representations are not directly applicable to clustering tasks. To address this limitation, we introduce a Clustering-Augmented Contrastive Module designed to facilitate effective clustering.

Initially, we employ the Student's t-distribution to measure the probability of assigning a spot i to a cluster j based on the combined latent representation H^L, as follows:

$$q_{ij} = \frac{\left(1 + \|H^L_i - \mu_j\|^2/\rho\right)^{-\frac{\rho+1}{2}}}{\sum_{j'}\left(1 + \|H^L_i - \mu_{j'}\|^2/\rho\right)^{-\frac{\rho+1}{2}}} \quad (12)$$

where μ represents the cluster center determined by K-means on learned representations, and H^L_i denotes the i-th column of H^L. The matrix $Q = [q_{ij}]$ represents the distribution of assignments for all samples, and ρ denotes the degrees of freedom in the Student's t-distribution.

Subsequently, we calculate a target distribution P as follows:

$$p_{ij} = \frac{\left(q^2_{ij}/s_j\right)}{\sum_{j'}\left(q^2_{ij'}/s_{j'}\right)}, \text{where } s_j = \sum_i q_{ij} \quad (13)$$

Each assignment in Q is squared and normalized to generate the target distribution P, which enhances the clustering by bringing the data representations closer to the cluster centers. This process enables SpaGT to learn more effective representations for clustering tasks. By minimizing the Kullback-Leibler (KL) divergence between the Q and P distributions, we increase the confidence of the assignments, leading to the following loss function:

$$\mathcal{L} = KL(P \parallel Q) = \sum_i \sum_j p_{ij}\, log\, \frac{p_{ij}}{q_{ij}} \quad (14)$$

Hyperparameter tunings and implementations

Similar to other deep learning models, SpaGT involves several hyperparameters, such as the number of heads in the structure-reinforced self-attention module, the number of neurons per layer, the learning rate, and the number of training epochs. However, SpaGT is relatively insensitive to hyperparameter settings. For all experiments, we used a single set of hyperparameters (default settings), adhering to scientific standards. Five multi-attention heads were sufficient to capture the complex topological structure in spatial transcriptomics, consistently demonstrating excellent performance across all experiments. For training, we set the learning rate to 1e-3 and the number of epochs to 1000, both of which resulted in stable convergence. For large-scale datasets, we recommend a mini-batch strategy, adjusting the number of input samples per batch based on the available GPU memory.

The tuning parameter "alpha (α)," a non-negative number, to proportionally control the expression mixing of the centered spot and its surrounding neighbors. Intuitively, alpha is set to 0 for the original expression values, while a larger alpha will aggregate more information from the spatial neighborhood, enabling more spatially local similarity that can impact the identification of tissue functional regions. Here, we evaluate the applicable range of alpha using a 10x Visium Invasive Ductal Carcinoma (IDC) slice. We varied alpha from 0 to 2 with increments of 0.5 to generate a series of normalized profiles and assessed the clustering performance and spatial coherence of spatial domains (Supplementary Fig. 10). As previously proposed, larger alphas improve the smoothness of spatial domains and reduce discrete subdomains. However, the clustering purity exhibited a trend from rising to decline (alpha = 1.0 with the highest purity = 0.867, Supplementary Fig. 10), which suggests some molecular difference, especially at the domain boundaries, might be overwhelmed if excessive spatial smoothing was implemented. Therefore, for the data with gold standard, the optimal parameter α can be trained in a cross-validation manner. For data without gold standard, the optimal parameter α can be judged based on the smoothness of the in situ staining results by floating up and down on the basis of the default parameters $\alpha = 1$.

Spatial domain identification, gene expression denoising and survival analysis

SpaGT integrates spatial information to update expressions, resulting in spatially aware graph representation $G\left(H^L, E^L\right)$. It employs clustering algorithms such as Mclust and Leiden to detect spatial domains, with the 'resolution' parameter being adjustable to align with the number of manual annotations. Note that Mclust requires the specification of the number of clusters and assumes an underlying probabilistic distribution, making it particularly suitable for specific types of data. In our study, we applied Mclust to the DLPFC datasets by adopting the current mainstream

processing strategy, while the Leiden algorithm was employed for spatial clustering of other SRT datasets.

Additionally, SpaGT utilizes the topological structure information to denoise gene expression data, resulting in $\widetilde{X}_1 = X_1 E^L$, which enhances spatial expression patterns and domain specificity. We also conducted Kaplan-Meier (KM) survival analyses, assessing the prognostic significance of domain-specific marker genes by combining extensive expression profile data with patient survival information in cancer studies.

Performance evaluation

For SRT datasets with available ground truth, we quantified the accuracy of spatial domain identification using two metrics: the Adjusted Rand Index (ARI)[39–45] and clustering purity[13].

$$cluster\ purity = \frac{1}{N}\sum_{y\in Y}\max_{f\in F}\left|f \cap y\right| \tag{13}$$

where Y is the set of clusters or spatial domains; F is the set of reference groups. In cases where ground truth was not available, we evaluated clustering performance based on other reliable references, such as known structural diagrams, histopathological annotations from H&E-stained images, and marker gene analysis.

