

SpaCross: Deciphering Spatial Structures and Correcting Batch Effects of Multi-slice Spatially Resolved Transcriptomics

Donghai Fang, Fangfang Zhu, Wenwen Min

Contents

| | |
|--------------------|-----------|
| 1 Table | 2 |
| 2 Algorithm | 3 |
| 3 Text | 4 |
| 4 Figure | 10 |

1 Table

Table S1: Comparative analysis of computational resource consumption, including Model Runtime (MR/Seconds), GPU Memory Usage (GMU/MB), and Memory Caching (MC/MB), across various models.

| Method | DLPFC | | | ME | | | MOB | | | MHP | | | MSC | | |
|------------|--------------|--------------|-------------|--------------|--------------|-------------|--------------|--------------|-------------|--------------|--------------|-------------|-------------|--------------|------------|
| | MR | GMU | MC | MR | GMU | MC |
| SpaCross | 28.55 | 109.8 | 2626 | 100 | 165.5 | 28918 | 90.47 | 182.3 | 22684 | 101.8 | 131.7 | 9562 | 13.19 | 72.65 | 376 |
| GraphST | 77.8 | 2307 | 3034 | 81.17 | 9689 | 12376 | 68.06 | 7759 | 10126 | 56.47 | 6256 | 7284 | 5.64 | 158.2 | 258 |
| STAaligner | 84.76 | 265.7 | 1428 | 125.9 | 608.8 | 5872 | 50.46 | 517 | 4652 | 54.31 | 113.68 | 5244 | 5.57 | 20.14 | 504 |
| SEDR | 17.93 | 140.37 | 2632 | 63.25 | 81.31 | 9284 | 61.62 | 416.3 | 10962 | 25.46 | 59.37 | 3864 | 6.39 | 73.83 | 246 |

Table S2: Summary of the datasets in this study.

| Platform | Tissue | Section | #Domains | #Spots | #Genes | Related figures |
|-------------|--|-------------------|----------|---------|--------|----------------------------|
| 10x Visium | Human dorsolateral prefrontal cortex (DLPFC) | 151507 | 7 | 4,226 | 33,538 | |
| | | 151508 | 7 | 4,384 | 33,538 | |
| | | 151509 | 7 | 4,789 | 33,538 | |
| | | 151510 | 7 | 4,634 | 33,538 | |
| | | 151669 | 5 | 3,661 | 33,538 | |
| | | 151670 | 5 | 3,498 | 33,538 | Figure 2, 4, |
| | | 151671 | 5 | 4,110 | 33,538 | Figure S1, S2 |
| | | 151672 | 5 | 4,015 | 33,538 | |
| | | 151673 | 7 | 3,639 | 33,538 | |
| | | 151674 | 7 | 3,673 | 33,538 | |
| | | 151675 | 7 | 3,592 | 33,538 | |
| | | 151676 | 7 | 3,460 | 33,538 | |
| | Human breast cancer (BRCA) | \ | 20 | 3,798 | 36,601 | Figure 3, Figure S3, S4 |
| | Mouse brain (MB) | Section1-Anterior | | 2,695 | 32,285 | Figure S5 |
| Stereo-seq | Mouse embryo (ME) | E9.5 | 12 | 5,913 | 23,015 | |
| | | E10.5 | 18 | 8,494 | 22,385 | Figure 5, Figure S6 |
| | | E11.5 | 19 | 30,124 | 26,854 | |
| Slide-seqV2 | Mouse olfactory bulb (MOB) | \ | 7 | 19,109 | 27,106 | Figure 6 |
| | Mouse olfactory bulb (MOB) | \ | 9 | 20,139 | 21,220 | Figure 6 |
| STARmap | Mouse primary visual cortex (MVC) | \ | 7 | 1,207 | 1,020 | Figure 3 |
| osmFISH | Mouse somatosensory cortex (MSC) | \ | 11 | 4,839 | 33 | Figure 3 |
| MERFISH | Mouse hypothalamic preoptic (MHP) | Bregma-0.04 mm | 8 | 5,488 | 155 | |
| | | Bregma-0.09 mm | 8 | 5,557 | 155 | |
| | | Bregma-0.14 mm | 8 | 5,926 | 155 | |
| | | Bregma-0.19 mm | 8 | 5,803 | 155 | |
| | | Bregma-0.24 mm | 8 | 5,543 | 155 | |
| ST | Mouse whole brain | 01A-35A | 3-12 | 152-620 | 23,371 | Figure 7 |

2 Algorithm

Algorithm S1. Spatial domain identification algorithm

The SpaCross spatial domain identification algorithm is described as follows:

Algorithm 1: SpaCross Algorithm for Spatial Domain Identification

Input: Gene expression matrix, spatial location matrix, number of iterations T , hybrid neighbors update step size S , masking rate ρ , graph encoder \mathcal{F}_g , graph predictor \mathcal{F}_p , graph decoder \mathcal{F}_d , discriminator \mathcal{D} .

Output: Spatial domain identification results.

```

1 Obtain the preprocessed feature matrix  $X$ ;
2 Construct the spatial adjacency matrix  $A$ ;
3 for  $t = 0$  to  $T - 1$  do
4   if  $t \bmod S = 0$  then
5     | Update the spatial-semantic hybrid nearest neighbors  $\mathcal{N}^F$  using Eq. (13);
6   end
7   Get masked spot set  $\mathcal{V}_m$  and its complementary masked spot set  $\mathcal{V}_{cm}$ ;
8   Compute the masked feature  $X_m$  and its complementary masked feature  $X_{cm}$ ;
9   Compute graph embeddings  $Z_g = \mathcal{F}_g(A, X_m)$  and  $Z_{cg} = \mathcal{F}_g(A, X_{cm})$ ;
10  Obtain the remasked representation  $Z_m$  from  $Z_g$  via the remasking procedure;
11  Compute the predicted representation  $Z_p = \mathcal{F}_p(A, Z_m)$ ;
12  Reconstruct the feature  $\hat{X} = \mathcal{F}_d(A, Z_p)$ ;
13  Calculate the reconstruction loss  $\mathcal{L}_{SCE}(X, \hat{X})$  according to Eq. (5);
14  Calculate the consistency loss  $\mathcal{L}_{NCE}(Z_{cg}, Z_p)$  according to Eq. (7);
15  Obtain the aggregated summary  $\mathcal{S}$  using Eq. (14);
16  Calculate the contrastive loss  $\mathcal{L}_{BCE}(Z_g)$  using Eq. (15);
17  Compute the total loss  $\mathcal{L}$  as defined in Eq. (17);
18  Update the parameters of  $\mathcal{F}_g$ ,  $\mathcal{F}_p$ ,  $\mathcal{F}_d$ , and  $\mathcal{D}$  by minimizing  $\mathcal{L}$ ;
19 end
20 Compute the final embedding  $Z = \mathcal{F}_g(A, X)$ ;
21 Apply Mclust to  $Z$ ;
22 return Clustering results, where each cluster corresponds to a spatial domain;
```

3 Text

Text S1. Ablation studies

To investigate the contribution of different components to SpaCross, we conducted comprehensive ablation studies by designing four architectural variants and evaluating their impacts on model performance using the DLPFC dataset. The first variant, designated as **w/o-Mask**, disables the mask mechanism by substituting the masked node set \mathcal{V}_m with all nodes \mathcal{V} during loss computation, thereby deactivating the Cross-Masked Latent Consistency (CMLC) module. The **w/o-CMLC** variant eliminates both the CMLC module and its associated consistency loss \mathcal{L}_{NCE} . The third variant, **w/o-AHSG**, removes the Adaptive Hybrid Spatial-Semantic Graph (AHSG) component and its contrastive learning framework. The final variant, **w/o-Predictor**, bypasses the graph predictor \mathcal{F}_p by directly utilizing graph embeddings in the CMLC module.

