Unveiling spatial domains from spatial multi-omics data using dual-graph regularized ensemble learning

Check for updates

Ying Li[1], Guangchang Cai[1], Fuqun Chen[1], Kepei Wen[1] & Le Ou-Yang [2] ✉

With the rapid advancement of spatial multi-omics technologies, the simultaneous analysis of molecular profiles and spatial locations has provided unprecedented insights into cellular heterogeneity and tissue microenvironments. However, data sparsity and the diversity of data distributions hinder the effective integration and analysis of spatial multi-omics data. In this study, we propose a novel ensemble learning framework based on dual-graph regularized anchor concept factorization, named SMODEL, for detecting spatial domains from spatial multi-omics data. SMODEL employs an element-wise weighted ensemble strategy to integrate multiple base clustering results, and leverages anchor concept factorization and dual-graph regularization to learn robust spatial consensus representations. We evaluated the performance of SMODEL on both real and simulated spatial multi-omics datasets, encompassing various technologies, tissue types, and species. Experimental results demonstrate that SMODEL not only outperforms existing methods in spatial domain identification but also effectively captures tissue structure, thereby enhancing the understanding of cellular heterogeneity.

Cell heterogeneity is a fundamental characteristic of the tissue microenvironment and plays a key role in physiological and pathological processes[1]. Accurate identification of cell types is a crucial step in understanding cell heterogeneity and serves as the foundation for downstream analyses. With the development of single-cell sequencing and spatial omics sequencing technologies, we can use clustering analysis methods to categorize cells in tissues into different clusters based on their molecular characteristics. These advanced techniques enable us to systematically reveal the distribution patterns of cells within tissues and their potential biological significance[2].

The advancement of single-cell sequencing techniques has enabled the simultaneous analysis of multiple molecular modalities, including mRNA, chromatin accessibility, and DNA methylation, within individual cells[3]. Compared to single-cell mono-omics data, single-cell multi-omics data provide a more comprehensive view of cellular heterogeneity, offering valuable insights into the complexity of biological systems[4–6]. Different omics data provide complementary insights on cells and their microenvironments. For example, transcriptomics focuses on gene transcription levels, including mRNA and non-coding RNA, reflecting the gene expression status of cells. Proteomics analyzes the composition, abundance, modifications, and interactions of proteins, elucidating protein functions.

Metabolomics investigates the types and quantities of small molecular metabolites within cells, reflecting the metabolic state of cells and their responses to environmental stimuli. Epigenomics, on the other hand, explores DNA methylation, histone modifications, and chromatin accessibility, revealing the regulatory mechanisms of gene expression. However, relying solely on single-omic data often fails to comprehensively and accurately capture the subtle differences and heterogeneity among cells.

Researchers have developed various computational methods for clustering multi-omics data, which can be broadly categorized into traditional methods and deep learning-based methods. Traditional methods primarily rely on statistical models for modeling. For example, MOFA+ is a statistical framework designed for integrating single-cell multimodal data, using variational inference to reconstruct low-dimensional representations of the data, thereby enabling joint modeling of variations across multiple sample groups and data modalities[7]. Seurat v5 leverages graph theory to connect cells with similar patterns by learning cell-specific weights, constructing a weighted nearest-neighbor graph to achieve multi-omics data fusion[8]. Other traditional methods include GRMEC-SC[9] and Mowgli[10]. On the other hand, deep learning has emerged as a powerful tool for integrating multi-omics data. For instance, totalVI uses a modeling strategy similar to scVI[11] to integrate CITE-seq data, which encompasses both RNA and cell surface

[1]Guangdong Key Laboratory of Intelligent Information Processing, College of Electronic and Information Engineering, Shenzhen University, Shenzhen, 518060 Guangdong, China. [2]SMBU-MSU-BIT Joint Laboratory on Bioinformatics and Engineering Biology, Shenzhen MSU-BIT University, Shenzhen, 518172 Guangdong, China. ✉e-mail: leouyang@smbu.edu.cn

protein information, providing a comprehensive and scalable framework for multi-omics data integration through variational inference[12]. MultiVI integrates multimodal information from single-cell transcriptomics, chromatin accessibility, and other molecular characteristics based on deep generative modeling principles, creating a joint latent space for paired and unpaired samples[13]. scMIC integrates cell attribute information and structural relationships between cells from both local and global perspectives based on deep multi-level information fusion principles, reducing redundant information between different omics at the cellular and feature levels[14]. Other deep learning-based methods for multi-omics data integration include scMDC[15], BABEL[16], and scMM[17]. However, these methods are designed for single-cell multi-omics data and overlook the crucial spatial information related to tissue structure and the microenvironment context.

Spatial location information plays a critical role in cell type identification, as cells of the same type often exhibit specific spatial distribution patterns closely related to particular functional domains[18]. These spatial distribution patterns, known as spatial domains, are regions within a tissue where cells with similar molecular profiles and functions are spatially organized[19,20]. Spatial domains can reflect the functional states of cells and their interactions within the tissue microenvironment. For example, in the immune microenvironment, the spatial distribution patterns of different immune cell subpopulations may directly reflect their functional states. Compared to single-cell sequencing technology, spatial multi-omics sequencing integrates molecular profiles with spatial location, enabling researchers to investigate cellular heterogeneity within the spatial context of tissues[21]. Currently, various technologies can simultaneously measure spatial multi-omics data on the same tissue section, such as DBiT-seq[22], CUT&Tag-RNA-seq[23], SPOTS[24], MERSCOPE[25], and Nanostring CosMx[26]. Recently, various computational methods have been developed for cell clustering and spatial domain identification in spatial multi-omics data. For example, SpatialGlue uses graph neural networks with dual attention mechanisms to integrate data modalities and reveal histologically relevant structures[27]. PRAGA uses dynamic graphs and prototype contrastive learning for spatial data integration[28]. COSMOS utilizes graph neural networks to extract complementary information from two omics, enabling the integration of spatial multi-omics data to identify spatial domains[29]. MISO can integrate multiple modalities from various spatial omics datasets and serves as a versatile algorithm for feature extraction and clustering[30]. However, despite the progress made by these methods, finding a single approach that consistently performs well across all scenarios remains challenging. While ensemble learning shows promise for improving algorithm performance and stability, existing approaches rarely consider the reliability of results from different clustering methods, leaving them susceptible to being influenced by the outcomes of poor-performing methods[31].

To address these challenges, we propose a novel ensemble learning framework, named SMODEL, for identifying spatial domains from spatial multi-omics data. First, to ensure robust performance across diverse scenarios, we introduce an element-wise weighted ensemble strategy that comprehensively integrates the results from multiple base clustering methods, enhancing the accuracy and robustness of spatial domain identification. To reduce excessive reliance on base clustering outcomes, we project multi-omics data with varying features into a shared low-dimensional representation using anchor concept factorization. This approach reduces redundant information caused by noise and adaptively captures the diversity and complementary relationships among different data modalities. Finally, to address the complex manifold structure of spatial data and the nonlinear relationships between cells[32,33], we incorporate high-order information between cells and preserve the geometric structure of the original data manifold through graph regularization. This effectively mitigates issues related to spatial smoothing and ensures the integrity of spatial relationships.

We evaluated the robustness and versatility of SMODEL on diverse spatial multi-omics datasets from various platforms, including spatial transcriptomics, proteomics, and epigenomics. Experimental results demonstrate that SMODEL outperforms existing methods in identifying medullary structures within human lymph node datasets. By effectively leveraging complementary information from spatial proteomics and transcriptomics, SMODEL provides deeper insights into the tissue microenvironment, particularly in applications such as breast cancer tissue analysis. In addition, SMODEL enhances the interpretation of spatial gene expression patterns in spatial epigenomic-transcriptomic datasets and effectively integrates spatial triple-omics data while maintaining interpretability. It excels at uncovering consistent and complementary information across different omics. These results highlight the remarkable capability of SMODEL in integrating spatial multi-omics data, revealing cellular heterogeneity, and advancing our understanding of tissue structure and function.

Results
Overview of the SMODEL framework
In this study, we developed SMODEL, an ensemble learning algorithm that integrates spatial multi-omics data in a spatially informed manner to analyze cellular heterogeneity. The schematic representation of SMODEL is illustrated in Fig. 1 (detailed descriptions are provided in the "Methods" section). SMODEL is based on the ensemble learning framework, employing the spatial neighborhood structure between cells and the clustering results of various methods as graph constraints, thereby enabling low-dimensional embeddings to effectively preserve positional information as well as base clustering results. The input data for SMODEL comprise the expression matrix of spatial multi-omics data, spatial coordinates, and multiple base clustering results, which are utilized to construct the dual-graph regularized anchor conceptual decomposition integration model. To ensure optimal performance across diverse scenarios, we employed a dual-graph regularization that simultaneously incorporates multiple base clustering results and spatial location information. This approach guarantees that the learned spatial consensus representations integrate the strengths of different methodologies while preserving the geometric structure of the original data manifold. Based on the Euclidean distance between any two points in low-dimensional representation, the spatial pseudo-expression (SPE) can be calculated using the 15 nearest neighbors[34]. SMODEL can be applied to a variety of downstream analysis tasks, which facilitates the elucidation of organizational heterogeneity.

Benchmarking SMODEL against existing methods on the human lymph node dataset
To evaluate the spatial domain identification performance of SMODEL, we applied it on human lymph node datasets that contain spatial transcriptomics and protein expression profiles of two samples. The datasets were generated from 5-μm thick sequential sections of formalin-fixed, paraffin-embedded (FFPE) lymph nodes using the CytAssist Visium platform by 10x Genomics. Ground truth annotations from SpatialGlue[27] were employed to assess spatial domain identification accuracy and compare SMODEL against seven representative multi-omics methods (Supplementary Note 1).

