Briefings in Bioinformatics, 2026, 27, bbag214
https://doi.org/10.1093/bib/bbag214
Published: 11 May 2026
Case Study

OXFORD

Cross-modality representation and multi-sample integration of spatially resolved omics data

Zhen Li[1], Xuejian Cui[2], Xiaoyang Chen[1], Zijing Gao[1], Yuyao Liu[1], Yan Pan[1], Shengquan Chen[3], Hairong Lv[1], Lei Zhai[4,*], Rui Jiang[1,*]

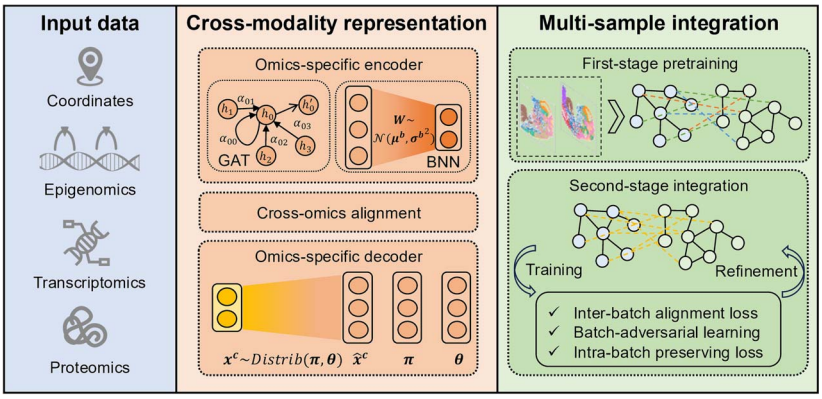
[1]Ministry of Education Key Laboratory of Bioinformatics, Bioinformatics Division at the Beijing National Research Center for Information Science and Technology, Center for Synthetic and Systems Biology, Department of Automation, Tsinghua University, Shuangqing Road, Haidian District, Beijing 100084, China
[2]Faculty of Computing, Harbin Institute of Technology, Xidazhi Street, Nangang District, Harbin, Heilongjiang 150001, China
[3]School of Mathematical Sciences and LPMC, Nankai University, Weijin Road, Nanjiao District, Tianjin 300071, China
[4]Department of Critical Rehabilitation, Beijing Bo'ai Hospital, China Rehabilitation Research Center, Jiangmen North Road, Fengtai District, Beijing 100068, China

*Corresponding authors. Rui Jiang, Ministry of Education Key Laboratory of Bioinformatics, Bioinformatics Division at the Beijing National Research Center for Information Science and Technology, Center for Synthetic and Systems Biology, Department of Automation, Tsinghua University, Shuangqing Road, Haidian District, Beijing 100084, China. E-mail: ruijiang@tsinghua.edu.cn; Lei Zhai, Department of Critical Rehabilitation, Beijing Bo'ai Hospital, China Rehabilitation Research Center, Jiangmen North Road, Fengtai District, Beijing 100068, China. E-mail: drzhailei@163.com.

Abstract

Spatially resolved sequencing technologies have revolutionized our understanding of biological regulatory processes within tissue microenvironments by simultaneously capturing the states of genomic regions, genes, and proteins alongside the spatial organization of cells. However, inherent heterogeneity across modalities and samples poses substantial challenges for the integrative analysis of spatial omics data, underscoring the urgent need for advanced computational methods. In this study, we propose PRESENT, a contrastive learning-based integrative framework for cross-modality representation of spatial multi-omics data. PRESENT employs omics-specific encoders consisting of graph attention networks and Bayesian neural networks coupled with distribution-aware decoders to model distinct modalities, and an inter-omics alignment module for multi-omics integration. By effectively incorporating spatial dependencies with multi-omics information across diverse species and technologies, PRESENT facilitates the accurate identification of spatial domains and the elucidation of underlying regulatory mechanisms. Furthermore, PRESENT can be extended to multi-sample integration via a two-stage training workflow, which incorporates inter-batch alignment loss, intra-batch preserving loss, batch-adversarial learning, and cyclic graph refinement strategies to eliminate batch effects while retaining biological signals. Extensive experiments on tissue samples across different anatomical regions and developmental stages demonstrate that PRESENT enables the characterization of hierarchical tissue structures from a spatiotemporal perspective.

Graphical abstract



Keywords spatial multi-omics data, cross-modality representation, multi-sample integration, spatial domain identification

Received: November 12, 2025. **Revised:** February 16, 2026. **Accepted:** April 10, 2026

Introduction

Single-cell sequencing technologies have enabled the characterization of states and activities within individual cells, promoting molecular cell biology and disease research in areas such as cell atlas production, cell lineage differentiation, cancer genetics and immunology [1]. However, an indispensable step of single-cell sequencing techniques involves the dissociation of cells from the tissue of interest, resulting in the loss of spatial information [2]. Complex biological regulatory processes and cellular activities within the microenvironment are driven by intricate cell–cell interactions mediated by ligand-receptor partners, which heavily depend on the spatial distance between cells [3–6]. Recently, rapid advancements in spatially resolved sequencing technologies have enabled the simultaneous profiling of spatial and omics information, including spatially resolved transcriptomics (SRT) [7, 8], spatial epigenomic assay for transposase-accessible chromatin sequencing (ATAC-seq) [9, 10] and spatial multi-omics approaches such as spatial CITE-seq [11], SPOTS [12], and MISAR-seq [13], granting us a new perspective to decipher biological regulatory mechanisms within the spatial context.

However, several significant challenges impede the exploration of spatial regulatory patterns based on spatially resolved omics data. First, spatial coordinate data and omics data exhibit significant modality gaps. Spatial coordinates anchor the relative positions of cells within tissues, while omics data depicts the states of chromatin regions, genes and proteins involved in the gene regulatory processes within cells. Second, different layers of omics data present distinct patterns and distributions [14–17]. For example, compared with RNA-seq, ATAC-seq datasets suffer from larger dimensionality, more severe sparsity and higher dropout rates, while proteomics datasets, mainly sequenced by antibody-derived tags (ADTs), exhibit opposite characteristics, requiring targeted modeling for each omics layer. Third, due to technical limitations, such as size restrictions of the captured area during data acquisition, individual spatial sequencing samples typically focus on a specific dissected region or developmental stage [18, 19]. Therefore, discrepancies in dissection and placement prevent the direct comparison of spatial coordinates between different samples [19]. Additionally, the intricate coupling between batch effects and biological variations also hinders the integrative analysis of multiple tissues across different conditions [19–21].

Some computational methods have been proposed to address these challenges, particularly for SRT data [22], including single-sample representation algorithms like SpaGCN [23], STAGATE [24], and PAST [25], as well as multi-sample integration frameworks like STAligner [26], Stitch3D [27], and GraphST [18]. Recently, algorithms have also been developed for the analysis of spatial epigenomics and multi-omics data. INSTINCT was proposed to integrate spatial epigenomic signals across multiple tissue samples through stochastic domain translation [28]. SpatialGlue was developed for the representation of spatial multi-omics data using an attentional feature fusion mechanism [29]. Tian *et al.* proposed a dependency-aware generative variational autoencoder framework for the representation of SRT, spatial ATAC-seq, and spatial CITE-seq data [30]. Despite the advances, most existing approaches are designed either for multi-sample integration within a single omics modality or for multi-omics integration within a single tissue sample, but rarely address both simultaneously. Consequently, a unified framework that can jointly model spatially resolved multi-omics signals across multiple samples is still lacking.

In this article, we propose PRESENT, an effective contrastive learning framework for the multi-omics and multi-sample integration of s**P**atially **RE**solved omics data, including spatial multi-omic**S, E**pigenomics, tra**N**scrip**T**omics (Fig. 1). Through comprehensive experiments on datasets generated by various spatial sequencing technologies, including spatial multi-omics [12, 13, 31], epigenomics [9, 10], and transcriptomics [7, 8], we demonstrated that PRESENT can simultaneously capture spatially coherent patterns and complementary multi-omics information, obtaining interpretable representations for various downstream analyses, including spatial visualization, spatial domain identification, domain-specific marker detection and discovery of domain-related functions. Furthermore, PRESENT can be extended for the integrative analysis of spatial multi-omics samples across different dissection regions and developmental stages, effectively eliminating batch effects while maintaining biological variations, paving the way to explore the hierarchical tissue structures and sophisticated cellular activities.

Materials and methods

Cross-modality representation of spatial multi-omics data

Using a spatial multi-omics dataset that measures spatial coordinates, open chromatin regions, gene expressions, and proteomic profiles, we illustrate the model architecture of PRESENT in Fig. 1a. Built on a multi-view autoencoder, PRESENT models each omics layer via an omics-specific encoder consisting of graph attention neural networks (GATs) [32] and a Bayesian neural network (BNN) [33]. The GAT module aims to incorporates spatial dependencies and omics information, while the BNN module, known for its effectiveness in non-linear regression problems [33], brings the potential to leverage reference data from existing sequencing data [25]. Furthermore, PRESENT fuses omics-specific embeddings from the omics-specific encoders using a cross-omics alignment module which implements a multi-modal contrastive learning loss function, namely the inter-omics alignment (IOA) loss [34]. Finally, to capture the distinct patterns of various omics layers, PRESENT employs omics-specific decoders which model the counts of transcriptomic and epigenomic data via zero-inflation negative binomial (ZINB) distribution [35] and zero-inflation Poisson (ZIP) distribution [36], respectively.

Graph attention network

Given the omics feature vector $\boldsymbol{h_i^{in}}$ and the spatial neighborhood $\mathcal{N}_i$ of spot i, the output embeddings of a graph attention layer for spot i is depicted as:

$$\boldsymbol{h_i^{out}} = \sigma\left(\sum_{j\in\mathcal{N}_i}\alpha_{ij}\boldsymbol{W_g}\boldsymbol{h_j^{in}}\right)$$

$$\alpha_{ij} = \frac{\sigma_{LeakyReLU}\left(\boldsymbol{a}^T\left(\boldsymbol{W_g}\boldsymbol{h_i^{in}}\oplus\boldsymbol{W_g}\boldsymbol{h_j^{in}}\right)\right)}{\sum_{k\in\mathcal{N}_i}\sigma_{LeakyReLU}\left(\boldsymbol{a}^T\left(\boldsymbol{W_g}\boldsymbol{h_i^{in}}\oplus\boldsymbol{W_g}\boldsymbol{h_k^{in}}\right)\right)}$$

where α_{ij} is the attention score between spot i and j, $\boldsymbol{W_g}$ and $\boldsymbol{a}$ are learnable parameters of the graph attention layer, σ indicates the combination of layer normalization and exponential linear unit (ELU) activation function, $\sigma_{LeakyReLU}$ denotes the LeakyReLU activation function with a negative slope of 0.2, and $\oplus$ refers to the concatenation function. In each GAT module, there is one fully connected network (FCN) followed by two graph attention layers.