Our experimental findings reveal three key insights: (1) The **w/o-Mask** variant exhibits significant performance degradation, demonstrating that the mask mechanism crucially enhances representation robustness and reconstruction efficacy through selective node masking. (2) While the **w/o-CMLC** variant shows marginally superior accuracy to **w/o-Mask**, both underperform the full model, suggesting that the CMLC module effectively mitigates convergence to local optima through cross-mask consistency regularization. (3) The **w/o-AHSG** variant achieves the second-best performance, highlighting AHSG’s dual role in harmonizing spatial continuity with semantic coherence to generate discriminative contrastive samples. Notably, the **w/o-Predictor** variant reveals the graph predictor’s importance in enabling indirect supervision of latent representations through \mathcal{L}_{NCE} computation, with its removal causing suboptimal feature learning.

These systematic ablation analyses quantitatively validate the complementary strengths of SpaCross’s components: The mask mechanism and CMLC module jointly ensure robust representation learning, while AHSG effectively bridges spatial and semantic information hierarchies. The graph predictor further enhances this framework by enabling progressive refinement of latent embeddings through contrastive learning objectives.

To systematically evaluate the design efficacy of core components in SpaCross, we conducted critical analyses on the Adaptive Hybrid Spatial-Semantic Graph (AHSG), Cross-Masked Latent Consistency (CMLC) module, and loss functions. First, for the AHSG architecture, we validated the importance of dual-layer neighborhood integration through three controlled experiments: (1) **AHSG w/o-Spatial** retained only semantic neighbors while removing spatial adjacency; (2) **AHSG w/o-Semantic** preserved spatial connections but eliminated semantic similarity constraints; (3) **Rnd-Walk-Neg.** replaced the original corrupt function-based negative sampling with random walk sampling. Experimental results showed comparable performance between the **Rnd-Walk-Neg.** variant and the baseline model, indicating minimal impact from random negative sampling—a phenomenon attributable to the substantial excess

of spot counts over neighborhood sizes in the dataset, which ensures high probability of selecting non-adjacent nodes through random sampling. In contrast, both **w/o-Spatial** and **w/o-Semantic** variants exhibited significant declines in ARI and ACC metrics, conclusively demonstrating that the hybrid spatial-semantic graph structure enhances spatial domain recognition through synergistic optimization of spatial proximity and semantic consistency.

For the CMLC module, we investigated the impacts of masking mechanisms and negative sample selection: (1) **CMLC w/o-Mask** disabled node masking and directly utilized full-node embeddings for latent space supervision; (2) **CMLC w/o-Neg.** excluded negative sample contrastive terms in loss computation. Both variants caused substantial performance fluctuations, manifested as dramatic variance increases in ARI and ACC metrics across all 12 tissue slices. Additionally, the graph predictor variant **Pre-w/o-remask** without remasking technology was found to disrupt consistent latent representation learning.

Finally, we systematically assessed the influence of different loss functions: (1) Replacing binary cross-entropy (BCE) loss with triplet loss in AHSG; (2) Substituting noise contrastive estimation (NCE) loss with mean squared error (MSE) loss in CMLC; (3) Employing MSE loss instead of the original design for mask reconstruction tasks. Results revealed that loss function selection directly affects the model’s discriminative capacity for spatial domain features, where contrastive learning-based NCE loss and task-adapted BCE loss more effectively guide the learning of discriminative representations.

Text S2. Alignment of multiple consecutive slices

SpaCross extends single-slice data to continuous multi-slice data by aligning multiple slices through spatial point registration, facilitating batch correction using SpaCross. Prior to alignment, edge points in each slice are detected based on the number of neighboring points. For instance, in datasets acquired using the Visium platform, non-edge points are expected to have six neighboring points. For robustness, we define an edge point as one that has fewer than five neighboring points and more than one. Subsequently, SpaCross employs the Iterative Closest Point (ICP) algorithm to align edge points from two slices in pairs, assuming that edge points are typically located at the boundaries of tissues¹. Let $P = \{p_1, p_2, \dots, p_m\}$ represent the set of edge point locations in the source slice, and $Q = \{q_1, q_2, \dots, q_k\}$ denote the set of edge point locations in the target slice, where $p_i, q_i \in \mathbb{R}^2$. The ICP algorithm aligns the points by iteratively executing the following steps until convergence:

1. For each point in the source point cloud P , find its closest point in the target point cloud Q to form a new set Q' . Then, $Q' = \{q'_1, q'_2, \dots, q'_m\}$ containing points in the target point cloud is obtained.
2. Find the transformation with the optimal rotation and translation by solving the least squares problem.

$$\hat{R}, \hat{v} = \arg \min_{R \in \mathbb{R}^{2 \times 2}, v \in \mathbb{R}^2} \sum_{i=1}^m \| (Rp_i + v) - q'_i \|^2, \quad (1)$$

subject to $R^T R = I$.

3. Apply the transformation $\hat{R}p_i + \hat{v}$ to the source point cloud P .

For multiple slices, SpaCross performs pairwise alignment with the ICP algorithm sequentially.

Text S3. Details on comparison with other methods

In this study, we selected a range of representative state-of-the-art methods. For spatial domain identification in single slices, we chose two methods that leverage histological image features combined with spatial expression and coordinate information: SpaGCN and DeepST. Additionally, we included two reconstruction-based approaches: SEDR, which utilizes a deep autoencoder structure, and STAGATE, which employs an adaptive attention mechanism. We also incorporated two contrastive learning-based methods: CCST, based on the Deep Graph Infomax (DGI) approach, and GraphST, which uses DGI while also reconstructing the original features. Finally, we included DiffusionST, a method based on a diffusion model. For spatial domain identification across multiple slices, we adopted STAaligner, Splane, STitch3D, and SPIRAL as baseline methods. The experimental setup is as follows:

- **SpaGCN** ² integrates gene expression data, histological images, and spatial location information through Graph Convolutional Networks (GCNs) to identify variable genes within spatial domains. It employs unsupervised clustering algorithms to detect distinct spatial expression patterns. SpaGCN was applied to the entire dataset following the recommended parameters outlined in the package vignette.
- **DeepST** ³ is a spatial clustering method that utilizes graph autoencoders, taking gene expression data, histological images, and spatial information as inputs. It employs GCNs as encoders to reconstruct the input graph topology and identify features pertinent to spatial domain delineation. We retained the default parameters of DeepST, adjusting the parameter k to 12 for computing the adjacency matrix within the *get_graph* function. Additionally, because the human melanoma dataset lacks corresponding histological image data, we excluded histological information during testing.
- **CCST** ⁴ integrates spatial structure and gene expression information by stacking multiple layers of GCNs. It constructs positive and negative pairs using the maximum mutual information graph contrastive learning method to facilitate spatial transcriptomics analysis. CCST preprocessed the initial gene expression matrix using Principal Component Analysis (PCA) on all spatial transcriptomics data, selecting the top 200 principal components as input dimensions. The selection of the radius ensured that each spot had between 5 and 10 neighbors. The overall CCST model utilized reference parameters provided in the original paper. Finally, the latent representations obtained were subjected to PCA to select the top 30 principal components, followed by the application of the mclust method to evaluate clustering performance.