The manual annotations for human lymph node sample A1 encompassed 10 structural categories, including the pericapsular adipose tissue and capsule as the outer layers, and the cortex and medulla along with their associated sinuses, cords, and vessels as core internal structures (Fig. 2A). The visualization of spatial domain identification results across the eight methods for sample A1 (Fig. 2B) highlights the robust performance of SMODEL. The capsule, a fibrous and collagenous membrane enclosing the lymph node, was identified by all algorithms except for scMIC and COSMOS. The cortex, which is critical for immune responses mediated by B cells and T cells, was accurately recognized by SpatialGlue, Seurat, totalVI, PRAGA, and SMODEL, whereas scMIC and MOFA+ failed to detect it. SMODEL effectively discriminated between these spatial domains, thereby enhancing our understanding of their distinct biological roles and spatial
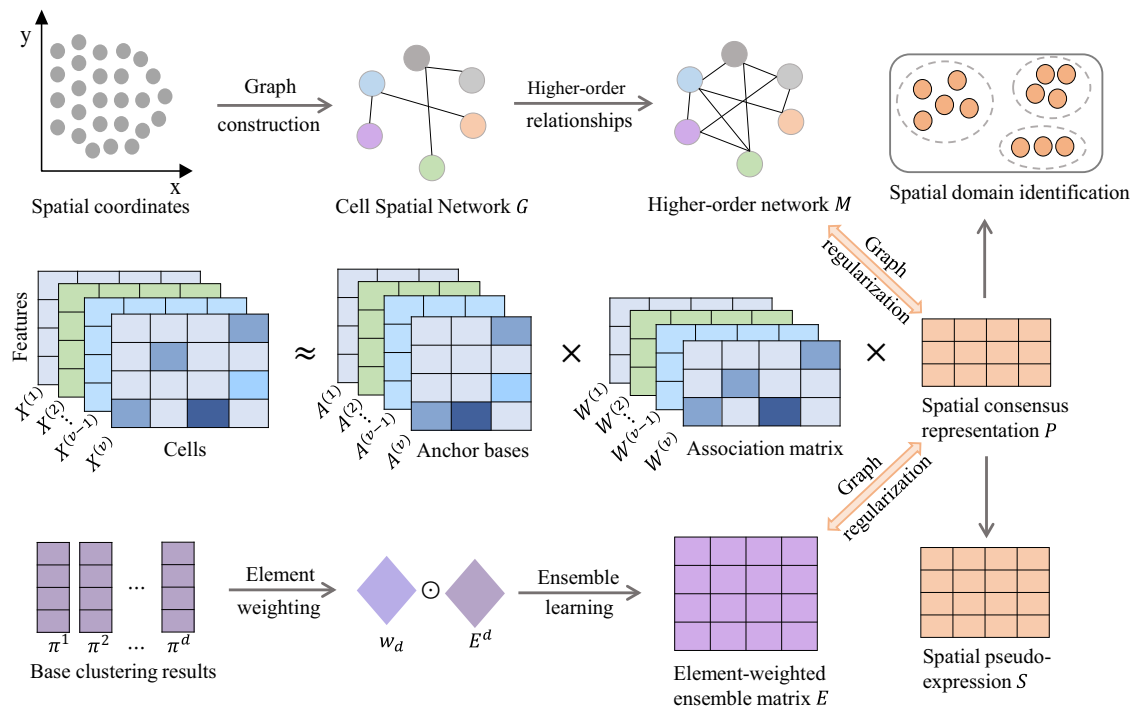
Fig. 1 | Schematic overview of SMODEL. SMODEL is a novel dual-graph regularized ensemble learning algorithm designed for integrating spatial multi-omics data. It integrates anchor bases learning, ensemble learning, and spatial consensus representation learning into a unified model. Additionally, SMODEL incorporates high-order information between cells, thereby preserving the local and global similarities of the original data in the low-dimensional manifold. Finally, an element-wise weighting ensemble strategy is employed to integrate multiple base clustering results, achieving optimal performance in multiple scenarios. The output of SMODEL can be readily utilized for various downstream analysis tasks.

organization. Medulla cords are deeply embedded in the innermost region of the medulla of the lymph node, while medulla sinus is a luminal space similar to blood vessels in the medulla, and the two are intertwined with each other in a staggered layout, which makes it difficult to distinguish between them morphologically[35]. SMODEL effectively distinguished between these spatial domains, thereby enhancing our understanding of their distinct biological roles and spatial organization.

To quantitatively assess the clustering performance, we employed multiple quantitative metrics[36]. These metrics included five supervised evaluation methods: accuracy (ACC), Normalized Mutual Information (NMI), Purity, F-score, and Adjusted Rand Index (ARI)[37]. Higher values of these metrics indicate better performance of the method for spatial domain identification. In Fig. 2C, we can see that SMODEL outperforms other methods on all the supervised metrics, and its position in the radar graph is closest to the edge, which indicates that it achieves high scores on all these supervised metrics. In addition, we introduced two unsupervised evaluation metrics: the Average Silhouette Width (ASW) and the Davies-Bouldin Index (DBI). ASW evaluates clustering quality by measuring the intra-cluster cohesion of data points and their separation from other clusters[38]. A higher ASW value indicates a better clustering performance. DBI quantifies the ratio between the tightness and separation of clusters, with lower DBI values implying better clustering[39]. Figure 2D further demonstrates the superior performance of SMODEL on the ASW metric, with a value of 0.184, which is the highest among all the baseline methods, implying that SMODEL performs the best in maintaining cluster tightness and separation. Meanwhile, SMODEL achieves the lowest DBI value of 1.457 among all baseline methods, further confirming its superiority in spatial domain identification (Fig. 2E).

Similar results were obtained by applying the same method to another human lymph node sample D1 (Supplementary Fig. 1-3), indicating the generalization ability and robustness of SMODEL. Overall, these results demonstrate that SMODEL can efficiently identify medullary structures in human lymph node datasets and outperform the existing methods.

Dissecting breast cancer tissue microenvironment using SMODEL

To test whether SMODEL can fully utilize the complementary information provided by spatial proteomics and spatial transcriptomics to deepen the understanding of the tissue microenvironment. We evaluated SMODEL against seven baseline methods (SpatialGlue, scMIC, totalVI, Seurat, MOFA+, PRAGA, and COSMOS) using spatial multi-omics data from mouse breast cancer, comprising 1978 spots, 18,932 genes, and 32 proteins. From Fig. 3A, we can observe that SpatialGlue and PRAGA exhibited strong spatial consistency during clustering but showed considerable noise in boundary transition areas. scMIC displays numerous discontinuous regions and fails to effectively capture the spatial integrity of individual domains. The clustering results of totalVI were spatially more dispersed compared to other methods. The clustering results produced by Seurat exhibit overly uniform cell population distributions at tissue boundaries, failing to accurately capture boundary-specific cellular characteristics. MOFA+ demonstrates good regional continuity in spatial clustering, but struggles with unclear boundaries in transitional areas. The results of COSMOS reveal relatively regular, block-shaped spatial domains. Notably, SMODEL exhibits a significant advantage in terms of clustering performance. Compared to other methods, SMODEL can clearly identify Cluster 1 (purple) and Cluster 2 (green). This suggests that SMODEL can reveal specific features that other methods fail to identify, which may be crucial for understanding the biological properties and functions of cell groups.

To further investigate the biological significance of the clusters identified by SMODEL, we conducted differential gene expression analysis to identify spatially variable genes (SVGs) within each cluster (Fig. 3B). Using the Wilcoxon rank sum test in combination with the FindAllMarkers function, we identified SVGs with a $\log_2$ fold change threshold of 0.25. Further filtering of the marker genes was conducted based on an adjusted p-value threshold of $p < 0.05$. The volcano plot revealed that *Arg1*, which encodes arginase 1, was significantly upregulated in cluster 1. The gene expression signature of cluster 1 aligns with the established transcriptional
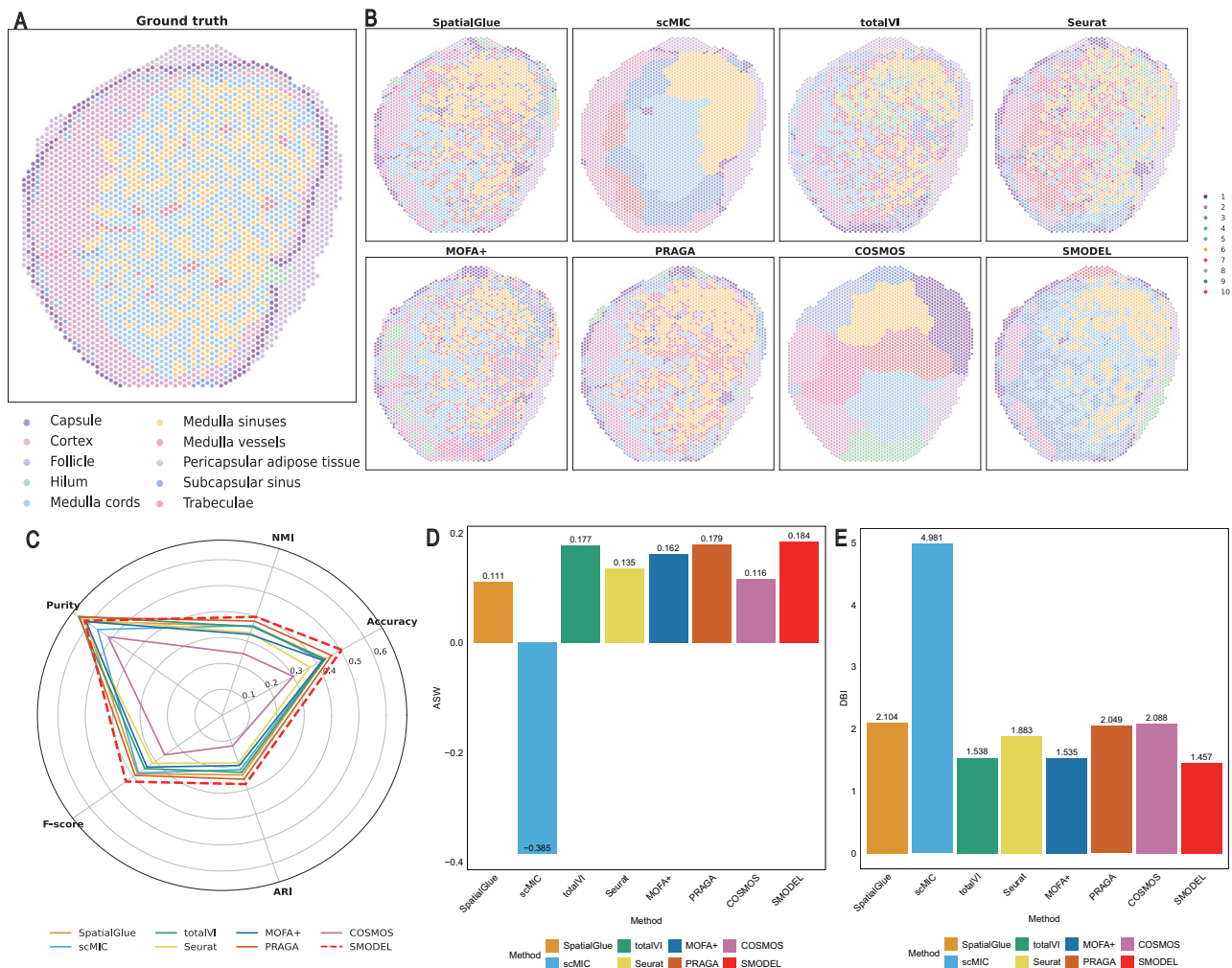
Fig. 2 | SMODEL improves the spatial domain detection performance on the human lymph node dataset. **A** Ground truth annotation of human lymph node sample A1. **B** Spatial plots generated by SpatialGlue, scMIC, totalVI, Seurat, MOFA +, PRAGA, COSMOS, and SMODEL in lymph node sample A1. **C** Radar chart of the Accuracy, NMI, Purity, F-score, ARI for lymph node sample A1 predicted by SpatialGlue, scMIC, totalVI, Seurat, MOFA+, PRAGA, COSMOS, and SMODEL. **D** Bar plots displaying ASW values, where a higher value indicates a better spatial domain detection accuracy. **E** Bar plots displaying DBI values, where a lower value indicates a better spatial domain detection accuracy.

