Mosaic integration of spatial multi-omics with SpaMosaic

Xuhua Yan[1,2], Min Li[1], Kok Siong Ang[2], Lynn van Olst[3], Alex Edwards[3], Thomas Watson[3], Ruiqing Zheng[1], Rong Fan[4], David Gate[3], Jinmiao Chen[2,5,6*]

[1] Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha 410083, China

[2] Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, Matrix Building, Level 7, Singapore 138671, Singapore

[3] The Ken & Ruth Davee Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[4] Department of Biomedical Engineering, Yale University, New Haven, CT, USA

[5] Centre for Computational Biology and Program in Cancer and Stem Cell Biology, Duke-NUS Medical School, 8 College Road, Singapore 169857, Singapore

[6] Immunology Translational Research Program, Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore (NUS), 5 Science Drive 2, Blk MD4, Level 3, Singapore 117545, Singapore

[*]Corresponding author. Email: micchenj@nus.edu.sg

Abstract

With the advent of spatial multi-omics, we can mosaic integrate such datasets with partially overlapping modalities to construct higher dimensional views of the source tissue. SpaMosaic is a spatial multi-omics mosaic integration tool that employs contrastive learning and graph neural networks to construct a modality-agnostic and batch-corrected latent space suited for analyses like spatial domain identification and imputing missing omes. Using simulated and experimentally acquired datasets, we benchmarked SpaMosaic against single-cell multi-omics mosaic integration methods. The experimental spatial omics data encompassed RNA and protein abundance, chromatin accessibility or histone modifications, acquired from brain, embryo, tonsil, and lymph node tissues. SpaMosaic achieved superior performance over existing methods in identifying known spatial domains while reducing noise and batch effects. We also integrated a set of five mouse brain datasets of RNA and different epigenomic modalities, and imputed the missing omes. We found the genes in the imputed omes enriched in the correct tissue specific biological processes, supporting the imputation accuracy.

Keywords: spatial multi-omics, mosaic integration, graph neural network, contrastive learning

Introduction

Spatial omics are powerful techniques for studying cellular heterogeneity in its spatial context within tissues[1]. Current developments in spatial multi-omics are expanding the scope to simultaneously acquire multiple omics from a single tissue section[2]. For example, spatial ATAC-RNA-seq captures chromatin accessibility and RNA expression[2], spatial CUT&Tag-RNA-seq captures histone modifications and RNA expression[2], and SPOTS measures RNA and protein abundance[3]. The latest advancements, spatial ATAC-RNA-Protein-seq and spatial CUT&Tag-RNA-Protein-seq[4], have made it possible to acquire tri-omics data. By capturing different omics profiles of the same cell, these complementary views can be used to construct more comprehensive pictures and better dissect phenotypic niches of samples[5]. However, spatial omics experiments are technically challenging, involving trade-offs between resolution, sensitivity and throughput, and acquiring multiple omics concurrently from the same section exacerbates these challenges. At present, the most commonly available data are from single-omics modality experiments on different and sometimes adjacent tissue sections. Thus, spatial multi-omics with shared but also complementary omes can be computationally integrated to obtain a more coherent, complete, and higher-dimensional picture of tissue biology. Such integration, termed mosaic integration, can help consensus domain identification, enhance signal to noise ratio, impute missing assays in different sections, and dissect cross-linking between omics, which will be crucial for constructing a multi-omics spatial tissue atlas.

In essence, mosaic integration seeks to integrate multiple batches with different multi-modal compositions[6]. However, unlike diagonal integration where there is no shared batches or modalities, mosaic integration leverages a mosaic of shared modalities between batches to serve as bridges to accomplish data integration. Therefore, mosaic integration needs to accomplish both vertical integration of harmonizing modality heterogeneity for the same cells and horizontal integration of reducing technical variations in the same modality across batches[7]. In the context of mosaic integrating spatial omics data, the spatial context is an additional layer of information that should be utilized to accomplish a more coherent and accurate integration.

There are currently several single-cell mosaic integration methods available, such as Cobolt[8], scMoMaT[7], StabMap[9] and MIDAS[10]. As these methods are tailored for single-cell mosaic integration, they do not consider the spatial context and thus may struggle with mosaic spatial multi-omics data. Spatially aware integration methods like SEDR[11], GraphST[12], STAligner[13] and STitch3D[14] are designed for horizontal integration, while SpatialGLUE[5] performs vertical integration. Due to the respective assumptions on data availability for these integrations, none of these methods are usable for spatial mosaic integration. Another recently proposed method, PRESENT[15], is designed for spatial vertical integration and horizontal integration, but it is not able to handle the general case of spatial mosaic integration.

In this work, we present SpaMosaic for spatial mosaic integration. Using a contrastive learning framework[16] with graph neural networks (GNNs), SpaMosaic projects each section into a modality-agnostic latent space in a spatially aware manner while minimizing batch variations. In particular, we adopted a weighted light graph convolution network to encode the similarity between spots. The output latent embedding is usable for imputation of missing data modalities and downstream analyses like spatial domain identification and joint analysis of multiple sections. We first employed simulated datasets to demonstrate SpaMosaic's capabilities at integrating mosaic datasets while reducing technical noise and recovering hidden spatial cluster patterns. We then benchmarked SpaMosaic with experimentally acquired data of both human and mouse tissues with data modalities encompassing RNA, ATAC, histone modifications, and protein. Compared to other mosaic integration methods, SpaMosaic showed superior performance in terms of modality alignment, batch integration and spatial domain identification. Furthermore, we also demonstrated SpaMosaic's ability to integrate a large number of different modalities to reconstruct a multi-omics map of mouse brain with five modalities comprising RNA, ATAC, and three different histone modifications. SpaMosaic was able to capture spatial domains consistent with the reference anatomic annotations and impute the missing modalities of each section.

Results

The SpaMosaic model

SpaMosaic is designed to integrate mosaic spatial omics data acquired from multiple tissue sections to output a joint latent space for downstream analyses (**Fig. 1a**). SpaMosaic takes in the mosaic feature-by-spot (or cell) count matrices of different data modalities from multiple sections and associated spatial coordinates (**Fig. 1b**). The sections are acquired from the same biological tissue type and can be serial sections from the same tissue or sections that form a time series such as embryo development. SpaMosaic is acquisition technology agnostic so the modalities can be acquired with different technologies. To capture the links between modalities, SpaMosaic requires modalities to be directly or indirectly connected through one or multiple sections. A direct connection appears in the same section, while an indirect connection requires multiple intermediary direct connections. Within each section and each modality, SpaMosaic first performs dimension reduction on the highly variable features with SVD to obtain a low-dimensional representation. We then use Harmony[17] to remove the batch effects within each modality. Thereafter, we construct a heterogeneous adjacency graph for each modality. Each adjacency graph records two types of adjacencies, spatial distance-based adjacency and expression-based adjacency. The spatial distance-based adjacency denotes adjacency of spots with their nearest spatial neighbors from the same section, while the expression-based adjacency denotes adjacency between spots and their mutual nearest neighbors (MNNs)[18] in other sections.

In the next step, SpaMosaic builds a modality-agnostic latent space using graph neural networks (GNNs) in a contrastive learning framework. For each modality, a GNN is employed to encode modality-specific information from the modality inputs and spatial adjacency graph into a set of embeddings. To encode the heterogeneous graphs, we explored various graph network architectures, including the graph attention network (GAT)[19,20], heterogeneous graph transformer (HGT)[21] and two variants of light graph convolution network (LGCN). We selected the weighted LGCN (**Supp. Fig. S1a**) for the final model due to its efficiency and robust comprehensive performance (**Supplementary Information**). We then employ contrastive learning to align modality-specific embeddings. The training pairs required are formed with the embeddings in sections with more than one modality, where we consider modality-specific embeddings from the same spot to constitute a positive pair and pairs of embeddings from different spots to constitute negative pairs. In conjunction, the contrastive learning objective is formulated to minimize the distance between positive pairs and maximize the distance between negative pairs. This seeks to align the different modalities while preserving variation between spots. After aligning the modalities, the average of the modality-specific embeddings will be used as the final embedding for sections with more than one modality. For sections with a single modality, we use a feature reconstruction objective to preserve inter-spot variations. SpaMosaic outputs the integrated spot embeddings across sections with imputed feature matrices for the missing omics. The output is usable in downstream analyses such as spatial domain identification and analysis feature correlations across modalities. To impute data for missing modalities, SpaMosaic uses expression information from the nearest neighbors identified in the modality-aligned latent space (details are found in the Methods section and **Supp. Fig. S1b**).

Full recovery of spatial factors from simulated datasets with SpaMosaic

In this first example, we employed simulated data to evaluate the performance of SpaMosaic, an alternate version of SpaMosaic (non-spatial) that does not use spatial information, and five state-of-the-art single-cell mosaic integration methods, namely Cobolt[8], scMoMaT[7], StabMap[9], CLUE[22] and MIDAS[10]. For Cobolt, scMoMaT and StabMap, we followed their publications' recommendations of adding an additional postprocessing step for batch integration, which we maintained for all examples presented. For this postprocessing step, we used Harmony for consistency. To create the simulated data, we followed the procedure of Townes et al.[23], using five sets of different simulation parameters to each generate three replicates of spatial multi-omics datasets with different random seeds (**Supp. Fig. S2a**). For the different datasets, we varied the erosion parameter s^{rna} and dropout probability p^{rna} to modify the RNA expression data distribution and varied the background noise parameter b^{rna} to modify the protein data expression distribution (Methods). Each dataset has three sections where the first section consists of RNA and protein modalities while the second only has RNA expression and the third only has protein expression (**Supp. Fig. S2b**). We kept all modalities at the same dimensions, namely 1,296 spots arranged in a 36×36 grid. The simulated data has a ground truth of four spatial factors and a background (**Fig. 2a**). For all datasets, we reduced the RNA expression of factor 1 and protein expression of factor 4 to simulate a scenario where these factors were missing in their modalities. This aims to test for the ability to recover all factors from integrating the two modalities. We also simulated batch effects between sections 1 and 2 in the RNA modality and sections 1 and 3 in the protein modality by varying the feature expression mean values in sections 2 and 3 (Methods). We then evaluated the methods in three aspects: preservation of spatial factors using ARI and Moran's *I* [24], batch mixing using iLISI[25] and modality alignment using the F1 score computed for label transfer accuracy between the second and the third section. Higher scores indicate better performance for all metrics.

We first examined the *k*-means clustering on the raw data of one replicate from one dataset (dataset id=1, first replicate) as an example (**Fig. 2b**). For section 1, the clustering captured spatial factors 2, 3 and 4 in the RNA modality, factors 1, 2 and 3 in the protein modality, and all four factors from clustering on the concatenated RNA and protein data. For sections 2 (RNA only) and 3 (protein only), the clustering also captured the same respective factors in section 2 (factors 2, 3 and 4) and section 3 (factors 1, 2 and 3). The UMAP[26] plots (**Fig. 2c**) were consistent with factor 1 merged into the cluster of factor 0 (background) in the RNA modality, and factor 4 merged into factor 0 in the protein modality. Batch effects were also visible in the UMAP plots. Next, we examined the clustering results on SpaMosaic's output embeddings (**Fig. 2d**). Unsurprisingly, all four factors were recovered from section 1 along with factors 2, 3 and 4 from section 2, and factors 1, 2 and 3 from section 3, with most of the noise removed. In addition, SpaMosaic enabled the recovery of factor 1 in section 2 and factor 4 in section 3. We also used UMAPs to visualize SpaMosaic's output (**Fig. 2e**). Visualizing with the section labels, we see the integration of all three sections within each factor's cluster (bottom) while the spatial clusters show clear separation (upper left), and this compares well with the ground truth (upper right).

Using the quantitative metrics, we compared SpaMosaic and competing methods on all five sets of simulated data with replicates. For each method, we compared the mclust[27] and *k*-means clustering algorithms and found that all methods performed better with *k*-means except SpaMosaic and SpaMosaic (non-spatial) (**Supp. Fig. S2c**). We then used mclust for SpaMosaic and SpaMosaic (non-spatial) and used *k*-means for the other methods. This was retained for all subsequent experiments. In terms of iLISI for assessing batch integration, Cobolt is the best method with SpaMosaic being a close second (**Fig. 2f**). scMoMaT ranked the third on datasets 2 to 5, followed by SpaMosaic (non-spatial). For the ARI metric, SpaMosaic was clearly the top performer on average (**Fig. 2g**). Among the different sections, MIDAS was a close second on section 1 but SpaMosaic significantly outperformed all other methods on sections 2 and 3. For the Moran's *I*, SpaMosaic was the top performer for all five datasets. MIDAS was second on datasets 1 and 5, while CLUE was second on datasets 2, 3, and 4 (**Fig. 2h**). From the perspective of sections, MIDAS was a close second on section 1 but SpaMosaic was significantly better than the rest on sections 2 and 3 (**Fig. 2i**). Here we believe that SpaMosaic's noise removal capabilities and its incorporation of spatial information achieved spatially smooth factors that contributed to its enhanced performance on section 1, while the recovered missing factors drove SpaMosaic's outperformance on sections 2 and 3. We also visualized the results of other methods with UMAP and spatial plots (**Supp. Fig. S2d, e**). SpaMosaic (non-spatial) achieved poorer separation of the factors than the full SpaMosaic, illustrating the contribution of spatial information on SpaMosaic performance. Of the competing methods, most achieved varying degrees of data integration, except for MIDAS with clearly visible batch differences, which was also reflected in its low iLISI scores. In terms of factor separation, we could also observe factors (e.g. factor 4) mixing with the background (factor 0) for most competing methods. Finally, SpaMosaic also outperformed all other methods at label transfer with the highest median F1-score for all datasets (**Fig. 2j**). We further investigated SpaMosaic's label transfer by comparing the assigned factors to the ground truth with a Sankey plot for the first replicate of dataset 1 (**Fig. 2k**). We found that most of the incorrect assignments are spots of the spatial factors (1 to 4) being assigned to the background (factor 0) or vice versa, which is consistent with the spatial clustering obtained (**Fig. 2c, e**).

SpaMosaic accurately captures structures in embryonic mouse brain (RNA and ATAC integration)

We next benchmarked SpaMosaic and competing methods on a mouse embryonic brain dataset acquired with MISAR-seq[28] that jointly profiles gene expression (RNA) and chromatin accessibility (ATAC). We selected the E13.5, E15.5 and E18.5 sections and removed the ATAC modality of E13.5 and RNA modality of E18.5 to simulate mosaic data availability. The original data has modality-specific clusters, clusters of concatenated modalities and the associated manual annotations (**Fig. 3a**), which we used as the ground truth in quantitative benchmarking. We also added the embryonic mouse brain anatomical references from the Allen Brain Atlas for comparison. Examination of the UMAPs of each modality showed the presence of strong batch effects between the different time points (**Supp. Fig. S3a**). Following the same procedure as in other examples, we applied the methods to integrate the data. Thereafter, the outputs of Cobolt, scMoMaT and StabMap were postprocessed with Harmony for batch effect removal, and SpaMosaic used mclust algorithm while the other methods used *k*-means to cluster the output embeddings.

Among the spatial plots of results obtained, SpaMosaic's spatial clusters had the least noise and best matched the original annotation and the Allen Brain Atlas reference in terms of the brain regions captured (**Fig. 3b**). For the competing methods, the outputs are much noisier for all sections. In the E13.5 section, SpaMosaic's cluster 3 matched Cartilage 2, 5 matched the Diencephalon, 6 and 7 overlapped with the DPallm, 10 matched the midbrain, and 11 matched Muscle in the original annotation. Cluster 2 preserved the RNA-specific variations and matched the original RNA-specific cluster 9. In Cobolt's output, we could see that it matched the DPallm (cluster 6), Diencephalon (clusters 13) and Hind brain (cluster 8) in the original annotation. For MIDAS, we observed that cluster 6 and 7 overlapped with the DPallm region. For the remaining methods, we could not unambiguously identify clusters that matched the annotation. In the E15.5 section, SpaMosaic also matched many annotated structures including Cartilage 2 (cluster 3), Cartilage 3 (cluster 4), DPallm (cluster 6), DPallv (cluster 7), Mid brain (cluster 10) and Thalamus (cluster 15). The other methods captured fewer regions, namely DPallm and DPallv (except StabMap that merged them together), and Thalamus by CLUE. Again, the clusters are noisier than SpaMosaic's. In the E18.5 section, all methods except StabMap captured the DPallm and DPallv, while the Subpallium was captured by SpaMosaic, Cobolt, scMoMaT and CLUE. SpaMosaic also captured the Diencephalon and Hind brain (cluster 8), and parts of the Mesenchyme region (cluster 9).

From SpaMosaic's results of the E13.5 section, we were intrigued by cluster 6 and cluster 7 that overlapped with the DPallm and DPallv regions of the original annotation. There, DPallm appeared to the top and left of the hole in the section and DPallv appear below the hole. In SpaMosaics' output, cluster 6 mostly overlapped with DPallm but cluster 7 surrounded the hole. We then made further comparisons with the Allen Brain Atlas which showed greater concordance with SpaMosaic's clusters (**Fig. 3c-e)**. To further verify SpaMosaic's clusters, we visualized the markers *Emx1*, *Emx2*, *Id4*, *Nr2f1* and *Pax6* that regulate mouse neocortex development[29–31] and are highly expressed in the DPallv or pallium region. Here we find them highly expressed in cluster 7, supported by both the RNA count and the gene activity scores derived from the ATAC modality (**Fig. 3c**). Further visualization of their expression in the spatial plot and comparison with ISH staining samples from the Allen Brain Atlas confirmed cluster 7's accuracy in capturing the DPallv (**Fig. 3d-e**).