- **SEDR**⁵ employs a deep autoencoder network to learn gene representations while simultaneously embedding spatial information using a variational graph autoencoder. We executed SEDR for all experiments using the recommended parameters from the online tutorial (*run_SEDR_DLPFC_all_data.py*). Specifically, the parameter k was selected within the range of 6 to 10 based on clustering performance, and the number of epochs was set to 200.
- **STAGATE**⁶ utilizes adaptive graph attention autoencoders to integrate spatial information with gene expression, thereby enhancing the accuracy of spatial domain identification. During preprocessing, STAGATE selected the top 3000 highly variable genes for each dataset and configured the radius to ensure that each spot had between 6 and 10 neighbors. All other parameters were consistent with those specified in the original publication.
- **GraphST**⁷ employs a graph-based self-supervised contrastive learning approach that combines graph neural networks with self-supervised contrastive learning. By minimizing the embedding distance between spatially adjacent spots, it learns enriched and discriminative spot representations, and vice versa. Other parameters were set following the original paper’s guidelines, with platform-specific models selected for datasets from platforms like Stereo and 10x Visium. After training, the latent representations for each spot are obtained, and the *refine_label()* function is used to optimize the clustering results. Additionally, GraphST provides processing methods for clustering across continuous multi-slice datasets, which we implemented consistently with the original paper’s procedures.
- **DiffusionST**⁸ utilizes a graph convolutional network model combined with a newly designed loss function and employs the zero-inflated negative binomial (ZINB) distribution for data denoising, followed by data augmentation using a diffusion model. DiffusionST is applied for spatial transcriptomics data imputation, with the imputed results validated by clustering performance. Initially, raw gene expression data undergoes filtering and normalization, with a threshold of 4,096 selected for feature genes. After obtaining the preprocessed gene expression count matrix, DenoiseST is used for feature learning, where a GCN model integrates each spot’s features with those of its neighbors, and a ZINB model learns features for each gene to better fit the ZINB distribution. The data is then augmented using diffusion. Following clustering, the *refine_label()* function is also applied to optimize the clustering results.
- **stDCL**⁹ employs a dual graph contrastive learning framework that integrates graph embedding autoencoders with adaptive fusion of gene expression and spatial information. By constructing dual graphs to model both gene regulation networks and spatial proximity, stDCL learns low-dimensional embeddings that preserve structural coherence and cluster consistency through contrastive alignment. The method enforces similarity between embeddings of spatially adjacent spots

while maximizing dissimilarity for non-adjacent ones, ensuring the learned representations capture nuanced spatial heterogeneity. Gene expression and spatial graphs are processed through a shared autoencoder, with adaptive attention mechanisms weighting their contributions to the final embeddings. Parameters were configured according to the original study, including platform-specific graph construction protocols for Stereo-seq, 10x Visium, and Slide-seq data. After training, the embeddings are clustered using a graph-aware refinement module (*cluster_refine()*) to resolve boundary ambiguities and enhance domain coherence.

- **STAligner** ¹⁰ utilizes STAGATE to learn latent representations and identifies MNN-pairs (mutual nearest neighbors) from this latent space. By applying a triplet loss function, it mitigates batch effects across multiple slices, enhancing multi-slice clustering performance. STAligner first integrates multiple slices and performs pretraining with STAGATE to learn the latent features of each spot. It then applies the MNN algorithm to construct positive and negative anchor pairs. Using triplet loss, it pulls similar anchor pairs closer across slices, thereby reducing batch effects between slices.
- **Splane** ¹¹ uses Spoint to obtain the composition of labeled cell types from scRNA-seq data, which then serves as input for Splane. Spoint constructs a variational autoencoder (VAE) composed of encoder and decoder layers. Cells from the scRNA-seq dataset are used to simulate pseudo-spots, with each pseudo-spot’s cell-type composition determined by its constituent cells. The deconvolution results from Spoint are subsequently input into Splane, which utilizes a combination of a GCN and adversarial learning to address batch effects across multiple slices. Splane operates in two stages: first, a discriminator is trained to develop discriminative capabilities, and second, features from different batches are mixed to mitigate batch effects across slices.
- **STitch3D** ¹ maximizes overlap across multiple slices using ICP or PASTE algorithms, reconstructing gene expression from inferred cell compositions via scRNA-seq data. Input includes aligned slices with matching scRNA-seq, forming 3D spatial coordinates and a global neighborhood graph. STitch3D is then trained to integrate data across slices, creating a shared latent space that captures biological variations and corrects batch effects. In this latent space, each spot representation supports spatial domain recognition and cell-type deconvolution. STitch3D maps gene expression and spatial data into this shared space using a graph attention network. A discriminator further infers cell-type proportions from latent representations, enabling gene expression reconstruction by combining these proportions with cell-type-specific gene profiles.
- **SPIRAL** ¹² integrates data from both feature space (including low-dimensional embeddings and high-dimensional gene expression) and physical space. It merges GraphSAGE and domain adaptation into a unified model to learn corrected embeddings and expression profiles by combining transcriptomic profiles with spatial context. GraphSAGE functions as the encoder, embedding

gene expression and spatial coordinates into a low-dimensional latent space. A noise classifier and a biological discriminator are then designed to decompose these low-dimensional embeddings into two components: the noise component, used to identify batch-specific noise, and the biological component, used to integrate signals across batches. A decoder network reconstructs gene expression from these embeddings. The biological component of the low-dimensional embeddings is further grouped to reveal spatial domains across multiple samples.