profile of M2 macrophages, which is characterized by high levels of *Arg1* expression (Supplementary Fig. 4). The pronounced expression of *Arg1* within this cluster not only serves as a molecular marker of M2 macrophage polarization but also underscores its potential as a therapeutic target for modulating the immunosuppressive components of the tumor microenvironment[40]. Fig. 3A also shows that cluster 1, as M2 macrophages, may establish an immunosuppressive boundary at the tumor edge, which is related to immune depletion and the inability of the body to effectively fight the tumor, which is consistent with the results of previous studies[41]. The boxplot shows the average $\log_2$ fold change in the SVGs for Cluster 2, comparing the data before and after SMODEL denoising. Notable differences are observed between the two conditions. The average $\log_2$ fold change after denoising with SMODEL showed improved performance compared with the raw data (Fig. 3C), higher average $\log_2$ fold change values suggest that genes are more easily detectable. SMODEL increased the fold change between genes, enabling the identification of additional differential genes that would otherwise remain undetected. This highlights SMODEL's potential to reveal biomarkers that may be overlooked by other methods. Kaplan-Meier plots for *Svopl* and *Serpini1* showed significant differences in overall survival between high-expression and low-expression groups. Patients with high *Svopl* expression, identified as an optimistic prognostic

biomarker, exhibited better survival outcomes ($p < 0.05$) (Fig. 3D), whereas high *Serpini1* expression, considered a pessimistic prognostic biomarker, correlated with worse survival ($p < 0.05$) (Fig. 3E). We also performed functional analysis of the SVGs (Supplementary Table 1), which were enriched in pathways closely related to macrophage polarization and cell state transitions.

Additionally, these SVGs are significantly enriched in the JAK/STAT signaling pathway, which is closely associated with macrophage polarization and interactions within the tumor microenvironment. Tumor-associated macrophages, one of the most common immune infiltrating cells in the tumor microenvironment, play a pivotal role in tumor initiation and progression. M0 macrophages can differentiate into the classically activated pro-inflammatory M1 phenotype and the alternatively activated anti-inflammatory M2 phenotype[42]. M1 macrophages collaborate with Th1 cells and innate lymphoid cells (ILCs) to drive the type 1 immune response. This response mechanism involves the release of cytokines, such as tumor necrosis factor-alpha (TNF-α), activation of cytotoxic T cells to produce reactive oxygen species (ROS), induction of antibody-mediated cytotoxic responses, and direct phagocytosis of tumor cells, demonstrating significant anti-tumor effects[43–45]. After M1 macrophages release anti-inflammatory cytokines in damaged areas, M2
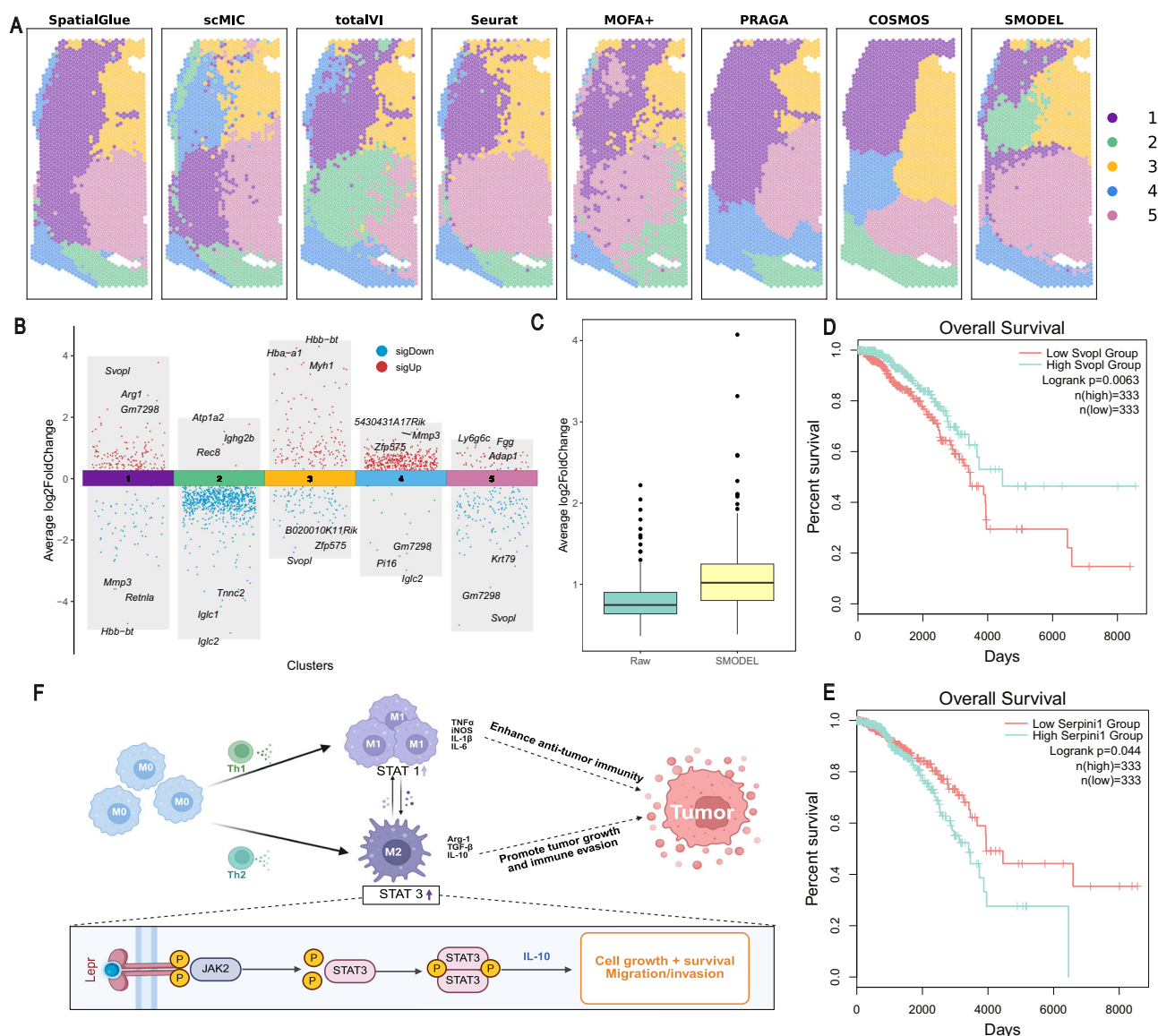
Fig. 3 | SMODEL can reveal heterogeneity of tissue microenvironment in breast cancer samples. A Spatial plots generated by SpatialGlue, scMIC, totalVI, Seurat, MOFA+, PRAGA, COSMOS, and SMODEL in a tissue section of MMTV-PyMT breast cancer, with each color representing a cluster. **B** Volcano plot displaying the significantly upregulated and downregulated genes, with the top three markers per cluster highlighted. Upregulated genes are shown in red, while downregulated genes are marked in blue. **C** Boxplot comparing the average $\log_2$ fold change values of SVGs before and after SMODEL denoising. **D**, **E** Survival curves of SVGs identified by SMODEL in breast cancer samples obtained from the TCGA database. **F** Schematic of the macrophage polarization mechanism in breast cancer. Some icons were created by https://biorender.com/.

macrophages are attracted to the primary tumor site, where substances produced by type 2 ILCs and Th2 responses, such as interleukin (IL)-10, are present. These substances further trigger a type 2 immune response, activate M2 macrophages, promote the accelerated growth, invasion, and metastasis of tumor cells, increase angiogenesis and lymphangiogenesis, and enhance the immunosuppressive state of the tumor[46]. Further investigation of the underlying signaling mechanisms showed that the STAT3 pathway plays a critical role in M2 polarization. SMODEL uncovered *Lepr* as an upstream regulator of the STAT3 pathway, crucial for M2 macrophage polarization (Fig. 3F). Specifically, activation of *Lepr* leads to the phosphorylation of JAK2, which then phosphorylates and promotes the dimerization of STAT3 protein, activating the STAT3 pathway and boosting cell growth, survival, migration, and invasion. The increase in IL-10 induced by *Lepr* signaling contributes to the formation of an immunosuppressive microenvironment that facilitates tumor immune evasion. These findings indicate that *Lepr* could serve as a potential therapeutic target, as inhibiting this signaling pathway may

counteract the pro-tumor effects of M2 macrophages and boost anti-tumor immunity.

In conclusion, SMODEL not only shows superior performance in spatial domain identification but also offers mechanistic insights into tumor microenvironment heterogeneity. The ability of SMODEL to integrate spatial transcriptomics and spatial proteomics provides a robust framework for the identification of clinically relevant biomarkers. It offers valuable insights into the complexity of the tissue microenvironment and holds significant potential for advancing targeted clinical therapies and improving prognostic assessments.