Spatial continuity and expression specificity of marker genes. The Moran's I statistic is utilized to assess the spatial autocorrelation of gene expression both before and after the denoising process. Furthermore, we evaluate the expression specificity by comparing the log2 fold change (Log2FC) values of the top marker genes associated with each spatial domain before and after denoising.

Statistics and reproducibility

All statistical analyses were conducted using the SCANPY package in Python. Comparisons between two independent groups were performed using the Wilcoxon rank sum test on log-normalized data, with a false discovery rate (FDR) of less than 0.05 considered statistically significant. For the real SRT data, each experiment was run 20 times to ensure that SpaGT's results were not affected by random initialization and to confirm the stability of the outcomes. Additionally, we have developed a dedicated documentation website (https://spagt-tutorial.readthedocs.io/) to provide a comprehensive guide on the usage of SpaGT and to facilitate the reproduction of the relevant experimental results.

Reporting summary

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

Data availability

The human dorsolateral prefrontal cortex (DLPFC) datasets[22] are available in the spatialLIBD package at http://research.libd.org/spatialLIBD/. The osmFISH dataset[26] of mouse somatosensory cortex is available at https://github.com/drieslab/spatial-datasets. The mouse hippocampus data by Slide-seq V2[4] is available at https://singlecell.broadinstitute.org/single_cell/study/SCP815. Please note that account registration is required prior to downloading the data. Triple-negative breast cancer (TNBC) data[27] from ST platform is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE213688. Mouse embryo data by Stereo-seq[5] is available at https://db.cngb.org/stomics/datasets/STDS0000058.

Code availability

Python source code of SpaGT, under the open-source BSD 3-Clause license, is available at https://github.com/xy428/SpaGT. The documentation website provides the installation guide, tutorials, and API references, which are available at https://spagt-tutorial.readthedocs.io/.

Received: 12 September 2024; Accepted: 28 March 2025;
Published online: 06 April 2025

References

1. Zhang, C., Wang, L. & Shi, Q. Computational modeling for deciphering tissue microenvironment heterogeneity from spatially resolved transcriptomics. *Comput. Struct. Biotechnol. J.* **23**, 2109–2115 (2024).
2. Crosetto, N., Bienko, M. & Van Oudenaarden, A. Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* **16**, 57–66 (2015).
3. Marx, V. Method of the year: spatially resolved transcriptomics. *Nat. Methods* **18**, 9–14 (2021).
4. Stickels, R. R. et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).
5. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e1721 (2022).
6. Hu, Y. et al. Spatially contrastive variational autoencoder for deciphering tissue heterogeneity from spatially resolved transcriptomics. *Brief. Bioinform.* **25**, bbae016 (2024).
7. Zhang, C., Li, X., Huang, W., Wang, L. & Shi, Q. Spatially aware self-representation learning for tissue structure characterization and spatial functional genes identification. *Brief. Bioinform.* **24**, bbad197 (2023).
8. Huang, W. et al. Spatially aligned graph transfer learning for characterizing spatial regulatory heterogeneity. *Brief. Bioinform.* **26**, bbaf021 (2025).
9. Wang, L., Bai, X., Zhang, C., Shi, Q. & Chen, L. Spatially aware domain adaptation enables cell type deconvolution from multi-modal spatially resolved transcriptomics. *Small Methods* https://doi.org/10.1002/smtd.202401163 (2024).
10. Bai, X., Bao, X., Zhang, C., Shi, Q. & Chen, L. SpaDCN: deciphering spatial functional landscape from spatially resolved transcriptomics by aligning cell–cell communications. *Small Methods* https://doi.org/10.1002/smtd.202402111 (2025).
11. Li, X., Huang, W., Xu, X., Zhang, H.-Y. & Shi, Q. Deciphering tissue heterogeneity from spatially resolved transcriptomics by the autoencoder-assisted graph convolutional neural network. *Front. Genet.* **14**, 1202409 (2023).
12. Dries, R. et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* **22**, 1–31 (2021).
13. Zhao, E. et al. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* **39**, 1375–1384 (2021).
14. Shang, L. & Zhou, X. Spatially aware dimension reduction for spatial transcriptomics. *Nat. Commun.* **13**, 7203 (2022).
15. Hu, J. et al. SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).
16. Dong, K. & Zhang, S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat. Commun.* **13**, 1739 (2022).
17. Xu, C. et al. DeepST: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Res.* **50**, e131 (2022).
18. Long, Y. et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat. Commun.* **14**, 1155 (2023).
19. Hussain, M. S., Zaki, M. J. & Subramanian, D. Global self-attention as a replacement for graph convolution. In *Proc. 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining)* (2022).
20. Li, Z., Chen, X., Zhang, X., Jiang, R. & Chen, S. Latent feature extraction with a prior-based self-attention framework for spatial transcriptomics. *Genome Res.* **33**, 1757–1773 (2023).
21. Traag, V. A., Waltman, L. & Van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

22. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).
23. Xu, H. et al. Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Med.* **16**, 12 (2024).
24. Tang, Z. et al. SiGra: single-cell spatial elucidation through an image-augmented graph transformer. *Nat. Commun.* **14**, 5618 (2023).
25. Bao, F. et al. Integrative spatial analysis of cell morphologies and transcriptional states with MUSE. *Nat. Biotechnol.* **40**, 1200–1209 (2022).
26. Codeluppi, S. et al. Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* **15**, 932–935 (2018).
27. Coutant, A. et al. Spatial transcriptomics reveal pitfalls and opportunities for the detection of rare high-plasticity breast cancer subtypes. *Lab. Invest.* **103**, 100258 (2023).
28. Liang, Y. et al. PROST: quantitative identification of spatially variable genes and domain detection in spatial transcriptomics. *Nat. Commun.* **15**, 600 (2024).
29. Rana, P. S. et al. YB1 is a major contributor to health disparities in triple negative breast cancer. *Cancers* **13**, 6262 (2021).
30. Mo, C.-h et al. The clinicopathological significance of UBE2C in breast cancer: a study based on immunohistochemistry, microarray and RNA-sequencing data. *Cancer Cell Int.* **17**, 1–17 (2017).
31. Wang, L. et al. Multi-modal domain adaptation for revealing spatial functional landscape from spatially resolved transcriptomics. *Brief. Bioinform.* **25**, bbae257 (2024).
32. Sunkin, S. M. et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res.* **41**, D996–D1008 (2012).
33. Xiong, Z.-g et al. Enhanced calcium transients in glial cells in neonatal cerebellar cultures derived from S100B null mice. *Exp. Cell Res.* **257**, 281–289 (2000).
34. Brewer, G. J. et al. Toward a self-wired active reconstruction of the hippocampal trisynaptic loop: DG-CA3. *Front. Neural Circuits* **7**, 165 (2013).
35. Peng, H. et al. Morphological diversity of single neurons in molecularly defined cell types. *Nature* **598**, 174–181 (2021).
36. Alvi, N. H., Turkstani, B. A., Ashi, A. S., Alzahrani, A. M. & Tawffeq, A. M. COQ8B-related steroid-resistant nephrotic syndrome in Saudi Arabia: a case report. Cureus **14**, e31922 (2022).
37. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 1–5 (2018).
38. Cover, T. & Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13**, 21–27 (1967).
39. Hubert, L. & Arabie, P. Comparing partitions. *J. Classif.* **2**, 193–218 (1985).
40. Shi, Q., Li, X., Peng, Q., Zhang, C. & Chen, L. scDA: single cell discriminant analysis for single-cell RNA sequencing data. *Comput. Struct. Biotechnol. J.* **19**, 3234–3244 (2021).
41. Shi, Q. et al. Pattern fusion analysis by adaptive alignment of multiple heterogeneous omics data. *Bioinformatics* **33**, 2706–2714 (2017).
42. Xu, J., Zhang, A., Liu, F., Chen, L. & Zhang, X. CIForm as a transformer-based model for cell-type annotation of large-scale single-cell RNA-seq data. *Brief. Bioinform.* **24**, bbad195 (2023).
43. Zhang, C. et al. Integration of multiple heterogeneous omics data. In *Proc. 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM))* (IEEE, 2016).
44. Zhang, C., Liu, J., Shi, Q., Zeng, T. & Chen, L. Differential function analysis: identifying structure and activation variations in dysregulated pathways. *Sci. China Inf. Sci.* **60**, 012108 (2017).
45. Zhang, C., Liu, J., Shi, Q., Zeng, T. & Chen, L. Comparative network stratification analysis for identifying functional interpretable network biomarkers. *BMC Bioinform.* **18**, 1–12 (2017).

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Nos. 62202120, T2341024), Natural Science Foundation of Wuhan (No. 2024040801020300), Zhejiang Provincial Natural Science Foundation of China under Grant No. LZ22C060001, the research funds of Hangzhou Institute for advanced study, UCAS (No. 2023HIAS-Y024, 2022ZZ01013, 2022ZZ01016).

Author contributions

Q.S. and C.Z. conceived and designed the framework and the experiments. X.Y.B. performed the experiments. X.Y.B., X.S.B., and C.Z. analyzed the data and wrote the paper. Q.S., C.Z. and X.L. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08015-w.

Correspondence and requests for materials should be addressed to Xiaoping Liu, Qianqian Shi or Chuanchao Zhang.

Peer review information *Communications Biology* thanks Cecilia Cisar and Ziyang Tang for their contribution to the peer review of this work. Primary handling editors: Shan E Ahmed Raza and Aylin Bircan, Jasmine Pan. A peer review file is available.