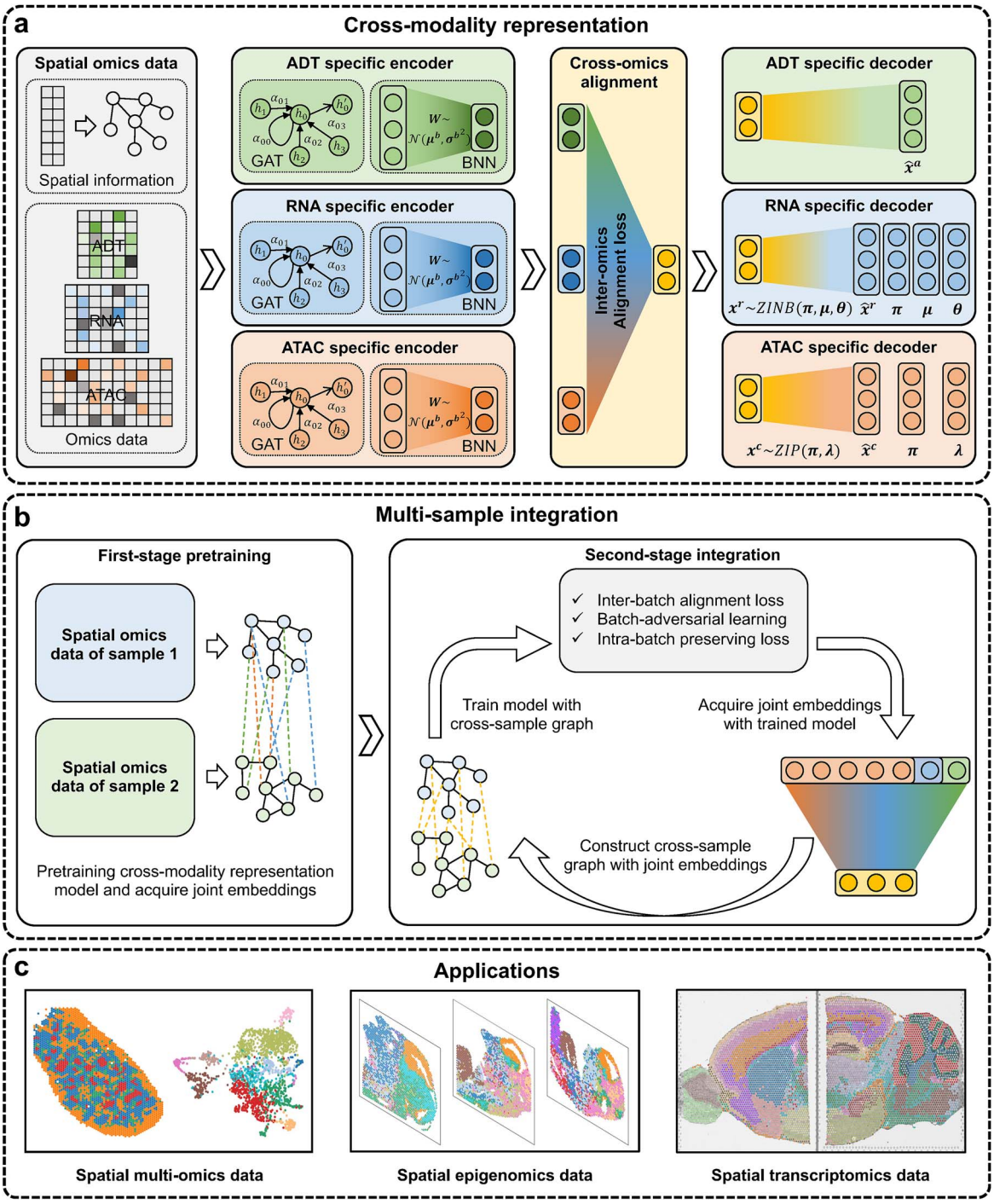


Figure 1 The framework of PRESENT. (a) Cross-modality representation of spatial multi-omics data. PRESENT is capable to integrate spatial coordinate data and multi-omics data, including proteomic data sequenced by ADTs, transcriptomic data by RNA sequencing (RNA), and epigenomic data generated by the assay for transposase-accessible chromatin with sequencing (ATAC). Firstly, an omics-specific encoder is used to extract specific biological variations contained in each omics layer. Each omics-specific encoder is composed of a GAT and a BNN module. All the omics-specific feature matrices extracted by encoders are fused by a cross-omics alignment module, which integrates multi-omics information via the contrastive learning IOA loss function. The output of cross-omics alignment module refers to the final representation matrix of PRESENT. The distinct patterns of different omics layers are addressed via omics-specific decoders, which take the output of cross-omics alignment module as input and model the count matrices of RNA and ATAC data using the ZINB and the ZIP distribution, respectively. (b) Multi-sample integration of spatial omics data. PRESENT adopts a two-stage workflow for the integration of multiple spatial samples. In the first pretraining stage, PRESENT is pretrained using omics data and the cross-sample graph obtained from all the samples. The intra-sample edges of the cross-sample graph are built using spatial coordinates, while inter-sample connections are established based on the similarity of omics features across samples. In the second integration stage, PRESENT implements a cyclic iterative strategy to refine the cross-sample graph using the joint embeddings obtained by the trained model. PRESENT further introduces a contrastive learning-based IBA loss, a batch-adversarial learning strategy and an IBP loss in the second stage to eliminate batch effects while preserving biological signals. (c) Applications of PRESENT in spatial omics data. PRESENT can be easily adapted from spatial multi-omics to spatial single-omics data by simplifying the multi-view autoencoder into a single-view autoencoder. The joint embeddings obtained by PRESENT can be applied to various downstream analyses, including spatial visualization, domain identification and the integrative analysis of samples across conditions and developmental stages.

Bayesian neural network

In BNN module, we assume the parameters of weight matrix, denoted as $\boldsymbol{W}$, follow Gaussian distributions. Each element of $\boldsymbol{W}$ is governed by a Gaussian distribution defined by the mean matrix $\boldsymbol{\mu^b}$ and variance matrix $\boldsymbol{\sigma^{b^2}}$:

$$\boldsymbol{W_{ij}} \sim \mathcal{N}\left(\boldsymbol{\mu_{ij}^b}, \boldsymbol{\sigma_{ij}^{b^2}}\right)$$

Additionally, $\boldsymbol{W_{ij}}$ is governed by a Gaussian prior defined by mean $\boldsymbol{\mu_{ij}^p}$ and variance $\boldsymbol{\sigma_{ij}^{p^2}}$. Inheriting our previous research [25], we calculate the principal component analysis (PCA) projection weight matrix of target omics data as the mean matrix $\boldsymbol{\mu^p}$ and utilize the standard deviation of the mean matrix $\boldsymbol{\mu^p}$ to construct the standard deviation matrix $\boldsymbol{\sigma^p}$. The distance between the prior and the parameters in BNN module is restricted via Kullback–Leibler divergence (KLD):

$$\mathcal{L}_{BNN} = KLD\left(\mathcal{N}\left(\boldsymbol{\mu^b}, \boldsymbol{\sigma^{b^2}}\right), \mathcal{N}\left(\boldsymbol{\mu^p}, \boldsymbol{\sigma^{p^2}}\right)\right)$$

Notably, the parameters of prior Gaussian distributions can also be constructed with external reference data, enabling the incorporation of prior information from existing single-cell or spatial omics data generated by various technologies and platforms. The output embeddings of spot i from BNN module, i.e. $\boldsymbol{h_i^{bnn}}$, is denoted as:

$$\boldsymbol{h_i^{bnn}} = \boldsymbol{W h_i^{in}} = \left(\boldsymbol{\mu^b} + \boldsymbol{\eta} \bigodot \boldsymbol{\sigma^b}\right) \boldsymbol{h_i^{in}}$$

where $\boldsymbol{h_i^{in}}$ denotes the omics feature vector of spot i, $\boldsymbol{\eta}$ is the standard Gaussian noise matrix and $\bigodot$ denotes element-wise product.

Finally, the output embedding of spot i from the omics-specific encoder is the concatenation of output embeddings from the GAT and BNN module.

Cross-omics alignment module

Given the omics-specific embeddings of spot i obtained by omics-specific encoders, including $\boldsymbol{h_i^{ATAC}}$ for ATAC, $\boldsymbol{h_i^{RNA}}$ for RNA and $\boldsymbol{h_i^{ADT}}$ for ADT, we utilized a multi-layer perceptron (MLP) module to fuse the embeddings:

$$\boldsymbol{e_i} = f_{MLP}\left(\boldsymbol{h_i^{ATAC}} \bigoplus \boldsymbol{h_i^{RNA}} \bigoplus \boldsymbol{h_i^{ADT}}\right)$$

where $\boldsymbol{e_i}$ represents the final cross-omics representation of spot i and f_{MLP} denotes the MLP module, which is composed of three FCNs. Furthermore, we design an IOA loss to minimize the multi-omics directional KLD over the output distributions between the joint embeddings and other omics-specific embeddings to improve the multi-modal feature fusion [34], denoted as:

$$\mathcal{L}_{IOA} = \frac{1}{M}\left[\frac{1}{n}\sum_{i=1}^{n} KLD\left(q_{ed}^{(i)}(\boldsymbol{e}), q_{ed}^{(i)}\left(\boldsymbol{h^{ATAC}}\right)\right) + \frac{1}{n}\sum_{i=1}^{n} KLD\left(q_{ed}^{(i)}(\boldsymbol{e}), q_{ed}^{(i)}\left(\boldsymbol{h^{RNA}}\right)\right) + \frac{1}{n}\sum_{i=1}^{n} KLD\left(q_{ed}^{(i)}(\boldsymbol{e}), q_{ed}^{(i)}\left(\boldsymbol{h^{ADT}}\right)\right)\right]$$

where M represents the number of omics layers and $q_{ed}^{(i)}(*)$ represents the embedding distribution of spot i over other spots, defined as:

$$q_{ed}^{(i)}(\boldsymbol{e}) = \left(\frac{\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_1})}{\sum_{j=1}^{n}\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_j})}, \frac{\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_2})}{\sum_{j=1}^{n}\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_j})}, \ldots, \frac{\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_n})}{\sum_{j=1}^{n}\delta_{ed}(\boldsymbol{e_i}, \boldsymbol{e_j})}\right)$$

$$\delta_{ed}(\boldsymbol{u}, \boldsymbol{v}) = exp\left(\frac{\boldsymbol{u}^T\boldsymbol{v}}{\tau}\right)$$

where τ is a temperature parameter and set to 0.1 according to the original study [34].

Omics-specific decoder

Taking the spatial ATAC-seq data as an example, we fit the fragment counts $\boldsymbol{X^{ATAC}}$ with the ZIP distribution to accommodate the lower capture rate of spatial technologies [10, 13, 36]. Using the fragment count vector of spot i, i.e. $\boldsymbol{x_i^{ATAC}}$, as an example, the omics-specific decoder for ATAC data is described as:

$$\boldsymbol{\pi_i} = \sigma_{sigmoid}\left(f_{MLP}^{ATAC,\pi}(\boldsymbol{e_i})\right) \quad \boldsymbol{\rho_i} = f_{MLP}^{ATAC,\rho}(\boldsymbol{e_i}) \quad \boldsymbol{\hat{x}_i^c} = f_{MLP}^{ATAC,x}(\boldsymbol{e_i})$$

$$\boldsymbol{\omega_i} = softmax\left(\boldsymbol{\rho_i} + \boldsymbol{r^c}\right)$$

$$\boldsymbol{\lambda_i} = l_i \cdot \boldsymbol{\omega_i}$$

where $\sigma_{sigmoid}$ represents the sigmoid activation function, $\boldsymbol{r^c}$ is designed to capture region specific bias such as peak length, $l_i = \sum_{j=1}^{r} x_{ij}^c$ refers to the total fragment counts of spot i, $f_{MLP}^{ATAC,\pi}(*)$, $f_{MLP}^{ATAC,\rho}(*)$ and $f_{MLP}^{ATAC,x}(*)$ are three output layers of $f_{MLP}^{ATAC}(*)$, the MLP module of ATAC omics-specific decoder. $f_{MLP}^{ATAC,\pi}(*)$ and $f_{MLP}^{ATAC,\rho}(*)$ output parameters of the ZIP probability model, while $f_{MLP}^{ATAC,x}(*)$ outputs reconstructed normalized profiles of ATAC data. Then for each region $j = 1, 2, \ldots, r$, the fragment count of i-th spot x_{ij}^{ATAC} satisfies:

$$x_{ij}^{ATAC} \sim ZIP\left(x_{ij}^{ATAC}; \pi_{ij}, \lambda_{ij}\right)$$

$$ZIP\left(x_{ij}^{ATAC}; \pi_{ij}, \lambda_{ij}\right) = \pi_{ij}\delta_0\left(x_{ij}^{ATAC}\right) + (1 - \pi_{ij})\, Poisson\left(x_{ij}^{ATAC}; \lambda_{ij}\right)$$

$$\delta_0\left(x_{ij}^{ATAC}\right) = \begin{cases} 1 & if\ x_{ij}^{ATAC}\ equals\ 0 \\ 0 & otherwise \end{cases} \qquad Poisson\left(x_{ij}^{ATAC}; \lambda_{ij}\right) = \frac{\lambda_{ij}^{x_{ij}^{ATAC}} e^{-\lambda_{ij}}}{x_{ij}^{ATAC}!}$$

Therefore, the negative log likelihood (NLL) loss function of the ZIP probability model is:

$$\mathcal{L}_{NLL}^{ATAC} = \frac{1}{nr}\sum_{i=1}^{n}\sum_{j=1}^{r}\left[-logZIP\left(x_{ij}^{ATAC}; \pi_{ij}, \lambda_{ij}\right) + \gamma\pi_{ij}^2\right]$$

where the $\gamma\pi_{ij}^2$ refers to a regularization term and γ is set to 0.5 by default. The reconstruction loss of ATAC omics layer is measured with mean squared error (MSE):

$$\mathcal{L}_{MSE}^{\mathbf{ATAC}} = \frac{1}{nr}\sum_{i=1}^{n}\sum_{j=1}^{r}\left(\hat{x}_{ij}^{ATAC} - \tilde{x}_{ij}^{ATAC}\right)^2$$

where $\tilde{x}_{ij}^{ATAC}$ and $\hat{x}_{ij}^{ATAC}$ denote the predicted and real normalized profiles of spot i in region j.

Similarly, we also model the counts of RNA data $\boldsymbol{X^{RNA}} \in R^{n \times g}$ using the ZINB distribution according to a recent single-cell RNA-seq research [35], as shown in detail in Supplementary Text S1. Compared

with ATAC and RNA data, lower dropout rates, dimensionality and sparsity make the analysis of ADT data less challenging. Therefore, the ADT omics-specific decoder is built on the MLP module with one output layer to reconstruct the normalized expression level of proteins. The final negative log likelihood loss and reconstruction loss are the sum of all modalities.

Loss functions

Based on the above descriptions, the overall loss function of PRESENT for cross-modality representation consists of four parts, including the loss of BNN module, the IOA loss, the negative log likelihood loss of ZIP and ZINB model, and the MSE loss:

$$\mathcal{L}_{repre} = \alpha_1 \mathcal{L}_{BNN} + \alpha_2 \mathcal{L}_{IOA} + \alpha_3 \mathcal{L}_{NLL} + \alpha_4 \mathcal{L}_{MSE}$$

where $\alpha_1, \alpha_2, \alpha_3$ and α_4 are coefficients of the corresponding loss term. The default values of $\alpha_1, \alpha_2, \alpha_3$ and α_4 are all set to 1.

Additional details regarding the model architecture for multi-omics representation learning, including the omics-specific encoders, the cross-omics alignment module, and the omics-specific decoders, have been provided in Supplementary Text S1.

Multi-sample integration of spatial multi-omics data

Multi-sample integration offers a comprehensive perspective for deciphering the structures of cellular organizations [18, 19, 26, 37]. For example, integrated analysis of samples under different stages can reveal developmental dynamics [26, 38], while integrating disease and normal samples facilitates the identification of disease-specific domains and promotes the investigation of disease mechanisms [26, 39]. PRESENT can be extended to the multi-sample integration of spatial multi-omics data by adopting a two-stage training workflow, as shown in Fig. 1b. Before integration, the feature sets of different samples need to be unified, and a cross-sample graph is constructed from the concatenated omics data, as detailed in Supplementary Text S2.