SMODEL enhances gene expression patterns in spatial epigenome-transcriptome data
We applied SMODEL to the epigenomic-transcriptomic dataset of mouse brains at postnatal day 22, derived from spatial-ATAC-RNA-seq[23]. This dataset demonstrates a high degree of consistency with anatomical annotations based on Nissl staining, accurately reflecting the regional
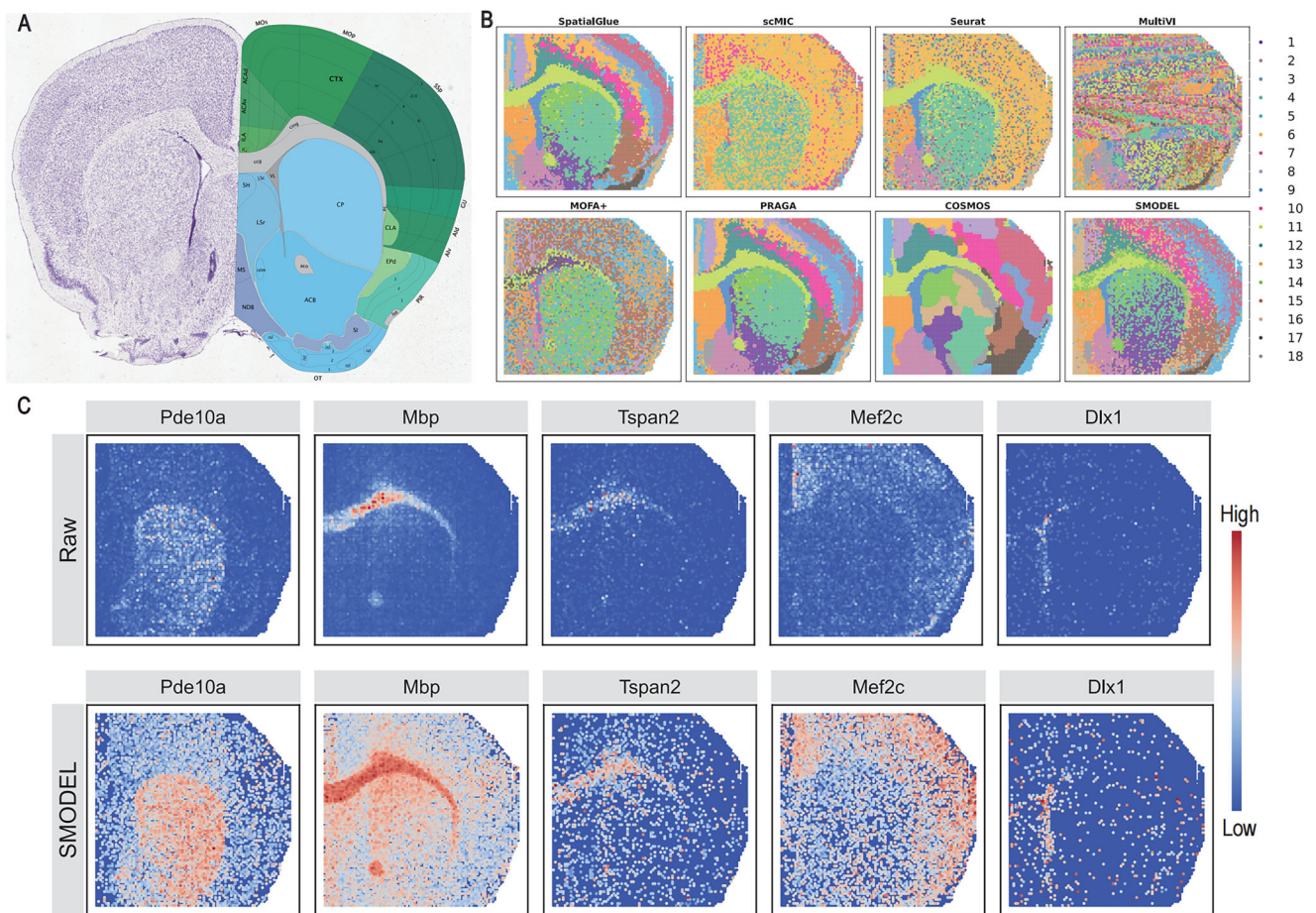
Fig. 4 | Integration results of SMODEL on samples with spatial epigenome-transcriptome data from the mouse brain dataset. A Anatomically annotated mouse brain images were obtained from the Allen Mouse Brain Atlas. **B** Spatial plots generated by SpatialGlue, scMIC, MultiVL, Seurat, MOFA+, PRAGA, COSMOS, and SMODEL for the P22 mouse brain sample acquired using spatial-ATAC-RNA-seq, with each color representing a distinct cluster. **C** Spatial distribution of layer-specific genes in the coronal section of the P22 mouse brain identified using SMODEL, alongside raw data for comparison.

characteristics of the mouse brain. In this study, we referenced the adult mouse coronal brain anatomical annotations provided by the Allen Brain Atlas[47], which primarily encompasses key brain regions such as the cortex (CTX), corpus callosum (CC), lateral ventricle (VL), striatum (Str), anterior commissure (aco), and lateral septal nucleus (LS) (Fig. 4A). To visually compare the performance of the different methods in integrating the mouse brain epigenome-transcriptome dataset, we visualized the results of SpatialGlue, scMIC, MultiVL, Seurat, MOFA+, PRAGA, COSMOS, and SMODEL (Fig. 4B). It was observed that, except for MultiVI, all other baseline methods were able to identify some major regions, such as the core structure (CP) of the striatum in CTX and Str. However, not all methods perfectly corresponded with the anatomical annotations from Allen Brain. In contrast, SpatialGlue and SMODEL could clearly distinguish different layered structures in CTX, which was highly consistent with the anatomical annotations.

The development of the mouse brain is a highly complex and finely regulated biological process that depends on the coordinated expression and precise silencing of numerous genes. During this developmental process, distinct spatial domains are established within the brain, reflecting not only the functional differentiation of brain regions, but also the spatial patterns of gene expression. To further understand and validate the potential of SMODEL in revealing the spatial patterns of gene expression, we analyzed the spatial expression patterns of specific marker genes. SMODEL computed the 15 nearest neighbor cells for each cell based on the spatial consensus representation P, and the average gene expression data of these nearest neighbor cells were used as SPE for that cell. Compared with the raw

expression data, SPE significantly enhanced the expression levels of certain layer-specific genes (Fig. 4C). In the Str, we found that the expression level of *Pde10a* was significantly higher than in other brain regions, and the role of this gene in the striatum is closely related to neuronal signaling and regulation. *Pde10a* affects neuronal excitability by regulating intracellular cAMP levels[48]. The CC contains structures such as the genu of the cgenu of corpus callosum (ccg) and the external capsule (ec), serving as the primary connection pathway between the two hemispheres of the brain[49]. The high expression of *Mbp* and *Tspan2* is closely associated with the myelination process in this region. *Mbp* plays a crucial role in the stability and function of myelin, while *Tspan2* may be involved in intercellular interactions and signaling[50,51]. The CTX, located in the peripheral area of brain tissue and occupying the largest area, is the center of higher cognitive functions, exhibiting significant organizational layering characteristics[52]. We found that the expression pattern of *Mef2c* in CTX was particularly distinct. The VL is a cavity structure filled with cerebrospinal fluid, and the distribution of cell types within its cavity and surrounding areas provides valuable information for studying gene expression in the regions surrounding the ventricles[53]. Notably, the expression pattern of *Dlx1* exhibited clear regional specificity for VL.

Using SMODEL, we significantly enhanced the quality of spatial transcriptomics data by improving marker expression levels, enabling a more accurate representation of expression profiles across different brain regions. In summary, these findings highlight the remarkable ability of SMODEL to effectively integrate spatial epigenomic-transcriptomic data and improve the visualization of spatial expression patterns.
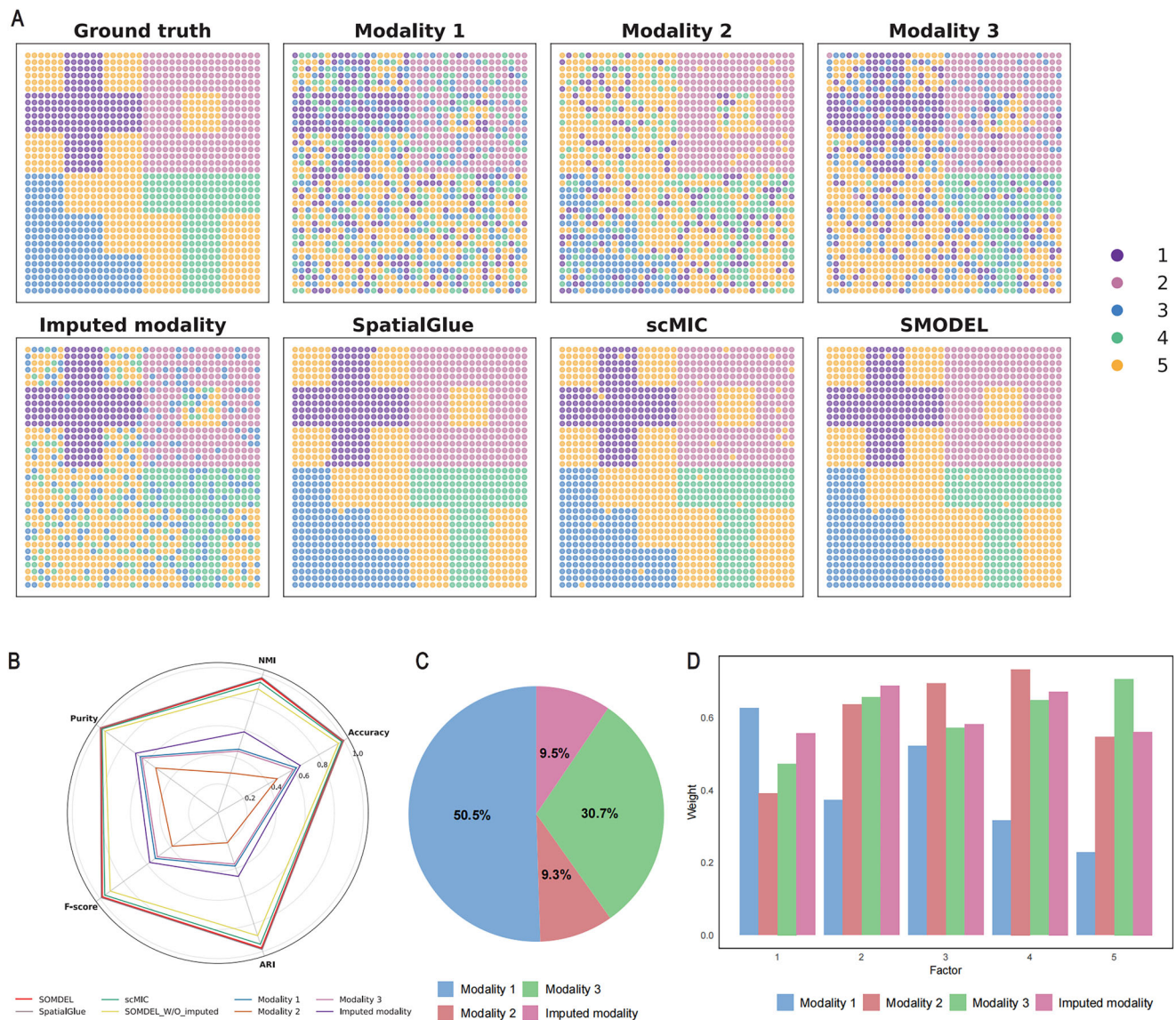
Fig. 5 | Integration of spatial triple omics of simulated data. A Ground truth and spatial plots of modality 1, modality 2, modality 3, imputed modality, SpatialGlue, scMIC, and SMODEL. **B** Radar plot of clustering performance for different methods. **C** Pie chart of modality weights. **D** Bar plots showing the influence of each modality on each cluster, with different colors corresponding to different modalities.

Performance on spatial triple omics integration

Leveraging the anchor concept factorization framework for multi-omics integration, SMODEL is capable of processing not only spatial data from two omics but also extending its functionality to more than three omics. As a case study, we applied SMODEL to integrate spatial triple-omics data simulated using nonnegative spatial factorization decomposition across three omics modalities: RNA, ATAC, and ADT.