We next quantitatively evaluated SpaMosaic's performance in three aspects: spatial domain identification (ARI and Moran's *I*), batch mixing (iLISI) and modality alignment (FOSCTTM[32] and MS[22,33]). Higher scores indicate better performance for all metrics except for FOSCTTM which is reversed. In terms of spatial domain identification, SpaMosaic outperformed the other methods for all sections (**Fig. 3f, g**), presenting more accurate (ARI) and spatially smoother (Moran's *I*) clusters. For batch mixing, MIDAS and Cobolt closely ranked first and second, respectively, followed by scMoMaT and SpaMosaic (**Fig. 3h**). We investigated further by examining the UMAPs colored by section, clusters and the original annotation (**Supp. Fig. S3b**). MIDAS and Cobolt indeed mixed the three sections well, but the original annotation also showed substantial mixing, which explained their lower ARI and Moran's *I* scores. We also assessed modality alignment with FOSCTTM and MS. To calculate the FOSCTTM and MS metrics, we created a new dataset using the E15.5 and E18.5 sections and split the RNA and ATAC modalities of the E18.5 section to create two new sections (**Supp. Fig. S3c, d**). The ground-truth pairing between the RNA and ATAC modalities of the E18.5 section was used to evaluate modality alignment. SpaMosaic ranked top for both metrics (**Fig. 3i**, lower is better for FOSCTTM and higher is better for MS), highlighting its capabilities in cross-modality integration.

Identifying matching structures across mouse brain sections with SpaMosaic while removing batch effects (RNA and histone modification integration)

We also benchmarked SpaMosaic and the other methods on postnatal mouse brain transcriptome and histone modification profiles compiled from different studies (**Fig. 4a**). The compiled data has three sections. The first section consists of the transcriptome and histone modification (H3K4me3) profiles of a P22 mouse brain section in a capture area with 100×100 pixels, each 20 µm in diameter[2]. In the second section, the original experiment acquired the spatial transcriptome and histone modification (H3K27ac) of a P22 mouse brain section with a capture array of 100×100 pixels, each 20 µm in diameter[2], of which we only retained the spatial transcriptome. For the third section, we selected the spatial histone modification (H3K4me3) profile of a P21 mouse brain section, acquired on a 50×50 array with 20 µm pixels[1]. These three sections were obtained from mouse brains of different ages and varying capture area sizes, which posed additional challenges for data integration. With this design of the dataset, there is a significant batch effect between section 3 and the other two. This is observable in the UMAP plots where there is little batch effect visible between sections 1 and 2 in the RNA modality, but significant batch effects are present between the histone modification profiles of sections 1 and 3 (**Supp. Fig. S4a**). Like the previous example, we added the postprocessing step of Harmony to Cobolt, scMoMaT and StabMap's workflows to handle the batch effects.

Among the methods, SpaMosaic's output again had the least noise, which we attribute to the use of spatial information (**Fig. 4b**). SpaMosaic's output was also able to capture more brain structures than competing methods. Using the Allen Brain Atlas as reference (**Fig. 4c**), we were able to annotate the clusters (**Fig. 4b**). In sections 1 and 2, we identified the cortex (clusters 1, 4 and 5), lateral ventricle (VL)/cortex (cluster 2), nucleus accumbens (ACB, cluster 3), genu of corpus callosum (CCG) and anterior commissure of olfactory limb (ACO) in cluster 6, caudoputamen (CP, cluster 7), and cluster 8 capturing the lateral septal complex (LS). In section 3, we could identify the VL (cluster 2), CCG (cluster 6), cortex (clusters 1 and 5), and CP (clusters 3 and 7). In contrast, the competing methods produced noisier clusters and captured the brain structures to varying degrees. Cobolt and CLUE differentiated more cortex layers (clusters 1, 4 and 5) than SpaMosaic in sections 1 and 2 but did not separate the ACB and CP. scMoMaT and MIDAS also did not separate the ACB and CP, while also failing to clearly capture the cortex. For section 3, we could not unambiguously match the competing methods' output clusters to the reference.

In addition to comparing clusters to the Allen Brain Atlas for annotation, we also plotted the RNA expression of marker genes of six brain regions for sections 1 and 2 (**Fig. 4d**). We observed that these markers were highly expressed in the corresponding SpaMosaic's clusters. For instance, the striatum markers *Pde10a*[2], *Adcy5*[2] and *Foxp1*[34], were highly expressed in cluster 7 (CP), which is part of the striatum. We also observed the markers of CCG, *Sox10*[2], *Mbp*[2] and *Tspan2*[2], to be highly expressed in cluster 6 that contains the CCG structure. For section 3 that lacked the RNA modality, we visualized the gene activity scores derived from the histone modification modality using spatial plots (**Fig. 4e**). The gene activity scores of known marker genes showed high levels of noise, but we could observe spatial correlation of genes to their respective regions like *Satb2*[2] to the cortex and *Pde10a* to the CP/striatum. To further verify the identified brain structures, we integrated section 3 with a published scRNA-seq mouse nervous system dataset[35]. We transferred the annotated cell types in the scRNA-seq dataset to section 3 using Seurat[36] and plotted four cell types in the spatial plots (**Fig. 4f**). We could see that the mature oligodendrocytes (MOL1) were enriched in the identified CCG region and the medium spiny neurons (MSN2) were located only within the CP/striatum, which is consistent with existing studies[1,35]. We also observed that the TEGLU8 excitatory neurons were enriched in cluster 5 (cortex) that consists of the upper cortical layer, and the TEGLU3 excitatory neurons were enriched in cluster 1 (cortex) that encompassed the deep cortical layer[35].

We next quantitatively compared SpaMosaic's performance with the other methods in two aspects: spatial domain identification (Moran's *I* metric) and batch integration (iLISI). We could not evaluate modality alignment metrics due to the lack of another section with paired RNA and H3K4me3 histone modification measurements. In terms of spatial domain identification, SpaMosaic significantly outperformed the competing methods for all three sections (**Fig. 4g**). For batch mixing, Cobolt ranked the first while SpaMosaic ranked the second and tied with scMoMaT (**Fig. 4h**). Notably, Cobolt and scMoMaT were employed with Harmony integration as the postprocessing step, which may be a contributing factor to the good data integration performance. We further examined the UMAP visualizations colored by the clustering labels and section labels (**Supp. Fig. S4b**). For the cluster labelled plots, the consistent grouping and clear separation of the SpaMosaic, Cobolt and CLUE's clusters corroborated with their Moran's *I* scores, especially for sections 1 and 2. In the visualization with section labels, the even mixing achieved by Cobolt matched its iLISI score, while SpaMosaic and CLUE achieved better mixing than the remaining methods but with small local clusters of batch specific spots still visible.

To ensure the robustness of our results, we tested more scenarios with three more postnatal mouse brain datasets and one mouse embryo dataset with different modality combinations[1,2,37]. The modality combinations tested were RNA and ATAC, RNA and H3K27me3 histone modification, and RNA and H3K27ac histone modification. Similar to the previous results, SpaMosaic achieved the highest Moran's *I* scores and second at iLISI for most datasets. For datasets on which we could compute the modality alignment scores (FOSCTTM and MS), SpaMosaic was the top performer (**Supp. Fig. S5-S12, Supplementary Information**). Overall, SpaMosaic's performance was consistently excellent across the different combinations of mouse brain and mouse embryo data. This increases confidence in the performance being tissue and modality independent.

SpaMosaic demarcates different immune cell enriched regions in human lymphoid organs (RNA and protein integration)

In this example, we benchmarked SpaMosaic and other methods on an in-house generated human tonsil dataset. The human tonsils are lymphoid organs which play a key role in human immunity. Tonsils are irregularly structured with B cell follicles embedded in T cell enriched zones that are further surrounded by epithelial and connective tissues. Within the follicles, transient germinal centers form to facilitate B cell activation, proliferation and somatic hypermutation during an immune response. The dataset consists of three sections generated using the 10x Genomics Visium RNA and protein co-profiling technology. The first and second sections are two sequential sections, and transcriptome and protein (ADT) profiles were acquired for all three sections. To create a mosaic dataset, we retained only the RNA and protein profiles for the second and third sections, respectively (**Fig. 5a**). We then examined the UMAP plots for batch effects. We did not find significant batch effects in the RNA modality, namely sections 1 and 2 which were sequential, but strong batch effects were present in the protein modality for sections 1 and 3 which were non-sequential sections (**Supp. Fig. S13a**). We also employed expert annotation of the accompanying H&E images to assist in annotating the subsequent analyses (**Supp. Fig. S13b**). This human tonsil dataset offers a different challenge from the previous examples with intricate cell type centric structures and different data modalities, namely RNA and protein.

Among the methods, we again observed the least noise in SpaMosaic's clusters (**Fig. 5b**), demonstrating that SpaMosaic can effectively utilize spatial information regardless of modality compositions. Even with this spatial smoothing, SpaMosaic was still able to capture the follicle mantle zones (cluster 1) and germinal centers (cluster 2) that are comparatively small structures within the tissue slide. Other SpaMosaic clusters captured the tonsillar parenchyma (cluster 3), connective tissue (cluster 6) and other epithelial tissues. For the competing methods, Cobolt's output was the closest to SpaMosaic, but it was less smooth than SpaMosaic's. CLUE's output was also similar to SpaMosaic, but with differences such as the different segregation of the connective & epithelial tissues. For the other methods, only MIDAS was unable to capture the germinal centers, while scMoMaT and StabMap did not correctly separate the connective tissues from other tissues.

To annotate the clusters, we plotted the relevant cell type marker expression in the RNA and protein (**Fig. 5c**) modalities. With the B cell markers *CD19* and *MS4A1*[38,39], we identified B cell enriched regions, namely clusters 1 (follicles) and 2 (germinal centers). We also found RNA expression of *IGHD* (naïve B cells)[38] and BCL2 protein (naive and memory B cells)[40] in cluster 1. For cluster 2, we found high RNA expression of *BCL2A1* and *LMO2* that mark germinal center B cells[41], expression of PDCD1 protein that regulates B cell survival in germinal centers[41], as well as both RNA and protein expression of *CR2* which is a marker for mature B cells and follicular dendritic cells[38]. To identify the T cell enriched zones, we plotted the canonical T cell markers of *CD3E*, *CD4* and *CD8A*. Here we found their RNA and protein to be highly expressed in cluster 3. For annotating epithelial tissues, we examined the expression of epithelial markers, namely KRT5 which was highly expressed in clusters 4 and 5, and EPCAM in cluster 7. We also note the RNA expression of *CD9* in clusters 4, 5, and 7; it identifies B cells and plasma cell subsets within tonsillar tissues, including the crypts. Finally, we found the expression of *ACTA2* that marks myofibroblasts and VIM that marks fibroblasts in connective tissues (cluster 6).

We next quantitatively evaluated SpaMosaic's performance in three aspects: spatial domain identification (ARI and Moran's *I*), batch mixing (iLISI) and modality alignment (FOSCTTM and MS). In terms of ARI, SpaMosaic, CLUE and scMoMaT achieved the highest overall (**Fig. 5d**). For Moran's *I* metric, SpaMosaic was better than other methods for sections 1 and 2 and tied with scMoMaT for section 3 (**Fig. 5e**). For batch mixing assessment with iLISI, SpaMosaic was the leading method (**Fig. 5f**). We also examined the UMAPs of the method output colored by the section labels, cluster labels and manual annotation (**Supp. Fig. S13c**). SpaMosaic and Cobolt showed high levels of section mixing, which was also reflected in their high iLISI scores (first and second, respectively). SpaMosaic's output also maintained good separation of the manually annotated clusters, which was in concordance with the ARI and Moran's *I* metrics. To compute the FOSCTTM and MS metrics, we created a new mosaic dataset using the first and second sections, where we split the RNA and protein profiles of the second section to obtain two new sections (**Supp. Fig. S13d**). For both metrics, SpaMosaic was again the best performing method (**Fig. 5g**).

We also applied the methods to a dataset of another organ from the human immune system, the lymph node. The same 10x Genomics Visium platform was again used to acquire the dataset with RNA and protein modalities. We selected the available modalities to be RNA and protein in the first section, RNA only in the second section and protein only in the third. The full results are presented in the **Supplementary Information** and **Supp. Fig. S14** and **S15**. In brief, we first successfully annotated SpaMosaic's clusters with known markers of tissues found within the lymph node. We also compared the results of all methods with a manual annotation derived from the H&E images. SpaMosaic was successful at demarcating many of the lymph node's tissue structures including the hilium that the other methods could not separate from the nearby adipose tissue. Using the metrics, SpaMosaic was competitive with other methods in terms of spatial domain identification (ARI and Moran's *I*) and batch mixing (iLISI), while being superior at modality alignment (FOSCTTM and MS).

SpaMosaic performs five-modal mosaic integration and imputation of postnatal mouse brain data

In this final example, we increased the number of sections to four and integrated sections with ATAC and different histone modification measurements as the second modality in addition to the RNA modality bridging all samples. We constructed this dataset using the P22 mouse brain sections of Zhang et al.[2]. All sections have a capture area of 100×100 pixels, each 20 μm in diameter (**Fig. 6a**). In addition to the transcriptome, the first three sections measured the H3K27me3, H3K4me3 and H3K27ac histone modifications, respectively, and the fourth section captured ATAC-seq profiles. We evaluated SpaMosaic and the other methods except for MIDAS as it cannot simultaneously integrate ATAC with other epigenome modalities. In our pre-integration analysis, we observed batch effects present within the RNA profiles (**Supp. Fig. S16a**).

We first examined the clusters obtained with the different methods (**Fig. 6b**). Like the previous examples, SpaMosaic's outputs had the least noise. It also clearly captured the different brain regions that we could compare against the Allen Brain Atlas (**Fig. 6c**). For example, SpaMosaic captured the diagonal band nucleus (NDB, cluster 1), CCG and ACO (cluster 2), CP (cluster 3), LS (cluster 7), ACB (cluster 9) and VL (cluster 12). Moreover, SpaMosaic's clusters captured more cortical layers (clusters 4, 5, 8, 10 and 11) than previously achieved with the third example (Figure 4) which only contained the RNA and H3K4me3 modalities. This improvement can be attributed to the inclusion of three additional epigenomic modalities: ATAC, H3K27me3 and H3K27ac. We also plotted the RNA expression of relevant markers to verify our annotation (**Fig. 6d**). We again found the expression of markers in their expected clusters, such as *Mbp* and *Tspan2* in the CCG, *Stard5* n the ACB, and *Trpc4* and *Dgkg* in the LS. For the cortex layers, we found the expression of markers like *Fezf2*[2] in the deep cortical layers (clusters 4 and 8) and *Cux2*[2] in the upper layers (clusters 5 and 10).

Of the competing methods, Cobolt and CLUE captured many of the brain regions like the CCG and ACO, CP, LS, ACB and VL, but not the NDB. Additionally, Cobolt and CLUE's clusters are much noisier than SpaMosaic's. This noise strongly affected the cortex layer clusters captured by Cobolt and CLUE, especially in the case of CLUE's cortex layers in section 4. For scMoMaT's output, the clusters were even noisier, which makes most of its clusters unidentifiable. We then employed the metrics for quantitative assessment. In terms of spatial domain identification (Moran's *I*), SpaMosaic's performance was superior for all sections (**Fig. 6e**). For batch mixing assessed with iLISI, SpaMosaic's performance was third behind Cobolt and scMoMaT (**Fig. 6f**), which is similar to our previous mouse brain example. This is unsurprising as the tissue type and assay technologies employed are the same, albeit capturing different histone modifications. We also visually inspected the UMAPs (**Supp. Fig. S16b**), where we see that batch mixing broadly reflects the metrics.

In addition to data integration and learning modality agnostic latent embeddings in a spatially aware manner, SpaMosaic can also impute the missing data modalities. We first established SpaMosaic's imputation performance by benchmarking it together with MIDAS, BABEL[42], TotalVI[43] and MultiVI[44] on four datasets, and measured their performance using the Pearson correlation coefficient, correlation matrix distance[45] and area under the receiver operating characteristic. We found that SpaMosaic outperformed other methods in preserving spot-spot correlations and effectively maintained feature-feature correlations. Full details on the employed datasets, metrics, and results can be found in the **Supplementary Information**. Here we use the five-modal dataset of mouse brain sections to further demonstrate the imputation capabilities of SpaMosaic. In particular, we used SpaMosaic to impute the gene activity scores (GAS) or chromatin silencing scores (CSS)[46] of the missing epigenomic profiles. We then compared the correlation between RNA expression and measured or imputed GAS or CSS of each section in a cluster-specific manner (**Supp. Fig. S17**). Taking the CCG as an example, the measured H3K27me3 CSS had anticorrelation with the RNA modality on section 1, while the imputed H3K27me3 CSS also showed similar anticorrelation with the RNA modality on the other sections (**Supp. Fig. S17a**). For section 2, the measured H3K4me3 GAS were positively correlated with the RNA modality and the imputed H3K4me3 GAS of the other sections showed similar correlations. We also examined the correlations of other measured and imputed epigenomic profiles with the RNA in other clusters like the CP (cluster 3), cortex (cluster 10) and VL (cluster 12) (**Supp. Fig. S17a-S17d**), and found them to be consistent. This shows that SpaMosaic's latent representation captured the relationships between modalities and preserved them in the imputation.