First-stage pretraining

In the first stage, we pretrain the PRESENT framework using the cross-sample graph. To incorporate batch information, we concatenate the one-hot encoded batch labels with the output of the cross-omics alignment module, which are then treated as input to the omics-specific decoders. The objective function and training process during the first pretraining stage are identical to those used for representation learning of individual samples, as described in the previous section.

Second-stage integration

The technical noise between different samples, i.e. batch effects, have not been fully addressed in the first stage. Therefore, we incorporate contrastive learning and batch-adversarial learning strategies in the second stage to eliminate the batch effects across samples while preserving the biological signals simultaneously, as detailed in the following subsections.

Inter-batch alignment

We designed a contrastive learning-based inter-batch alignment (IBA) loss [40] to align spots with similar omics features across different samples. Specifically, for each spot i, we first identify the mutual nearest neighbors from other samples as the positive set, denoted as N_i^{pos}, which refers to the inter-sample neighbors in the cross-sample graph. Next, we randomly select 1/4 of the spots from the same sample as the negative set, depicted as N_i^{neg}. The IBA loss of spot i is defined as:

$$\mathcal{L}_{IBA}^{(i)} = -log \frac{\sum_{j \in N_i^{pos}} exp\left(\delta_{sim}\left(\boldsymbol{e_i}, \boldsymbol{e_j}\right)\right)}{\sum_{j \in N_i^{pos}} exp\left(\delta_{sim}\left(\boldsymbol{e_i}, \boldsymbol{e_j}\right)\right) + \sum_{j \in N_i^{neg}} exp\left(\delta_{sim}\left(\boldsymbol{e_i}, \boldsymbol{e_j}\right)\right)}$$

where $\boldsymbol{e_i}$ denotes the latent embedding of spot i and δ_{sim} denotes the cosine similarity.

Batch-adversarial learning

Inspired by a domain-adversarial learning research [41] and its success in multi-batch integration of single-cell ATAC-seq data [42], we implement the batch-adversarial learning strategy to remove batch effect while preserving biological signals. Specifically, we introduce a batch discriminator to predict the batch index of each spot from its latent embedding. The omics-specific encoders and cross-omics alignment module serve as the generator, aiming to produce latent embeddings without batch effects, thereby confusing the discriminator in an adversarial manner. The batch-adversarial (ADV) loss is defined as:

$$\mathcal{L}_{ADV} = \max_{G} \min_{D} CE\left(P, T\right)$$

where G refers to the generator, D denotes the batch-discriminator, CE represents the cross-entropy loss, P denotes the predicted batch indices and T refers to the ground truth batch labels.

Intra-batch preserving

A key challenge of multi-sample integration lies in eliminating batch effects while simultaneously preserving the biological signals within each individual sample. Therefore, we proposed an intra-batch preserving (IBP) loss to maintain the biological pattern of latent representations within each sample during training process, denoted as:

$$\mathcal{L}_{IBP} = \frac{1}{S} \sum_{k=1}^{S} \frac{1}{\mid s_k \mid} \sum_{i \in s_k} KLD\left(q_{s_k}^{(i)}\left(\boldsymbol{e^{(t)}}\right), q_{s_k}^{(i)}\left(\boldsymbol{e^{(t-1)}}\right)\right)$$

where S represents the number of samples, s_k denotes the spots in sample k, $\boldsymbol{e^{(t)}}$ and $\boldsymbol{e^{(t-1)}}$ refer to the joint embeddings of all spots in current epoch and the previous epoch, respectively. $q_{s_k}^{(i)}(*)$ represents the embedding distribution of spot i over other spots within sample k, defined as:

$$q_{s_k}^{(i)}\left(\boldsymbol{e^{(t)}}\right) = \left(\frac{\delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(1)}}\right)}{\sum_{j=1}^{|s_k|} \delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(j)}}\right)}, \frac{\delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(2)}}\right)}{\sum_{j=1}^{|s_k|} \delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(j)}}\right)}, \ldots, \times \frac{\delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(|s_k|)}}\right)}{\sum_{j=1}^{|s_k|} \delta_{ed}\left(\boldsymbol{e_i^{(t)}}, \boldsymbol{e}^{(t)}_{s_k^{(j)}}\right)} \right)$$

$$\delta_{ed}\left(\boldsymbol{u}, \boldsymbol{v}\right) = exp\left(\frac{\boldsymbol{u}^T \boldsymbol{v}}{\tau}\right)$$

where $s_k^{(j)}$ denotes the j-th spot in sample k, τ is a temperature parameter and set to 0.1 by default [34]. The IBP loss is designed to preserve the overall manifold structure of spots within each sample.

Training process of the second stage

The second-stage training follows a cyclic iterative process. First, we obtain the joint embeddings of multiple samples using the trained PRESENT model. Then, we construct a cross-sample graph, where intra-sample connections are based on spatial coordinates and inter-sample edges on cosine similarity between joint embeddings from different samples. Finally, we train the model for one epoch using the newly constructed graph. This iterative process is repeated until model convergence, which allows the model to progressively differentiate between technical noise and biological signals, reducing the risk of misaligning due to initial noisy graph connections. When updating the PRESENT model in each epoch, we first update the parameters of the batch-discriminator using the cross-entropy loss $\mathcal{L}_{ADV}^{disc} = \beta_3 CE(P, G)$, with parameters of the generator and omics-specific decoders fixed. Next, we fix the parameters of batch-discriminator and update the parameters of generator and omics-specific decoders based on an overall loss function, including the representation learning loss, the IBA loss, the IBP loss and the negative cross-entropy loss for the batch-adversarial learning strategy:

$$\mathcal{L}_{integ} = \mathcal{L}_{repre} + \beta_1 \mathcal{L}_{IBA} + \beta_2 \mathcal{L}_{IBP} + \beta_3 \mathcal{L}_{ADV}^{gen}$$

$$\mathcal{L}_{ADV}^{gen} = -CE(P, G)$$

where β_1, β_2 and β_3 are coefficients of the corresponding loss term and set to 1 by default.

Applications of PRESENT to spatial single-omics data

PRESENT can be seamlessly adapted from spatial multi-omics data to spatial single-omics data, such as SRT and spatial epigenomics data, by simplifying the multi-view autoencoder framework into a single-view autoencoder, which is also detailed in Supplementary Text S3.

Implementation and deployment of PRESENT

PRESENT is implemented based on the Pytorch and PyG framework and further supports mini-batch training via ClusterGCN algorithm [43]. To ensure fairness and robustness, we adopted a fixed global hyperparameter setting across all the experiments, including the hidden dimension, latent dimension, learning rate, and weight decay. These global settings are detailed in Supplementary Text S4.

Datasets collection and evaluation of PRESENT

To evaluate the performance of PRESENT, we have collected comprehensive spatially resolved omics datasets, including spatial multi-omics data generated by spatial RNA-Protein and spatial RNA-ATAC techniques, as well as spatial single-omics data sequenced by spatial ATAC-seq and SRT data. The collection and description of these datasets are shown in Supplementary Text S5. To ensure fairness and transparency, we provide detailed experiment settings for PRESENT versus baseline methods on these datasets in Supplementary Text S6. The quantitative metrics are illustrated in Supplementary Text S7.

Downstream applications of PRESENT

Through comprehensive experiments on spatial datasets generated by spatial multi-omics, epigenomics and transcriptomics technologies, we have demonstrated that the joint embeddings obtained by PRESENT facilitate various downstream applications, including spatial visualization, domain detection, differential gene, or protein expression analysis, gene ontology (GO) analysis and the integrative analysis of tissue samples from different dissection regions or developmental stages (Fig. 1c). The description of downstream analyses is illustrated in Supplementary Text S8.

Results

PRESENT facilitates accurate spatial domain identification in spatial RNA-ADT data

Spatial domain identification, which aims to decipher tissue regions with distinct anatomical structures, cell-type compositions, and cell–cell interactions [44, 45], is a fundamental step in revealing gene regulatory processes and tissue functions in spatial omics data [23, 24]. To evaluate the spatial domain detection performance of PRESENT, we collected a spatial RNA-ADT human lymph node dataset generated by 10x Genomics Visium technology [29], which simultaneously assessed transcriptomic and proteomic expression as well as the spatial locations of spots and was manually annotated by the clinician group according to the histological image [29] (Fig. 2a). We benchmarked PRESENT against five methods including SpatialGlue [29], spaMultiVAE [30], totalVI [46], CiteFuse [47], and Seurat-v4 [48] and implemented the Leiden algorithm [49] to acquire the same number of clusters as the manually annotated domains. Four commonly used metrics [50, 51], including adjusted rand index (ARI), adjusted mutual information (AMI), normalized mutual information (NMI) and homogeneity (Homo), were utilized to quantitatively evaluate the performance of spatial domain detection. PRESENT achieved the best performance across all the four metrics for spatial domain identification (Fig. 2b). As a qualitative evaluation and intuitive illustration of detected domains, we also demonstrated the uniform manifold approximation and projection (UMAP) manifold structure [52] and spatial visualization results of different methods, which is consistent with the quantitative results (Fig. 2d–f). When PRESENT was applied individually to the RNA or ADT modality of the human lymph node dataset (denoted as RNA-only and ADT-only, respectively; Fig. 2c), its performance declined markedly, emphasizing the importance of spatial multi-omics integration in achieving optimal results. We also demonstrated the robustness and reliability of PRESENT against practical data disruptions through extensive experiments, such as signal sparsity and contamination in ADT channels (Supplementary Text S9 and Supplementary Fig. 1).

As the largest secondary lymphoid organ in the body, spleen plays a significant role in hematopoiesis and the clearance of erythrocytes, contributing significantly to immune functions [53]. We further applied PRESENT to a SPOTS mouse spleen dataset [12], which measured spatially transcriptomic and proteomic profiles simultaneously, to decipher the functionality of mouse spleen in the context of cellular organizations. PRESENT identified four main spatial domains (Fig. 2g), denoted as the B cell-enriched domain, the T cell-enriched domain and two macrophage-enriched subdomains, i.e. Mac1-enriched and Mac2-enriched domains, according to our differentially expressed analysis (Fig. 2h and i). We compared the protein expression level of each domain with those of the rest domains based on the Wilcoxon rank-sum test [54] (Fig. 2h), identifying protein markers distinguishing the T cell-enriched domain (CD3, CD4, CD8), the B cell-enriched domain (CD19, B220, lgD) and the macrophage-enriched domain (F4/80, CD68),