Compared to the baseline methods (SpatialGlue, scMIC), SMODEL demonstrates superior performance in matching the ground truth (Fig. 5A). Individual modalities (modality 1, modality 2, and modality 3) exhibited incomplete or noisy patterns. Specifically, the determination of factor 1 was facilitated by the synergistic contributions of modality 1 and modality 3. Factor 2 was identified through a composite analysis involving modality 1, modality 2, and modality 3. Factor 3 was exclusively identified by modality 2, and factor 4 was uniquely identified by modality 3. The imputed modality, derived from the sparse RNA modality, improves the initial patterns, but does not match the accuracy and resolution of other methods. Among the baseline methods, SMODEL is one of the two methods that can accurately

identify four factors with minimal noise. The poor performance of scMIC may be attributed to its failure to leverage spatial information, underscoring the importance of spatial information in spatial multi-omics analysis. A radar plot summarizing the evaluation metrics, including NMI, ARI, purity, accuracy, and F-score, showed that SMODEL generally outperformed SpatialGlue, scMIC, and other modalities. Notably, the performance of SMODEL across these metrics highlights its superior clustering and spatial integration capabilities (Fig. 5B). We quantified the contribution of each modality to the final integration results (Fig. 5C). SMODEL assigns weights to each modality, with modality 2 contributing the most (50.5%), followed by modality 3 (30.7%), modality 1 (9.5%), and imputed modality (9.3%). This distribution highlights the ability of SMODEL to effectively prioritize relevant information across modalities. The bar plot illustrates the weights assigned to each modality for different factors (Fig. 5D). Modality 3 maintained a relatively balanced weight across all factors, with a slight increase in factor 5, suggesting that this modality provides complementary information that enhances the accuracy of the integrated dataset. This adaptive weighting mechanism ensures that the integrated results capture
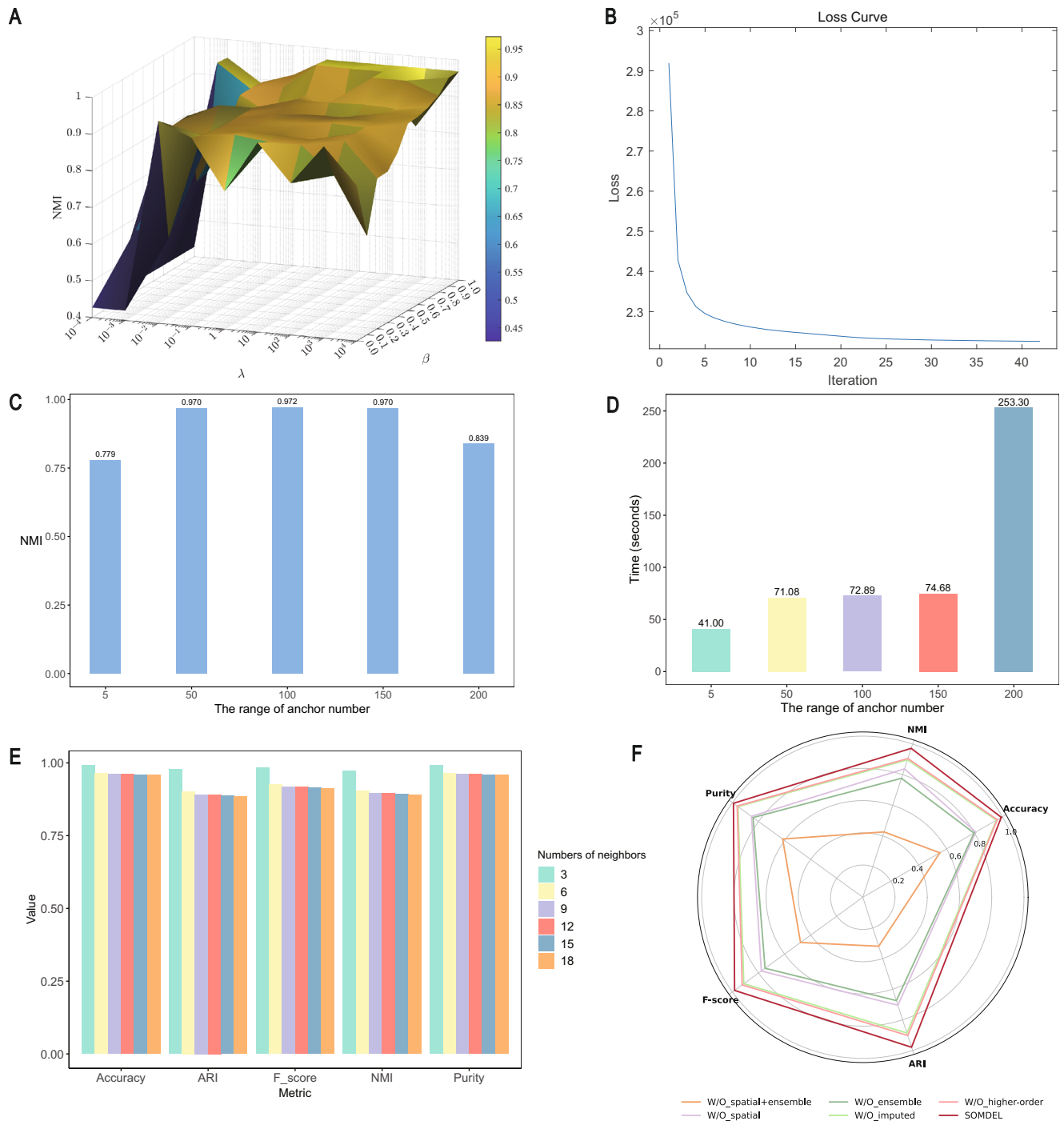
Fig. 6 | Assessment of hyperparameter selection and ablation analysis of the simulated data. A Sensitivity analysis of the hyperparameters. **B** Objective value variation as a function of iteration steps, showcasing the convergence behavior of SMODEL. **C** NMI performance for different anchor numbers. **D** Comparison of time costs for different numbers of anchors. **E** Performance comparison for different numbers of neighbors. **F** Radar plot of clustering performance for ablation comparisons.

the essential spatial patterns embedded in the spatial triple omics data. Moreover, to evaluate noise robustness, we introduced varying levels of sparsity (0.1–0.9) to the non-zero elements of simulated RNA, ADT, and ATAC. As shown in Supplementary Fig. 5, SMODEL demonstrated superior stability characteristics under sparse conditions when compared to other methods.

Overall, these results highlight the robustness and interpretability of SMODEL in integrating spatial triple-omics data to uncover spatial relationships.

Evaluation of Hyperparameter Selection and Ablation Analysis

To evaluate the impact of hyperparameters on the performance of our method, we conducted sensitivity analyses and ablation experiments. Ablation studies further assess the contributions of individual components, providing insights into the reliability and robustness of the computational model. The 3D hyperparameter landscape (Fig. 6A) reveals significant variations in the NMI score across different values of λ and β, with moderate values indicating a region of stability. Based on a comprehensive analysis of the hyperparameter landscape, it is recommended to set λ to 1 and β to 0.9

for datasets with fewer than 1500 cells. Conversely, for datasets with more than 1500 cells, the default parameters should be adjusted to $\lambda = 10$ and $\beta = 0.2$. This differentiation is critical to accommodate the increased complexity and variance associated with larger datasets, thereby ensuring that the model maintains its effectiveness and efficiency. Furthermore, the loss curve (Fig. 6B) demonstrates a rapid decline in the objective value during the initial optimization steps, with convergence occurring within 40 iterations. This highlights the efficiency and stability of the SMODEL optimization process. Building on this, an analysis of the NMI performance across varying anchor numbers underscores the impact of anchor selection on clustering accuracy. The results indicate that an anchor number of 100 offers the best trade-off between the accuracy and computational efficiency (Fig. 6C). As the number of anchors increases, the computational time increases substantially, with 200 anchors significantly increasing the time costs compared with 100 anchors (Fig. 6D). This trade-off highlights the importance of selecting an appropriate anchor number to balance performance and efficiency in spatial domain reconstruction. Therefore, we recommend a default anchor value of 100. The evaluation of performance metrics across different numbers of neighbors (Fig. 6E) reveals a consistently stable performance, with the best results achieved when using three neighbors. Furthermore, the ablation study (Fig. 6F) highlights the importance of each component within the SMODEL framework, as the complete model consistently outperforms all ablated versions across all metrics.

In conclusion, we recommend a default parameter configuration ($\lambda = 10$, $\beta = 0.2$, anchor = 100, KNN = 12), which is generally suitable for a wide range of data types and applications. Users may further adjust these parameters based on specific requirements to achieve optimal performance.

Discussion

Spatial multi-omics sequencing technology enables the simultaneous analysis of multiple omics data, such as transcriptomics, proteomics, and chromatin accessibility, within individual cells. This approach provides a comprehensive molecular view of tissues while revealing the spatial locations of cells, offering a more complete representation of cellular states. Spatial domain identification is a critical step in uncovering cellular heterogeneity from spatial omics data, facilitating a deeper understanding of tissue organization and functional distribution. With the accumulation of spatial multi-omics data, integrating various types of omics data for spatial domain identification has become essential for systematically elucidating the molecular mechanisms underlying cellular heterogeneity. However, developing a universally optimal method for spatial domain identification remains challenging due to the diverse distributions of data, which can significantly affect model performance across different scenarios.

In this study, we introduce SMODEL, an ensemble learning framework designed to integrate spatial multi-omics data for spatial domain identification. Through extensive analysis on five spatial multi-omics datasets, including real and simulated data across different techniques, tissues, and species, we demonstrate the effectiveness of SMODEL in identifying spatial domains. The results highlight the robustness and interpretability of SMODEL in capturing complex spatial relationships while fully leveraging the complementary information provided by spatial multi-omics data. This capability enables in-depth analysis of tissue microenvironments and enhances the understanding of spatial gene expression patterns.

Unlike graph neural network models commonly used for analyzing spatial multi-omics data[27–30], SMODEL leverages anchor concept factorization to extract low-dimensional manifold features. This approach emphasizes the semantic relevance between anchor bases and cell cluster centers, enabling projection learning of concepts defined by the association matrix derived from omics expression data with varying adaptive confidence levels. This process yields a unified spatial consensus representation. Furthermore, SMODEL integrates spatial location information and spatial domain identification results from existing methods, demonstrating exceptional performance across diverse scenarios. Specifically designed for spatial multi-omics data, SMODEL incorporates preprocessing techniques tailored to the unique characteristics of spatial multi-omics data and is

extensible for analyzing additional omics types in practical applications. Its key advantages include modularity, interpretability, robustness, and flexibility, making it a powerful tool for spatial multi-omics data analysis. While SMODEL demonstrates superior performance in terms of accuracy and robustness, we acknowledge that its iterative optimization framework introduces additional computational cost compared to some deep learning-based methods. Spatial domain identification is typically performed as an offline analysis task, where the primary objective is to achieve high accuracy in delineating biologically meaningful domains. Given its non-real-time nature, improving the accuracy of domain identification has been our main focus. However, improving computational efficiency remains an important direction for future research.