4 Figure

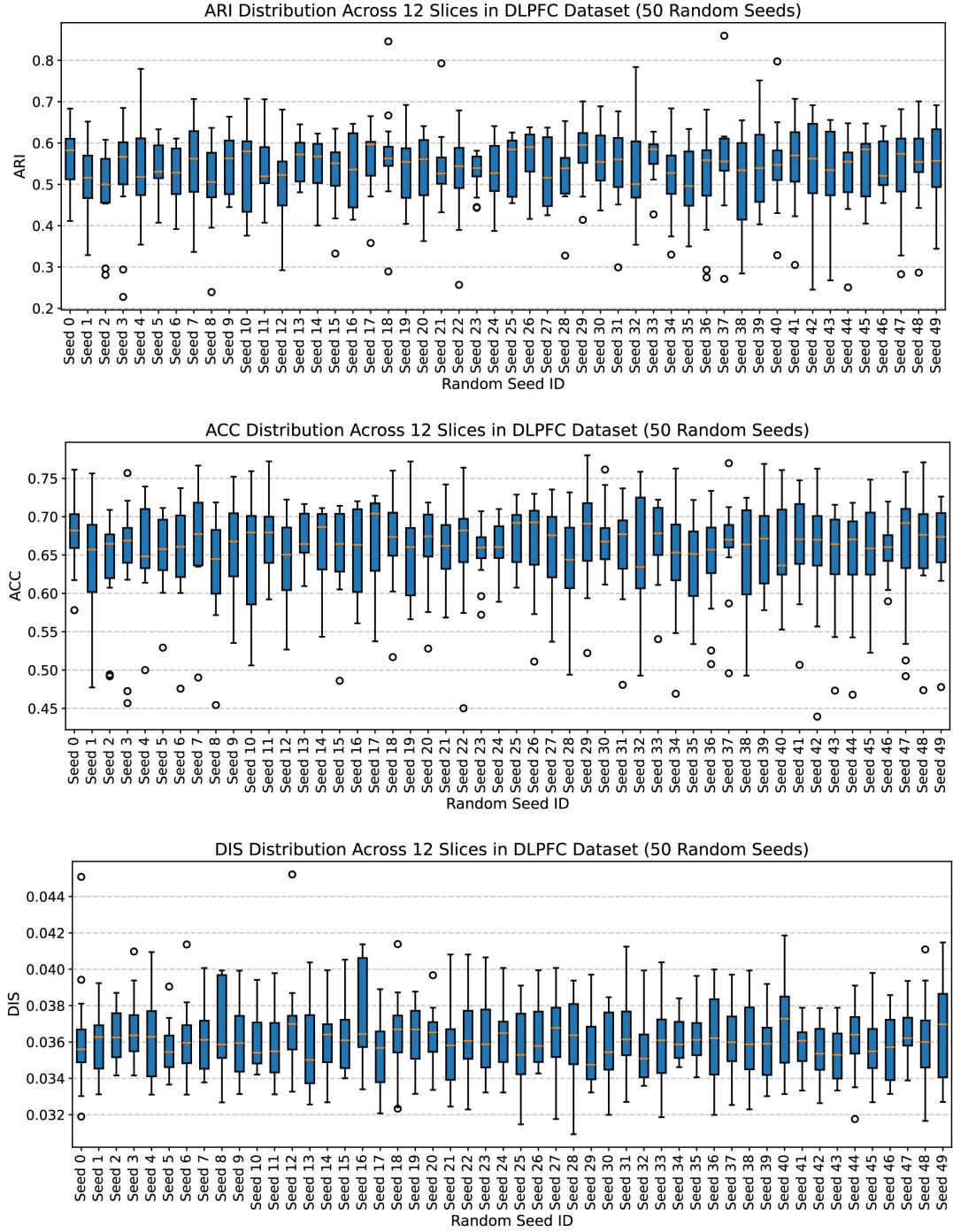
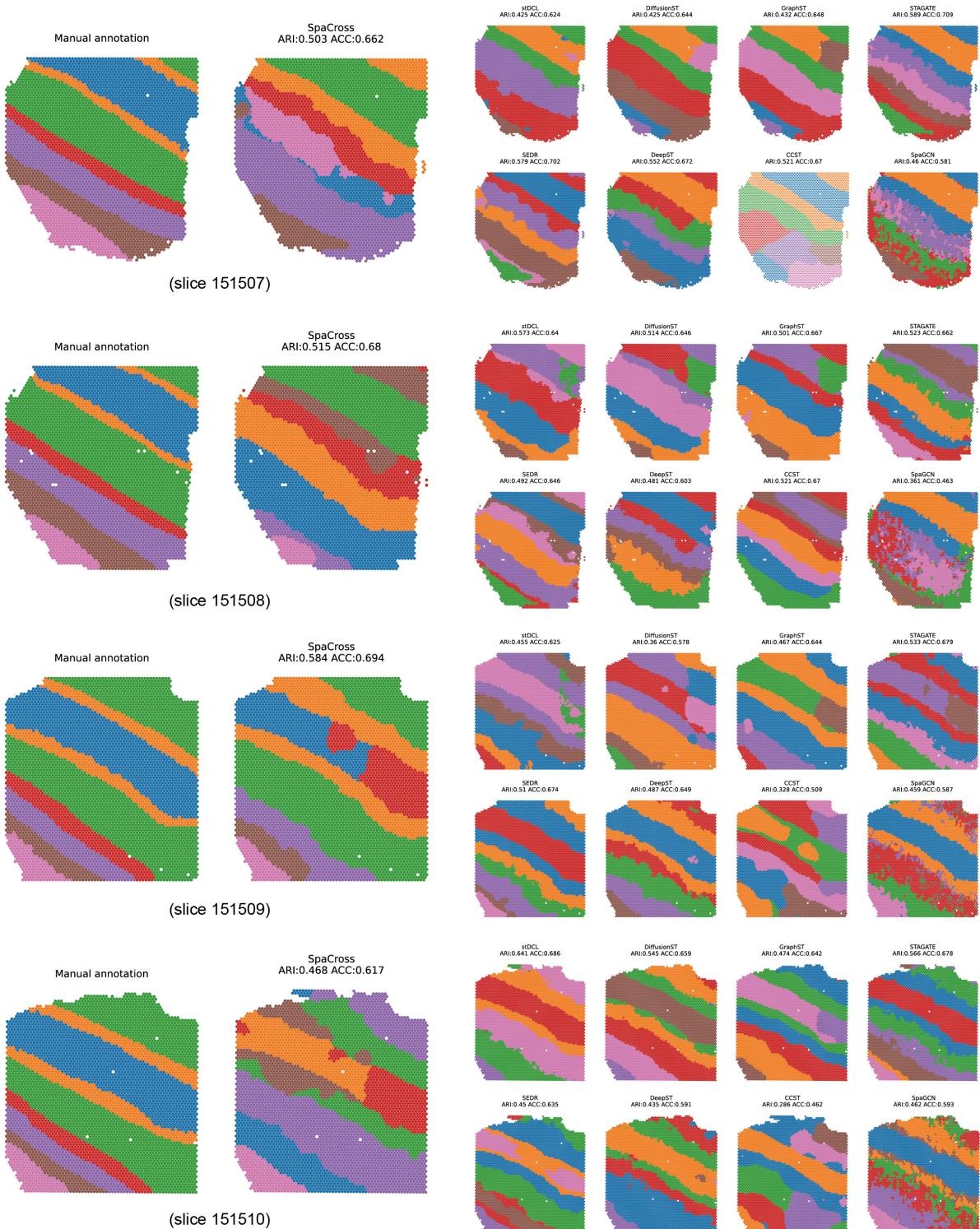
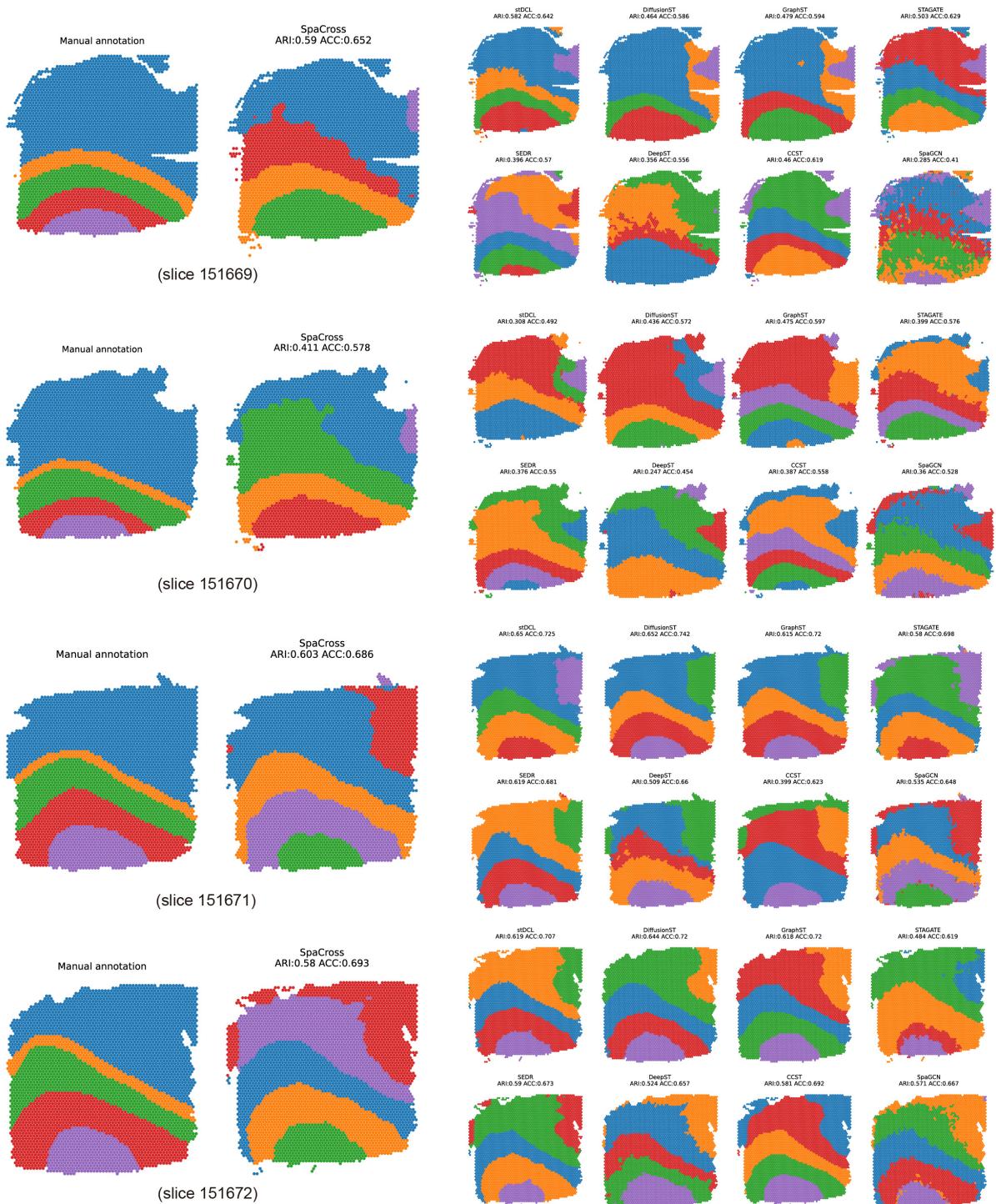