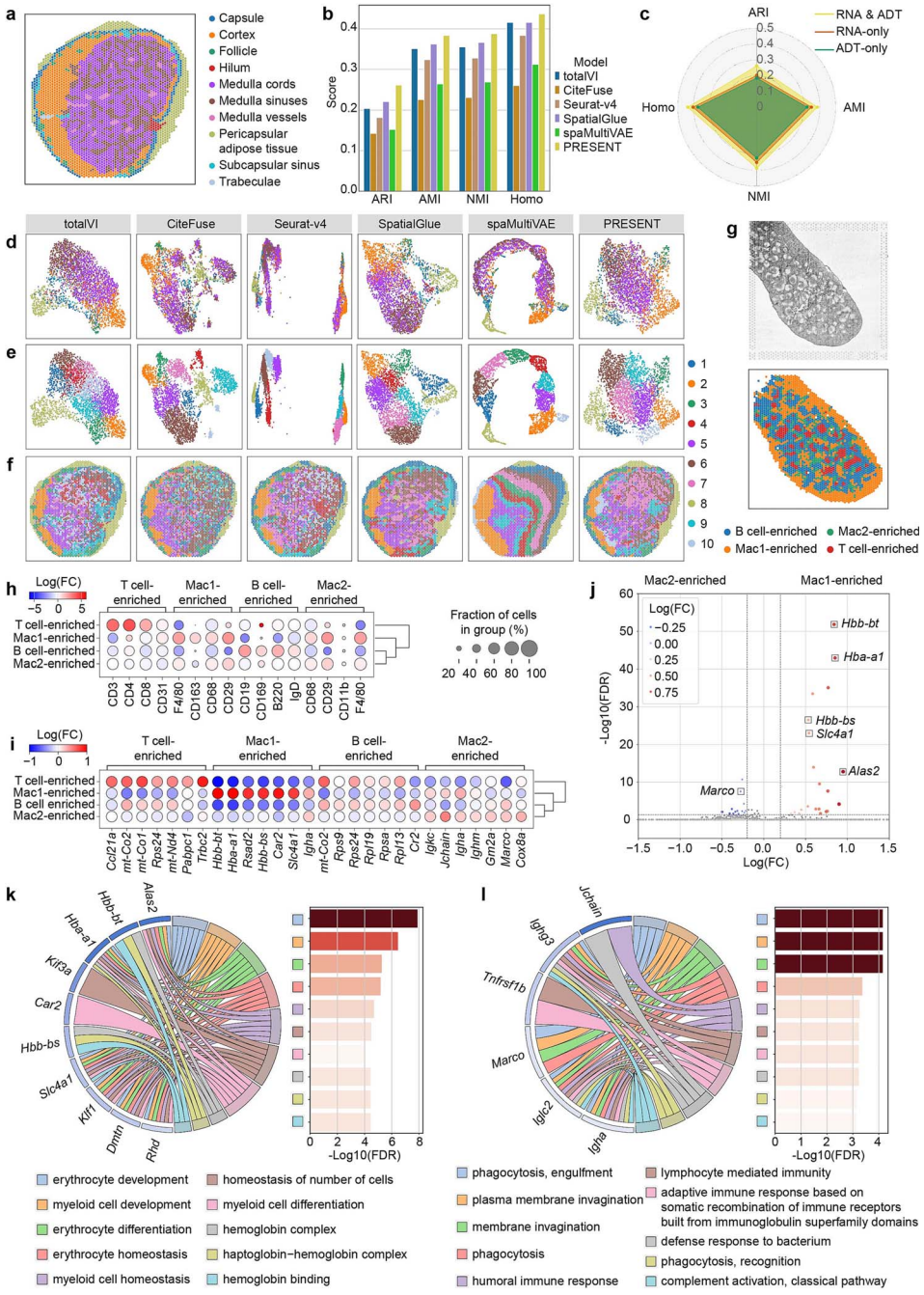


Figure 2 PRESENT facilitates accurate spatial domain identification in spatial RNA-ADT data. (a) spatial visualization of the 10x Genomics Visium RNA-protein human lymph node sample colored by ground truth domain labels. (b) Quantitative comparison of spatial domain identification performance between PRESENT and other baseline methods, shown as a bar plot for the human lymph node dataset. (c) Quantitative comparison between PRESENT utilizing both RNA and ADT data (RNA & ADT) and PRESENT using only RNA (RNA-only) or ADT data (ADT-only), shown as a radar plot in the human lymph node dataset. (d) UMAP visualization of latent embeddings from different methods, colored by ground truth domain labels in the human lymph node dataset. (e) UMAP visualization of latent embeddings, colored by cluster labels in the human lymph node dataset. (f) Spatial visualization of spots colored by cluster labels in the human lymph node dataset. The cluster labels in e and f were derived from latent embeddings of different methods using the Leiden algorithm. (g) Histology image of the SPOTS mouse spleen dataset and spatial visualization of spots colored by cluster labels in the SPOTS mouse spleen dataset. The cluster labels were derived from latent embeddings of PRESENT using Leiden algorithm. (h) Differentially expressed proteins of each spatial domain through one-versus-all Wilcoxon rank-sum test, shown as a dot plot. (i) DEGs of all the spatial domains through one-versus-all Wilcoxon rank-sum test, shown as a dot plot. (j) Volcano plot showing the DEGs of Mac1-enriched domain and Mac2-enriched domain through Mac1-versus-Mac2 Wilcoxon rank-sum test, where the x axis denotes the log(fold-change) (log(FC)), while the y axis denotes the significance measured by -log10(false discovery rate) ($-\log10(FDR)$). The vertical dashed line represents the threshold for log(FC)=$\pm$0.2, while the horizontal dashed line denotes the threshold for -log10(FDR) = 0.05. (k) Chord plot demonstrating the linkage of DEGs in the Mac1-enriched domain and the corresponding enriched pathways. The left semicircle represents DEGs while the right semicircle denotes the enriched biological processes. Bar plot is employed to demonstrate the significance of each pathway (x axis, $-\log10(FDR)$). (l) The linkage of DEGs in the Mac2-enriched domain and corresponding pathways as well as the significance of each enriched pathway.

consistent with the original study [12]. Specifically, the Mac1-enriched and Mac2-enriched domains had almost identical protein markers (Fig. 2h) as the macrophage markers identified in the original study [12], making it hard to distinguish the two subdomains purely by ADT data. Therefore, we further conducted differentially expressed gene (DEG) analysis based by comparing the gene expression levels of each domain with the rest domains (Fig. 2i), revealing the transcriptomic difference of each domain. We also directly compared the gene expression levels of Mac1-enriched domain with that of Mac2-enriched domain (Fig. 2j). Results showed that Mac1-enriched domain could be marked by genes like *Hbb-bt, Hbb-a1, Hbb-bs, Alas2*, and *Slc4a1*, while Mac-2 enriched domain was recognized by *Marco, Jchain,* and *Iglc2*. According to previous studies, *Hbb-bt, Hbb-a1,* and *Hbb-bs* are classical hemoglobin genes [55]. *Alas2* is essential for hemoglobin formation by erythroid cells [56], while *Slc4a1* plays a critical role in maintaining erythrocyte membrane stability [57]. In contrast, *Marco, Jchain,* and *Iglc2* were all reported to be related to immune response [58–60]. *Marco* encodes a scavenger receptor class-A protein, which is expressed on the cell surface of macrophages and is related to cell uptake [58]. GO enrichment analysis also validated the functional difference between the two macrophage-enriched domains (Fig. 2k–l, Supplementary Text S10 and Supplementary Tables S1–S2). We further adopted other baseline methods in this dataset for comparison, highlighting the capability of PRESENT in uncovering biological insights that remain hidden in simpler or unimodal approaches (Supplementary Text S11, Supplementary Figs 2–4). In conclusion, PRESENT effectively integrates complementary multi-omics information and spatial patterns in spatial RNA-ADT data, facilitating accurate spatial domain identification and systematic exploration of biological implications in the identified domains.

PRESENT integrates spatial multi-omics biological signals in spatial RNA-ATAC data

Next, we collected a MISAR-seq mouse brain sample at the E15.5 developmental stage [13], which simultaneously profiles epigenomic and transcriptomic signals, yielding paired spatial RNA-ATAC profiles (Fig. 3a). Compared with spatial RNA-ADT co-profiling data, spatial RNA-ATAC datasets suffer from higher dimension and sparsity. For systematic evaluation, we compared the performance of PRESENT in spatial domain identification with a spatial multi-omics method SpatialGlue [29] and five single-cell multi-omics methods, including MOFA+ [61], scAI [62], scMDC [63], scMVP [64], and Seurat-v4 [48]. PRESENT ranked among the best in this quantitative comparison (Fig. 3b). We further demonstrated the UMAP visualization of the latent representations obtained by different methods in this MISAR-seq mouse brain sample (Fig. 3c and d). Most baseline methods were able to roughly separate domains such as DPallm, DPallv, diencephalon & hindbrain, subpallium 2, and muscle in the UMAP scatter plot. However, MOFA+ and scMVP failed to distinguish the thalamus and midbrain from other domains, while scAI, scMDC, Seurat-v4, and SpatialGlue struggled to discriminate among mesenchyme, cartilage 2, and cartilage 3. In contrast, PRESENT not only identified domains including DPallm, DPallv, diencephalon & hindbrain, subpallium 2, and muscle, but also accurately resolved more challenging domains such as thalamus, mesenchyme, cartilage 2, and cartilage 3. Moreover, spatial domains identified by SpatialGlue and PRESENT exhibited enhanced spatial smoothness, underscoring the importance of integrating spatial information (Fig. 3e).

Furthermore, we have conducted ablation experiments to investigate the impact of biological signals from different omics layers on model performance to provide deeper insights into the model's ability to extract multi-omics biological signals. Specifically, we used PRESENT to obtain latent representations based on only ATAC data or RNA data, denoted as ATAC-only and RNA-only, respectively. As a result, leveraging only ATAC or RNA omics information led to decreased performance, highlighting the importance of integrating multi-omics information in deciphering biological variations in the context of cellular organizations (Fig. 3f). Specifically, compared with the failure of PRESENT using only transcriptomic data (RNA-only) on the E15.5 brain sample, the model based on the epigenomic data (ATAC-only) contributed more to the differentiation between cartilage 2, cartilage 3, and mesenchyme domains (blue circles in Fig. 3g). Therefore, relying solely on a single omics layer might lead to a biased interpretation of biological signals in spatial domain detection. When taking both epigenomic and transcriptomic signals as input, PRESENT successfully captured biological variations among those domains in E15.5 mouse brain sample, providing significantly better performance (Fig. 3f and g), highlighting the capabilities of PRESENT in leveraging transcriptomic & epigenomic multi-omics information for accurate identification of spatial domains.

PRESENT eliminates batch effects while preserving biological signals in spatial multi-omics data

We further collected four MISAR-seq mouse brain samples from four developmental stages, including E11.0, E13.5, E15.5, and E18.5. These mouse brain samples simultaneously measured the transcriptomic and epigenomic signals of the developmental mouse brain in the context of cellular organizations [13]. However, varying experimental conditions and sequencing batches introduce potential technical biases, resulting in significant discrepancies across different samples. These discrepancies, commonly referred to as batch effects in single-cell studies, pose challenges to the integrative analysis of samples from different conditions and developmental stages [20, 37]. For illustration, we applied joint PCA and UMAP algorithm to raw RNA and ATAC data of the four MISAR-seq mouse brain samples, revealing pronounced batch effects (Fig. 4a). Moreover, during the development of mouse brain (Fig. 4b), several functional structures might be specific to certain developmental stages (e.g. primary brain), while others may persist throughout the entire developmental process (e.g. midbrain and muscle). Therefore, the key challenges in integrating the MISAR-seq mouse brain samples lied not only in the incorporation of spatial multi-omics information and the correction of batch effects across samples, but also in the conservation of shared and sample-specific biological variations.

Due to the lack of multi-sample integration methods specially designed for spatial multi-omics data, we compared the integration performance of PRESENT with five methods developed for single-cell data (MNN [65], Scanorama [21], Harmony [20], SCALEX [66], and SnapATAC2 [67]) and three recently proposed methods designed for SRT data (STAligner [26], Stitch3D [27], and GraphST [18]). Specifically, MNN, Scanorama, Harmony, and SCALEX could be applied to either RNA-seq or ATAC-seq data, denoted as RNA-$*$ and ATAC-$*$, respectively, where $*$ represents these methods. SnapATAC2 provides detailed guidance for incorporating multi-omics (i.e. transcriptomics & epigenomics) information, serving as an integrative workflow for single-cell multi-omics data. To intuitively evaluate the batch

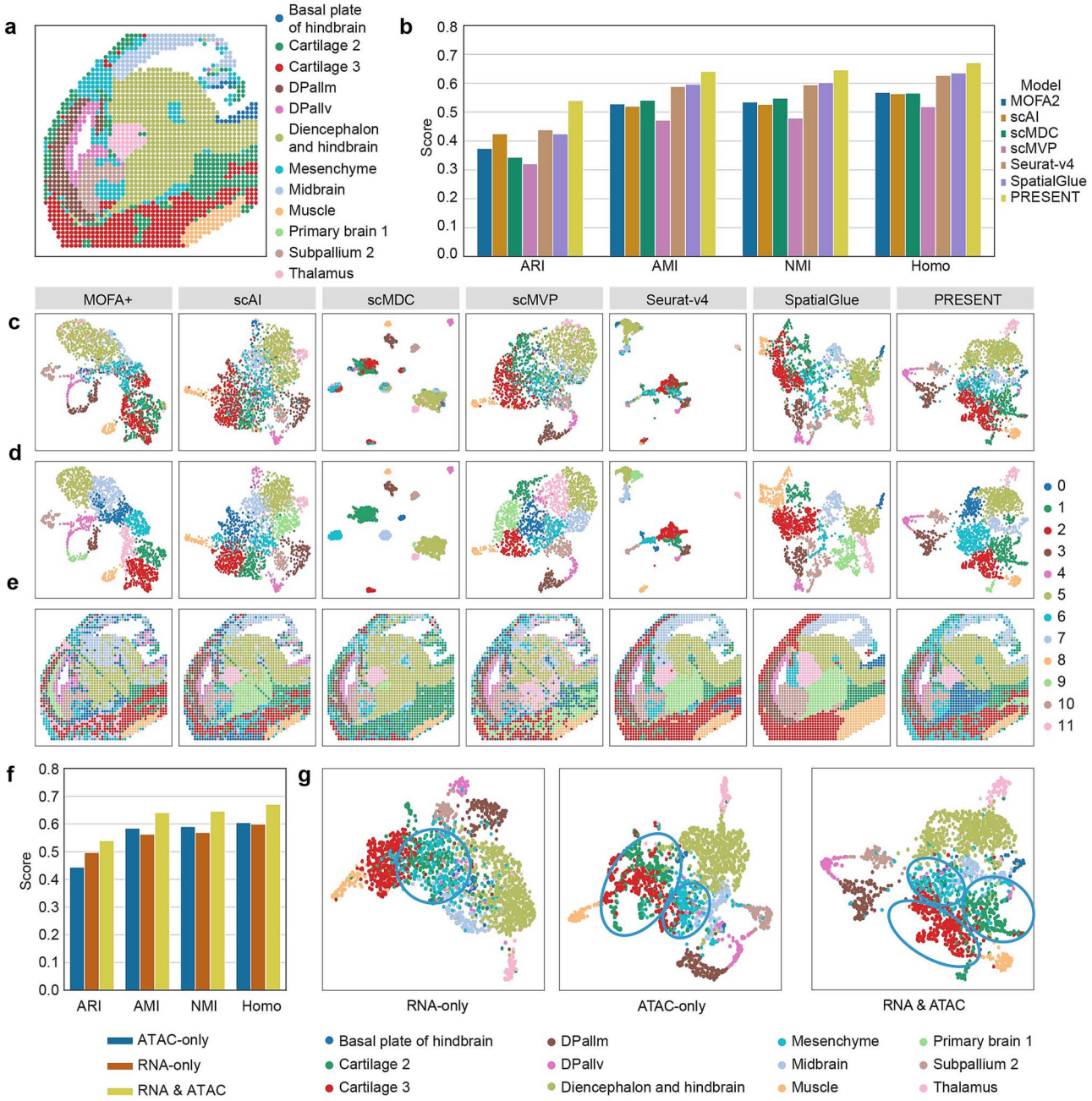


Figure 3 PRESENT integrates spatial multi-omics biological signals in spatial RNA-ATAC data. (a) Spatial visualization of spots in the MISAR-seq E15.5 mouse brain sample, colored by ground truth domain labels. (b) Quantitative performance of PRESENT and other baseline methods for spatial domain identification, illustrated by a bar plot. (c) UMAP visualization of latent embeddings obtained by different methods, colored by ground truth domain labels. (d) UMAP visualization of latent embeddings obtained by different methods, colored by cluster labels. (e) Spatial visualization of spots colored by cluster labels in the E15.5 mouse brain sample. The cluster labels in d and e were derived from latent embeddings of different methods using the Leiden algorithm. (f) Quantitative performance of PRESENT utilizing both RNA and ATAC data (RNA & ATAC) and PRESENT using only ATAC (ATAC-only) or RNA (RNA-only) data. (g) UMAP visualization of latent embeddings obtained by PRESENT using only RNA data, only ATAC data and both RNA & ATAC data, colored by ground truth domain labels in the E15.5 mouse brain sample.

integration and biological conservation capabilities of PRESENT and other baseline methods, we systematically benchmarked them using 14 metrics, evenly divided into two categories: batch effect removal and biological variance conservation, following a recent benchmark framework for single-cell omics data integration [68]. PRESENT, which simultaneously integrated spatial dependency and multi-omics information of all the samples, ranked among the best in both batch effect removal and biological variance conservation among all the 14 methods (Fig. 4c). Among these baseline methods, SnapATAC2 also provided satisfactory performance, stressing the importance of incorporating multi-omics signals.