To further investigate the imputation results of the CCG cluster, we selected section 4 to compare the correlations between the measured RNA with GAS computed from the measured ATAC modality and the imputed histone modification modalities (**Fig. 6g**). We observed that the imputed histone modifications' (H3K4me3 and H3K27ac) GAS of the fourth section had more genes correlated with the RNA modality compared to the measured ATAC GAS (**Fig. 6g**). We then compared the genes and found that the correlated genes (in quadrant I and III) for the measured ATAC assay were also found in the imputed histone modification modalities (**Fig. 6h** and **Supp. Fig. S17e**). This showed that the imputed GAS preserved the measured gene signals. We then investigated the other genes in imputed histone modification modalities that were not found in the measured ATAC assay. We first selected the quadrant I correlated genes unique to the imputed H3K27ac-RNA comparison for GO enrichment analysis with clusterProfiler[47], and found that these genes were associated with myelination, ensheathment of neurons and axon ensheathment processes (**Supp. Fig. 17f**), which is accordance with previous findings[2]. We also performed similar analysis with the quadrant I correlated genes unique to the imputed H3K4me4 RNA comparison and obtained similar enriched processes (**Supp. Fig. S17g**). Furthermore, the imputed histone modification GAS of CCG marker genes showed enhanced signals compared to the measured ATAC modality, enabling clearer segregation of the CCG region (**Supp. Fig. S18**). We believe this is due to the information aggregation from other sections, as well as the incorporation of spatial information in the imputation process, which fuses gene signals within each region and enhances them. In addition to the CCG cluster, we also examined the correlations for other brain regions in section 4, namely the CP (cluster 3), cortex (cluster 10) and VL (cluster 12), and we likewise observed similar signal enhancement (**Supp. Fig. S17a-17d** and **S17h**).

In addition to the correlations between the transcriptomic and epigenomic modalities, we also investigated the relationships among different epigenomic modalities (e.g., imputed histone modifications versus measured ATAC). We again considered the CCG cluster in section 4, focusing on the correlations between the imputed histone modification assays and the measured ATAC GAS. Specifically, H3K27me3 CSS was anticorrelated with ATAC GAS while H3K27ac GAS and H3K4me3 GAS were positively correlated with ATAC GAS (**Fig. 6i**). We then selected the anticorrelated genes in quadrant IV (H3K27me3) and correlated genes in quadrant I (H3K4me3 and H3K27ac) and performed GO enrichment for each set. The enriched processes included oligodendrocyte differentiation, oligodendrocyte development and myelination, which are shared by all three sets (**Fig. 6j**). This analysis was also extended to the CP (cluster 3) and cortex (cluster 10) which showed similar correlation patterns, and we again observed significant overlap in the GO terms obtained between the different sets of genes (**Supp. Fig. S19a-d**). These results give confidence on the imputation's accuracy.

In our investigations, we also found that combining both imputed and measured epigenomic modalities in each section markedly improved gene expression prediction accuracy in terms of decreasing mean squared error and increasing the Pearson correlation coefficient (**Supp. Fig. S20a**). This suggests that the imputed modalities can provide useful complementary information.

Discussion

Spatial omics technologies are powerful techniques for interrogating tissue sections. By measuring different omes, we can obtain complementary views of the tissue. Coupled with the spatial information, such data can enable greater insights into the underlying biology. However, technological limitations and high costs pose challenges to acquiring multiple high quality omics profiles from a single section. Therefore, an alternative approach is to integrate data acquired of different omes from different sections. In this work, we presented SpaMosaic for such spatial mosaic integration. It leverages a contrastive learning framework to construct modality-agnostic latent spaces and employs graph neural networks to encode spatial proximity and expression similarity into the embeddings, achieving spatial-aware integration. The contrastive learning framework bridges the information gap between modalities, which has been validated as a simple and effective strategy for modality alignment and semantic preservation in various domains[48–50]. Moreover, it is also highly flexible and allows easy scaling to include more data modalities. As shown in our last example, SpaMosaic is able to perform five-modal mosaic integration. To correct for technical variations between sections, SpaMosaic does not use complex alignment strategies such as adversarial learning[51] or discrepancy-based learning[52]. Instead, SpaMosaic employs graph neural networks and weighted adjacency matrices to connect spots between sections.

SpaMosaic's latent embedding output enables accurate and robust consensus spatial clustering, outperforming other methods in our benchmarking with different tissue types and different technology platforms. By integrating information from different sections and different omics modalities, SpaMosaic reduces noise in sections to enhance spatial clustering results. SpaMosaic is not restricted to integrating data from adjacent sections; it can integrate data from different time points, different studies, or even different technology platforms. Alongside cross-modality integration, it simultaneously achieves cross-batch integration to remove technical effects while preserving biological variations. This enables the identification of consensus cell types and tissue structures across samples, as well as the capture of subtle differences between sample groups. In addition to data integration, SpaMosaic can impute the missing molecular layers within a mosaic dataset. To accomplish imputation, SpaMosaic employs a *k*NN-based approach to estimate the missing modalities, avoiding the direct use of predictions from neural networks and thereby enhancing its interpretability.

Due to the high cost of simultaneously profiling multiple omics from the same tissue section, most existing data have been generated with different single-omics modalities from separate sections. In this context, diagonal integration is necessary, and image registration is often required to align the sections for most currently available integration algorithms. However, aligning sections can be challenging and even impossible in cases where the sections are highly distinct. In such instances, SpaMosaic offers alignment-free mosaic integration by leveraging the limited available spatial multi-omics data as bridging samples. This feature makes it widely applicable and particularly crucial for constructing spatial multi-omics atlases that requires the integration of diverse datasets with varying modality compositions. Furthermore, SpaMosaic's cross-modality integration and imputation also enable the dissection of relationships between omes that can't be co-profiled, revealing new connections between molecular layers that would otherwise remain undetectable.

Future developments of SpaMosaic include enhancing its scaling with large datasets and avoiding model retraining with new data. By enhancing SpaMosaic' scaling, it can handle larger datasets with more tissue sections, which is critical for constructing spatial multi-omics atlases. As graph neural networks are more computationally expensive than traditional multiple layer perceptron (MLP) networks, we propose to adopt techniques such as graph sampling and caching of intermediate outputs to improve performance[53]. We also plan to investigate continual learning[54] and transfer learning[55] approaches to enable SpaMosaic to analyze new sections without the need to retrain on the entire dataset. Another benefit is the offline analysis of private datasets with a relevant pretrained SpaMosaic model, thus addressing potential data privacy issues. In the future, we will also incorporate accompanying image data such as staining images as an additional modality to enhance the accuracy of cell type and tissue domain identification.

Methods

Data Preprocessing

For spot (cell) $i \in \{1, \ldots, N_b\}$ from section id $b \in \{1, \ldots, B\}$, let $c_i^m \in R^{d^m}$ of size d^m be the expression count vector of spot i from modality $m \in U = \{m_1, \ldots, m_M\}$ where U is the set of all RNA, epigenomic, and protein modalities. For each modality, we first normalize the count matrix using log-normalization for RNA, TF-IDF transformation for epigenomic modality, and CLR normalization for protein. Thereafter, we select the top 5,000 highly variable features for RNA and top 100,000 for epigenome, followed by SVD decomposition into low-dimensional representations. To correct batch effects within a modality, we use Harmony to integrate these low-dimensional representations. For spot i with modality m, we obtain the batch-corrected representation x_i^m, which is then used as input for model training.

Construction of the spot-spot graph

We first create a spot-spot graph for each modality, where each graph contains two types of edges: spatial-based edges and expression-based edges. Spatial-based edges connect spots within the same section while expression-based edges connect spots across different sections. For spatial-based edges, we first compute the Euclidean distance between spots of the same section based on their spatial coordinates. We then identify each spot's k nearest neighbors (*k*NN) using the Annoy python package (https://github.com/spotify/annoy) for efficient (*k*NN) searching. The neighbors with the shortest distance are considered spatial neighbors. Finally, we construct an undirected spatial neighbor graph of all spots in modality m (all modalities share the same spatial neighbor graph) in the form of adjacency matrix $A^{intra-m}$, where $A_{ij}^{intra-m} = 1$ if spots i and j are spatial neighbors.

For the expression-based edges, we use the batch-corrected low-dimensional representation to find the mutual nearest neighbors (MNNs)[18] between sections. Specifically, for each spot i in section b_1, we find its k nearest neighbors in another section b_2, and for each spot j in b_2, we find its k nearest neighbors in b_1. If spots i and j are in each other's nearest neighbors set, then i and j form a pair of MNNs. We then construct an undirected neighbor graph for each modality m denoted by adjacency matrix $A^{inter-m}$, where $A_{ij}^{inter-m} = 1$ if i and j belong to a pair of MNNs. For each modality m, we now combine the $A^{intra-m}$ and $A^{inter-m}$ into one graph:

$$A^m = A^{intra-m} + w_g A^{inter-m}$$

where w_g is a positive constant, denoting the weight for $A^{inter-m}$.

Graph embedding

The spot-spot graph encodes information about the spatial proximity and expression similarity between spots which we integrate with the expression profiles into spot embeddings via graph neural networks. However, our spot-spot graph contains two types of edges, each with different semantics. This differs from traditional homogeneous graphs and thus we explored various approaches to address this situation, namely the graph attention network (GAT) [19,20], heterogeneous graph transformer (HGT)[21], and two variants of light graph convolution network (LGCN)[56,57] (**Supplementary Information**). In our final model, we adopted the weighted LGCN, which is defined as:

$$\tilde{X}^m = LGCN(A^m, X^m) = Concat(X^m, \hat{A}^m X^m, {\hat{A}^m}^2 X^m, \ldots, {\hat{A}^m}^l X^m)$$

where m denotes the modality and l denotes the number of layers. $\hat{A}^m = {\tilde{D}^m}^{-\frac{1}{2}}(A^m + I){\tilde{D}^m}^{-\frac{1}{2}}$, $\tilde{D}^m$ is the diagonal degree matrix of $A^m + I$. The embeddings are then transformed by a multiple layer perceptron (MLP):

$$Z^m = f_{mlp}^m(\tilde{X}^m)$$

where m denotes the modality. Finally, the graph embedding function is defined as g:

$$Z^m = g^m(A^m, X^m) = f_{mlp}^m(LGCN^m(A^m, X^m))$$

By setting w_g, we make A^m a weighted graph to control the importance between spatial proximity and expression similarity. From our testing, we set $w_g = 0.8$, giving more weight to spatial proximity than expression similarity, as we consider the former to be more reliable due to measurement noise present in the expression profiles.

Contrastive learning

Within SpaMosaic, the goal of the model is to learn a set of modality-agnostic low-dimensional embeddings, thereby aligning multiple sections with different modalities. SpaMosaic employes contrastive learning to leverage the input data's multi-modality nature to achieve this goal. In the contrastive learning framework, two views of one sample constitute a positive pair and views from different samples constitute negative pairs[16]. By pushing positive pairs together while separating negative pairs in the latent space, contrastive learning achieves alignment between the different views and preserve variation between samples. For SpaMosaic, two modalities of one spot constitute a positive pair and modalities from different spots constitute negative pairs. Thus, for two modalities m_1 and m_2, there is a sampled mini-batch of spots $\{x_1^{m_1}, \ldots, x_n^{m_1}\}$ and $\{x_1^{m_2}, \ldots, x_n^{m_2}\}$ that form positive pairs and the associated contrastive learning objective is:

$$L_{con}^{m_1,m_2} = -\frac{1}{2n}\sum_{i=1}^{n} \log\frac{\exp\left(s_{i,i}^{1,2}\right)}{\sum_{l=1}^{n}\exp\left(s_{i,l}^{1,2}\right) + \sum_{l\neq i}^{n}\exp\left(s_{i,l}^{1,1}\right)} + \log\frac{\exp\left(s_{i,i}^{2,1}\right)}{\sum_{l=1}^{n}\exp\left(s_{i,l}^{2,1}\right) + \sum_{l\neq 1}^{n}\exp\left(s_{i,l}^{2,2}\right)}$$

where

$$s_{i,l}^{1,2} = \frac{1}{\tau}\cdot\frac{< z_i^{m_1}, z_l^{m_2} >}{\left\|z_i^{m_1}\right\|\cdot\left\|z_l^{m_2}\right\|} = \frac{1}{\tau}\cdot\frac{< g^{m_1}(x_i^{m_1}), g^{m_2}(x_l^{m_2}) >}{\left\|g^{m_1}(x_i^{m_1})\right\|\cdot\left\|g^{m_2}(x_l^{m_2})\right\|},$$

g^{m_1}, g^{m_2} denote the modality-specific graph encoder network, $z_i^{m_1}, z_l^{m_2}$ denote the modality-specific embeddings, $<,>$ denotes inner product of two vectors and τ is a positive constant. The numerator in the contrastive learning objective aims to bring closer the embeddings of each spot's two modalities, while the first term in the denominator aims to separate the embeddings of two modalities from different spots and the second term aims to separate the embeddings within the same modality but from different spots.

Combining embeddings from the measured modalities, the unique representation for spot l in section b is defined as:

$$z_l = \sum_{m \in T_b} w_m \cdot z_l^m$$

where T_b denotes the set of modalities that section b is measured with, $w_m \in (0, 1)$ denotes the weight for modality m and $\sum_{m \in T_b} w_m = 1$.

Feature reconstruction

Based on our formulation, contrastive learning can only be applied to those spots with multimodal profiles. For those spots without multimodal profiles, we add a feature reconstruction objective to include them in the model training. We employ a modality-specific MLP decoder, r^m, to reconstruct the inputs x_i^m based on the embeddings z_i^m. The reconstruction loss is defined as:

$$L_{rec}^{m} = \frac{1}{n} \sum_{i=1}^{n} \| r^m(z_i^m) - x_i^m \|_2$$

Thus, final loss function for the whole model is defined as:

$$L = \sum_{m_i, m_j \in U} L_{con}^{m_i, m_j} + \sum_{m \in U} L_{rec}^{m}$$

where $U = \{m_1, \dots, m_M\}$.

Implementation details

For each modality, the graph encoder network comprises two weighted-LGCN layers, followed by a fully connected layer with a dimension of 128, with leaky ReLU activation (slope=0.2)[58], batch normalization[59], dropout ($p = 0.2$)[60], and a final fully connected layer with a dimension of 32. The decoder comprises one fully connected layer with linear activation. We implemented the graph neural networks using the PyTorch geometric library[61]. To train the model, we used the Adam optimizer[62] with a learning rate of 0.01. The weight decay was set to 5e-4, the number of training epochs was set to 100 and the temperature parameter τ was set to 0.01. The parameter used to calculate the spot-spot graph was set to 10 by default and used for all datasets presented except for the human lymphoid dataset where we adjusted it to 2 to reduce over-smoothing.

Imputation

Let $Z^{m_j} = \{z_1^{m_j}, \dots, z_N^{m_j}\}$ denote the graph embeddings for all spots that were measured with modality m_j. Accordingly, $C^{m_j} = \{c_1^{m_j}, \dots, c_N^{m_j}\}$ denotes the feature profiles (counts, gene activity scores, etc.) for all spots that were measured with modality m_j. T_b denotes the set of modalities that section b was measured with and $\bar{T}_b = U - T_b$ denotes the set of modalities that are missing in section b. To impute the missing modality $m_j \in \bar{T}_b$ for$b$, we first perform cross-modal matching between Z^{m_j} and $\{z^{m_t} | m_t \in T_b\}$. Specifically, for each $z_l^{m_t}$ in section b, we find the indices of its k_2 nearest neighbors from Z^{m_j}, $Q = \{q_1, q_2, \dots, q_{k_2}\}$. The profiles $c_l^{m_j}$ imputed from modality m_t are thus defined as:

$$c_l^{m_t \to m_j} = \frac{1}{k_2} \sum_{q \in Q} c_q^{m_j}$$

where $m_t \to m_j$ denotes the profiles imputed based on modality m_j. For each modality in T_b, we repeat above imputation process and the final imputed profiles for m_j in section b are given by:

$$c_l^{m_j} = \frac{1}{|T_b|} \sum_{m_t \in T_b} c_l^{m_t \to m_j}$$

where $|T_b|$ denote the number of elements in T_b.

Data simulation

Following the procedure of Townes et al.[23], we simulated multivariate count data with the ‘ggblocks’ spatial pattern. The ground true spatial factors comprise four simple shapes in different spatial regions. For each section created, the spots were arranged in a 36×36 grid for 1296 spots. The number of RNA features was set to $J^{rna} = 200$, and the number of protein features was set to $J^{pro} = 60$. Each feature was randomly assigned to one of the four factors with uniform probability. For the RNA modality, its count matrix was drawn from a zero-inflated negative binomial (ZINB) distribution with the mean $M^{rna} = b^{rna} + S^{rna} F^{rna} W^{rna}$ where $S^{rna} \in R^{1296 \times 1296}$, $F^{RNA} \in R^{1296 \times 4}$ and $W^{rna} \in R^{4 \times 200}$. $F_{il}^{rna} = 1$ if spatial factor l was active in spatial location i and $F_{il}^{rna} = 0$ otherwise. By default, $W_{jl}^{rna} = 10$ if feature j was assigned to spatial factor l and $W_{jl}^{rna} = 0$ otherwise. b^{rna} represents the background expression mean value. To simulate a spatial factor l mixing with the background factor, we multiplied F^{rna} by a diagonal matrix S^{rna}. For those entries corresponding to factor 1 we set them to s^{rna}, while for factors 2 and 3, we randomly selected 40% entries to s^{rna} and the other entries were set to 1. The shape parameter for the negative binomial component was set to 20 and the dropout probability was set to p^{rna}. For the protein modality, its count matrix was drawn from a negative binomial distribution with a mean $M^{pro} = b^{pro} + S^{pro}F^{pro}W^{pro}$, where $S^{pro} \in R^{1296\times1296}$, $F^{pro} \in R^{1296\times4}$ and $W^{pro} \in R^{4\times60}$. $F_{il}^{pro} = 1$ if spatial factor l was active in location i and 0 otherwise. By default, $W_{jl}^{pro} = 12$ if feature j was assigned to pattern l and $W_{jl}^{pro} = 0$ otherwise. S^{pro} is a diagonal matrix and for those entries corresponding to factor 4 we set them to s^{pro}. For factors 2 and 3, we randomly selected 40% entries to be assigned s^{pro}, and the other entries were set to 1. The shape parameter was set to 10.