Figure S1: The clustering accuracy of SpaCross under the default hyperparameters with different random seeds.



See next page



See next page

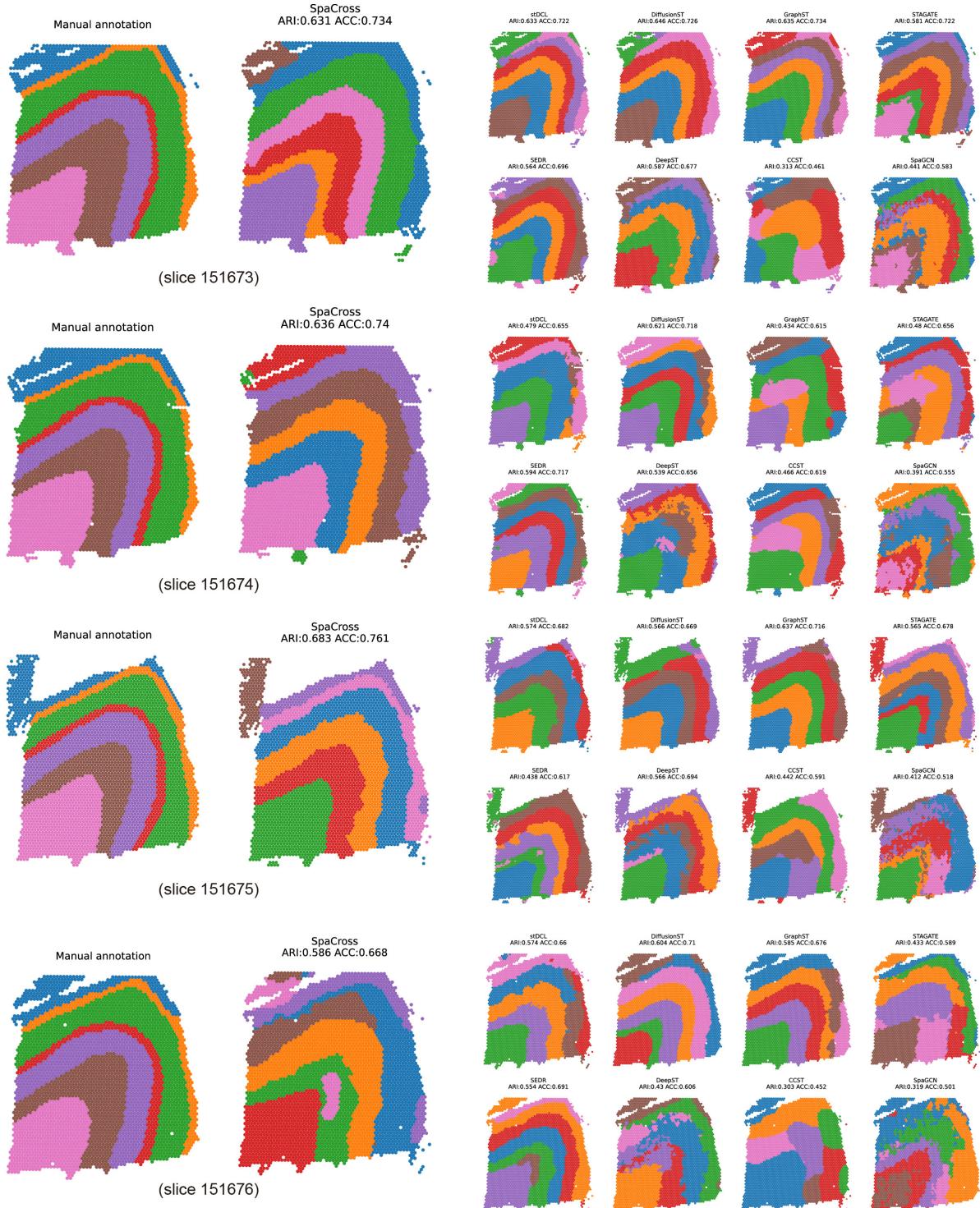


Figure S2: Comparison of spatial domains by clustering assignments via SpaCross, various methods, and manual annotation in all 12 sections of the DLPFC dataset.