We further applied UMAP algorithm to the joint embeddings generated by PRESENT and other baseline methods for visualization, with the spots colored by batch indices (Fig. 4d and Supplementary Fig. 5a)

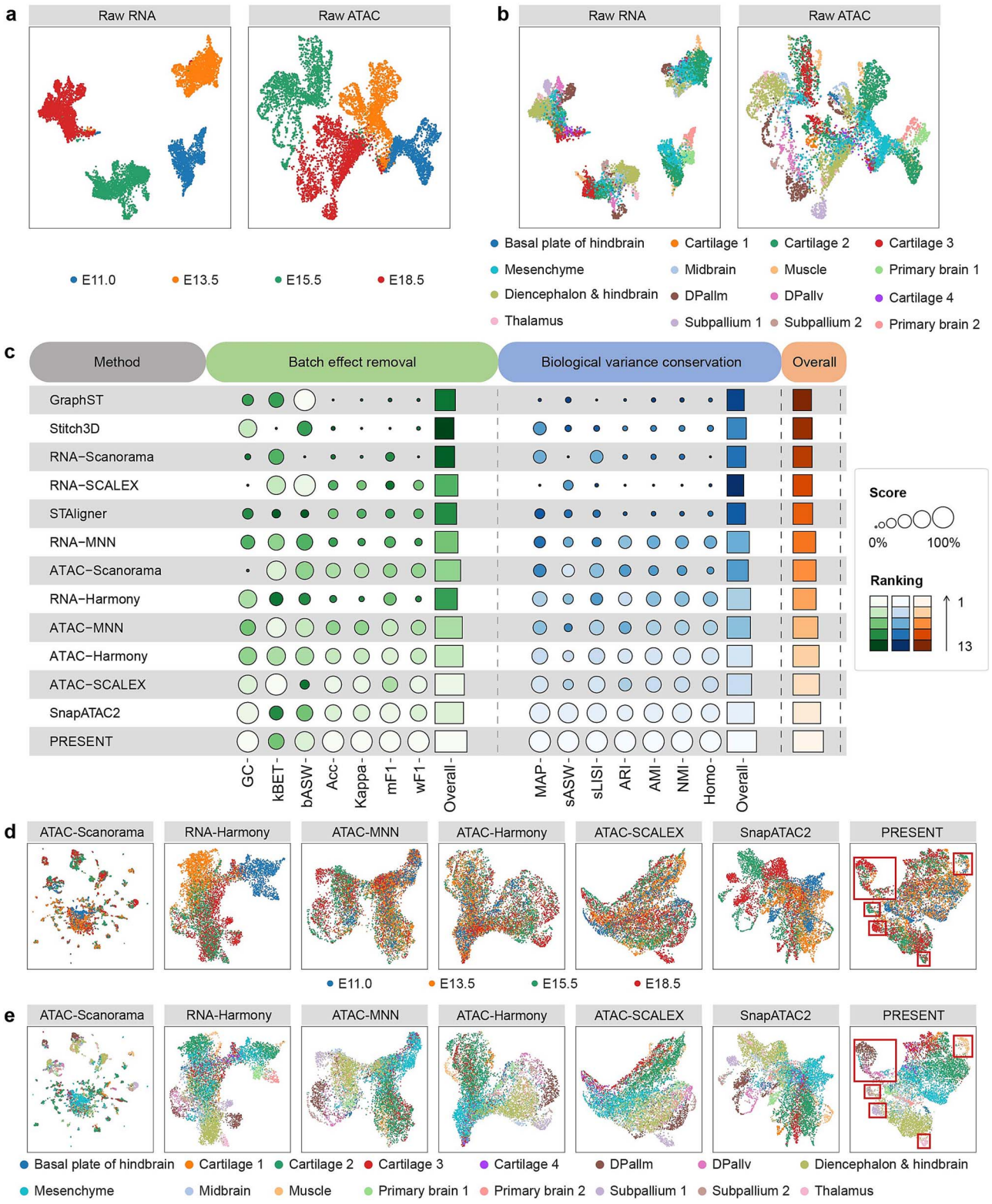


Figure 4 PRESENT eliminates batch effects while preserving biological signals in spatial multi-omics data. (a) UMAP visualization of spots across the four MISAR-seq mouse brain samples based on raw RNA or ATAC data colored by sample indices. (b) UMAP visualization of spots across the four MISAR-seq mouse brain samples based on raw RNA or ATAC data colored by ground truth domain labels. Note that prior to applying the UMAP algorithm to raw RNA or ATAC data, we first normalize the omics data and reduce the dimensionality to 50 using PCA algorithm. (c) Quantitative evaluation of different integration methods on the four MISAR-seq mouse brain samples using 14 metrics divided into two categories, namely batch effect removal and biological variance conservation. The category scores of these two aspects were calculated by averaging the metrics within each category. An overall score for each integration method was computed using a 40/60 weighted mean of the category scores for batch effect removal and biological variance conservation. (d) UMAP visualization of spots across the four MISAR-seq mouse brain samples based on the joint embeddings produced by different integration methods colored by sample indices. (e) UMAP visualization of spots across the four MISAR-seq mouse brain samples based on the joint embeddings produced by different integration methods colored by ground truth domain labels.

and ground truth domain labels (Fig. 4e and Supplementary Fig. 5b). PRESENT successfully integrated spots from different developmental stages (red boxes in Fig. 4d) while maintaining the biologically distinct patterns of DPallm, DPallv, subpallium 1 & 2, thalamus, and muscle (red boxes in Fig. 4e). Specifically, spots of muscle domain persisted throughout the entire life of mouse brain development, while subpallium 1&2, DPallm, DPallv, and thalamus only existed in specific stages. The success of PRESENT in distinguishing these domains illustrated its capabilities in eliminating batch effects while capturing shared and stage-specific biological variance. Additionally, we conducted joint spatial clustering across the four stages based on the joint embeddings obtained by different methods (Supplementary Figs 5c and 6). Except for SnapATAC2, all the baseline methods delivered limited clustering results. In contrast, PRESENT successfully detected not only large domains like DPallm (cluster 7), but also small domains including thalamus (cluster 16) and muscle (cluster 11). Notably, when performing single-sample analyses on the E18.5 MISAR-seq mouse brain based on PRESENT, it failed to capture the muscle domain as a distinct, coherent biological structure (Supplementary Text S12 and Supplementary Fig. S7), further underscoring the advantages of multi-sample integration. In conclusion, PRESENT eliminates batch effects through the IBA loss and batch-adversarial learning strategy, while preserving biological signals by incorporating the IBP loss (Materials and Methods), thereby facilitating the integrative analysis of multiple spatial multi-omics datasets.

PRESENT facilitates effective cross-modal representation in spatial single-omics data

PRESENT is highly extensive and can be seamlessly adapted from spatial multi-omics data to spatial single-omics data, such as spatial epigenomics and transcriptomics data, by simplifying the multi-view autoencoder into a single-view model (Supplementary Text S3). We first assessed the performance of PRESENT on spatial epigenomics data by collecting six spatial ATAC mouse embryo samples, which originated from three developmental stages [10], including E15.5, E13.5, and E12.5. Each developmental stage containing two replicated samples, denoted as $*$-S1 and $*$-S2, respectively. Initially, we evaluated the data quality of the six samples using logarithmic total fragment counts, namely log10(nFrags), and transcription starting site (TSS) enrichment scores. The former metric represents the cellular sequencing depth of samples [69], while the latter reflects the signal-to-noise ratio of datasets [70]. The samples of E15.5 stage, i.e. E15.5-S1 and E15.5-S2, achieved the highest sequencing depth and signal-to-noise ratio, while the E13.5-S1 sample and E12.5-S2 sample suffered from the lowest sequencing depth and signal-to-noise ratio, respectively (Fig. 5a and b). For spatial domain identification, we benchmarked PRESENT and other six baseline methods, including three designed for single-cell ATAC-seq data (SCALE [71], cisTopic [72], and SnapATAC2 [67]) and three for SRT data (SpaGCN [23], STAGATE [24], and PAST [25]). PRESENT achieved the highest performance, surpassing the second-best baseline by 8.60%, 26.13%, 25.91%, and 22.94% across the respective clustering metrics (Fig. 5c). Using the E15.5-S1 sample as an illustrative example, we visualized the UMAP and spatial embeddings of spots colored by manually annotated domains (Fig. 5d and e) and by the spatial cluster labels (Fig. 5f), which were consistent with the quantitative evaluations. Similar results were also observed across the remaining five samples (Supplementary Figs S8–S13).

Despite the superiority of PRESENT over existing baseline methods, low sequencing depth and signal-to-noise ratio stills constrained the effectiveness of computational models (Supplementary Fig. S14). Our previous study has shown that integrating external reference data through BNN module can effectively address issues related to high noise levels in SRT data [25]. Inspired by this, we collected a single-nucleus ATAC-seq mouse embryo dataset involving cells from three developmental stages, including E12, E13, and E15, which was sequenced in conjunction with the spatial ATAC mouse embryo data [10]. We utilized the single-nucleus ATAC-seq dataset (denoted as PRESENT-SN) and the other five spatial ATAC samples (denoted as PRESENT-SP) as the reference data for PRESENT when analyzing each spatial ATAC mouse embryo sample. The integration of prior information from reference data facilitated enhanced performance of PRESENT, especially in samples with lower sequencing depth and signal-to-noise ratio, such as the E13.5-S1 and E12.5-S2 mouse embryo samples (Fig. 5g and Supplementary Fig. S15). This finding was further validated by the UMAP visualization of the two low-quality samples (Fig. 5h, Supplementary Figs S10, S13). Specifically, PRESENT-SN and PRESENT-SP successfully identified spatial domains including liver, limb, forebrain, hindbrain, and midbrain, which were not distinguished by other baseline methods or by the original PRESENT model. We further increased the noise level of datasets by converting lower fragment counts to zero with growing dropout rates for each sample, according to a study that models the dropout events in single-cell data [73]. PRESENT-SN and PRESENT-SP gained greater advantages as the noise level increased, highlighting the improved benefits of integrating prior information from reference data with the growing dropout rates in spatial ATAC-seq datasets (Fig. 5i).

We also carried out extensive experiments on SRT axolotl brain datasets generated by Stereo-seq, which include six samples from different developmental stages, including stage 44, stage 54, stage 57, juvenile, adult, and metamorphosis [7]. The performance of PRESENT was benchmarked with several recently developed SRT methods, involving SpaGCN [23], STAGATE [24], PAST [25], and SpatialPCA [74]. Experimental results demonstrated the consistent superiority of PRESENT in cross-modal representation, spatial visualization and domain identification of SRT data (Supplementary Figs S16–S22). In addition, utilizing an SRT regenerated axolotl brain sample at 20 days post-injury as reference data [7], PRESENT-SP again gained promoted performance through leveraging prior information contained in reference data (Supplementary Fig. S16). In summary, PRESENT not only promotes effective cross-modal representation in both spatial ATAC-seq and SRT data, but also enables the integration of reference data to address the low sequencing depth and signal-to-noise ratio.

PRESENT integrates single-omics samples of multiple developmental stages or dissected areas

Integrated analysis of multiple samples across different developmental stages (vertically serial tissue samples) and dissected areas (horizontally adjacent tissue samples), commonly referred to as vertical and horizontal integration, provides a powerful approach for deciphering the hierarchical structures of cellular organizations [18]. We first evaluated the performance of PRESENT in the integrative analysis of vertically serial tissue samples, using the mouse embryo samples from three developmental stages (E15.5, E13.5, and E12.5) generated by the spatial ATAC technique [10]. We compared PRESENT with five single-cell methods (MNN [65], Scanorama [21], Harmony [20], SCALEX [66], and SnapATAC2 [67]) and two SRT methods (STAligner [26] and GraphST [18]).