To simulate batch effects between sections, we changed the W^{rna} and W^{pro} for the second and the third section, respectively. Specifically, if the RNA feature set $\{j_1^{rna}, j_2^{rna}, \ldots, j_{n_l}^{rna}\}$ was assigned to spatial factor l, we randomly picked half of these features and increased each W_{jl}^{rna} by a factor of 3. Similarly, if protein feature set $\{j_1^{rna}, j_2^{rna}, \ldots, j_{v_l}^{pro}\}$ was assigned to spatial factor l, we randomly picked half of these features and increased each corresponding W_{jl}^{pro} by a factor of 1.1, while for the remaining features, we increased each W_{jl}^{pro} by a factor 3.

We designed five sets of simulation parameters (**Supp. Fig. S2a**) and we created three datasets with different random seeds for each parameter set. Each dataset consists of three sections sharing the same simulation parameters. The first section contained RNA and protein expression, the second section only contained RNA expression, and the third section contained only protein expression (**Supp. Fig. S2b**).

Differential expression analysis and GO enrichment analysis

We used the Wilcoxon test as implemented in the rank_genes_groups function of the Scanpy package (version 1.9.6)[63] to identify DEGs of spatial clusters. We performed the GO enrichment analysis on cluster specific DEGs using the R package clusterProfile (version 4.6.2)[47].

Benchmarking metrics for integration

The benchmarking employed in this work encompassed three aspects: spatial domain identification, batch correction and modality alignment. Spatial domain identification was assessed using the adjusted rand index (ARI) and Moran's *I* metrics[24], batch correction was assessed using the graph integration local inverse Simpson's index (iLISI) metric[25], and modality alignment was assessed using Fraction of Samples Closer than True Match (FOSCTTM)[32], matching score (MS)[22] and label transfer F1 score.

Moran's I For a given feature (gene or label) and spatial coordinates of observations, Moran's *I* score assesses whether the pattern is clustered, dispersed or random. Moran's *I* statistic[24] is defined as:

$$I = \frac{n}{W}\frac{\sum_i^n \sum_j^n w_{i,j} v_i v_j}{\sum_{i=1}^n {v_i}^2}$$

where v_i is the deviation of the feature from the mean, $w_{i,j}$ is the spatial weight between observations, n is the number of spatial units, and $W = \sum_i^n \sum_j^n w_{i,j}$. Moran's *I* ranges from -1 to 1 where a higher value indicates greater smoothing of spatial patterns. We computed Moran's *I* score using the Squidpy package (version 1.2.3).

Graph iLISI The original iLISI[17] was initially conceived to assess the degree of batch mixing, and Graph iLISI has been extended to assess graph-based outputs. Graph iLISI uses a graph-based distance metric to determine the nearest neighbor list and avoids skews in graph-based integration outputs[10]. Its scaled output ranges from 0 to 1, where 0 indicates strong separation of batches while 1 indicates perfect batch mixing.

FOSCTTM FOSCTTM quantifies the degree of alignment between embeddings of different modalities for the same location. For each modality pair (m_1, m_2), FOSCTTM is defined as:

$$FOSCTTM = \frac{1}{2N}\left(\sum_i \frac{N_i^{m_1}}{N} + \sum_i \frac{N_i^{m_2}}{N}\right)$$

where N is the number of spots that are simultaneously measured in modality m_1 and m_2, $N_i^{m_1}$ denotes the number of spots in modality m_1 that are closer to $z_i^{m_2}$ than $z_i^{m_1}$ to $z_i^{m_2}$ and $N_i^{m_2}$ is the number of spots in modality m_2 that are closer to $z_i^{m_1}$ than $z_i^{m_2}$ to $z_i^{m_1}$. The computed value ranges from 0 to 1, where 1 indicates perfect modality matching.

MS Like FOSCTTM, MS quantifies the degree of modality alignment and is calculated using spots with multimodal measurements. For two modalities, m_1 and m_2, a cross-modality matching matrix P is first constructed by computing the Jaccard index of cross-modality nearest neighbors in the modal-aligned embedding space[22]:

$$P_{m_1,m_2}(i,j) = \frac{\left|\left(N^{m_2}\left(z_i^{m_1}\right) \cap N^{m_2}\left(z_j^{m_2}\right)\right) \cup \left(N^{m_1}\left(z_j^{m_2}\right) \cap N^{m_1}\left(z_i^{m_1}\right)\right)\right|}{\left|\left(N^{m_2}\left(z_i^{m_1}\right) \cup N^{m_2}\left(z_j^{m_2}\right)\right) \cup \left(N^{m_1}\left(z_j^{m_2}\right) \cup N^{m_1}\left(z_i^{m_1}\right)\right)\right|}$$

where $N^{m_2}\left(z_i^{m_1}\right)$ denotes the set of spot $z_i^{m_1}$'s nearest neighbors in modality m_2, $N^{m_2}\left(z_j^{m_2}\right)$ denotes the set of spot $z_j^{m_2}$'s nearest neighbors in modality m_2 (including itself), and vice versa for $N^{m_1}\left(z_j^{m_2}\right)$ and $N^{m_1}\left(z_i^{m_1}\right)$. Following the approach of CLUE, the number of nearest neighbors is set to 1 (i.e., $P(i,j) = 1$ if spot i and j are mutual nearest neighbors). Finally, MS is calculated by:

$$MS = \frac{1}{N}\sum_i \sum_j \tilde{P}_{m_1,m_2}(i,j) \cdot \delta_{i,j}$$

where $\tilde{P}$ denotes the row-normalized P, $\delta_{i,j} = 1$ if spot i and j represent the same spot and 0 otherwise, and N is the number of spots. The computed MS ranges between 0 and 1 where 1 indicates perfect alignment.

Label transfer F1 score This metric assesses the alignment of labeled cell types or spatial domains between sections and is not restricted to spots with multimodal measurements. Using this metric, we can test for the transfer accuracy of cell type or spatial domain labels from one section to another. For each pair of modalities, we first built a *k*NN graph between their embeddings and the transferred labels from one to another based on the *k*NN classifier (*k*=10). Thereafter, we calculated the F1 score by comparing the transferred labels with the ground truth. For each pair of modalities, we averaged the F1 scores from both transfer directions.

Benchmarking methods for mosaic integration

We benchmarked SpaMosaic and other state-of-the-art single-cell mosaic integration methods, CLUE, Cobolt, scMoMaT, StabMap and MIDAS. We also tested a non-spatial version of SpaMosaic which we refer to as CL. Here we describe the detailed settings used for each method.

CLUE. We employed CLUE from the scglue package (version 0.1.1) in Python. We followed the processing pipeline as described by the authors (https://github.com/openproblems-bio/neurips2021_multimodal_topmethods/tree/main/src/match_modality/methods/clue). Raw count matrices of each modality were used as input, which were then transformed by the customized preprocessing steps in CLUE. Thereafter, we applied the default model training pipeline. We set the $shared\_batches = False$ and use the recommended training parameters for different types of input data, namely RNA+ATAC and RNA+protein.

Cobolt. We ran the Cobolt package (version 1.0.1) following its tutorial (https://github.com/epurdom/cobolt/blob/master/docs/tutorial.ipynb). We first filtered for the highly variable features (top 5000 for RNA, top 100000 for epigenome) from the count data matrices before input to the model for training. Most of training hyperparameters were set to default values: training iteration=100, latent dimension=10, batch size=128. However, we modified the learning rate from 0.005 to 0.001 to avoid the numeric issues encountered with our datasets. When running the 'calc_all_latent' function in the five-modality integration task, we set the parameters $target = [True, False, False, False, False]$ for successful execution.

scMoMaT We employed the Python package scMoMaT (version 0.2.2) by following its tutorial (https://github.com/PeterZZQ/scMoMaT/blob/main/demo_scmomat.ipynb). We first selected the highly variable features (top 5000 for RNA, top 100000 for epigenome) in the count data matrices as input for model training. The model parameters were kept at their defaults except that we disabled the generation of pseudo-count RNA matrix from ATAC data for better performance with our examples.

StabMap We employed the StabMap package (version 0.1.8) in R, following its tutorial (https://github.com/MarioniLab/StabMap/tree/main/vignettes). We first used Seurat's functions for data normalization, namely the 'NormalizeData' function for RNA data, 'RunTFIDF' function for epigenomic data, and 'NormalizeData' function with parameter $normalize.method = 'CLR'$ for protein data. Thereafter we selected the highly variable features (top 5000 for RNA, top 100000 for epigenome) as input into StabMap. We adjusted the parameter $maxFeatures$ according to the number of total features for each dataset while the remaining parameters were left at their default values.

MIDAS We ran the MIDAS package following its tutorial (https://sc-midas-docs.readthedocs.io/en/latest/mosaic.html). The highly variable features (top 5000 for RNA, top 100000 for epigenome) were first selected, and count matrices were then used for training. For epigenomic input, the features were split by chromosome id to reduce model complexity. For all runs, the number of training epoch was set to 2,000.

SpaMosaic (non-spatial). SpaMosaic (non-spatial) is a modified version of SpaMosaic that uses an MLP as the encoder network without considering spatial information. The MLP encoder consists of two fully connective layers where the first layer is followed by a dropout layer (p=0.2) and non-linear activation ELU. The other settings are the same as those for SpaMosaic.

Data availability

All datasets used in this work were obtained from public repositories, except for the in-house generated human lymph node and tonsil datasets. The mouse embryo dataset was obtained from https://www.biosino.org/node/project/detail/OEP003285 and https://pan.baidu.com/s/15i0R5na2Jud_7EpkWSaZpA?pwd=oipj. For the mouse postnatal brain datasets and mouse embryo datasets, the multimodal sections were obtained from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE205055 or https://cells.ucsc.edu/?ds=brain-spatial-omics and the single-modal sections are available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165217 and https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171943. The simulated datasets, in-house human lymph node and tonsil datasets are available at https://zenodo.org/uploads/12654113. The scRNA-seq atlas of the mouse nervous system was obtained from http://mousebrain.org/adolescent/downloads.html. The relevant in situ hybridization and expression images were downloaded from the Allen Mouse Brain Atlas.

Code availability

An open-source python package of SpaMosaic is available at https://github.com/JinmiaoChenLab/SpaMosaic and https://zenodo.org/uploads/12654113.

Authors' contributions

J.C. and M.L. initiated the project and provided funding support. J.C. conceived the idea. X.Y., J.C. and R.Z. developed the method and designed the experiments. X.Y., J.C and K.S.A. performed the data analysis and wrote the manuscript. L.V.O., A.E., T.W. and D.G. provided the lymph node and tonsil datasets, and manually annotated the tissue structure. R.F. provided the postnatal mouse brain and mouse embryo datasets.

Acknowledgements

The research was supported by A*STAR under its BMRC Central Research Fund (CRF, UIBR) Award; AI, Analytics and Informatics (AI3) Horizontal Technology Programme Office (HTPO) seed grant (Spatial transcriptomics ST in conjunction with graph neural networks for cell–cell interaction #C211118015) from A*STAR, Singapore; Open Fund Individual

Research Grant (Mapping hematopoietic lineages of healthy and high-risk acute myeloid leukemia patients with FLT3-ITD mutations using single-cell omics #OFIRG18nov-0103) from Ministry of Health, Singapore; National Research Foundation (NRF), Award no. NRF-CRP26-2021-0001; the National Research Foundation, Singapore, and Singapore Ministry of Health's National Medical Research Council under its Open Fund-Large Collaborative Grant ("OF-LCG") (MOH-OFLCG18May-0003); Singapore National Medical Research Council (#NMRC/OFLCG/003/2018). M.L. is supported in part by the National Natural Science Foundation of China under Grant (No.62225209 to M.L.).

Competing Interests

The authors declare that there are no competing interests.

Figure legend

Fig. 1 Overview of SpaMosaic framework. a SpaMosaic is designed to integrate multiple tissue sections with different modality compositions to construct a more comprehensive view of tissues and impute missing modalities. **b** SpaMosaic first performs dimension reduction and applies Harmony to obtain batch-corrected low-dimension representations for each modality. Thereafter, a heterogeneous adjacency graph is constructed between spots within each modality. SpaMosaic employs graph neural networks to embed the modality-specific representations into modality-specific embeddings. The paired multi-modal embeddings are input for contrastive learning which aims to align embeddings from different modalities and constructs a modality-agnostic latent space. Finally, SpaMosaic infers the aligned embeddings for each section for downstream analysis (such as spatial domain identification) and imputation.

Fig. 2 Benchmarking on simulated datasets. a Ground truth of spatial factors and background of the three sections. **b** Clustering patterns obtained with the concatenated RNA and protein profiles, RNA profiles and protein profiles. Each section was clustered individually using *k*-means. ARI scores were evaluated by comparing the clustering with the ground truth. **c** UMAP plots for RNA profiles of sections 1 and 2, protein profiles of sections 1 and 3. Spots are colored by section and factor in each panel. **d** Clustering patterns obtained using SpaMosaic's embeddings for all three sections. **e** UMAP plots of SpaMosaic's embeddings. From left to right, spots are colored by section, clustering and factor. **f** iLISI scores of all benchmarked methods across five simulation sets. The error bars represent the 95% confidence interval around the mean of the scores. **g** ARI scores on the whole dataset, section 1, section 2 and section 3 for all methods. 'Raw' indicates the ARI scores that were calculated on the raw data of each section. In the box plot, the center lines indicate the median, boxes indicate the interquartile range, and whiskers indicate 1.5× interquartile range. **h** Average Moran's *I* scores on each dataset for all methods across the five datasets. **i** Average Moran's *I* score on individual sections for all methods. **j** Label transfer F1-scores of all methods. **k** Sankey plots showing the comparison of prediction and ground truth of factors on the first replicate of the first simulation setting.

Fig. 3 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on an embryonic mouse brain (RNA+ATAC) dataset. a Mosaic dataset and clustering patterns for all three sections: E13.5 without the ATAC modality, E15.5 with both RNA and ATAC expression profiles, and E18.5 without the RNA modality. The modality-specific clustering labels and manual annotations were obtained from the original study. The rightmost column shows the Allen Developing Mouse Brain Atlas reference annotations for the corresponding three developmental stages. **b** Spatial plots of the clusters obtained from all methods for the three sections. All methods share the color map. Manual annotation was performed to identify SpaMosaic's clusters. **c** RNA expression and gene activity scores of marker genes *Emx1*, *Emx2*, *Id4*, *Nr2f1* and *Pax6* in SpaMosaic's clusters. **d** Spatial heatmap of gene expression and gene activity scores of the marker genes on the E13.5 section. The bottom row shows the ISH images from the Allen Mouse Developing Brain Atlas for the corresponding markers. **e** Reference annotations of the E13.5 embryonic mouse brain obtained from the Allen Developing Mouse Brain Atlas. **f** ARI scores of the whole dataset, E15.5, E13.5 and E18.5 sections, respectively, for all methods. **g** Moran's I scores of all methods on each section. **h** iLISI scores of all methods. **e** Modality alignment scores (FOSCTTM and MS) of all methods.

Fig. 4 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a postnatal

mouse brain (RNA+H3K4me3) dataset. a Mosaic dataset configuration and clustering patterns obtained for all three sections. The RNA and H3K4me3 clusters were obtained by performing *k*-means clustering on their respective batch-corrected low-dimensional embeddings from Harmony. The multi-modal cluster labels of section 1 were obtained by performing *k*-means clustering on the concatenated batch-corrected low-dimensional embeddings of the RNA and H3K4me3 profiles. **b** Spatial plots of the clusters obtained from all methods for all three sections. All methods share the color map. Manually annotation was performed to identify SpaMosaic's clusters. **c** Allen Brain Atlas reference annotation for the postnatal mouse brain. **d** RNA expression of marker genes for the six brain regions in SpaMosaic's clusters. **e** Spatial heatmap showing the gene activity scores of selected markers on the third section. **f** Clusters overlaid on the image of the third section and cell types identified by label transfer. **g** Moran's I scores of all methods across the three sections. We varied the clustering numbers (n=4, 6, 8, 10) and calculated the Moran's I scores for each clustering result. In the box plots, the center lines indicate the median, boxes indicate the interquartile range, and whiskers indicate 1.5× interquartile range. **h** iLISI scores of all methods.