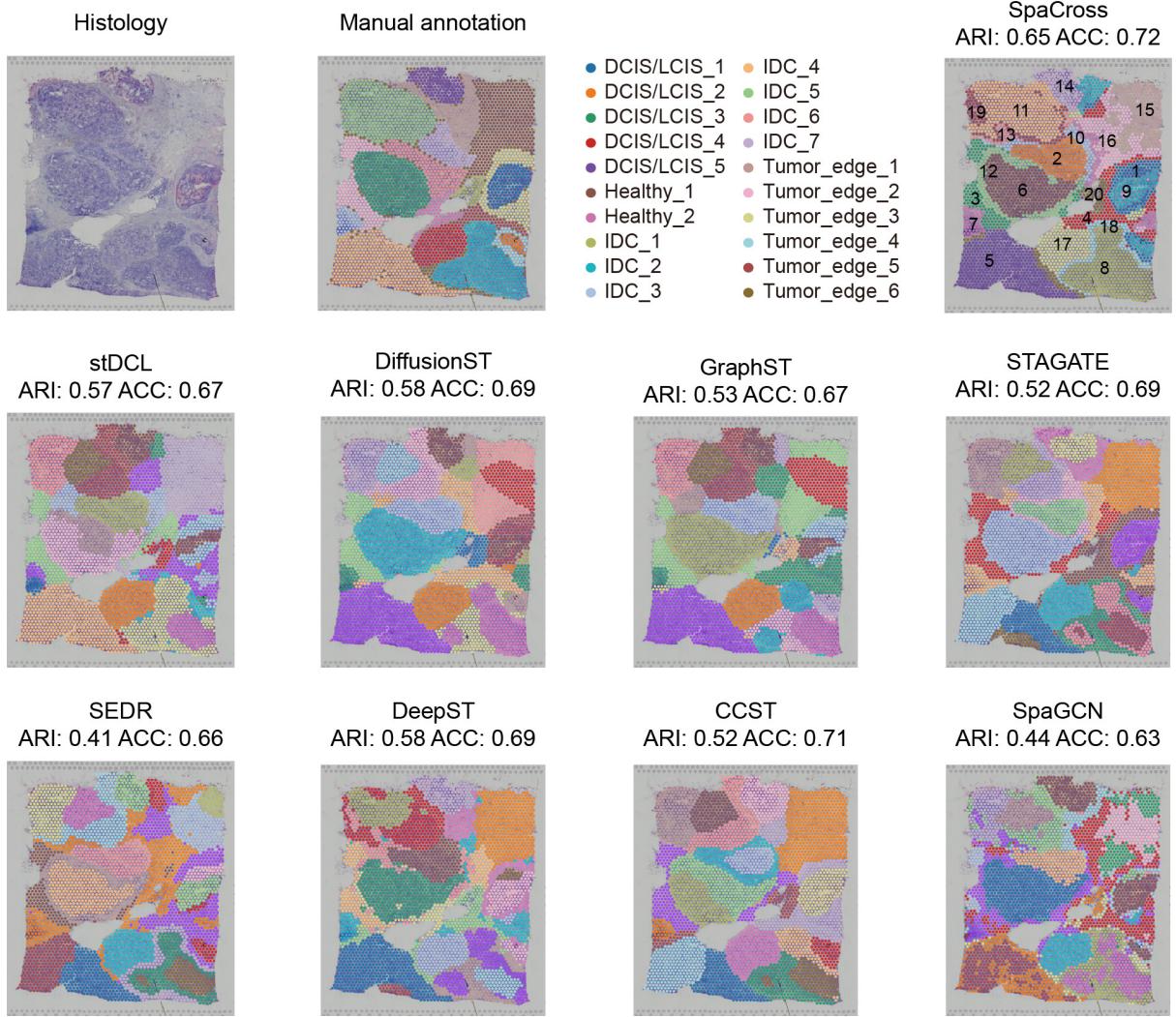


Figure S3: Comparison of spatial domains by clustering assignments via SpaCross, various methods, and manual annotation in the human brain cancer dataset.

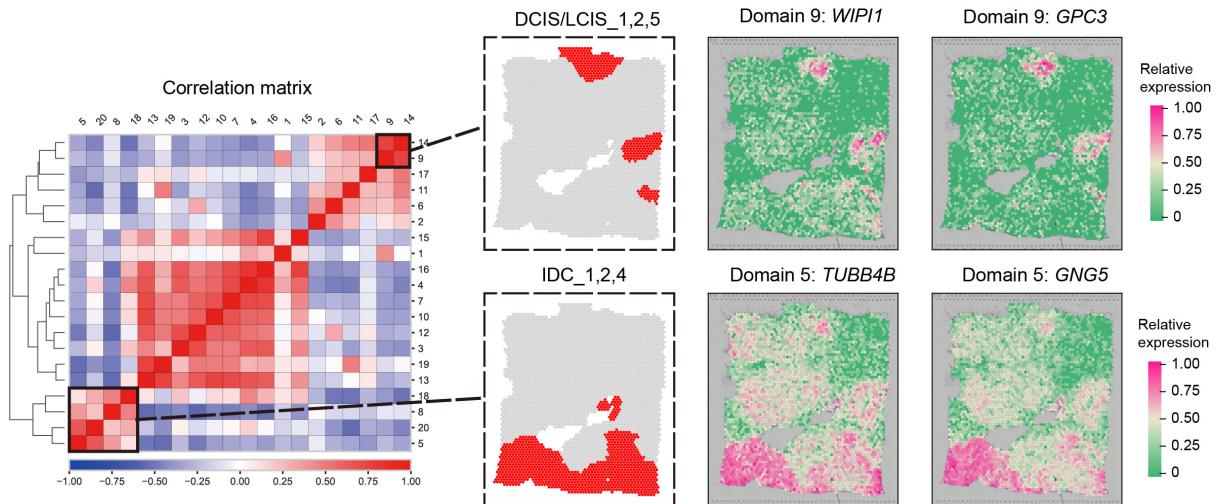


Figure S4: SpaCross identifies a correlation matrix between spatial domains and detects SVGs through visualization of regions exhibiting the highest correlations in the BRCA dataset. The analysis reveals strong concordance between the identified SVGs and the observed spatial domain patterns, demonstrating the method's capability in uncovering biologically meaningful spatial relationships.