The rapid advancement of spatial multi-omics technologies has enabled the acquisition of multi-dimensional data, offering diverse and complementary views of the same cell at an unprecedented resolution and scale. We plan to extend SMODEL to integrate additional types of omics data, thereby enabling more comprehensive biological insights and improving its scalability to large-scale and high-resolution datasets. Building on this foundation, we will incorporate image data generated by spatial sequencing technologies, particularly hematoxylin and eosin (H&E) staining images, to integrate spatial multi-omics data from a high-resolution perspective. Additionally, we will develop algorithms for integrating spatial multi-omics data across continuous tissue sections, thereby improving the efficiency of spatial domain identification and enhancing the potential to uncover cellular heterogeneity.

Methods
Data preprocessing
Spatial multi-omics datasets were downloaded from the Gene Expression Omnibus (GEO) Data Center (https://www.ncbi.nlm.nih.gov/geo/). The data were preprocessed using the SCANPY package[54]. For all datasets, cells outside the primary tissue region were removed. For the spatial transcriptome, raw expression data were filtered and log-transformed based on library size. Only the top 3000 genes were retained for downstream analysis based on the differences in gene expression. For the spatial epigenome, latent semantic indexing (LSI) dimensionality reduction was performed. Spatial proteomes were preprocessed by normalization and scaling. The expression matrix of different omics is denoted by $X^{(v)} \in \mathbb{R}^{P_v \times n}(v = 1, 2, \cdots V)$, the spatial position coordinates are denoted by $Y \in \mathbb{R}^{n \times 2}$. Detailed descriptions of each dataset are provided in Supplementary Note 2 and Supplementary Table 2.

The sparse omics imputation
Due to technical limitations, most spatial transcriptomics technologies still result in incomplete expression measurements and a high number of missing values. To address these issues, imputation methods are essential for filling in missing data, thereby improving the overall interpretability of the data. The expression of gene j within cell i can be estimated through a process that considers both the spatial proximity to neighboring cells and the expression similarity to nearby genes. Specifically, the imputed expression value X_{ij}^c is calculated using the spatial similarity between cells based on Euclidean distance, effectively reflecting the spatial relationships among cells:

$$X_{ij}^c = \begin{cases} \frac{1}{Q}\sum_{d=1}^{p} \beta_d X_{id} & \text{if } X_{ij} = 0 \\ X_{ij} & \text{otherwise} \end{cases} \quad (1)$$

Here, the weight β_j is determined by the decay factor ρ, which ranges from 0 to 1 with a default value of 0.5, and the cell spatial similarity G^c based on Euclidean distance:

$$\beta_j = \rho^{j-1} \times G^c(X_{ij}, X_{id}) \quad (2)$$

and Q normalization terms:

$$Q = \sum_{d=1}^{p} G^c(X_{ij}, X_{id}) \tag{3}$$

In addition to spatial similarity, gene expression similarity is also considered. X_{ij}^g is the imputed expression value based on the cosine similarity G^g between genes, which accounts for the genetic relationships that may not be captured by spatial proximity alone.

Finally, the imputed value for gene j in cell i is selected as the maximum of the cell-based imputation X_{ij}^c and the gene-based imputation X_{ij}^g, ensuring that the most reliable estimate is used:

$$\widehat{X}_{ij} = \max(X_{ij}^c, X_{ij}^g) \tag{4}$$

Anchor concept factorization

Spatial multi-omics data offer a comprehensive view of individual cells, revealing their intricate biological processes from multiple perspectives. To extract the intrinsic consistency within these spatial multi-omics datasets and obtain a more accurate spatial consensus representation for cell heterogeneity analysis, we introduce an anchor concept factorization[55]:

$$\min_{A^{(v)}, W^{(v)}, P, \alpha_v} \quad \sum_{v=1}^{V} \alpha_v \parallel X^{(v)} - A^{(v)} W^{(v)} P \parallel_F^2 \tag{5}$$
$$\text{s.t.} \quad A^{(v)^T} A^{(v)} = I_m$$

where α_v can adaptively learn the corresponding weights from the spatial multi-omics data[56]. $\parallel \cdot \parallel_F$ is the Frobenius norm. $A^{(v)} \in \mathbb{R}^{P_v \times m}$ is the anchor base of a specific omic, on which orthogonality constraints are applied to make different omics anchor bases independent from each other and more discriminative. Anchors are strategically chosen as the cluster centroids derived from the original data space through the execution of the k-means algorithm, ensuring that they encapsulate the inherent structure of the data. $W^{(v)} \in \mathbb{R}^{m \times k}$ denotes the association matrix with conceptually related anchor bases, a connector between spatial multi-omics data, anchor bases, and spatial consensus representation that collects interaction information and correctly indicates semantic cues. $P \in \mathbb{R}^{k \times n}$ is the spatial consensus representation matrix of cells, recording cell expression to the conceptual projections defined by $W^{(v)}$. The anchor concept factorization represents each individual concept as a linear combination of cells, which in turn allows each cell to be expressed as a linear combination of concepts[57]. Due to the clear semantics of the linear coefficients, it is straightforward to learn the consensus information of the same cell across multi-omics datasets. According to equation (5), this framework can simultaneously investigate three key components: a high-quality set of anchor bases, a set of coefficient representations that elucidate the correlations between these anchors and the underlying concepts, and a spatial consensus representation that integrates the diverse contributions of spatial multi-omics data. It is important to note that these components are not operating in isolation; rather, they are interwoven in a way that fosters a synergistic interaction. This interplay is crucial as it allows the framework to leverage the collective strength of each component, enhancing the overall analytical power and providing a more comprehensive understanding of the complex biological systems under study.

Element-wise weighted ensemble strategy

In order to utilize the different clustering results obtained from the currently available models, we introduce an ensemble learning strategy, and denote the result of the d-th base clustering as $B = \{\pi^1, \cdots, \pi^d\}$, with $cls_d(x_i)$ denoting the category that the i-th cell belongs to in the d-th cluster, and construct the

initial ensemble matrix E^d of the d-th base clustering as:

$$E_{ij}^d = \begin{cases} 1, & cls_d(x_i) = cls_d(x_j) \\ 0, & cls_d(x_i) \neq cls_d(x_j) \end{cases} \tag{6}$$

The ensemble matrix E_0 is constructed by averaging over E^d:

$$E_0 = \frac{1}{D} \sum_{d=1}^{D} E^d \tag{7}$$

Considering the remarkable diversity among cells, we have found that different basic clustering methods perform unevenly when dealing with different types of cells. Some methods may outperform on some types of cells, but not as well on others. In order to more fully exploit the advantages of each base clusterings, we decided to weight the base clusterings element by element at the cell level. We construct the set of τ nearest neighbors of cell i for the ensemble matrix E_0 as follows:

$$\tau_i = \{x_{i_1}, \cdots, x_{i_\tau}\} \tag{8}$$

where $i_1, \cdots, i_\tau$ are the column numbers corresponding to the largest τ elements in the i-th row of E_0.

A high similarity between the set of τ-nearest neighbors of a cell and the cluster to which the cell belongs may imply that the algorithm used is able to accurately identify similarities and differences between cells, thus effectively grouping similar cells together. This similarity may reflect cellular interactions, common biological processes, or similar microenvironments. We use the Jaccard coefficient to obtain the correlation weight matrix w_{id} to analyze the similarity between the τ-neighborhood sets and clusters of cells, which helps to gain a deeper understanding of the biological properties and functions of cells.

$$w_{id} = \frac{|\tau_i \cap cls_d(x_i)|}{|\tau_i| + |cls_d(x_i)| - |\tau_i \cap cls_d(x_i)|} \tag{9}$$

where $|\cdot|$ represents the number of elements in the set. The better the d-th base clustering is for the i-th cell, the larger the w_{id}. The element-weighted ensemble graph E is calculated as:

$$E = \sum_{d=1}^{D} (w_d w_d^T)^{\frac{1}{2}} \odot E^d, \tag{10}$$

where $\odot$ represents the matrix dot product. To achieve a stable representation of E, we iteratively update the weighted matrix $W_d = \{w_1, w_2, \cdots, w_D\}$. This process involves following the procedures outlined in equation (8) through equation (9), which guide the generation of the updated weight matrix and its corresponding values. The iterative updating continues until the condition $\mu < 10^{-3}$, where:

$$\mu = \frac{\parallel E_{i+1} - E_i \parallel_F}{\parallel E_i \parallel_F} \tag{11}$$

The element-wise weighted ensemble strategy in SMODEL is inherently flexible and modular, allowing users to select and combine base clustering methods based on dataset characteristics and specific analytical objectives. There are no strict requirements on which methods must be used, as long as they generate valid clustering results that can be incorporated into the ensemble graph (Supplementary Note 3). Notably, SMODEL does not depend on any single base method. Instead, it leverages the diversity and complementarity of multiple methods, with the ensemble mechanism effectively integrating their strengths while compensating for individual method limitations.

Construction of spatial location graph

Valuable insights into high-dimensional data can be derived by analyzing samples embedded within lower-dimensional manifolds. Consequently, it is essential to thoroughly consider the geometric structure of the data within these manifolds. The Euclidean distance effectively captures the spatial relationships among data points and is less influenced by outliers. We use the Euclidean distance to compute the similarity matrix K of cell i at spatial location coordinates (Y_{ix}, Y_{iy}).

$$K_{ij} = \sqrt{\left(Y_{ix} - Y_{jx}\right)^2 + \left(Y_{iy} - Y_{jy}\right)^2},$$ (12)

In manifold learning and spectral graph theory, local invariance is simulated by discretizing the near-neighbor selection of data points to efficiently capture local geometries. Employing the p-nearest neighbor graph(p-NNG) approach[58] to get similar cells close enough in cell space to obtain the cell space network G. This approach ensures that similar cells remain proximate, thereby preserving the intrinsic structural information of the dataset.

$$N_{ij} = \begin{cases} 1, & \text{if } i \in N^p(j) \text{ and } j \in N^p(i) \\ 0, & \text{if } i \notin N^p(j) \text{ and } j \notin N^p(i) \\ 0.5, & \text{otherwise} \end{cases}$$ (13)

$$G_{ij} = N_{ij} \times K_{ij}$$ (14)

where $N^p(i)$ and $N^p(j)$ denote the sets of p nearest cells to cell i and cell j, respectively.