Fig. 5 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a human tonsil dataset. a Mosaic dataset configuration and clustering patterns obtained for all three sections. The RNA and protein clustering labels were obtained by performing *k*-means clustering on the batch-corrected low-dimension embeddings from Harmony. **b** Spatial plots of clustering labels from the methods for all three sections. All methods share the color map. Manually annotation was performed to identify SpaMosaic's clusters. **c** RNA and protein expression of region-specific markers in SpaMosaic's clusters. **d** ARI scores on the whole dataset, section 1, section 2 and section 3 for all methods. The number of clusters was varied (n=4, 6, 8, 10) and we calculate the ARI for each clustering result. **e** Moran's I scores for all methods across all sections. Similar to the ARI evaluation, the number of clusters was varied in the same fashion. **f** iLISI scores of all methods. **g** Modality alignment scores of all methods (FOSCTTM and MS).

Fig. 6 Application of SpaMosaic to five-modality mosaic integration. a Mosaic dataset configuration and clustering patterns of the four sections. The RNA cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimension embeddings from Harmony, and cluster labels for each epigenomic modality were obtained by performing *k*-means clustering on the principal components of the epigenomics profiles. The multi-modal clustering labels were obtained by performing *k*-means clustering on the concatenated embeddings of the RNA and epigenomics profiles. **b** Spatial plots of clusters obtained from the methods for all four sections. All methods share the color map. Manual annotation was performed to identify SpaMosaic's clusters. **c** Allen Brain Atlas reference anatomic annotations for the postnatal mouse brain. **d** RNA expression of brain region markers in SpaMosaic's clusters. **e** Moran's I scores for the four sections. The number of clusters was varied (n=6, 8, 10, 12). **f** iLISI scores of all methods. **g** Relation of the measured ATAC GAS, imputed H3K4me3 GAS and imputed H3K27ac GAS with the RNA in the CCG region of the fourth section. Only region-specific genes were kept for all modalities if their adjusted *p* value was less than 1e-5 and the absolute value of $\log_2$ fold change was larger than 0.5. **h** Upset plot of genes in all quadrant I of **g**. **i** Relation of imputed H3K27me3 GAS, imputed H3K4me3 GAS and imputed H3K27ac GAS with measured ATAC GAS in the CCG region of the fourth section. Only region-specific genes were kept if their adjusted *p* value was less than 1e-5 and the absolute value of $\log_2$ fold change was larger than 0.5. **j** Enriched GO terms for the genes in quadrant IV (**i**, H3K27me3 relation), genes in quadrant I (**i**, H3K4me3 relation) and genes in quadrant I (**i**, H3K27ac relation).

Supplementary figures

Supp. Fig. S1 Graph neural network architecture and imputation strategy. a Aggregation and propagation processes of weighted LGCN. The weight for spatial neighbors is set to 1 while the weight for expression neighbors is set to 0.8. In each layer, each node aggregates the information from its neighbors using the weighted average. After two layers, the information from each node is concatenated across three output layers and then processed by a multiple layer perceptron. **b** Imputation-based on modal-aligned latent space. Taking section 3 as an example, which only has the ATAC modality, we aim to impute its RNA and protein profiles. To impute the RNA profile of each spot in section 3, we find its RNA-specific neighbors from the other sections and then compute the average target expression profiles of these neighbors. Protein imputation employs the same process.

Supp. Fig. S2 Simulation settings and plots for competing methods. a Five sets of simulation parameters used for data generation. **b** Mosaic configuration of each dataset. A dotted box denotes that this profile is missing. **c** ARI scores of clusters obtained by *k*-means and mclust clustering of the embeddings from the different methods and the raw data. **d** UMAP plots of the embeddings from all benchmarked methods. Spots are colored by section (first row), cluster (second row) and factor (third row). **e** Spatial plots of the clustering results of all methods. All methods and sections share the same color map.

Supp. Fig. S3 Additional results from benchmarking with an embryonic mouse brain (RNA+ATAC) dataset. a UMAP plots of the raw RNA and ATAC profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all tested methods. Spots in the first row are colored by section, the second row by cluster and the third row by manual annotation. **c** Mosaic dataset for evaluation with modality alignment metrics. The RNA and ATAC profiles of the E18.5 section were split into two separate sections to simulate originating from different experiments. Cluster labels were obtained from the original study. **d** UMAP plots of the raw RNA and ATAC profiles of the E15.5 and E18.5 sections. Spots are colored by section.

Supp. Fig. S4 Additional results of a postnatal mouse brain dataset (RNA+H3K4me3) benchmarking. a UMAP plots of the raw RNA and H3K4me3 profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all methods. Spots in the first row are colored by section, the second row by cluster.

Supp. Fig. S5 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a postnatal mouse brain (RNA+H3K27me3) dataset. a Mosaic dataset configuration and clustering patterns obtained for all three sections. The RNA and H3K27me3 cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. The multi-modal cluster labels of section 1 were obtained by performing *k*-means clustering on the concatenated batch-corrected low-dimensional embeddings of the RNA and H3K27me3 profiles. **b** Spatial plots of the clusters from the methods for all three sections. All methods share the color map. Manually annotation was performed to identify SpaMosaic's clusters. **c** RNA expression of marker genes of the six brain regions in SpaMosaic's clusters. **d** Clusters overlaid on the image of the third section (top row) and spatial heatmap showing the gene activity scores of selected markers on the third section (middle and bottom rows). **e** Moran's I scores of all methods across the three sections. We varied the clustering numbers (n=4, 6, 8, 10) and calculated the Moran's I scores for each clustering result. **f** iLISI scores of all methods.

Supp. Fig. S6 Additional results of a postnatal mouse brain dataset (RNA+H3K27me3) benchmarking. a UMAP plots of the raw RNA and H3K27me3 profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all methods. Spots in the first row are colored by section, the second row by cluster.

Supp. Fig. S7 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a postnatal mouse brain dataset (RNA+ATAC). a Mosaic dataset configuration and clustering patterns obtained for all three sections. The RNA and ATAC cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. The multi-modal cluster labels of section 1 were obtained by performing *k*-means clustering on the concatenated batch-corrected low-dimensional embeddings of the RNA and ATAC profiles. **b** Spatial plots of the clusters obtained from the methods for all three sections. All methods share the color map. Manually annotation was performed to identify SpaMosaic. **c** Heatmap of the DEGs of SpaMosaic's clusters in section 1 and section 2. **d** Clusters overlaid on the image of the third section (left) and spatial heatmap depicting the gene activity scores of selected markers on the third section (middle and right). **e** Moran's I scores of all methods across the three sections. We varied the clustering numbers (n=4, 6, 8, 10) and calculated the Moran's I scores for each clustering result. **f** iLISI scores of all methods. **g** Mosaic dataset for evaluation with modality alignment metrics. Raw RNA and ATAC profiles of the second section were split into two separate sections to simulate originating from different experiments. Cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. **h** UMAP plots of the raw RNA profiles and ATAC profiles in **g**. Spots are colored by section. **i** Modality alignment scores (FOSCTTM and MS) of all methods.

Supp. Fig. S8 Additional results of a postnatal mouse brain dataset (RNA+ATAC) benchmarking. a UMAP plots of the raw RNA and ATAC profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all methods. Spots in the first row are colored by section, the second row by cluster.

Supp. Fig. S9 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a postnatal mouse brain (RNA+H3K27ac) dataset. a Mosaic dataset configuration and clustering patterns obtained for all three sections. The RNA and H3K27ac cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. The multi-modal cluster labels of section 1 were obtained by performing *k*-means clustering on the concatenated batch-corrected low-dimensional embeddings of the RNA and H3K27ac profiles. **b** Spatial plots of the clustering labels from the methods for all three sections. All methods share the color map. Manually annotation was performed to identify SpaMosaic's clusters. **c** Heatmap of the DEG of SpaMosaic's clusters in section 1 and section 2. **d** Clusters overlaid on the image of the third section (top left) and spatial heatmap showing the gene activity scores of selected markers on the third section. **e** Moran's **I** scores of all methods across the three sections. We varied the clustering numbers (n=4, 6, 8, 10) and calculated the Moran's **I** scores for each clustering result. **f** iLISI scores of all methods. **g** Mosaic dataset for evaluation with modality alignment metrics. Raw RNA and H3K27ac profiles of the second section were split into two separate sections to simulate originating from different experiments. Cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. **h** UMAP plots of the raw RNA and H3K27ac profiles in **g**. Spots are colored by section. **i** Modality alignment scores of all methods (FOSCTTM and MS).

Supp. Fig. S10 Additional results of the postnatal mouse brain dataset (RNA+H3K27ac) benchmarking. a UMAP plots of the raw RNA and H3K27ac profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all methods. Spots in the first row are colored by section, the second row by cluster.

Supp. Fig. S11 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a RNA+ATAC mouse embryo dataset. a Mosaic dataset configuration and clustering patterns obtained for all four sections. The RNA cluster labels were obtained by performing *k*-means clustering on the principal components of the RNA profiles, and ATAC cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimension embeddings from Harmony. The multi-modal cluster labels of section 1 were obtained by performing *k*-means clustering on the concatenated batch-corrected low-dimensional embeddings of the RNA and ATAC profiles. **b** Spatial plots of the clusters obtained from the methods for all four sections. All methods share the color map. **c** RNA expression and gene activity scores of marker genes for the different embryo organs and tissues. **d** Moran's **I** scores of all methods across four sections. We varied the clustering numbers (n=6, 8, 10, 12) and calculated the Moran's **I** scores for each clustering result. **e** iLISI scores of all methods.

Supp. Fig. S12 Additional results of a mouse embryo dataset (RNA+ATAC) benchmarking. a UMAP plots of the raw ATAC profiles. Spots are colored by section. **b** UMAP plots of the embeddings from all methods. Spots in the first row are colored by section, the second row by cluster.

Supp. Fig. S13 Additional results of benchmarking with a human tonsil dataset. a UMAP plots of the raw RNA and protein profiles. Cells are colored by section. **b** H&E images of the three sections. **c** UMAP plots of the embeddings from all methods. Spots are colored by section in the first row, colored by cluster in the second row and colored by manual annotation in the third row. **d** Mosaic dataset for evaluation with modality alignment metrics. Raw RNA and protein profiles of the second section were split into two separate sections to simulate originating from different experiments. For each modality, cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. **e** UMAP plots of the raw RNA and protein profiles in **d**. Spots are colored by section.

Supp. Fig. S14 Downstream analysis results obtained with SpaMosaic and quantitative benchmarking on a human lymph node dataset. **a** Mosaic dataset configuration and clustering patterns for all three sections. The RNA and protein clustering labels were obtained by performing *k*-means clustering on the batch-corrected low-dimension embeddings from Harmony. **b** Spatial plots of clusters obtained from all methods for the three sections. All methods share the color map. **c** RNA and protein expression of region markers in SpaMosaic's clusters. **d** ARI scores of the whole dataset, section 1, section 2 and section 3, respectively, for all methods. The number of clusters was varied (n=4, 6, 8, 10) and we calculated the ARI for each clustering result. **e** Moran's I scores of all methods across all three sections. Similar to the ARI evaluation, the number of clusters was varied in the same fashion. **f** iLISI scores of all methods. **g** Modality alignment scores of all methods (FOSCTTM and MS).

Supp. Fig. S15 Additional results in the human lymph node experiment. **a** UMAP plots of the raw RNA and protein profiles. Cells are colored by section. **b** H&E images of all three sections. **c** UMAP plots of the embeddings from all methods. Spots are colored in section in the first row, colored by cluster in the second row and colored by manual annotation in the third row. **d** Mosaic dataset for evaluation with modality alignment metrics. Raw RNA and protein profiles of the second section were split into two separate sections to simulate originating from different experiments. For each modality, cluster labels were obtained by performing *k*-means clustering on the batch-corrected low-dimensional embeddings from Harmony. **e** UMAP plots of the raw RNA and protein profiles in **d**. Spots are colored by section.

Supp. Fig. S16 UMAP plots for five-modality mosaic dataset. **a** UMAP plots of the raw RNA profiles. **b** UMAP plots of the embeddings from all methods. Spots are colored by section in the first row and colored by cluster in the second row.

Supp. Fig. S17 Correlation between RNA and GAS/CSS within different regions. All genes plotted have an adjusted *p* value less than 1e-5 and absolute value of $\log_2$ fold change larger than 0.5. The orange boxes in **a, b, c, d** denoted the measured relation on each section. **a** Correlation between RNA expression and GAS/CSS for CCG-specific (SpaMosaic's cluster 2) genes across the four sections. **b** Correlation between RNA expression and GAS/CSS for CP-specific (SpaMosaic's cluster 3) genes across all sections. **c** Correlation between RNA expression and GAS/CSS for cortex-specific (SpaMosaic's cluster 10) genes across all sections. **d** Correlation between RNA expression and GAS/CSS for the VL-specific (SpaMosaic's cluster 12) genes across all sections. **e** Upset plot of CCG-specific genes in section 4 from all quadrants III of measured, H3K4me3 imputed and H3K27ac imputed relation. **f** Enriched GO terms for CCG-specific genes in section 4 from quadrant I that are unique to the H3K27ac imputed relation compared to the measured relation. **g** Enriched GO terms for CCG-specific genes in section 4 from quadrant I that are unique to the H3K4me3 imputed relation compared to the measured relation. **h** Upset plot of other brain region-specific genes in section 4 from quadrant I and III.

Supp. Fig. S18 Spatial heatmap of measured and imputed gene expression/gene activity scores. Spatial heatmap of measured RNA, measured ATAC GAS, imputed H3K4me3 GAS and imputed H3K27ac GAS for CCG-specific genes *Bcas1*, *Car2* and *Ermn*.

Supp. Fig. S19 Relation between ATAC GAS and histone modification GAS/CSS in the CP and cortex regions. **a** Relation between imputed GAS/CSS and measured ATAC GAS for CP-specific genes on the fourth section. **b** Enriched GO terms for CP-specific genes from quadrant I and IV in **a**. **c** Relation between imputed GAS/CSS and measured ATAC GAS for cortex-specific genes on the fourth section. **d** Enriched GO terms for cortex-specific genes from quadrant I and IV in **c**.

Supp. Fig. S20 Imputed epigenomic information improves RNA prediction accuracy. **a** RNA prediction error and

Pearson correlation coefficient with different modality inputs. For example, 'H3K27me3 (section 1)' indicates the score of imputing RNA on the first section. 'Single-modality' indicates that the prediction model only uses the measured H3K27me3 CSS while 'four modalities' indicates that the measured H3K27me3 CSS and other imputed histone modification/ATAC GAS were all used for prediction. **b** Quality control plots of RNA profiles in the four sections. The first section (H3K27me3) has the highest level of 'n_gene_by_counts' (the number of genes expressed in the count matrix), the highest level of 'total_counts' (total counts per cell), and the lowest level of 'pct_counts_mt' (percentage of counts in mitochondrial genes), suggesting it has the highest data quality among the four sections, which explains its superior prediction performance.

Supp. Fig. S21 Imputation benchmarking of MIDAS, MultiVI, BABEL and SpaMosaic. a Benchmarking of ATAC imputation on the embryonic and postnatal mouse brain datasets. **b** Benchmarking of ATAC imputation with smoothed ground truth peak count matrices. **c** Benchmarking of protein imputation on the human lymph node and tonsil datasets.

Supp. Fig. S22 Benchmarking different graph neural network architectures. a UMAP plots of the embeddings generated with different GNNs on the first replicate of the first simulation setting. From left to right, the spots are colored by section, cluster and factor ground truth. **b** Spatial plots of clusters obtained with different GNNs. The spot colors are matched with the cluster colors in **a**. **c** Box plots show the ARI scores on the whole dataset, section 1, section 2 and section 3 for all GNNs. 'Raw' indicates ARI scores that were calculated on the raw data of each section. In the box plot, the center lines indicate the median, boxes indicate the interquartile range, and whiskers indicate 1.5× interquartile range. **d** Average Moran's I scores of all GNNs for each dataset with all replicates. **e** Average Moran's I scores of all GNNs on individual sections for all datasets. **f** iLISI scores of all GNNs for each dataset with all replicates. **g** Label transfer F1-scores of all GNNs for each dataset with all replicates.