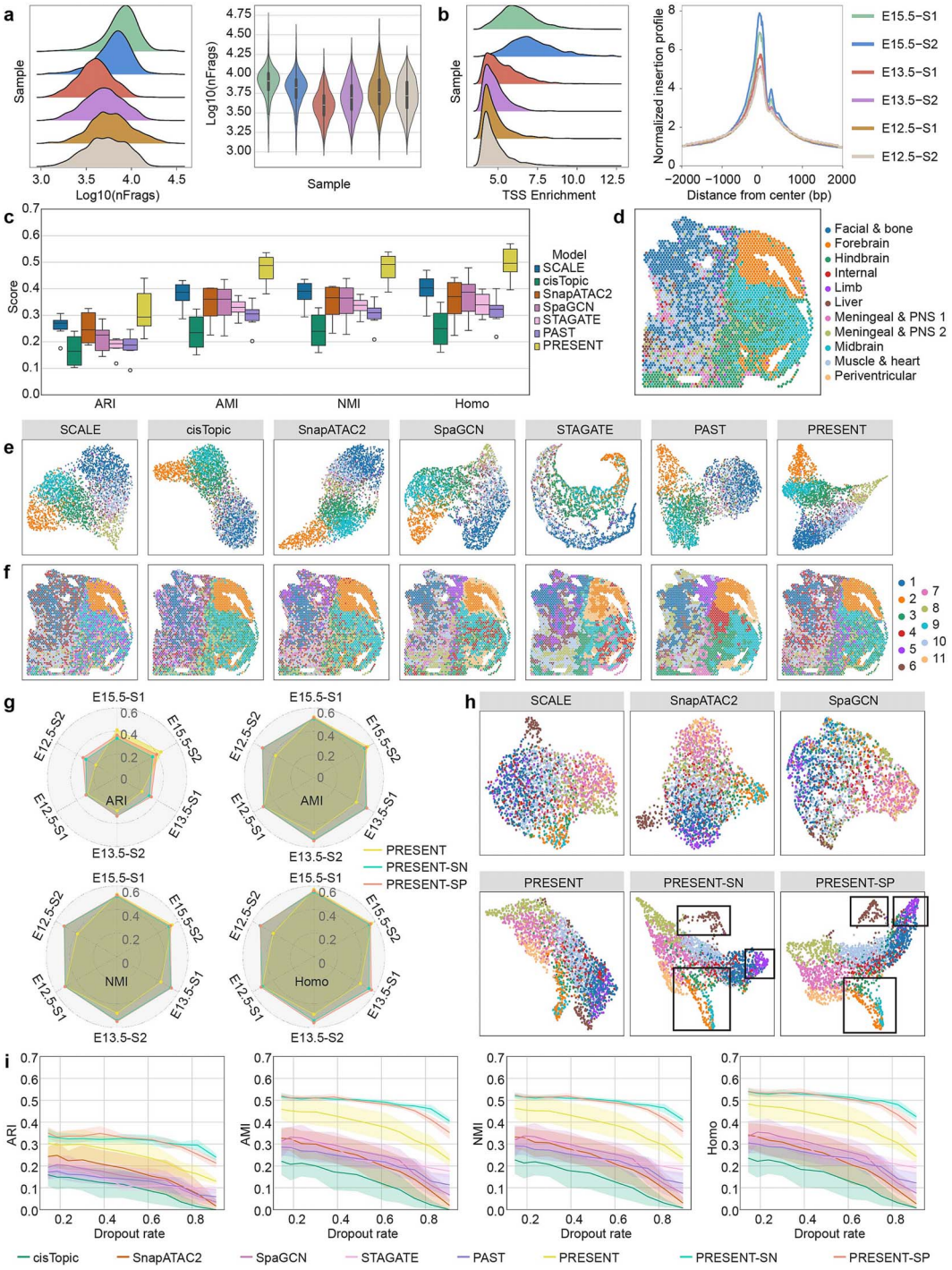


Figure 5 PRESENT facilitates effective cross-modal representation in spatial single-omics data. (a) Log10 total fragment counts, namely log10(nFrags), across the six spatial ATAC mouse embryo samples, illustrated by a ridge plot and a violin plot. (b) TSS enrichment scores for spots and the enrichment profiles of fragments around the TSS across the six mouse embryo samples. (c) Quantitative evaluation of spatial clustering performance of different methods on the six mouse embryo samples using four metrics, including ARI, AMI, NMI and Homo. The center line, box limits and whiskers in the box plot are the median, upper and lower quartiles, and 1.5× interquartile range, respectively. (d) Spatial visualization of spots from the E15.5-S1 mouse embryo sample, colored by ground truth domain labels. (e) UMAP visualization of the latent embeddings obtained by different methods on the E15.5-S1 mouse embryo sample, colored by ground truth domain labels. (f) Spatial visualization of spots from the E15.5-S1 mouse embryo sample, colored by spatial cluster labels. The cluster labels were derived from latent embeddings of different methods using the Leiden algorithm. (g) The detailed quantitative comparison between PRESENT and PRESENT-SN & SP on the six mouse embryo samples based on radar plots. (h) UMAP visualization of the latent representations obtained by different methods colored by ground truth domain labels on the E12.5-S2 mouse embryo sample. (i) Spatial clustering performance of baseline methods, PRESENT and PRESENT-SN & SP on the six mouse embryo samples under different levels of dropout rate. Here SCALE was omitted as it encountered error when the dropout rate was larger than 0.2.

Similarly, we employed the benchmarking framework consisting of 14 metrics to simultaneously assess the batch effect removal and biological variance conservation capabilities of these integration methods. PRESENT again ranked first in biological variance conservation and second in batch effects removal among all the integration methods (Fig. 6a). The UMAP visualization of spot embeddings across different developmental stages also demonstrated that, compared with other methods, PRESENT gained better performance in aggregating spots from the same domain but in different batches (Supplementary Fig. S23). We further performed spatial clustering based on the joint embeddings produced by different methods. Scanorama and STAligner could hardly distinguish different spatial domains across different developmental stages (Fig. 6b and c). MNN, Harmony, SCALEX, and SnapATAC2 provided improved results by roughly recognizing forebrain and liver in E12.5 and E13.5 stages, but they still failed in other domains. In comparison, PRESENT not only identified forebrain (cluster 1) and liver (cluster 5) in all the three stages, but also gained satisfactory results in domains including hindbrain (cluster 2) and midbrain (cluster 8), consistent with the quantitative results.

Furthermore, we collected two horizontally adjacent SRT mouse brain samples generated by the 10x Genomics Visium platform, which involves a sagittal anterior dissected section and a sagittal posterior dissected section. We applied two SRT integration methods, namely STAligner [26] and GraphST [18], and PRESENT to perform integrated analysis of these horizontally adjacent brain samples and utilized the anatomic annotation of the sagittal region in postnatal day 56 (P56) mouse brain provided by Allen Reference Atlas [75] to intuitively compare the performance of these methods (Fig. 6d). Compared with STAligner and GraphST, analytical results of PRESENT achieved improvements in shared and sample-specific domain identification (Fig. 6e–g). For example, PRESENT captured the classical hierarchical structures, i.e. from outer layer 1 to inner layer 6, of dorsal pallium in both sagittal anterior and posterior sections, as illustrated in the red rectangle of Fig. 6d. In addition, PRESENT successfully distinguished the typical cortex Ammonis (CA, cluster 18) and dentate gyrus (DG, cluster 25) structures simultaneously in the hippocampus (black frames in Fig. 6d), which have been proved to be associated with mood, anxiety and personality disorders caused by various factors [76]. In contrast, although STAligner and GraphST detected larger domains encompassing both the hippocampal CA and DG, they failed to capture the biological differences between them (Supplementary Figs S24, S25). We further conducted differentially expressed analysis for domain CA and domain DG, acquiring marker genes including *C1ql2* and *Hpca*, respectively. According to previous studies, *C1ql2* serves as a marker gene for dentate gyrus granule cells enriched in the DG structure [77], while *Hpca* is the encoding gene for a high-affinity calcium-binding protein restricted to the central nervous system, abundant in pyramidal cells of the hippocampus CA region [78].

Moreover, PRESENT also exhibited great potential in capturing sample-specific functional domains and uncovering biological implications in identified spatial domains. PRESENT successfully revealed the sophisticated substructures of the cerebellar vermis, which only exists in the sagittal posterior section of mouse brain and plays important roles in coordinating movements of the central body such as trunk, head, and proximal limbs [79]. Specifically, PRESENT successfully detected the molecular layer (cluster 3), Purkinje cell layer (cluster 22), and internal granular layer (cluster 17) of cerebellar vermis cortex and the white matter (cluster 20) of cerebellar vermis, while both STAligner and GraphST were unable to identify the Purkinje cell layer of the cerebellar vermis cortex (Supplementary Figs S26–S29). The DEGs in the corresponding domains have also been validated in previous studies, such as *S100b* [80], *Car8* [81], *Atp1a1* [82], and *Mbp* [83]. For example, *Car8* and *Atp1a1*, the mutation of which cause the motor coordination defect and dominant Charcot–Marie-Tooth type 2, were proved to be highly expressed in the granular layers and Purkinje cells in recent studies [81, 82], respectively. In summary, these results highlighted the superior performance of PRESENT in the integrative analysis of both vertically serial and horizontally adjacent samples generated by various technologies, thereby enabling the discovery of biological insights within identified spatial domains across multiple samples.

Discussion

Recent breakthroughs in spatial omics sequencing technologies have enabled the simultaneous acquisition of spatial coordinates and regulatory states of biomolecules, accumulating rich spatial omics data in repositories and databases [84, 85]. In this article, we proposed PRESENT, an effective contrastive learning framework, to address the multi-omics and multi-sample integration of spatial omics data. Through extensive experiments on different species involving human, mouse, and axolotl, we highlighted the superior performance of PRESENT in analyzing spatial omics data from various technologies, including spatial multi-omics techniques (spatial RNA-Protein [12, 29] and spatial RNA-ATAC [13]) and spatial single-omics technologies (spatial ATAC [10] and spatial transcriptomics [7]). Specifically, PRESENT effectively integrated the spatial dependency and complex multi-omics information contained in spatial multi-omics or single-omics techniques (Figs 2, 3, and 5). More importantly, PRESENT can be further extended to the integrative analysis of multiple spatial samples across distinct dissection areas and developmental stages, thereby promoting the detection of hierarchical structures and biological functions within cellular organizations (Figs 4 and 6).

The BNN offers the potential to leverage reference data from existing sequencing datasets. By default, PRESENT constructs the BNN prior weights using the target dataset itself, ensuring that the model remains fully functional and effective even in the absence of external reference data. Optionally, PRESENT can address challenges related to low sequencing depth and signal-to-noise ratio in spatial omics data by utilizing matched or even cross-platform external reference data (Fig. 5). However, we do not recommend using cross-species datasets as a reference source, due to at least three fundamental reasons. First, genomic feature alignment between different species is inherently complex and unreliable. For instance, differences in chromosome numbers and structures between humans and mice could introduce systematic biases during external reference integration. Second, non-coding regulatory elements, such as enhancers, evolve significantly faster than coding sequences, indicating that a cross-species prior might introduce irrelevant signals that are not present in the target species. Third, even orthologous genes often exhibit divergent spatial regulatory patterns across species due to evolutionary divergence in gene regulatory networks.

PRESENT incorporates several critical components tailored to the challenges of spatial multi-omics data. Based on the MISAR-seq mouse brain datasets [13], we performed comprehensive ablation studies to disentangle the contributions of key components, including the BNN and GAT module, the IOA loss for multi-omics integration, as well as the IBA loss, IBP loss, batch-adversarial strategy, and iterative graph refinement for multi-sample integration (Supplementary Text S13).

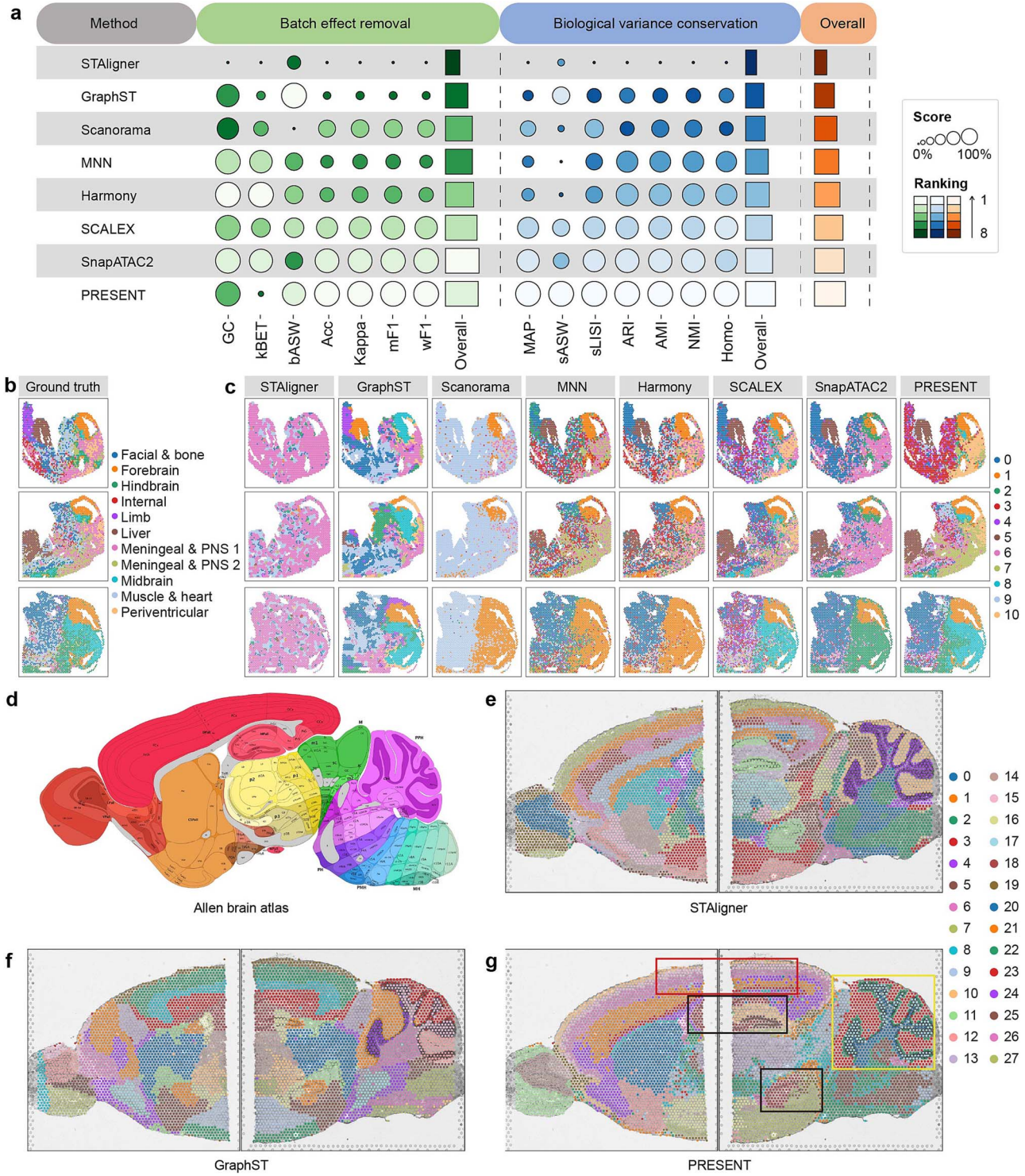


Figure 6 PRESENT integrates single-omics samples of multiple developmental stages or dissected areas. (a) The quantitative evaluation of different integration methods on the three spatial ATAC mouse embryo samples using 14 metrics divided into two categories, namely batch effect removal and biological variance conservation. The category scores of these two aspects were calculated by averaging the metrics within each category. An overall score for each integration method was computed using a 40/60 weighted mean of the category scores for batch effect removal and biological variance conservation. (b) The spatial visualization of spots across the three spatial ATAC mouse embryo samples colored by ground truth spatial domains. (c) The spatial visualization of spots across the three spatial ATAC mouse embryo samples colored by the spatial clusters identified based on different integration methods. The first, second and third row of b and c denotes the samples from E12.5, E13.5 and E15.5 stages, respectively. (d) The anatomic annotation of the sagittal region in P56 mouse brain provided by Allen Reference Atlas [13]. (e) The joint spatial clustering results based on the latent embeddings obtained by STAligner on the two horizontal mouse brain sagittal samples generated by the 10x Genomics Visium platform. (f) The joint spatial clustering results based on the latent embeddings obtained by GraphST on the two horizontal mouse brain sagittal samples generated by the 10x Genomics Visium platform. (g) The joint spatial clustering results based on the latent embeddings obtained by PRESENT on the two horizontal mouse brain sagittal samples generated by the 10x Genomics Visium platform.