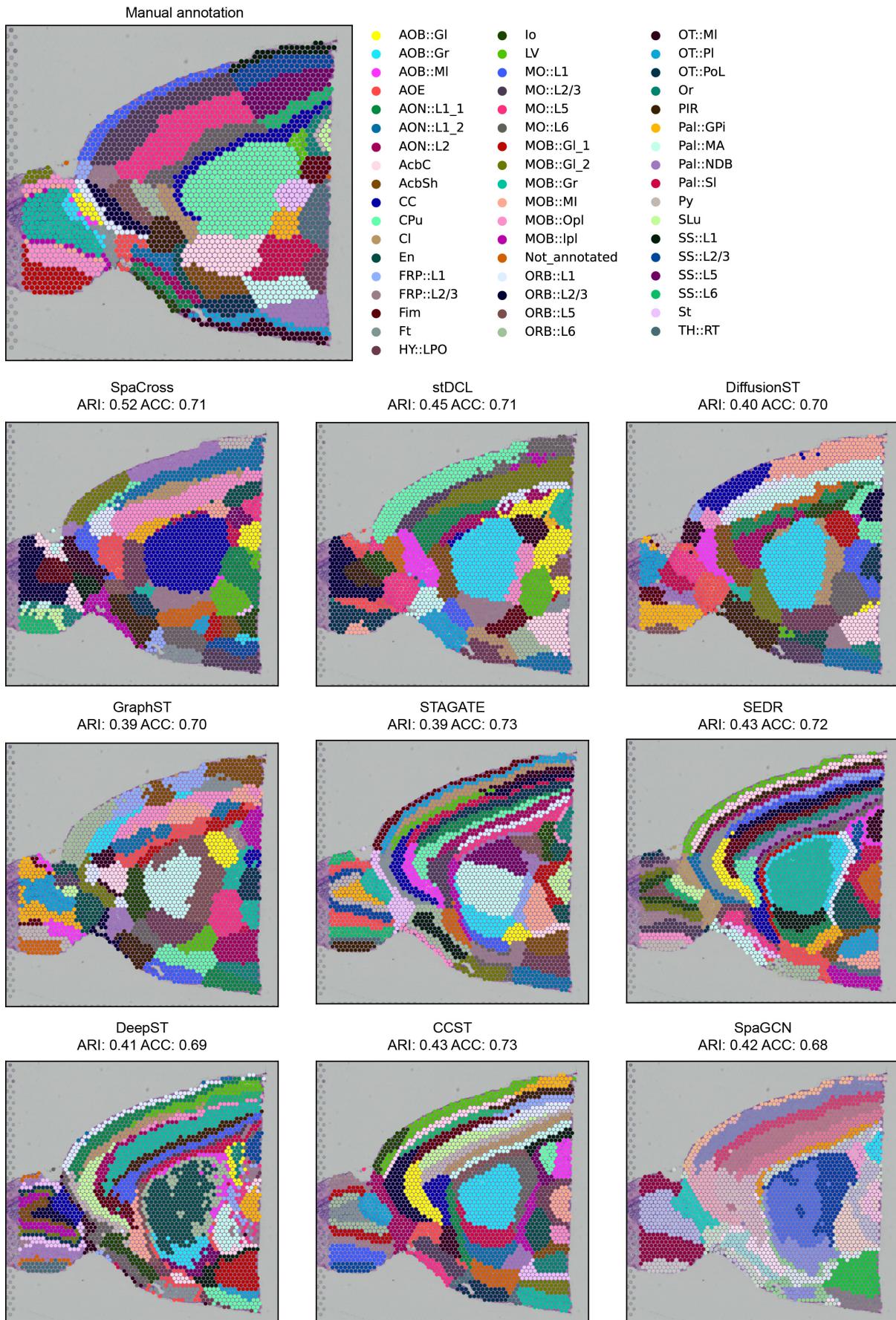


Figure S5: Comparison of spatial domains by clustering assignments via SpaCross, various methods, and manual annotation in the mouse brain dataset.

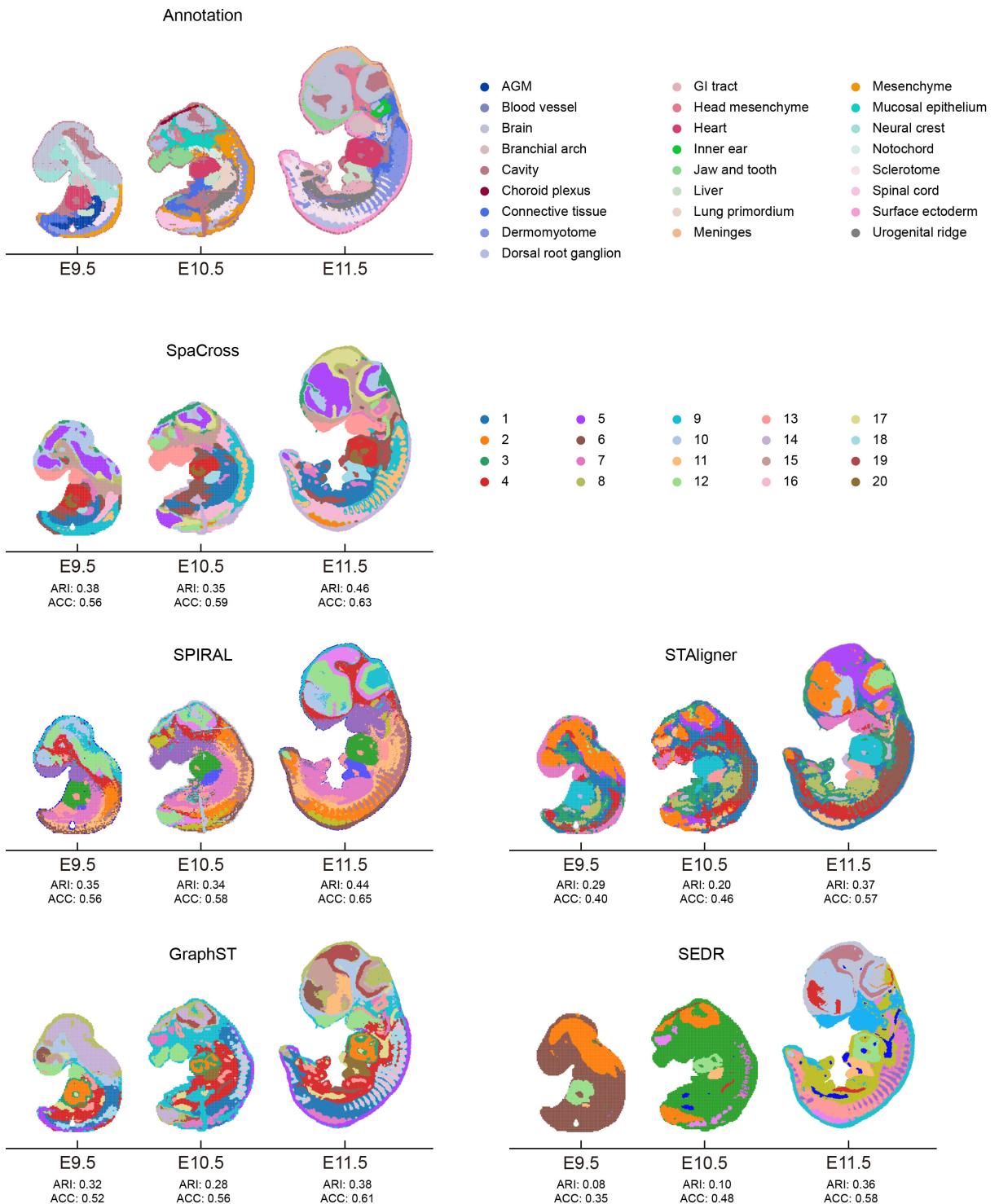


Figure S6: Comparison of spatial domains by clustering assignments via SpaCross, various methods, and manual annotation in the mouse embryo dataset.