Reference

1. Deng, Y. *et al.* Spatial-CUT&amp;Tag: Spatially resolved chromatin modification profiling at the cellular level. *Science (1979)* **375**, 681–686 (2022).
2. Zhang, D. *et al.* Spatial epigenome–transcriptome co-profiling of mammalian tissues. *Nature* **616**, 113–122 (2023).
3. Ben-Chetrit, N. *et al.* Integration of whole transcriptome spatial profiling with protein markers. *Nat Biotechnol* **41**, 788–793 (2023).
4. Zhang, D. *et al.* Spatial dynamics of mammalian brain development and neuroinflammation by multimodal tri-omics mapping. *bioRxiv* (2024) doi:10.1101/2024.07.28.605493.
5. Long, Y. *et al.* Deciphering spatial domains from spatial multi-omics with SpatialGlue. *Nat Methods* (2024) doi:10.1038/s41592-024-02316-4.
6. Argelaguet, R., Cuomo, A. S. E., Stegle, O. & Marioni, J. C. Computational principles and challenges in single-cell data integration. *Nat Biotechnol* **39**, 1202–1215 (2021).
7. Zhang, Z. *et al.* scMoMaT jointly performs single cell mosaic integration and multi-modal bio-marker detection. *Nat Commun* **14**, 384 (2023).
8. Gong, B., Zhou, Y. & Purdom, E. Cobolt: integrative analysis of multimodal single-cell sequencing data. *Genome Biol* **22**, 351 (2021).
9. Ghazanfar, S., Guibentif, C. & Marioni, J. C. Stabilized mosaic single-cell data integration using unshared features. *Nat Biotechnol* **42**, 284–292 (2024).
10. He, Z. *et al.* Mosaic integration and knowledge transfer of single-cell multimodal data with MIDAS. *Nat Biotechnol* (2024) doi:10.1038/s41587-023-02040-y.
11. Xu, H. *et al.* Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Med* **16**, 12 (2024).
12. Long, Y. *et al.* Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat Commun* **14**, 1155 (2023).
13. Zhou, X., Dong, K. & Zhang, S. Integrating spatial transcriptomics data across different conditions,

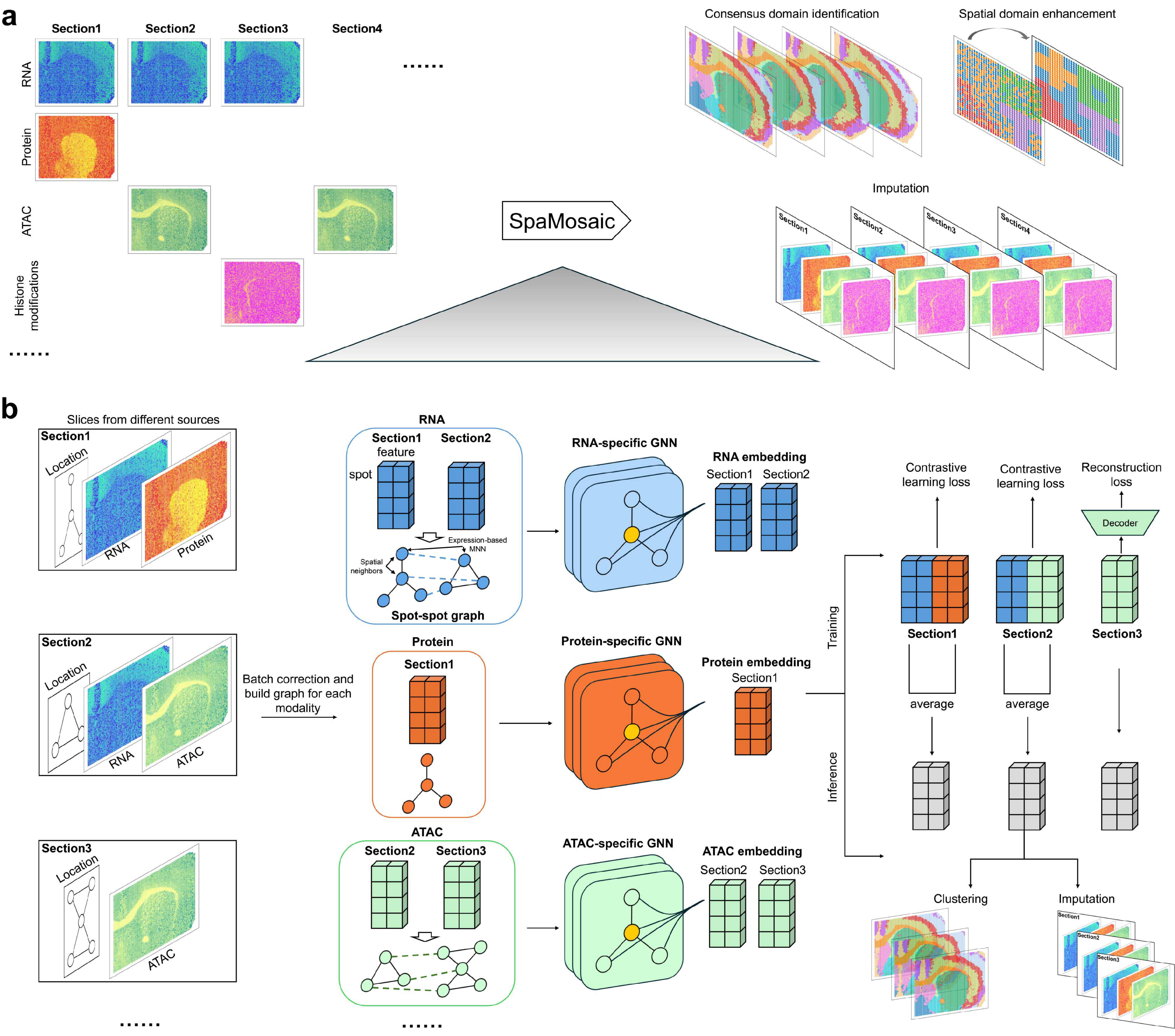
technologies and developmental stages. *Nat Comput Sci* **3**, 894–906 (2023).

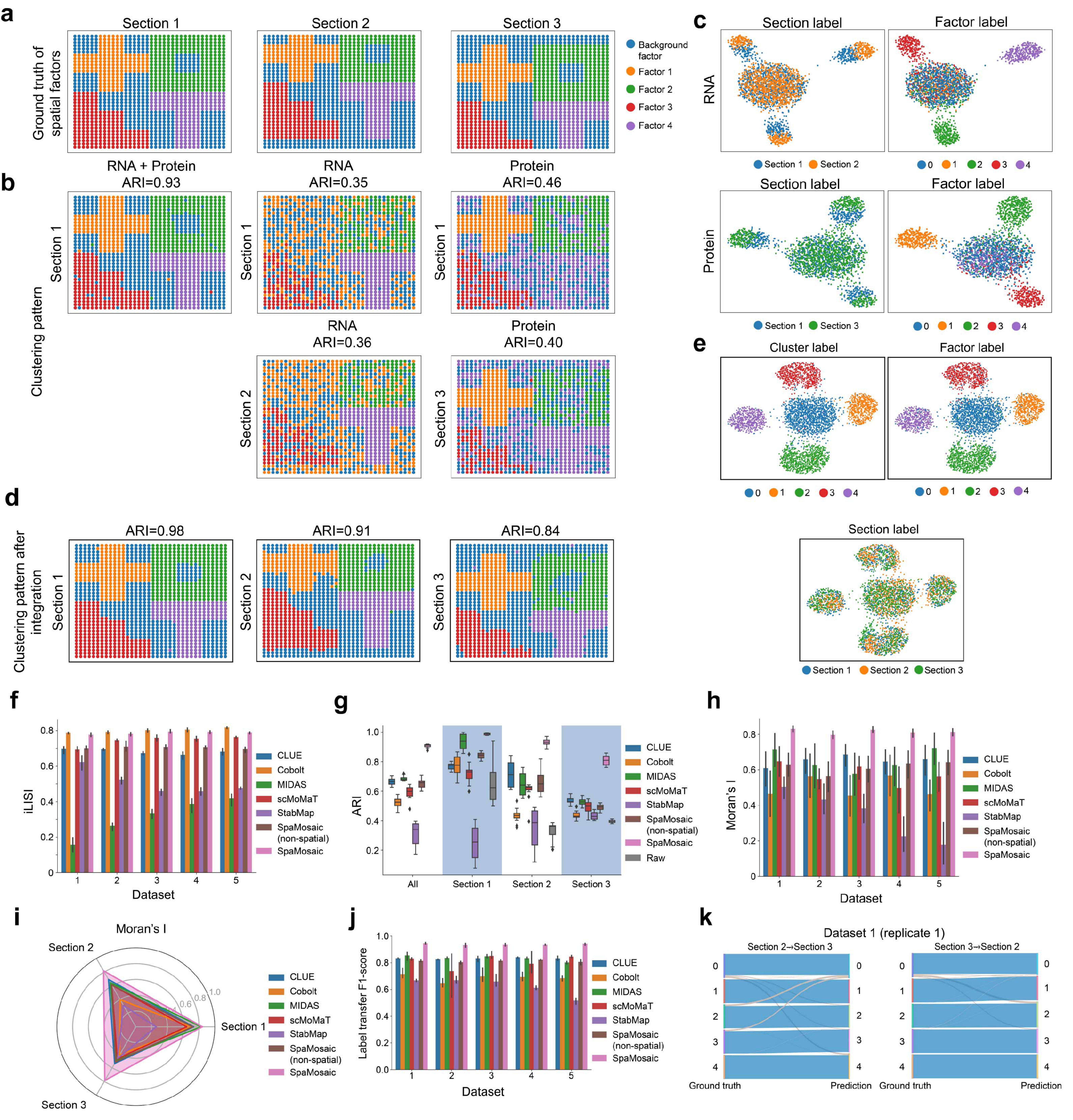
14. Wang, G. *et al.* Construction of a 3D whole organism spatial atlas by joint modelling of multiple slices with deep neural networks. *Nat Mach Intell* **5**, 1200–1213 (2023).
15. Li, Z. *et al.* Cross-modality representation and multi-sample integration of spatially resolved omics data. *bioRxiv* (2024) doi:10.1101/2024.06.10.598155.
16. Chen, T., Kornblith, S., Norouzi, M. & Hinton, G. A simple framework for contrastive learning of visual representations. in *International conference on machine learning* 1597–1607 (2020).
17. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).
18. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**, 421–427 (2018).
19. Veličković, P. *et al.* Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
20. Dong, K. & Zhang, S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat Commun* **13**, 1739 (2022).
21. Hu, Z., Dong, Y., Wang, K. & Sun, Y. Heterogeneous graph transformer. in *Proceedings of the web conference 2020* 2704–2710 (2020).
22. Tu, X., Cao, Z.-J., Mostafavi, S., Gao, G. & others. Cross-linked unified embedding for cross-modality representation learning. *Adv Neural Inf Process Syst* **35**, 15942–15955 (2022).
23. Townes, F. W. & Engelhardt, B. E. Nonnegative spatial factorization applied to spatial genomics. *Nat Methods* **20**, 229–238 (2023).
24. Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nat Methods* **19**, 171–178 (2022).
25. Luecken, M. D. *et al.* Benchmarking atlas-level data integration in single-cell genomics. *Nat Methods* **19**, 41–50 (2022).
26. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**, 38–44 (2019).
27. Fraley, C. & Raftery, A. E. MCLUST: Software for model-based cluster and discriminant analysis. *Department of Statistics, University of Washington: Technical Report* **342**, 1312 (1998).
28. Jiang, F. *et al.* Simultaneous profiling of spatial gene expression and chromatin accessibility during mouse brain development. *Nat Methods* **20**, 1048–1057 (2023).
29. Bishop, K. M., Rubenstein, J. L. R. & O'Leary, D. D. M. Distinct Actions of *Emx1* , *Emx2* , and *Pax6* in Regulating the Specification of Areas in the Developing Neocortex. *The Journal of Neuroscience* **22**, 7627–7638 (2002).
30. Bertacchi, M. *et al.* NR2F1 regulates regional progenitor dynamics in the mouse neocortex and cortical gyrification in BBSOAS patients. *EMBO J* **39**, (2020).
31. Yun, K., Mantani, A., Garel, S., Rubenstein, J. & Israel, M. A. *Id4* regulates neural progenitor proliferation and differentiation in vivo. *Development* **131**, 5441–5448 (2004).
32. Singh, R. *et al.* Unsupervised manifold alignment for single-cell multi-omics data. in *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics* 1–10 (ACM, New York, NY, USA, 2020). doi:10.1145/3388440.3412410.
33. Luecken, M. D. *et al.* A sandbox for prediction and integration of DNA, RNA, and proteins in single cells. in *Thirty-fifth conference on neural information processing systems datasets and benchmarks track (Round 2)* (2021).
34. Khandelwal, N. *et al.* FOXP1 regulates the development of excitatory synaptic inputs onto striatal neurons and induces phenotypic reversal with reinstatement. *Sci Adv* **10**, (2024).
35. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999-1014.e22 (2018).
36. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
37. Deng, Y. *et al.* Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature* **609**, 375–383 (2022).
38. Liu, Y. *et al.* High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-

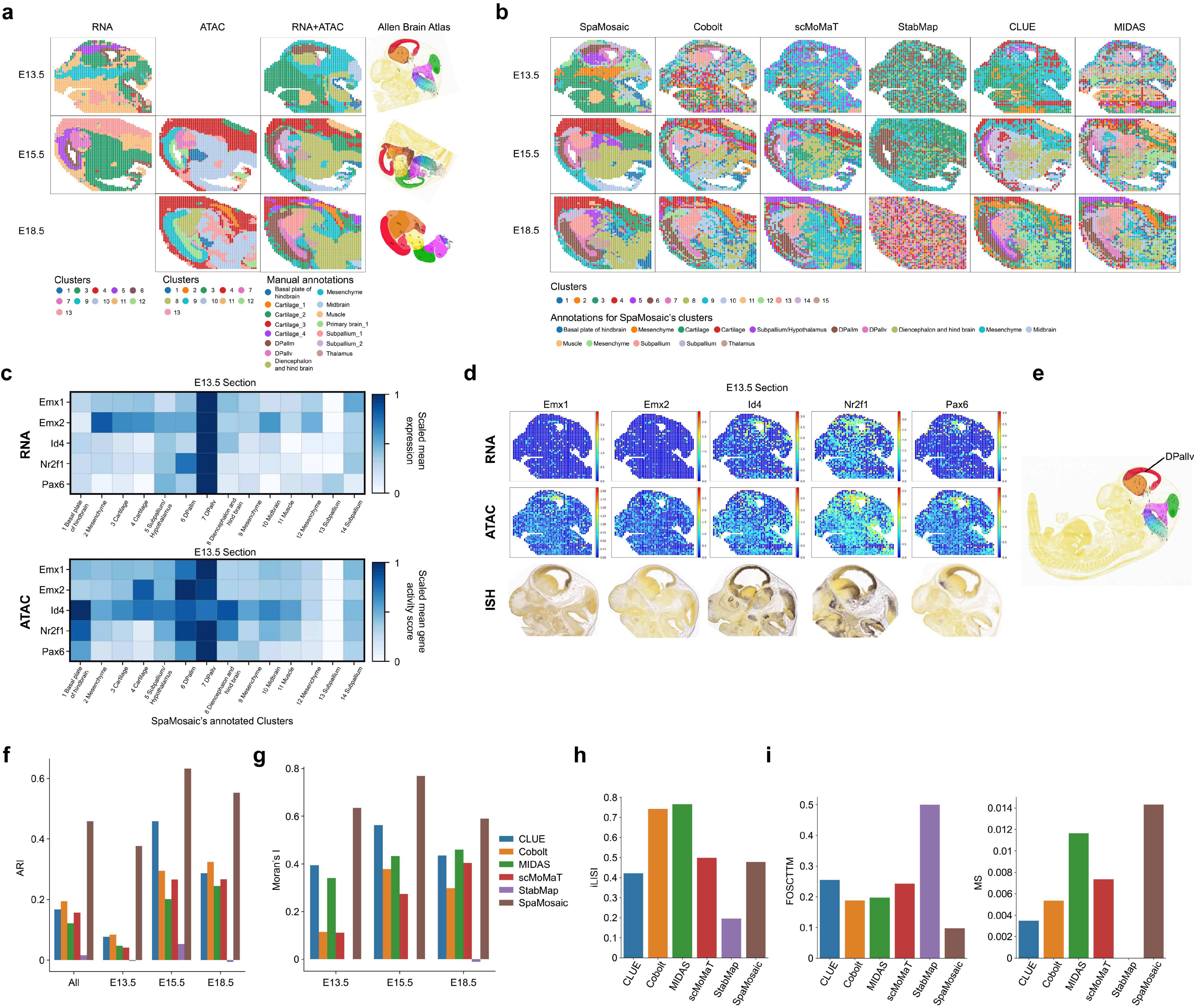
seq. *Nat Biotechnol* **41**, 1405–1409 (2023).

39. Liang, Y. & Tedder, T. F. Identification of a CD20-, FcεRIβ-, and HTm4-Related Gene Family: Sixteen New MS4A Family Members Expressed in Human and Mouse. *Genomics* **72**, 119–127 (2001).
40. Martinez-Valdez, H. *et al.* Human germinal center B cells express the apoptosis-inducing genes Fas, c-myc, P53, and Bax but not the survival gene bcl-2. *J Exp Med* **183**, 971–977 (1996).
41. Russell, A. J. C. *et al.* Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. *Nature* **625**, 101–109 (2024).
42. Wu, K. E., Yost, K. E., Chang, H. Y. & Zou, J. BABEL enables cross-modality translation between multiomic profiles at single-cell resolution. *Proceedings of the National Academy of Sciences* **118**, (2021).
43. Gayoso, A. *et al.* Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat Methods* **18**, 272–282 (2021).
44. Ashuach, T. *et al.* MultiVI: deep generative model for the integration of multimodal data. *Nat Methods* **20**, 1222–1231 (2023).
45. Hu, Y. *et al.* WEDGE: imputation of gene expression values from single-cell RNA-seq datasets using biased matrix decomposition. *Brief Bioinform* (2021) doi:10.1093/bib/bbab085.
46. Wu, S. J. *et al.* Single-cell CUT&amp;Tag analysis of chromatin modifications in differentiation and tumor progression. *Nat Biotechnol* **39**, 819–824 (2021).
47. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* **16**, 284–287 (2012).
48. Yang, M. *et al.* Contrastive learning enables rapid mapping to multimodal single-cell atlas of multimillion scale. *Nat Mach Intell* **4**, 696–709 (2022).
49. He, K., Fan, H., Wu, Y., Xie, S. & Girshick, R. Momentum contrast for unsupervised visual representation learning. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* 9729–9738 (2020).
50. Radford, A. *et al.* Learning transferable visual models from natural language supervision. in *International conference on machine learning* 8748–8763 (2021).
51. Tzeng, E., Hoffman, J., Saenko, K. & Darrell, T. Adversarial discriminative domain adaptation. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 7167–7176 (2017).
52. Long, M., Cao, Y., Wang, J. & Jordan, M. Learning transferable features with deep adaptation networks. in *International conference on machine learning* 97–105 (2015).
53. Waikhom, L. & Patgiri, R. Graph neural networks: Methods, applications, and opportunities. *arXiv preprint arXiv:2108.10733* (2021).
54. Hadsell, R., Rao, D., Rusu, A. A. & Pascanu, R. Embracing Change: Continual Learning in Deep Neural Networks. *Trends Cogn Sci* **24**, 1028–1040 (2020).
55. Weiss, K., Khoshgoftaar, T. M. & Wang, D. A survey of transfer learning. *J Big Data* **3**, 1–40 (2016).
56. He, X. *et al.* Lightgcn: Simplifying and powering graph convolution network for recommendation. in *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval* 639–648 (2020).
57. Xia, C.-R., Cao, Z.-J., Tu, X.-M. & Gao, G. Spatial-linked alignment tool (SLAT) for aligning heterogenous slices. *Nat Commun* **14**, 7236 (2023).
58. Maas, A. L., Hannun, A. Y., Ng, A. Y. & others. Rectifier nonlinearities improve neural network acoustic models. in *Proc. icml* vol. 30 3 (2013).
59. Ioffe, S. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* (2015).
60. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* **15**, 1929–1958 (2014).
61. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. *arXiv preprint arXiv:1903.02428* (2019).
62. Kingma, D. P. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

63. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018).