Table 1 Summary of strengths, limitations, and scope of applicability for PRESENT.

Category	Key characteristics
Strengths and recommended use cases of PRESENT	Cross-modal representation learning of spatial multi-omics data, including epigenomic, transcriptomic, and proteomic data Multi-sample integration of spatial multi-omics data, including epigenomic, transcriptomic, and proteomic data Spatial domain identification and downstream analyses based on the integrative embeddings obtained by PRESENT Analyses of spatial datasets with low sequencing depth and signal-to-noise ratio by leveraging matched reference data
limitations and non-recommended use cases of PRESENT	Analyses of spatial datasets by leveraging cross-species reference data Multi-sample integration of unpaired spatial samples, such as unpaired spatial ATAC-seq and spatial transcriptomics data Alignment of spatial coordinates from multiple samples Integration of dissociated single-cell data and spatial resolved omics data Integration of histology image and spatial omics information

The results demonstrate that each component contributes to the superior performance of PRESENT (Supplementary Figs S30–S32). Furthermore, we conducted exploratory experiments to assess the interpretability of the GAT and BNN modules using the SPOTS mouse spleen dataset, marking an initial step toward model interpretability (Supplementary Text S14 and Supplementary Fig. S33). To assess the robustness of PRESENT, we also conducted sensitivity analyses for key hyperparameters, including latent space dimensionality, the number of neighbors, the temperature parameter τ in the IOA loss, and the coefficients of loss functions (Supplementary Text S15). PRESENT consistently achieved the best or comparable performance compared to the second-best baseline method, demonstrating its robustness across different hyperparameter settings (Supplementary Figs S34, S35). Moreover, scalability with large-scale datasets are crucial aspects of integrative frameworks. Through systematic benchmarking on datasets with varying modalities and spot numbers, we demonstrate PRESENT can efficiently handle large-scale data without exponential increases in runtime or memory usage (Supplementary Text S16 and Supplementary Fig. S36).

Nevertheless, several avenues remain for further improvement of PRESENT. First, Due to technical and cost limitations, the number of unpaired spatial samples exceeds that of paired spatial samples [86], emphasizing the importance of unpaired integration. While PRESENT can anchor shared spots to analyze paired multi-omics sequencing data or leverage shared genomic features to integrate multiple spatial samples, it still struggles with the integration of unpaired samples with distinct feature space [86]. Second, beyond integrating samples into the latent space, the alignment of spatial coordinates in physical space also plays an important role in joint analysis of multiple spatial samples [19, 37]. Third, integrating single-cell and spatial data is crucial to compensate for the loss of spatial coordinates in the former and the coarse resolution of the latter. For example, Li *et al.* applied a style transfer learning paradigm, SpaIM [87], to impute spatial transcriptomics using single-cell data, which effectively reconstructs sparse gene signals and provides an enhanced foundation for downstream tasks. Finally, integrating histological morphology from high-resolution imaging, such as 10x Genomics Xenium, remains a promising frontier. For instance, xSiGra [88] jointly models multi-channel images and gene expression through a hybrid graph transformer. By leveraging an interpretable gradient-weighted mapping, this approach identifies pivotal genes and cellular interactions, demonstrating that histological morphology can significantly improve our understanding of tissue heterogeneity. To conclude, we also provide a concise summary of the strengths, limitations, and scope of applicability for PRESENT in Table 1, allowing readers to better understand the boundaries of PRESENT and avoid overgeneralization.

Key Points

- We propose PRESENT, a contrastive learning framework grounded in a multi-view autoencoder, for cross-omics and multi-sample integration of spatial multi-omics data, including epigenomic, transcriptomic, and proteomic modalities.
- To capture the unique patterns of individual omics layers, PRESENT employs omics-specific encoders consisting of graph attention networks and Bayesian neural networks, as well as omics-specific decoders that model transcriptomic and epigenomic data via zero-inflation negative binomial and zero-inflation Poisson distributions, respectively.
- PRESENT incorporates a contrastive learning-based inter-omics alignment loss to integrate multi-omics information inherent in spatial multi-omics data.
- PRESENT adopts a contrastive learning-based inter-batch alignment loss and a batch-adversarial learning strategy to eliminate batch effects, along with an intra-batch preserving loss to retain biological signals for integrative analysis across multiple samples.
- PRESENT enables the simultaneous integration of spatial multi-omics data across multiple tissue samples, facilitating the detection of hierarchical structures from a spatiotemporal view.

Author contributions

R.J. and L.Z. conceived the study and supervised the project. Z.L. designed, implemented, and validated the PRESENT framework. X. Cui, X. Chen, Z.Gao, Y.L., Y.P., S.C., and H.L. helped analyzing the results. Z.L. and R.J. wrote the manuscript, with input from all the authors.

Supplementary material

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

Conflicts of interest

The authors declare that they have no competing interests.

Funding

This work was supported by the National Key Research and Development Program of China (grant nos. 2025YFC3409300, 2023YFF1204802), the National Natural Science Foundation of China

(grant nos. 32550616, 62273194), Beijing Natural Science Foundation (grant no. L242026) and Tsinghua University Initiative Scientific Research Program (grant no. 20257020014).

Code availability

The code as well as the detailed tutorials for implementing PRESENT can be found at the Github website https://github.com/lizhen18THU/PRESENT.

Data availability

The single-nucleus ATAC-seq and spatial ATAC mouse embryo datasets [10] are available at Gene Expression Omnibus (GEO) using the accession code GSE214991. The Stereo-seq axolotl brain datasets [7] as well as the reference dataset, i.e. axolotl brain data at 20 days post-injury, can be obtained through their database https://db.cngb.org/stomics/artista/. The 10x Visium mouse brain datasets, including a sagittal anterior section and a sagittal posterior section, are accessible at the 10x Genomics websites https://www.10xgenomics.com/datasets/mouse-brain-serial-section-2-sagittal-anterior-1-standard and https://www.10xgenomics.com/datasets/mouse-brain-serial-section-2-sagittal-posterior-1-standard, respectively. The SPOTS mouse spleen data [12] is accessible at GEO with accession code GSE198353. The 10x Visium RNA & protein human lymph node dataset [29] can be downloaded from the Zenodo database https://zenodo.org/records/10362607. The MISAR-seq mouse brain data [13] is available from the National Genomics Data Center via accession number OEP00003285 (https://www.biosino.org/node/project/detail/OEP00003285). The Spatial ATAC-RNA-seq mouse brain dataset [89] is deposited in the Gene Expression Omnibus with accession code GSE205055. The Stereo-seq mouse olfactory bulb dataset [38] is accessible from https://db.cngb.org/stomics/mosta/download/.

References

1. Baysoy A, Bai ZL, Satija R *et al.* The technological landscape and applications of single-cell multi-omics. *Nat Rev Mol Cell Biol* 2023;**24**:695–713. https://doi.org/10.1038/s41580-023-00615-w
2. Adema K, Schon MA, Nodine MD *et al.* Lost in space: what single-cell RNA sequencing cannot tell you. *Trends Plant Sci* 2024;**29**:1018–28. https://doi.org/10.1016/j.tplants.2024.03.010
3. Armingol E, Officer A, Harismendy O *et al.* Deciphering cell-cell interactions and communication from gene expression. *Nat Rev Genet* 2021;**22**:71–88. https://doi.org/10.1038/s41576-020-00292-x
4. Cang ZX, Zhao YX, Almet AA *et al.* Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nat Methods* 2023;**20**:218–28. https://doi.org/10.1038/s41592-022-01728-4
5. Shao X, Li CY, Yang HH *et al.* Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk. *Nat Commun* 2022;**13**:4429. https://doi.org/10.1038/s41467-022-32111-8
6. Pan Y, Gao Z, Cui X *et al.* collectNET: a web server for integrated inference of cell–cell communication network. *Database* 2024;**2024**:baae098. https://doi.org/10.1093/database/baae098
7. Wei X, Fu S, Li H *et al.* Single-cell stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science* 2022;**377**:eabp9444. https://doi.org/10.1126/science.abp9444
8. Moses L, Pachter L. Museum of spatial transcriptomics. *Nat Methods* 2022;**19**:534–46. https://doi.org/10.1038/s41592-022-01409-2
9. Deng YX, Bartosovic M, Ma S *et al.* Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature* 2022;**609**:375–83. https://doi.org/10.1038/s41586-022-05094-1
10. Llorens-Bobadilla E, Zamboni M, Marklund M *et al.* Solid-phase capture and profiling of open chromatin by spatial ATAC. *Nat Biotechnol* 2023;**41**:1085–8. https://doi.org/10.1038/s41587-022-01603-9
11. Liu Y, DiStasio M, Su GH *et al.* High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq. *Nat Biotechnol* 2023;**41**:1405–9. https://doi.org/10.1038/s41587-023-01676-0
12. Ben-Chetrit N, Niu X, Swett AD *et al.* Integration of whole transcriptome spatial profiling with protein markers. *Nat Biotechnol* 2023;**41**:788–93. https://doi.org/10.1038/s41587-022-01536-3
13. Jiang FQ, Zhou X, Qian YY *et al.* Simultaneous profiling of spatial gene expression and chromatin accessibility during mouse brain development. *Nat Methods* 2023;**20**:1048–57. https://doi.org/10.1038/s41592-023-01884-1
14. Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007;**447**:407–12. https://doi.org/10.1038/nature05915
15. MacNeil LT, Walhout AJM. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome Res* 2011;**21**:645–57. https://doi.org/10.1101/gr.097378.109
16. Stoeckius M, Hafemeister C, Stephenson W *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017;**14**:865–8. https://doi.org/10.1038/nmeth.4380
17. Peterson VM, Zhang KX, Kumar N *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol* 2017;**35**:936–9. https://doi.org/10.1038/nbt.3973
18. Long Y, Ang KS, Li M *et al.* Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat Commun* 2023;**14**:1155. https://doi.org/10.1038/s41467-023-36796-3
19. Zeira R, Land M, Strzalkowski A *et al.* Alignment and integration of spatial transcriptomics data. *Nat Methods* 2022;**19**:567–75. https://doi.org/10.1038/s41592-022-01459-6
20. Korsunsky I, Millard N, Fan J *et al.* Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* 2019;**16**:1289–96. https://doi.org/10.1038/s41592-019-0619-0
21. Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol* 2019;**37**:685–91. https://doi.org/10.1038/s41587-019-0113-3
22. Yuan Z, Zhao F, Lin S *et al.* Benchmarking spatial clustering methods with spatially resolved transcriptomics data. *Nat Methods* 2024;**21**:712–22. https://doi.org/10.1038/s41592-024-02215-8
23. Hu J, Li XJ, Coleman K *et al.* SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat Methods* 2021;**18**:1342–51. https://doi.org/10.1038/s41592-021-01255-8
24. Dong KN, Zhang SH. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat Commun* 2022;**13**:1739. https://doi.org/10.1038/s41467-022-29439-6
25. Li Z, Chen XY, Zhang XG *et al.* Latent feature extraction with a prior-based self-attention framework for spatial transcriptomics. *Genome Res* 2023;**33**:1757–73. https://doi.org/10.1101/gr.277891.123
26. Zhou X, Dong KN, Zhang SH. Integrating spatial transcriptomics data across different conditions, technologies and developmental stages. *Nat Comput Sci* 2023;**3**:894–906. https://doi.org/10.1038/s43588-023-00528-w
27. Wang G, Zhao J, Yan Y *et al.* Construction of a 3D whole organism spatial atlas by joint modelling of multiple slices with deep neural