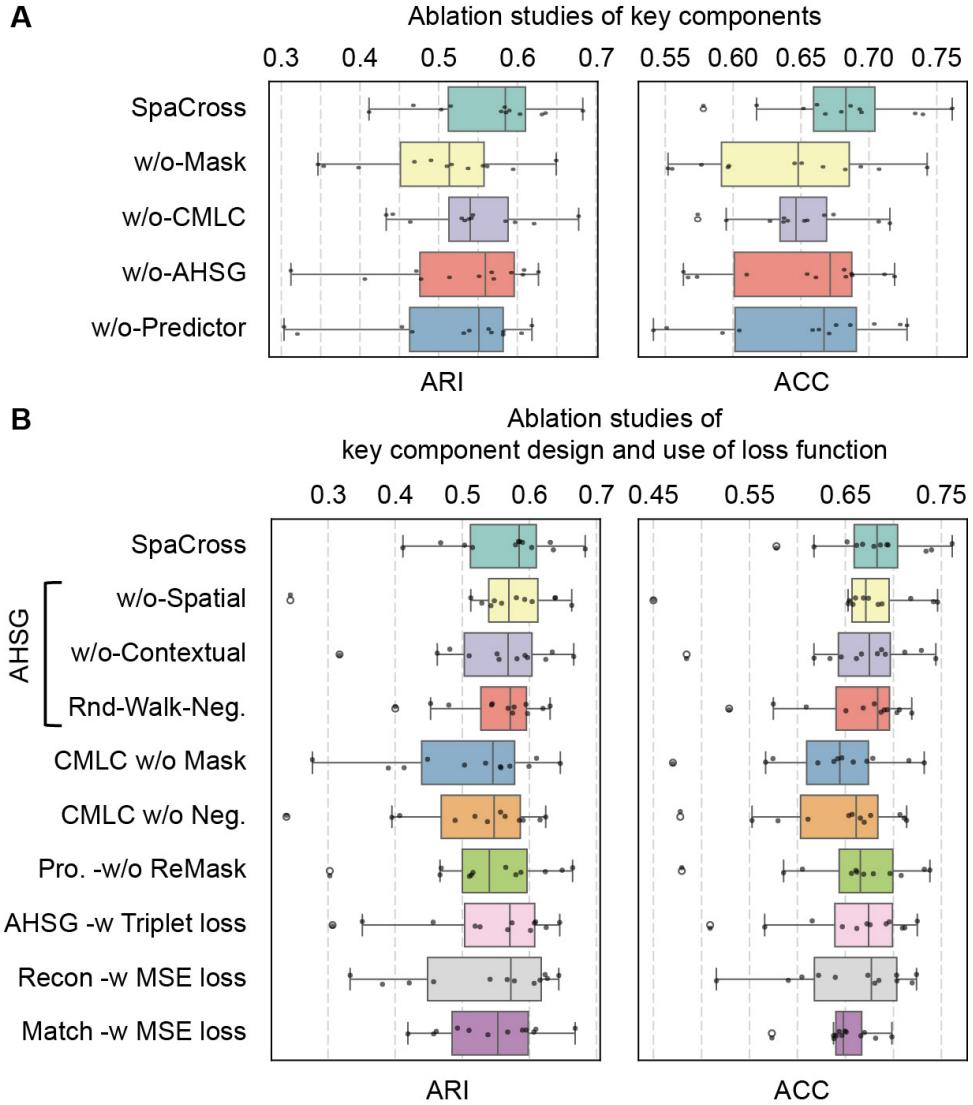


Figure S7: Ablation studies. (A). Ablation studies of core architectural components. Performance impacts of removing key modules (CMLC, AHSG, mask mechanism, graph predictor) across variants (w/o-Mask, w/o-CHSG, etc.), validating their complementary roles in representation learning and spatial domain discovery. (B). Component-wise analysis of module designs and loss functions. Evaluation of AHSG, CMLC masking/negative sampling strategies, and loss function choices (w/o-Spatial, w/o-Semantic, CMLC w/o-Mask, etc.), demonstrating critical dependencies on hybrid graph structures and contrastive learning objectives.

References

- [1] Gefei Wang, Jia Zhao, Yan Yan, Yang Wang, Angela Ruohao Wu, and Can Yang. Construction of a 3d whole organism spatial atlas by joint modelling of multiple slices with deep neural networks. *Nature Machine Intelligence*, 5(11):1200–1213, 2023.
- [2] Jian Hu, Xiangjie Li, Kyle Coleman, Amelia Schroeder, Nan Ma, David J Irwin, Edward B Lee, Russell T Shinohara, and Mingyao Li. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature Methods*, 18(11):1342–1351, 2021.
- [3] Chang Xu, Xiyun Jin, Songren Wei, Pingping Wang, Meng Luo, Zhaochun Xu, Wenyi Yang, Yideng Cai, Lixing Xiao, Xiaoyu Lin, et al. DeepST: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Research*, 50(22):e131–e131, 2022.
- [4] Jiachen Li, Siheng Chen, Xiaoyong Pan, Ye Yuan, and Hong-Bin Shen. Cell clustering for spatial transcriptomics data with graph neural networks. *Nature Computational Science*, 2(6):399–408, 2022.
- [5] Hang Xu, Huazhu Fu, Yahui Long, Kok Siong Ang, Raman Sethi, Kelvin Chong, Mengwei Li, Rom Uddamvathanak, Hong Kai Lee, Jingjing Ling, et al. Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Medicine*, 16(1):12, 2024.
- [6] Kangning Dong and Shihua Zhang. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nature Communications*, 13(1):1739, 2022.
- [7] Yahui Long, Kok Siong Ang, Mengwei Li, Kian Long Kelvin Chong, Raman Sethi, Chengwei Zhong, Hang Xu, Zhiwei Ong, Karishma Sachaphibulkij, Ao Chen, et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nature Communications*, 14(1):1155, 2023.
- [8] Yaxuan Cui, Yang Cui, Ruheng Wang, Kenta Nakai, Xiucai Ye, Tetsuya Sakurai, and Leyi Wei. DiffusionST: A diffusion model-based framework for enhancing spatial transcriptomics data quality and identifying spatial domains. *Available at SSRN 4894131*, 2024.
- [9] Zhuohan Yu, Yuning Yang, Xingjian Chen, Ka-Chun Wong, Zhaolei Zhang, Yuming Zhao, and Xiangtao Li. Accurate spatial heterogeneity dissection and gene regulation interpretation for spatial transcriptomics using dual graph contrastive learning. *Advanced Science*, 12(3):2410081, 2025.
- [10] Xiang Zhou, Kangning Dong, and Shihua Zhang. Integrating spatial transcriptomics data across different conditions, technologies and developmental stages. *Nature Computational Science*, 3(10):894–906, 2023.

- [11] Hao Xu, Shuyan Wang, Minghao Fang, Songwen Luo, Chunpeng Chen, Siyuan Wan, Rirui Wang, Meifang Tang, Tian Xue, Bin Li, et al. SPACEL: deep learning-based characterization of spatial transcriptome architectures. *Nature Communications*, 14(1):7603, 2023.
- [12] Tiantian Guo, Zhiyuan Yuan, Yan Pan, Jiakang Wang, Fengling Chen, Michael Q Zhang, and Xiangyu Li. SPIRAL: integrating and aligning spatially resolved transcriptomics data across different experiments, conditions, and technologies. *Genome Biology*, 24(1):241, 2023.