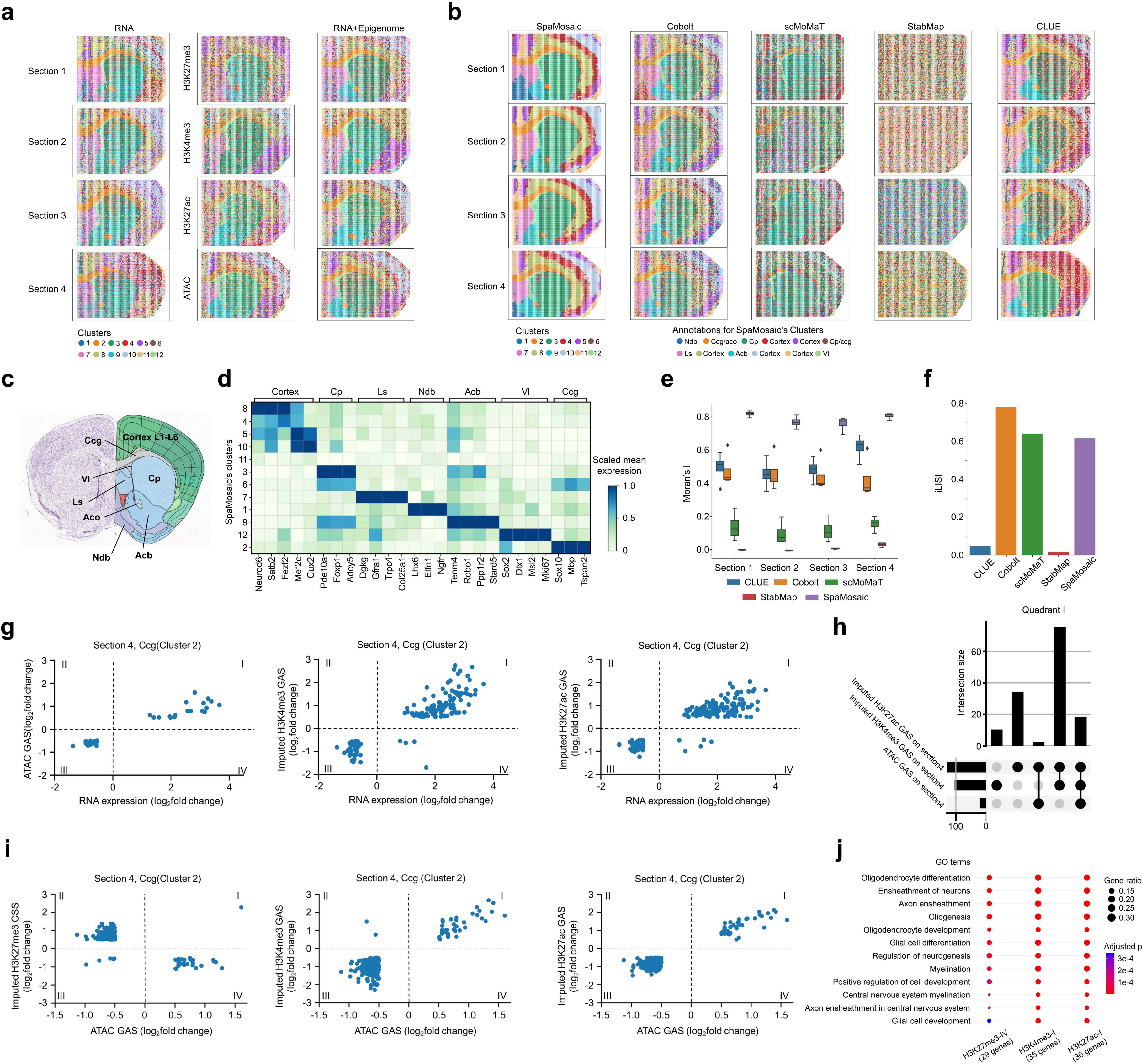

a
Section1
Section2
Section3
Section4
RNA
Protein
ATAC
Histone modifications
......
......
SpaMosaic
Consensus domain identification
Spatial domain enhancement
Imputation
Section1
Section2
Section3
Section4
b
Slices from different sources
Section1
Location
RNA
Protein
Section2
Location
RNA
ATAC
Section3
Location
ATAC
......
Batch correction and build graph for each modality
RNA
Section1
Section2
feature
spot
Expression-based MNN
Spatial neighbors
Spot-spot graph
Protein
Section1
ATAC
Section2
Section3
......
RNA-specific GNN
Protein-specific GNN
ATAC-specific GNN
RNA embedding
Section1
Section2
Protein embedding
Section1
ATAC embedding
Section2
Section3
Training
Inference
Contrastive learning loss
Contrastive learning loss
Reconstruction loss
Decoder
Section1
Section2
Section3
average
average
Clustering
Imputation
Section1
Section2
Section3

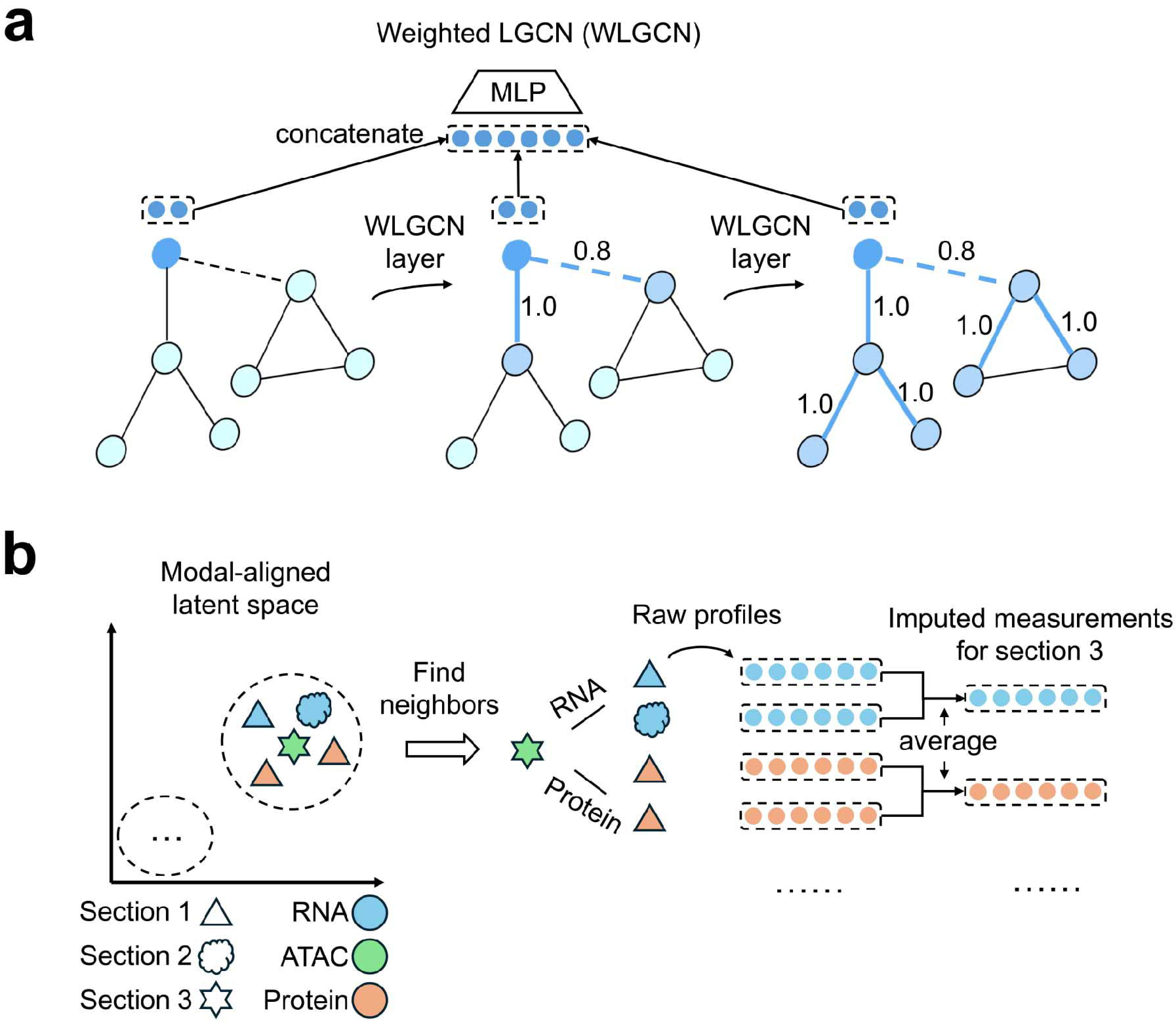

a
Ground truth of spatial factors
Section 1
Section 2
Section 3
Background factor
Factor 1
Factor 2
Factor 3
Factor 4
b
Clustering pattern
RNA + Protein
ARI=0.93
Section 1
RNA
ARI=0.35
Section 1
Protein
ARI=0.46
Section 1
RNA
ARI=0.36
Section 2
Protein
ARI=0.40
Section 3
c
RNA
Section label
Factor label
Section 1
Section 2
0 1 2 3 4
Protein
Section label
Factor label
Section 1
Section 3
0 1 2 3 4
d
Clustering pattern after integration
ARI=0.98
Section 1
ARI=0.91
Section 2
ARI=0.84
Section 3
e
Cluster label
Factor label
0 1 2 3 4
0 1 2 3 4
Section label
Section 1
Section 2
Section 3
f
iLISI
0 0.2 0.4 0.6 0.8
1 2 3 4 5
Dataset
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic (non-spatial)
SpaMosaic
g
ARI
0.2 0.4 0.6 0.8 1.0
All
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic (non-spatial)
SpaMosaic
Raw
h
Moran's I
0 0.2 0.4 0.6 0.8
1 2 3 4 5
Dataset
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic (non-spatial)
SpaMosaic
i
Moran's I
Section 1
Section 2
Section 3
0.6 0.8 1.0
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic (non-spatial)
SpaMosaic
j
Label transfer F1-score
0.0 0.2 0.4 0.6 0.8 1.0
1 2 3 4 5
Dataset
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic (non-spatial)
SpaMosaic
k
Dataset 1 (replicate 1)
Section 2→Section 3
Section 3→Section 2
0 1 2 3 4
0 1 2 3 4
Ground truth
Prediction
Ground truth
Prediction


a
RNA
ATAC
RNA+ATAC
Allen Brain Atlas
E13.5
E15.5
E18.5
Clusters
1 3 4 5 6 7 9 10 11 12 13
Clusters
1 2 3 4 7 8 9 10 11 12 13
Manual annotations
Basal plate of hindbrain
Cartilage_1
Cartilage_2
Cartilage_3
Cartilage_4
DPallm
DPallv
Diencephalon and hind brain
Mesenchyme
Midbrain
Muscle
Primary brain_1
Subpallium_1
Subpallium_2
Thalamus
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
E13.5
E15.5
E18.5
Clusters
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Annotations for SpaMosaic's clusters
Basal plate of hindbrain
Mesenchyme
Cartilage
Cartilage
Subpallium/Hypothalamus
DPallm
DPallv
Diencephalon and hind brain
Mesenchyme
Midbrain
Muscle
Mesenchyme
Subpallium
Subpallium
Thalamus
c
E13.5 Section
RNA
Emx1
Emx2
Id4
Nr2f1
Pax6
Scaled mean expression
1
0
E13.5 Section
ATAC
Emx1
Emx2
Id4
Nr2f1
Pax6
Scaled mean gene activity score
1
0
1 Basal plate of hindbrain
2 Mesenchyme
3 Cartilage
4 Cartilage
5 Subpallium/ Hypothalamus
6 DPallm
7 DPallv
8 Diencephalon and hind brain
9 Mesenchyme
10 Midbrain
11 Muscle
12 Mesenchyme
13 Subpallium
14 Subpallium
SpaMosaic's annotated Clusters
d
E13.5 Section
Emx1
Emx2
Id4
Nr2f1
Pax6
RNA
ATAC
ISH
e
DPallv
f
ARI
0
0.2
0.4
0.6
All
E13.5
E15.5
E18.5
g
Moran's I
0
0.2
0.4
0.6
0.8
E13.5
E15.5
E18.5
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
h
iLISI
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
i
FOSCTTM
0
0.1
0.2
0.3
0.4
0.5
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
MS
0
0.002
0.004
0.006
0.008
0.01
0.012
0.014
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

a
RNA
H3K4me3
RNA+ H3K4me3
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6 7 8
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6 7 8
Annotations for SpaMosaic's clusters
Cortex VI/cortex Acb Cortex Cortex Ccg/aco Cp Ls
c
Ccg
VI
Ls
Aco
Acb
Cortex
Cp
d
Section 1 and Section 2
6 2 3 8 7 5 4 1
Neurod6
Mef2c
Satb2
Cux2
Camk2a
Pde10a
Foxp1
Adcy5
Dgkg
Gfra1
Trpc4
Col25a1
Tenm4
Robo1
Ppp1r2
Stard5
Dlx1
Msi2
Mki67
Sox10
Mbp
Tspan2
RNA
Cortex
Cp
Ls
Acb
Vl
Ccg/aco
Scaled mean expression
0 0.5 1.0
e
Section 3
Satb2 (Cortex)
Sox10 (Ccg)
Pde10a (Cp)
Adcy5 (Cp)
Msi2 (Vl)
Dlx1 (Vl)
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75
0.00 0.25 0.50 0.75 1.00 1.25 1.50
0.0 0.5 1.0 1.5 2.0 2.5
0.0 0.5 1.0 1.5 2.0 2.5
0.0 0.5 1.0 1.5 2.0
0.0 0.5 1.0 1.5 2.0
f
Section 3
SpaMosaic clusters
Cortex, 5
Cortex, 1
Ccg, 6
Cp, 7
TEGLU8
TEGLU3
MOL1
MSN2
g
Moran's I
0 0.2 0.4 0.6 0.8
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
h
iLISI
0.0 0.1 0.2 0.3 0.4 0.5 0.6
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