networks. *Nat Mach Intell* 2023;**5**:1200–13. https://doi.org/10.1038/s42256-023-00734-1
28. Liu Y, Li Z, Chen X *et al.* INSTINCT: multi-sample integration of spatial chromatin accessibility sequencing data via stochastic domain translation. *Nat Commun* 2025;**16**:1247. https://doi.org/10.1038/s41467-025-56535-0
29. Long Y, Ang KS, Sethi R *et al.* Deciphering spatial domains from spatial multi-omics with SpatialGlue. *Nat Methods* 2024;**21**:1658–67. https://doi.org/10.1038/s41592-024-02316-4
30. Tian T, Zhang J, Lin X *et al.* Dependency-aware deep generative models for multitasking analysis of spatial omics data. *Nat Methods* 2024;**21**:1501–13. https://doi.org/10.1038/s41592-024-02257-y
31. Vandereyken K, Sifrim A, Thienpont B *et al.* Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet* 2023;**24**:494–515. https://doi.org/10.1038/s41576-023-00580-2
32. Veličković P, Cucurull G, Casanova A *et al.* Graph Attention Networks. In: Murray I, Ranzato MA, Vinyals O. (eds), *6th International Conference on Learning Representations*; 30 April, pp. 1–12. Canada: BC, 2018.
33. Blundell C, Cornebise J, Kavukcuoglu K *et al.* Weight uncertainty in neural network. In: Bach F, Blei D. (eds), *32nd International Conference on Machine Learning; 6 July*, pp. 1613–22. France: Lille, 2015.
34. Lin Z, Zhang Z, Wang M *et al.* Multi-modal contrastive representation learning for entity alignment. In: Calzolari N, Huang C-R, Kim H *et al.* (eds), *29th International Conference on Computational Linguistics; 12 October*, pp. 2572–84. South Korea: Gyeongju, 2022.
35. Eraslan G, Simon LM, Mircea M *et al.* Single-cell RNA-seq denoising using a deep count autoencoder. *Nat Commun* 2019;**10**:390. https://doi.org/10.1038/s41467-018-07931-2
36. Martens LD, Fischer DS, Yépez VA *et al.* Modeling fragment counts improves single-cell ATAC-seq analysis. *Nat Methods* 2023;**21**:28–31. https://doi.org/10.1038/s41592-023-02112-6
37. Guo TT, Yuan ZY, Pan Y *et al.* SPIRAL: integrating and aligning spatially resolved transcriptomics data across different experiments, conditions, and technologies. *Genome Biol* 2023;**24**:241. https://doi.org/10.1186/s13059-023-03078-6
38. Chen A, Liao S, Cheng MN *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 2022;**185**:1777–92. https://doi.org/10.1016/j.cell.2022.04.003
39. Chen S, Chang Y, Li L *et al.* Spatially resolved transcriptomics reveals genes associated with the vulnerability of middle temporal gyrus in Alzheimer's disease. *Acta Neuropathol Commun* 2022;**10**:188. https://doi.org/10.1186/s40478-022-01494-6
40. Chen T, Kornblith S, Norouzi M *et al.* A simple framework for contrastive learning of visual representations. In: Daumé III H, Singh A. (eds), *37th International Conference on Machine Learning; 12 July*, pp. 1597–607. Vienna: Austria, 2020.
41. Ganin Y, Ustinova E, Ajakan H *et al.* Domain-adversarial training of neural networks. *J Mach Learn Res* 2016;**17**:1–35.
42. Kopp W, Akalin A, Ohler U. Simultaneous dimensionality reduction and integration for single-cell ATAC-seq data using deep learning. *Nat Mach Intell* 2022;**4**:162–8. https://doi.org/10.1038/s42256-022-00443-1
43. Chiang W-L, Liu X, Si S *et al.* Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks. In: Li Y, Rosales R, Terzi E, Karypis G. (eds), *25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; 4 August, pp. 257–66. Association for Computing Machinery, Anchorage, AK, 2019.
44. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* 2012;**72**:4875–82. https://doi.org/10.1158/0008-5472.CAN-12-2217
45. Janiszewska M. The microcosmos of intratumor heterogeneity: the space-time of cancer evolution. *Oncogene* 2020;**39**:2031–9. https://doi.org/10.1038/s41388-019-1127-5
46. Gayoso A, Steier Z, Lopez R *et al.* Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat Methods* 2021;**18**:272–82. https://doi.org/10.1038/s41592-020-01050-x
47. Kim HJ, Lin YX, Geddes TA *et al.* CiteFuse enables multi-modal analysis of CITE-seq data. *Bioinformatics* 2020;**36**:4137–43. https://doi.org/10.1093/bioinformatics/btaa282
48. Hao YH, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021;**184**:3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048
49. Traag VA, Waltman L, Van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019;**9**:5233. https://doi.org/10.1038/s41598-019-41695-z
50. Chen S, Hua K, Cui H *et al.* VPAC: variational projection for accurate clustering of single-cell transcriptomic data. *BMC Bioinformatics* 2019;**20**:139–51. https://doi.org/10.1186/s12859-019-2742-4
51. Jiang R, Li Z, Jia Y *et al.* SINFONIA: scalable identification of spatially variable genes for deciphering spatial domains. *Cells* 2023;**12**:604. https://doi.org/10.3390/cells12040604
52. McInnes L, Healy J, Saul N, Grossberger L. UMAP: Uniform Manifold Approximation and Projection. *J Open Source Softw* 2018;**3**:861. https://arxiv.org/abs/1802.03426
53. Lewis SM, Williams A, Eisenbarth SC. Structure and function of the immune system in the spleen. *Sci Immunol* 2019;**4**:eaau6085. https://doi.org/10.1126/sciimmunol.aau6085
54. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods* 2018;**15**:255–61. https://doi.org/10.1038/nmeth.4612
55. Sugeno A, Piao W, Yamazaki M *et al.* Cortical transcriptome analysis after spinal cord injury reveals the regenerative mechanism of central nervous system in CRMP2 knock-in mice. *Neural Regen Res* 2021;**16**:1258–65. https://doi.org/10.4103/1673-5374.301035
56. Whatley SD, Ducamp S, Gouya L *et al.* C-terminal deletions in the ALAS2 gene lead to gain of function and cause X-linked dominant protoporphyria without anemia or iron overload. *Am J Hum Genet* 2008;**83**:408–14. https://doi.org/10.1016/j.ajhg.2008.08.003
57. Bertocchio J-P, Genetet S, Da Costa L *et al.* Red blood cell AE1/band 3 transports in dominant distal renal tubular acidosis patients. *Kidney Int Rep* 2020;**5**:348–57. https://doi.org/10.1016/j.ekir.2019.12.020
58. Kanno S, Hirano S, Sakamoto T *et al.* Scavenger receptor MARCO contributes to cellular internalization of exosomes by dynamin-dependent endocytosis and macropinocytosis. *Sci Rep* 2020;**10**:21795. https://doi.org/10.1038/s41598-020-78464-2
59. Chang Y-T, Tsai W-C, Lin W-Z *et al.* A novel IGLC2 gene linked with prognosis of triple-negative breast cancer. *Front Oncol* 2022;**11**:759952. https://doi.org/10.3389/fonc.2021.759952
60. Kawasaki K, Ohta Y, Castro CD *et al.* The immunoglobulin J chain is an evolutionarily co-opted chemokine. *Proc Natl Acad Sci* 2024;**121**:e2318995121. https://doi.org/10.1073/pnas.2318995121
61. Argelaguet R, Arnol D, Bredikhin D *et al.* MOFA+ : a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* 2020;**21**:1–17. https://doi.org/10.1186/s13059-020-02015-1
62. Jin SQ, Zhang LH, Nie Q. scAI: an unsupervised approach for the integrative analysis of parallel single-cell transcriptomic and epigenomic profiles. *Genome Biol* 2020;**21**:1–19. https://doi.org/10.1186/s13059-020-1932-8
63. Lin X, Tian T, Wei Z *et al.* Clustering of single-cell multi-omics data with a multimodal deep learning method. *Nat Commun* 2022;**13**:7705. https://doi.org/10.1038/s41467-022-35031-9

64. Li GY, Fu SL, Wang SG *et al.* A deep generative model for multi-view profiling of single-cell RNA-seq and ATAC-seq data. *Genome Biol* 2022;**23**:20. https://doi.org/10.1186/s13059-021-02595-6
65. Haghverdi L, Lun ATL, Morgan MD *et al.* Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* 2018;**36**:421–7. https://doi.org/10.1038/nbt.4091
66. Xiong L, Tian K, Li Y *et al.* Online single-cell data integration through projecting heterogeneous datasets into a common cell-embedding space. *Nat Commun* 2022;**13**:6118. https://doi.org/10.1038/s41467-022-33758-z
67. Zhang K, Zemke NR, Armand EJ *et al.* A fast, scalable and versatile tool for analysis of single-cell omics data. *Nat Methods* 2024;**21**:217–27. https://doi.org/10.1038/s41592-023-02139-9
68. Luecken MD, Büttner M, Chaichoompu K *et al.* Benchmarking atlas-level data integration in single-cell genomics. *Nat Methods* 2022;**19**:41–50. https://doi.org/10.1038/s41592-021-01336-8
69. Heumos L, Schaar AC, Lance C *et al.* Best practices for single-cell analysis across modalities. *Nat Rev Genet* 2023;**24**:550–72. https://doi.org/10.1038/s41576-023-00586-w
70. Zhang H, Rice ME, Alvin JW *et al.* Extensive evaluation of ATAC-seq protocols for native or formaldehyde-fixed nuclei. *BMC Genomics* 2022;**23**:214. https://doi.org/10.1186/s12864-021-08266-x
71. Xiong L, Xu K, Tian K *et al.* SCALE method for single-cell ATAC-seq analysis via latent feature extraction. *Nat Commun* 2019;**10**:4576. https://doi.org/10.1038/s41467-019-12630-7
72. González-Blas CB, Minnoye L, Papasokrati D *et al.* cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat Methods* 2019;**16**:397–400. https://doi.org/10.1038/s41592-019-0367-1
73. Pierson E, Yau C. ZIFA: dimensionality reduction for zero-inflated single-cell gene expression analysis. *Genome Biol* 2015;**16**:1–10. https://doi.org/10.1186/s13059-015-0805-z
74. Shang L, Zhou X. Spatially aware dimension reduction for spatial transcriptomics. *Nat Commun* 2022;**13**:7203. https://doi.org/10.1038/s41467-022-34879-1
75. Sunkin SM, Ng L, Lau C *et al.* Allen brain atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res* 2012;**41**:D996–1008. https://doi.org/10.1093/nar/gks1042
76. Teicher MH, Anderson CM, Polcari A. Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proc Natl Acad Sci* 2012;**109**:E563–72. https://doi.org/10.1073/pnas.1115396109
77. Matsuda K, Budisantoso T, Mitakidis N *et al.* Transsynaptic modulation of kainate receptor functions by C1q-like proteins. *Neuron* 2016;**90**:752–67. https://doi.org/10.1016/j.neuron.2016.04.001
78. Park S-Y, Yoon SN, Kang M-J *et al.* Hippocalcin promotes neuronal differentiation and inhibits astrocytic differentiation in neural stem cells. *Stem Cell Rep* 2017;**8**:95–111. https://doi.org/10.1016/j.stemcr.2016.11.009
79. Coffman KA, Dum RP, Strick PL. Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. *Proc Natl Acad Sci* 2011;**108**:16068–73. https://doi.org/10.1073/pnas.1107904108
80. Hachem S, Laurenson A-S, Hugnot J-P *et al.* Expression of S100B during embryonic development of the mouse cerebellum. *BMC Dev Biol* 2007;**7**:1–15. https://doi.org/10.1186/1471-213X-7-17
81. Hirasawa M, Xu X, Trask RB *et al.* Carbonic anhydrase related protein 8 mutation results in aberrant synaptic morphology and excitatory synaptic function in the cerebellum. *Mol Cell Neurosci* 2007;**35**:161–70. https://doi.org/10.1016/j.mcn.2007.02.013
82. Lassuthova P, Rebelo AP, Ravenscroft G *et al.* Mutations in ATP1A1 cause dominant Charcot-Marie-tooth type 2. *Am J Hum Genet* 2018;**102**:505–14. https://doi.org/10.1016/j.ajhg.2018.01.023
83. Pacey LK, Xuan IC, Guan S *et al.* Delayed myelination in a mouse model of fragile X syndrome. *Hum Mol Genet* 2013;**22**:3920–30. https://doi.org/10.1093/hmg/ddt246
84. Xu Z, Wang W, Yang T *et al.* STOmicsDB: a comprehensive database for spatial transcriptomics data sharing, analysis and visualization. *Nucleic Acids Res* 2024;**52**:D1053–61. https://doi.org/10.1093/nar/gkad933
85. Yuan ZY, Pan WT, Zhao X *et al.* SODB facilitates comprehensive exploration of spatial omics data. *Nat Methods* 2023;**20**:387–99. https://doi.org/10.1038/s41592-023-01773-7
86. Argelaguet R, Cuomo AS, Stegle O *et al.* Computational principles and challenges in single-cell data integration. *Nat Biotechnol* 2021;**39**:1202–15. https://doi.org/10.1038/s41587-021-00895-7
87. Li B, Tang Z, Budhkar A *et al.* SpaIM: single-cell spatial transcriptomics imputation via style transfer. *Nat Commun* 2025;**16**:7861. https://doi.org/10.1038/s41467-025-63185-9
88. Budhkar A, Tang Z, Liu X *et al.* xSiGra: explainable model for single-cell spatial data elucidation. *Brief Bioinform* 2024;**25**:bbae388. https://doi.org/10.1093/bib/bbae388
89. Zhang D, Deng Y, Kukanja P *et al.* Spatial epigenome–transcriptome co-profiling of mammalian tissues. *Nature* 2023;**616**:113–22. https://doi.org/10.1038/s41586-023-05795-1

Briefings in Bioinformatics, 2026, **27**, bbag214
https://doi.org/10.1093/bib/bbag214
Case Study