In order to explore the global information between cells on the basis of local information, we constructed a global co-correlation matrix to describe the higher-order similarity between cells. In the real world, if two people have a lot of common friends, it is very likely that these two people are acquaintances. Inspired by this phenomenon, We use the Pointwise Mutual Information (PMI) matrix[59] to capture the high-order topological structure M among cells in the network G, Specifically, the element m_{ij} of M is defined as follows:

$$m_{ij} = log \frac{G_{ij} \sum_l d_l}{d_i d_j} - log\kappa$$ (15)

where G_{ij} represent the weight between the i-th and j-th cells in the cell spatial network G. d_l denote the degree of the l-th cell in the network (i.e., the number of edges connected to it). d_i and d_j are the degrees of the i-th and j-th cells, respectively. The parameter κ is typically set to 1 by default. The construction of the PMI matrix is based on the statistical analysis of co-expression patterns among cells, which allows us to identify cell pairs that influence each other not just directly, but through a series of intermediate steps. M as a higher-order topology of the network G enhances the representation of similar relationships between cells. This provides a more accurate characterization of the topology of the network.

Dual-graph regularization

To achieve optimal results across multiple scenarios, we construct an element-weighted ensemble graph E which integrates clustering results from multiple methods, and a higher-order network graph M which captures the high-order relationships in spatial multi-omics data. These two graphs jointly guide the learning of the spatial consensus representation P. A dual-graph regularization term is added to the initial objective function to obtain the final objective function:

$$\min_{A^{(v)}, W^{(v)}, P, \alpha_v} \sum_{v=1}^{V} \alpha_v \parallel X^{(v)} - A^{(v)} W^{(v)} P \parallel_F^2$$
$$+ \lambda \operatorname{Tr}(P(\beta L_E + (1 - \beta) L_M) P^T)$$
$$\text{s.t. } A^{(v)^T} A^{(v)} = I_m$$ (16)

where L_E and L_M are the graph Laplacian matrices of E and M, respectively. For example, $L_E = D_E - W_E$, where D_E is the degree matrix and W_E is the adjacency matrix. These Laplacian matrices serve to preserve the local geometry of each graph structure in the spatial consensus representation P by penalizing large variations between connected nodes, thereby enforcing local smoothness and structural coherence. The parameters λ and β are hyperparameters that control the strength of the regularization and the balance between the two graph structures, with values ranging between 0 and 1. This problem can be addressed using an alternating optimization algorithm, which iteratively optimizes each variable while keeping the others fixed. After learning the final spatial consensus representation P, k-means is applied to identify specific cell types, thereby facilitating a deeper understanding of the functional relationships and heterogeneity within tissue regions.

SMODEL optimization

We employed an alternate minimizing algorithm strategy to address the problem in equation (16) mentioned above. Specifically, we decomposed the entire objective function into several subproblems and optimized one variable individually, while keeping the other variables fixed.

$A^{(v)}$-**subproblem**: To isolate the impact of $A^{(v)}$ on the objective function, we fixed all other variables at a constant level. Under these controlled conditions, the objective function can be reformulated in terms of $A^{(v)}$ alone as:

$$\min_{A^{(v)}} \sum_{v=1}^{V} \alpha_v \parallel X^{(v)} - A^{(v)} W^{(v)} P \parallel_F^2,$$
$$\text{s.t. } A^{(v)^T} A^{(v)} = I_m.$$ (17)

Given that each individual $A^{(v)}$ is independent with respect to different spatial omics, the equation (17) can be reconstructed as:

$$\max_{A^{(v)}} \operatorname{Tr}(A^{(v)T} B^{(v)}), \quad \text{s.t. } A^{(v)T} A^{(v)} = I_m$$ (18)

where $B^{(v)}$ is defined as $B^{(v)} = X^{(v)} P^T W^{(v)T} = X^{(v)} [W^{(v)} P]^T$. By employing the Singular Value Decomposition (SVD) method[60], $B^{(v)}$ is expressed as $U_B \Sigma_B V_B^T$. Consequently, $A^{(v)}$ is then computed as follows:

$$A^{(v)} = U_B^{(v)} V_B^{(v)T}$$ (19)

$W^{(v)}$-**subproblem:** By maintaining the other variables constant, $W^{(v)}$ can be updated through the solution of the following problem:

$$\min_{W^{(v)}} \sum_{v=1}^{V} a_v \parallel X^{(v)} - A^{(v)} W^{(v)} P \parallel_F^2$$ (20)

Similarly, the optimization problem with respect to each independent $W^{(v)}$ can be reformulated to:

$$\min_{W^{(v)}} \operatorname{Tr}(P^T W^{(v)T} W^{(v)} P) - 2\operatorname{Tr}(P^T W^{(v)T} A^{(v)T} X^{(v)})$$ (21)

By taking the derivative of equation (21) with respect to $W^{(v)}$ and setting it to zero[61,62], we derive the subsequent update rule for $W^{(v)}$:

$$W^{(v)} = \left(A^{(v)^T} X^{(v)} P^T\right)\left(PP^T\right)^{-1} \tag{22}$$

P-**subproblem:** By maintaining the other variables constant, P can be updated through the solution of the following optimization problem:

$$\min_P \sum_{v=1}^V \alpha_v \parallel X^{(v)} - A^{(v)}W^{(v)}P \parallel_F^2 + \lambda \operatorname{Tr}\left(P(\beta L_E + (1-\beta)L_M)P^T\right) \tag{23}$$

To obtain the update rule for P, we calculate the derivative of equation (23) with respect to P and equate it to zero. The detailed derivation is presented in the Supplementary Note 4, resulting in the following:

$$\sum_{v=1}^V \alpha_v W^{(v)T} A^{(v)T} A^{(v)} W^{(v)} P + P\lambda[\beta L_E + (1-\beta)L_M]$$
$$= \sum_{v=1}^V \alpha_v W^{(v)T} A^{(v)T} X^{(v)} \tag{24}$$

Equation (24) can be rewritten in the standard form of the Sylvester equation as follows:

$$XP + PY = Z \tag{25}$$

where

$$
\begin{aligned}
X &= \sum_{v=1}^V \alpha_v W^{(v)T} A^{(v)T} A^{(v)} W^{(v)} \\
Y &= \lambda[\beta L_E + (1-\beta)L_M] \\
Z &= \sum_{v=1}^V \alpha_v W^{(v)T} A^{(v)T} X^{(v)}
\end{aligned} \tag{26}
$$

When there are no shared eigenvalues between X and Y, equation (26) has a unique solution[63]. Thus, the update for P can be expressed as:

$$P = \operatorname{sylvester}(X, Y, Z) \tag{27}$$

Here, the sylvester ($\cdot$) function solves the Sylvester equation and returns matrix P as the solution.

α_v-**subproblem:** To quantify and optimize the adaptive confidence level weights α_v, we consider the reconstruction error of each omic. Initially, the weights for each omic α_v are set to be equal, i.e., $\alpha_v = \frac{1}{V}$, where V is the total number of omics. With the weights α_v fixed, the objective function is optimized by minimizing the weighted reconstruction error for each omic. The weights are updated based on the inverse of the reconstruction error, ensuring that omics with lower reconstruction errors receive higher weights. Specifically, the weights are updated as follows:

$$\alpha_v = \frac{1}{2 \parallel X^{(v)} - A^{(v)}W^{(v)}P \parallel_F} \tag{28}$$

where $\parallel \cdot \parallel_F$ is the Frobenius norm.

According to the nature of the alternating optimization process, the objective function value decreases monotonically with each iteration. This equation (28) implies that if the reconstruction error $\parallel X^{(v)} - A^{(v)}W^{(v)}P \parallel_F$ is small, indicating that the v-th omic is reliable and informative, the corresponding weight α_v will be large. Conversely, if the reconstruction error is large, suggesting that the v-th omic is less informative or noisy, a smaller weight will be assigned to it[56].

Complexity Analysis

Time complexity. The computational complexity of SMODEL consists of three parts, corresponding to the optimization of its key components. The first part is anchor concept factorization. For the v-th omic, computing the reconstruction term $A^{(v)}W^{(v)}P$ involves $\mathcal{O}(p_v mk + mkn)$ operations. Updating $A^{(v)}$ with orthogonality constraints requires $\mathcal{O}(p_v m^2)$, updating $W^{(v)}$ costs $\mathcal{O}(m^2 k)$, and updating α_v requires $\mathcal{O}(1)$. The second part is ensemble learning, which involves constructing an ensemble similarity matrix using d base clustering results and deriving a corresponding Laplacian matrix. This step requires $\mathcal{O}(dn^2)$ operations. The third part is dual-graph regularization. Computing the manifold Laplacian incurs $\mathcal{O}(n^2 k)$ operations, and combining Laplacians as well as solving for the consensus representation P adds another $\mathcal{O}(kn^2 + k^2 n)$ operations. Therefore, the total time complexity of SMODEL over T iterations is given by:

$$\mathcal{O}\big(T\big(V(p_{\max}mk + mkn + p_{\max}m^2 + m^2 k) + dn^2 + kn^2\big)\big),$$

where T denotes the number of iterations, p_v is the feature dimension of the v-th omic, $p_{\max} = \max(p_1, p_2, \ldots, p_V))$, m is the number of anchors, k denotes the latent dimensionality of the spatial consensus representation P, and n is the number of cells. Since typically $n \gg p_v, m, k$, the time complexity is dominated by $\mathcal{O}(T(d+k)n^2)$.

Space complexity. The main memory consumption of SMODEL comes from storing the input omic matrices $X^{(v)} \in \mathbb{R}^{p_v \times n}$, the anchor matrices $A^{(v)} \in \mathbb{R}^{p_v \times m}$, the weight matrices $W^{(v)} \in \mathbb{R}^{m \times k}$, and the low-dimensional representation $P \in \mathbb{R}^{k \times n}$, contributing a total of $\mathcal{O}(Vp_{\max}n + Vp_{\max}m + Vmk + kn)$ in space. Additional memory is required to store the ensemble similarity matrix E and the Laplacian matrices L_E and L_M, each of size $n \times n$, contributing an additional $\mathcal{O}(n^2)$ space. Thus, the total space complexity is:

$$\mathcal{O}(Vp_{\max}n + Vp_{\max}m + Vmk + kn + n^2),$$

which is dominated by the $O(n^2)$ term when dealing with large-scale datasets where the number of cells n is significantly larger than other dimensions.

Reporting summary

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

Data availability

All data analyzed in this paper are available in publicly available datasets. The in-house human lymph node datasets and spatial triplet omics data were acquired from SpatialGlue. The MMTV-PyMT breast cancer datasets can be found in the GEO repository (GSE198353, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE198353). The spatial-epigenome-transcriptome mouse brain datasets can be downloaded from AtlasXplore (https://web.atlasxomics.com/visualization/Fan). The source data of the main figures are at Supplementary Data 1.

Code availability

The code for the SMODEL algorithm is implemented in MATLAB, and the source code is accessible online at https://github.com/liying-1028/SMODEL. The SMODEL code has been archived on Zenodo and is also available at https://doi.org/10.5281/zenodo.15598653[64].

References

1. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

2. Park, J. et al. Spatial omics technologies at multimodal and single cell/subcellular level. *Genome Biol.* **23**, 256 (2022).

3. Lee, J., Hyeon, D. Y. & Hwang, D. Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.* **52**, 1428–1442 (2020).

4. Clark, S. J. et al. scnmt-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* **9**, 781 (2018).

5. Ma, S. et al. Chromatin potential identified by shared single-cell profiling of rna and chromatin. *Cell* **183**, 1103–1116 (2020).

6. Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).

7. Argelaguet, R. et al. Mofa+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* **21**, 1–17 (2020).

8. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

9. Chen, F., Zou, G., Wu, Y. & Ou-Yang, L. Clustering single-cell multi-omics data via graph regularized multi-view ensemble learning. *Bioinformatics* **40**, btae169 (2024).

10. Huizing, G.-J., Deutschmann, I. M., Peyré, G. & Cantini, L. Paired single-cell multi-omics data integration with mowgli. *Nat. Commun.* **14**, 7711 (2023).

11. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).

12. Gayoso, A. et al. Joint probabilistic modeling of single-cell multi-omic data with totalvi. *Nat. Methods* **18**, 272–282 (2021).

13. Ashuach, T. et al. Multivi: deep generative model for the integration of multimodal data. *Nat. Methods* **20**, 1222–1231 (2023).

14. Zhan, Y., Liu, J. & Ou-Yang, L. scmic: a deep multi-level information fusion framework for clustering single-cell multi-omics data. *IEEE Journal of Biomedical and Health Informatics* (2023).

15. Lin, X., Tian, T., Wei, Z. & Hakonarson, H. Clustering of single-cell multi-omics data with a multimodal deep learning method. *Nat. Commun.* **13**, 7705 (2022).

16. Wu, K. E., Yost, K. E., Chang, H. Y. & Zou, J. Babel enables cross-modality translation between multiomic profiles at single-cell resolution. *Proc. Natl. Acad. Sci. USA* **118**, e2023070118 (2021).

17. Minoura, K., Abe, K., Nam, H., Nishikawa, H. & Shimamura, T. A mixture-of-experts deep generative model for integrated analysis of single-cell multiomics data. *Cell Reports Methods* **1** (2021).

18. Wang, Y., Liu, Z. & Ma, X. Mnmst: topology of cell networks leverages identification of spatial domains from spatial transcriptomics data. *Genome Biol.* **25**, 133 (2024).

19. Wang, Y., Liu, Z. & Ma, X. Mucst: restoring and integrating heterogeneous morphology images and spatial transcriptomics data with contrastive learning. *Genome Med.* **17**, 21 (2025).

20. Zhang, M., Zhang, W. & Ma, X. St-scsr: identifying spatial domains in spatial transcriptomics data via structure correlation and self-representation. *Brief. Bioinform.* **25**, bbae437 (2024).

21. Zhang, M. et al. Spatial molecular profiling: platforms, applications and analysis tools. *Brief. Bioinforma.* **22**, bbaa145 (2021).

22. Liu, Y. et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* **183**, 1665–1681 (2020).

23. Zhang, D. et al. Spatial epigenome–transcriptome co-profiling of mammalian tissues. *Nature* **616**, 113–122 (2023).

24. Ben-Chetrit, N. et al. Integration of whole transcriptome spatial profiling with protein markers. *Nat. Biotechnol.* **41**, 788–793 (2023).

25. Liu, J. et al. Concordance of merfish spatial transcriptomics with bulk and single-cell RNA sequencing. *Life Science alliance* **6** (2023).

26. He, S. et al. High-plex imaging of rna and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* **40**, 1794–1806 (2022).

27. Long, Y. et al. Deciphering spatial domains from spatial multi-omics with spatialglue. *Nature Methods* 1–10 (2024).

28. Huang, X. et al. Praga: prototype-aware graph adaptive aggregation for spatial multi-modal omics analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, 326–333 (2025).

29. Zhou, Y. et al. Cooperative integration of spatially resolved multi-omics data with cosmos. *Nat. Commun.* **16**, 27 (2025).

30. Coleman, K. et al. Resolving tissue complexity by multimodal spatial omics modeling with miso. *Nature Methods* 1–9 (2025).

31. Huang, Y., Jiang, H. & Ching, W.-K. scwe: high-order element-wise weighted ensemble clustering for heterogeneity analysis of single-cell rna-sequencing data. *Brief. Bioinforma.* **25**, bbae203 (2024).

32. Wang, S. et al. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Trans. Image Process.* **31**, 556–568 (2021).

33. Wu, W., Chen, Y., Wang, R. & Ou-Yang, L. Self-representative kernel concept factorization. *Knowl.-Based Syst.* **259**, 110051 (2023).

34. Li, Y., Lu, Y., Kang, C., Li, P. & Chen, L. Revealing tissue heterogeneity and spatial dark genes from spatially resolved transcriptomics by multiview graph networks. *Research* **6**, 0228 (2023).

35. Vondenhoff, M. F. et al. Lymph sacs are not required for the initiation of lymph node formation. *Development* **136**, 1233–1242 (29–34).

36. Wang, C.-D., Chen, M.-S., Huang, L., Lai, J.-H. & Philip, S. Y. Smoothness regularized multiview subspace clustering with kernel learning. *IEEE Trans. Neural Netw. Learn. Syst.* **32**, 5047–5060 (2020).

37. Hubert, L. & Arabie, P. Comparing partitions journal of classification **2** 193–218 (1985).

38. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).

39. Bhola, R., Krishna, N. H., Ramesh, K., Senthilnath, J. & Anand, G. Detection of the power lines in uav remote sensed images using spectral-spatial methods. *J. Environ. Manag.* **206**, 1233–1242 (2018).

40. Su, X. et al. Breast cancer–derived gm-csf regulates arginase 1 in myeloid cells to promote an immunosuppressive microenvironment. *J. Clin. Investig.* **131** (2021).

41. Keren, L. et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* **174**, 1373–1387 (2018).

42. Martinez, F. O. & Gordon, S. The m1 and m2 paradigm of macrophage activation: time for reassessment. *F1000prime Rep.* **6** (2014).

43. Zhang, Z. et al. Anti-tumor macrophages activated by ferumoxytol combined or surface-functionalized with the tlr3 agonist poly (i: C) promote melanoma regression. *Theranostics* **8**, 6307 (2018).

44. Keeley, T., Costanzo-Garvey, D. L. & Cook, L. M. Unmasking the many faces of tumor-associated neutrophils and macrophages: considerations for targeting innate immune cells in cancer. *Trends Cancer* **5**, 789–798 (2019).

45. Munn, D. H. & Cheung, N. Phagocytosis of tumor cells by human monocytes cultured in recombinant macrophage colony-stimulating factor. *J. Exp. Med.* **172**, 231–237 (1990).

46. Mantovani, A., Marchesi, F., Malesci, A., Laghi, L. & Allavena, P. Tumour-associated macrophages as treatment targets in oncology. *Nat. Rev. Clin. Oncol.* **14**, 399–416 (2017).

47. 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).

48. Whiteley, E. L., Tejeda, G. S., Baillie, G. S. & Brandon, N. J. Pde10a mutations help to unwrap the neurobiology of hyperkinetic disorders. *Cell. Signal.* **60**, 31–38 (2019).

49. De León Reyes, N. S., Bragg-Gonzalo, L. & Nieto, M. Development and plasticity of the corpus callosum. *Development* **147**, dev189738 (2020).

50. Jörg, L. M., Schlötzer-Schrehardt, U., Lefebvre, V., Sock, E. & Wegner, M. Transcription factors sox8 and sox10 contribute with different importance to the maintenance of mature oligodendrocytes. *Int. J. Mol. Sci.* **25**, 8754 (2024).

51. Zhang, H. et al. Tspan protein family: focusing on the occurrence, progression, and treatment of cancer. *Cell Death Discov.* **10**, 187 (2024).
52. Wu, G., Cui, Z., Wang, X. & Du, Y. Unveiling the core functional networks of cognition: an ontology-guided machine learning approach. *bioRxiv* (2024).
53. Dani, N. et al. A cellular and spatial map of the choroid plexus across brain ventricles and ages. *Cell* **184**, 3056–3074 (2021).
54. Wolf, F. A., Angerer, P. & Theis, F. J. Scanpy: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 1–5 (2018).
55. Chen, M.-S., Wang, C.-D., Huang, D., Lai, J.-H. & Yu, P. S. Concept factorization based multiview clustering for large-scale data. *IEEE Trans. Knowl. Data Eng.* **72**, 1–14 (2024).
56. Nie, F. et al. Self-weighted multiview clustering with multiple graphs. In *IJCAI*, 2564–2570 (2017).
57. Xu, W. & Gong, Y. Document clustering by concept factorization. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, 202–209 (2004).
58. Jiang, H., Wang, M.-N., Huang, Y.-A. & Huang, Y. Graph-regularized non-negative matrix factorization for single-cell clustering in scrna-seq data. *IEEE Journal of Biomedical and Health Informatics* (2024).
59. Li, Y., Sha, C., Huang, X. & Zhang, Y. Community detection in attributed graphs: An embedding approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32 (2018).
60. Wang, S. et al. Multi-view clustering via late fusion alignment maximization. In *IJCAI*, 3778–3784 (2019).
61. Yang, B. et al. Fast multi-view clustering via nonnegative and orthogonal factorization. *IEEE Trans. image Process.* **30**, 2575–2586 (2020).
62. Kang, Z., Peng, C., Cheng, Q. & Xu, Z. Unified spectral clustering with optimal graph. In *Proceedings of the AAAI conference on artificial intelligence*, vol. 32 (2018).
63. Golub, G., Nash, S. & Van Loan, C. A hessenberg-schur method for the problem ax+ xb= c. *IEEE Trans. Autom. Control* **24**, 909–913 (1979).
64. Li, Y., Cai, G., Chen, F., Wen, K. & Ou-Yang, L. Unveiling spatial domains from spatial multi-omics data using dual-graph regularized ensemble learning (2025).

Author contributions
Y.L. and L.O.-Y. conceived and designed the study; Y.L. implemented the model; Y.L. wrote the manuscript; Y.L., G.C.C., F.Q.C., and K.P.W. collected and constructed the benchmark datasets; L.O.-Y. supervised and reviewed the manuscript; L.O.-Y. supported the funding. All authors have approved 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/s42003-025-08372-6.

Correspondence and requests for materials should be addressed to Le Ou-Yang.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Yang Zhang and Tobias Goris.

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

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