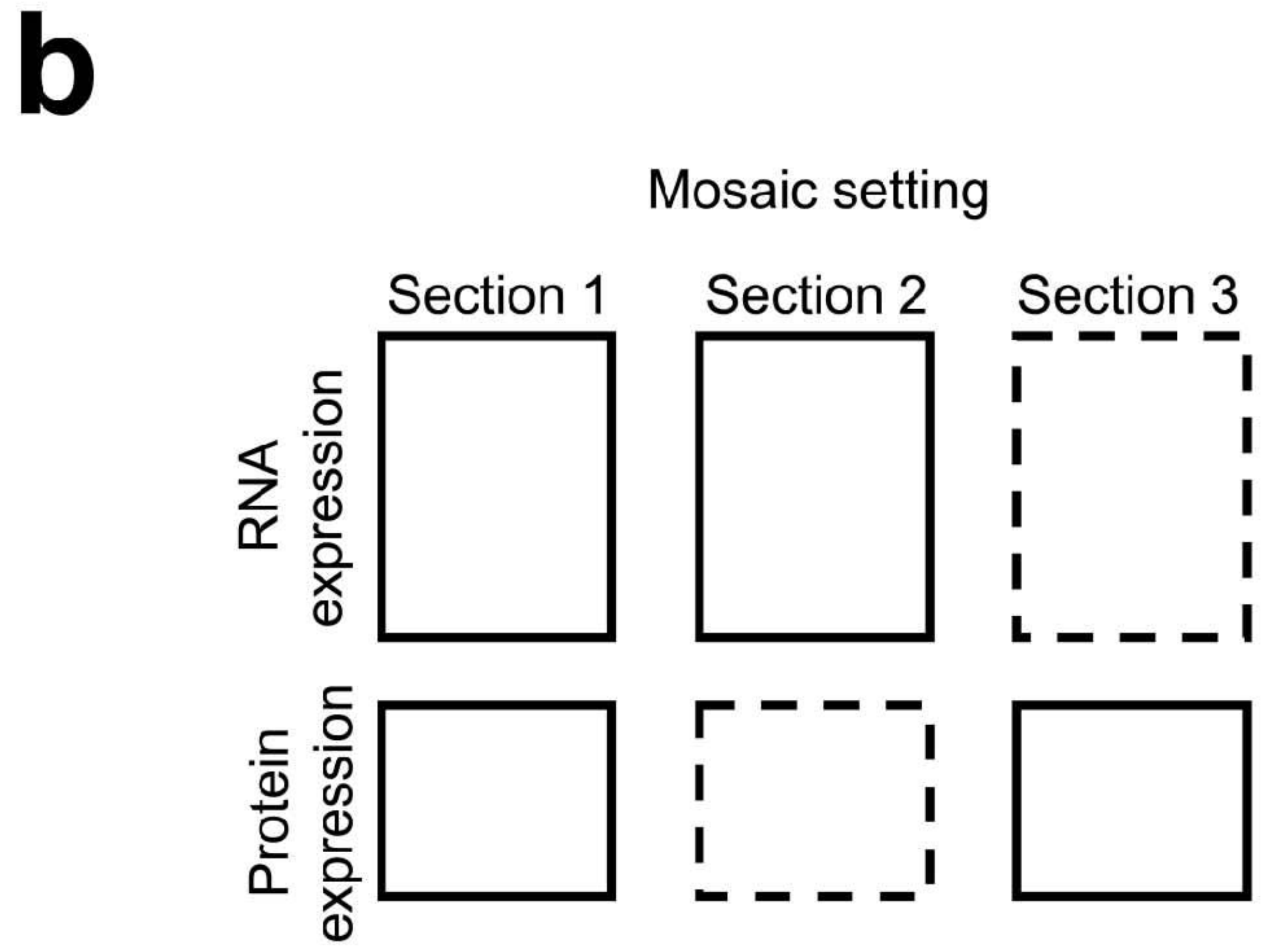
a
RNA
Protein
Manual annotations
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6 7
Connective & epithelial tissue
Germinal center
Lymphoid follicle
Tonsillar parenchyma
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6 7
Annotation of SpaMosaic's Clusters
Follicle mantle zone
Germinal center
T cell enriched zone
Epithelial
Epithelial
Connective tissue
Epithelial
c
RNA (Sections 1 and 2)
Follicle Mantle zone
Germinal center
T cell enriched zone
Epithelial and connective tissue
SpaMosaic's clusters
1 2 3 4 5 6 7
CD19 MS4A1 IGHD CR2 BCL2A1 LMO2 CD3E CD4 CD8A KRT5 CD9 ACTA2 EPCAM
Scaled mean expression
0 0.5 1.0
Protein (Sections 1 and 3)
Follicle Mantle zone
Germinal center
T cell enriched zone
Epithelial and connective tissue
SpaMosaic's clusters
1 2 3 4 5 6 7
CD19 BCL2 CR2 PDCD1 CD8A CD3E CD4 EPCAM VIM ACTA2 KRT5
Scaled mean expression
0 0.5 1.0
d
ARI
0 0.05 0.10 0.15 0.20
All
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
e
Moran's I
0.5 0.6 0.7
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
f
iLISI
0 0.2 0.4 0.6 0.8
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
g
FOSCTTM
0 0.1 0.2
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
MS
0 0.05 0.10 0.15
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

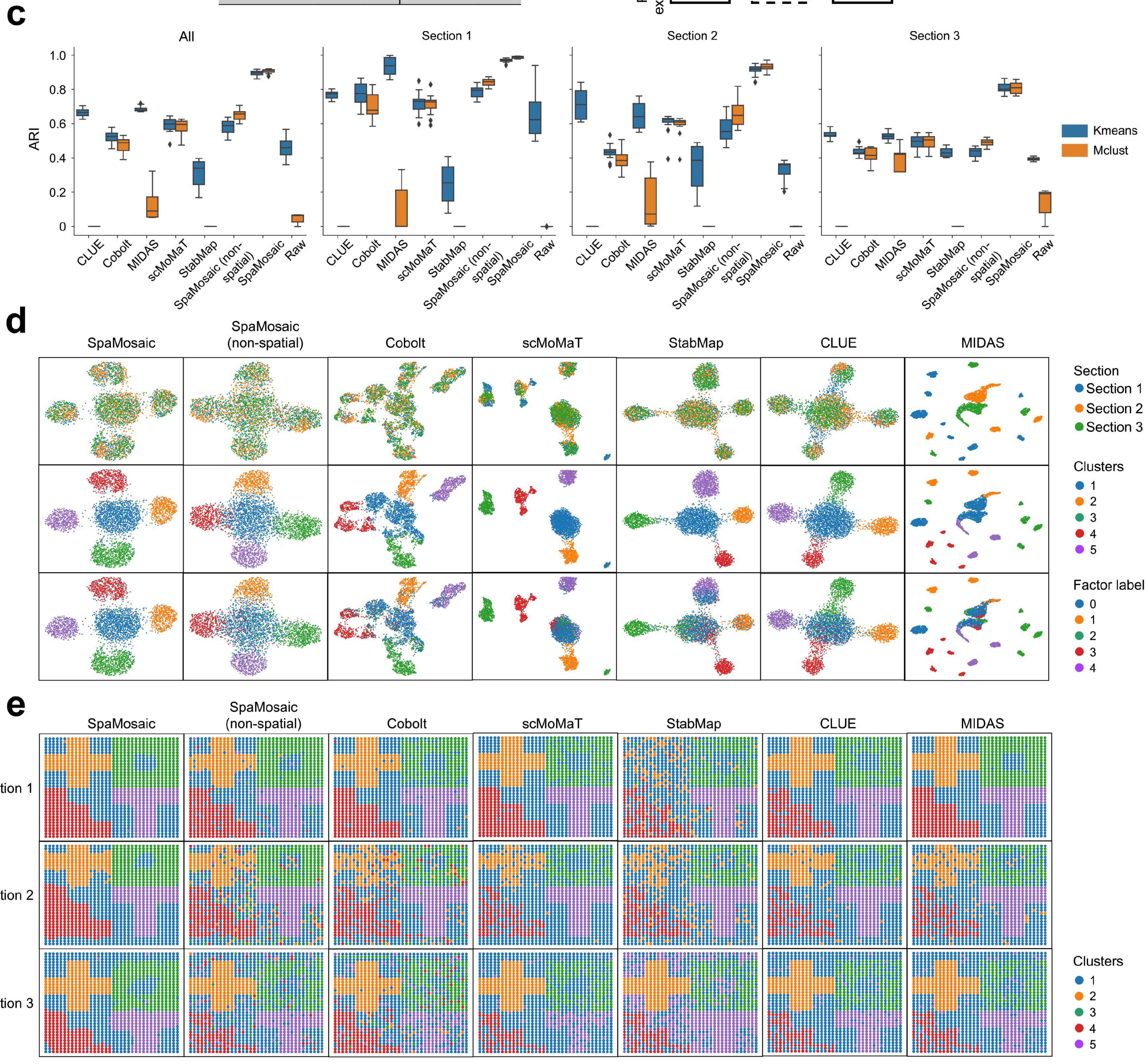

a
RNA
RNA+Epigenome
Section 1
Section 2
Section 3
Section 4
H3K27me3
H3K4me3
H3K27ac
ATAC
Clusters
1 2 3 4 5 6
7 8 9 10 11 12
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
Section 1
Section 2
Section 3
Section 4
Clusters
1 2 3 4 5 6
7 8 9 10 11 12
Annotations for SpaMosaic's Clusters
Ndb Ccg/aco Cp Cortex Cortex Cp/ccg
Ls Cortex Acb Cortex Cortex Vl
c
Ccg
Cortex L1-L6
Vl
Cp
Ls
Aco
Ndb
Acb
d
Cortex
Cp
Ls
Ndb
Acb
Vl
Ccg
SpaMosaic's clusters
8 4 5 10 11 3 6 7 1 9 12 2
Neurod6 Satb2 Fezf2 Mef2c Cux2 Pde10a Foxp1 Adcy5 Dgkg Gfra1 Trpc4 Col25a1 Lhx6 Elfn1 Ngfr Tenm4 Robo1 Ppp1r2 Stard5 Sox2 Dlx1 Msi2 Mki67 Sox10 Mbp Tspan2
Scaled mean expression
1.0
0.5
0
e
Moran's I
0.8
0.6
0.4
0.2
0.0
Section 1
Section 2
Section 3
Section 4
CLUE
Cobolt
scMoMaT
StabMap
SpaMosaic
f
iLISI
0.8
0.6
0.4
0.2
0.0
CLUE
Cobolt
scMoMaT
StabMap
SpaMosaic
g
Section 4, Ccg(Cluster 2)
Section 4, Ccg (Cluster 2)
Section 4, Ccg (Cluster 2)
ATAC GAS($\log_2$fold change)
Imputed H3K4me3 GAS ($\log_2$fold change)
Imputed H3K27ac GAS ($\log_2$fold change)
RNA expression ($\log_2$fold change)
I
II
III
IV
3 2 1 0 -1 -2
-2 0 2 4
h
Quadrant I
Intersection size
60
40
20
0
Imputed H3K27ac GAS on section4
Imputed H3K4me3 GAS on section4
ATAC GAS on section4
100
0
i
Section 4, Ccg(Cluster 2)
Section 4, Ccg (Cluster 2)
Section 4, Ccg (Cluster 2)
Imputed H3K27me3 CSS ($\log_2$fold change)
Imputed H3K4me3 GAS ($\log_2$fold change)
Imputed H3K27ac GAS ($\log_2$fold change)
ATAC GAS ($\log_2$fold change)
I
II
III
IV
3 2 1 0 -1 -2 -3
-1.5 -1.0 -0.5 0 0.5 1.0 1.5
j
GO terms
Oligodendrocyte differentiation
Ensheathment of neurons
Axon ensheathment
Gliogenesis
Oligodendrocyte development
Glial cell differentiation
Regulation of neurogenesis
Myelination
Positive regulation of cell development
Central nervous system myelination
Axon ensheathment in central nervous system
Glial cell development
H3K27me3-IV (29 genes)
H3K4me3-I (35 genes)
H3K27ac-I (36 genes)
Gene ratio
0.15
0.20
0.25
0.30
Adjusted p
3e-4
2e-4
1e-4


a
Weighted LGCN (WLGCN)
MLP
concatenate
WLGCN layer
0.8
1.0
WLGCN layer
0.8
1.0
1.0
1.0
1.0
1.0
b
Modal-aligned latent space
...
Find neighbors
RNA
Protein
Raw profiles
Imputed measurements for section 3
average
......
......
Section 1
Section 2
Section 3
RNA
ATAC
Protein

a

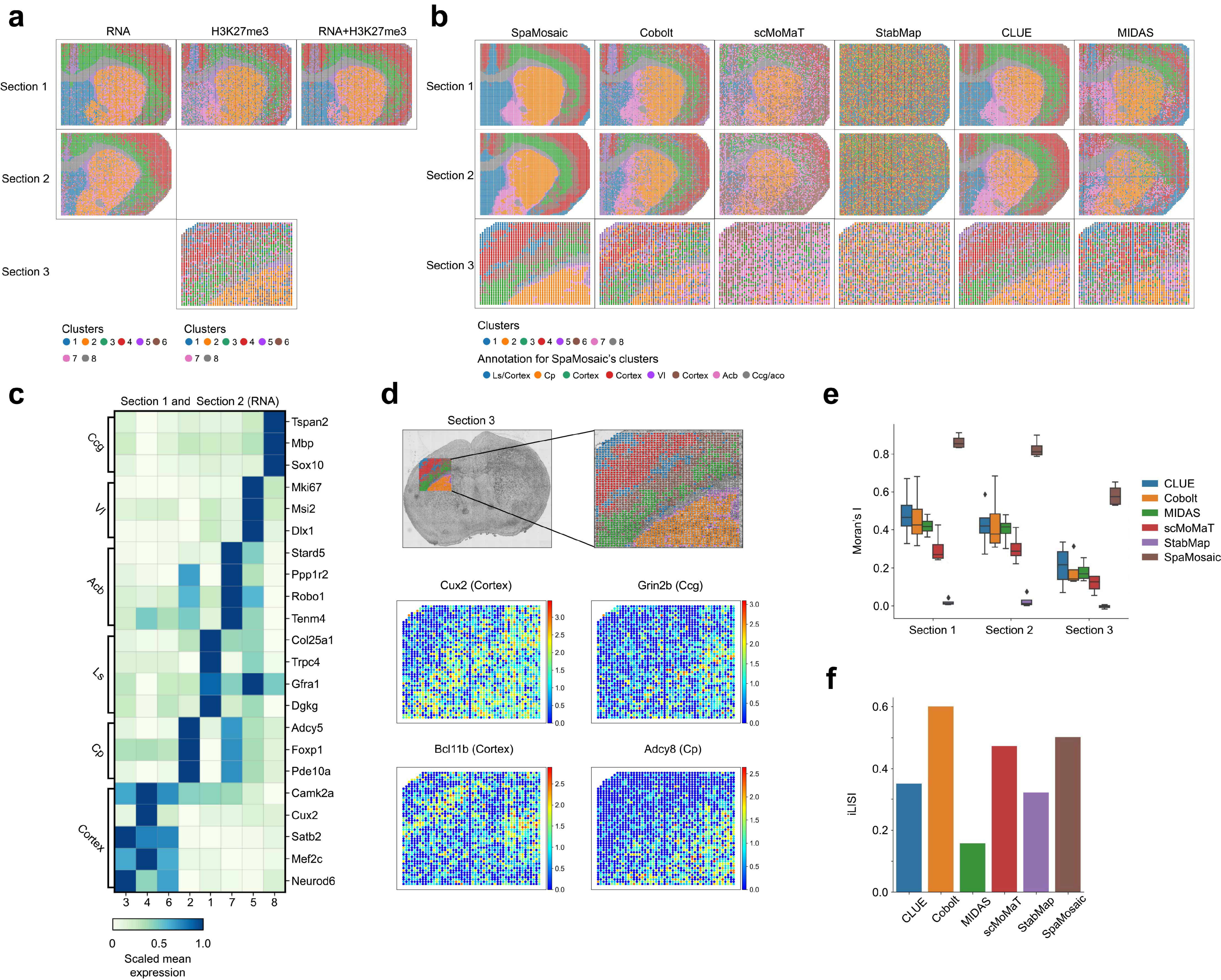
	RNA			Protein	
	ZINB distribution			NB distribution	
id	b^{rna}	s^{rna}	p^{rna}	b^{pro}	s^{pro}
1	4	1/12	0.1	2.1	1/20
2	4	1/10	0.2	2.2	1/20
3	4	1/7	0.3	2.3	1/20
4	4	1/5	0.4	2.4	1/20
5	4	1/4	0.5	2.5	1/20


b
Mosaic setting
Section 1
Section 2
Section 3
RNA expression
Protein expression



a
Section
RNA
ATAC
● E13.5 ● E15.5 ● E18.5
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
● E13.5
● E15.5
● E18.5
Clusters
● 1 ● 9
● 2 ● 10
● 3 ● 11
● 4 ● 12
● 5 ● 13
● 6 ● 14
● 7 ● 15
● 8
Manual annotations
● Basal plate of hindbrain ● Mesenchyme
● Cartilage_1 ● Midbrain
● Cartilage_2 ● Muscle
● Cartilage_3 ● Primary brain_1
● Cartilage_4 ● Subpallium_1
● DPallm ● Subpallium_2
● DPallv ● Thalamus
● Diencephalon and hind brain
c
RNA
ATAC
E15.5
E18.5-RNA
E18.5-ATAC
Clusters
● 1 ● 2 ● 3 ● 4 ● 5
● 6 ● 7 ● 8 ● 9 ● 10
● 11 ● 12 ● 13
Clusters
● 1 ● 2 ● 3 ● 4 ● 7
● 8 ● 9 ● 10 ● 11 ● 12
● 13
d
Section
● E15.5-RNA
● E18.5-RNA
● E15.5-ATAC
● E18.5-ATAC

a
Section
RNA
H3K4me3
Section 1
Section 2
Section 3
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6
7
8


a
RNA
H3K27me3
RNA+H3K27me3
Section 1
Section 2
Section 3
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6 ● 7 ● 8
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6 ● 7 ● 8
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6 ● 7 ● 8
Annotation for SpaMosaic's clusters
● Ls/Cortex ● Cp ● Cortex ● Cortex ● Vl ● Cortex ● Acb ● Ccg/aco
c
Section 1 and Section 2 (RNA)
Ccg
Vl
Acb
Ls
Cp
Cortex
Tspan2
Mbp
Sox10
Mki67
Msi2
Dlx1
Stard5
Ppp1r2
Robo1
Tenm4
Col25a1
Trpc4
Gfra1
Dgkg
Adcy5
Foxp1
Pde10a
Camk2a
Cux2
Satb2
Mef2c
Neurod6
3 4 6 2 1 7 5 8
0 0.5 1.0
Scaled mean expression
d
Section 3
Cux2 (Cortex)
Grin2b (Ccg)
Bcl11b (Cortex)
Adcy8 (Cp)
0.0 0.5 1.0 1.5 2.0 2.5 3.0
e
Moran's I
0.0 0.2 0.4 0.6 0.8
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
f
iLISI
0.0 0.2 0.4 0.6
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

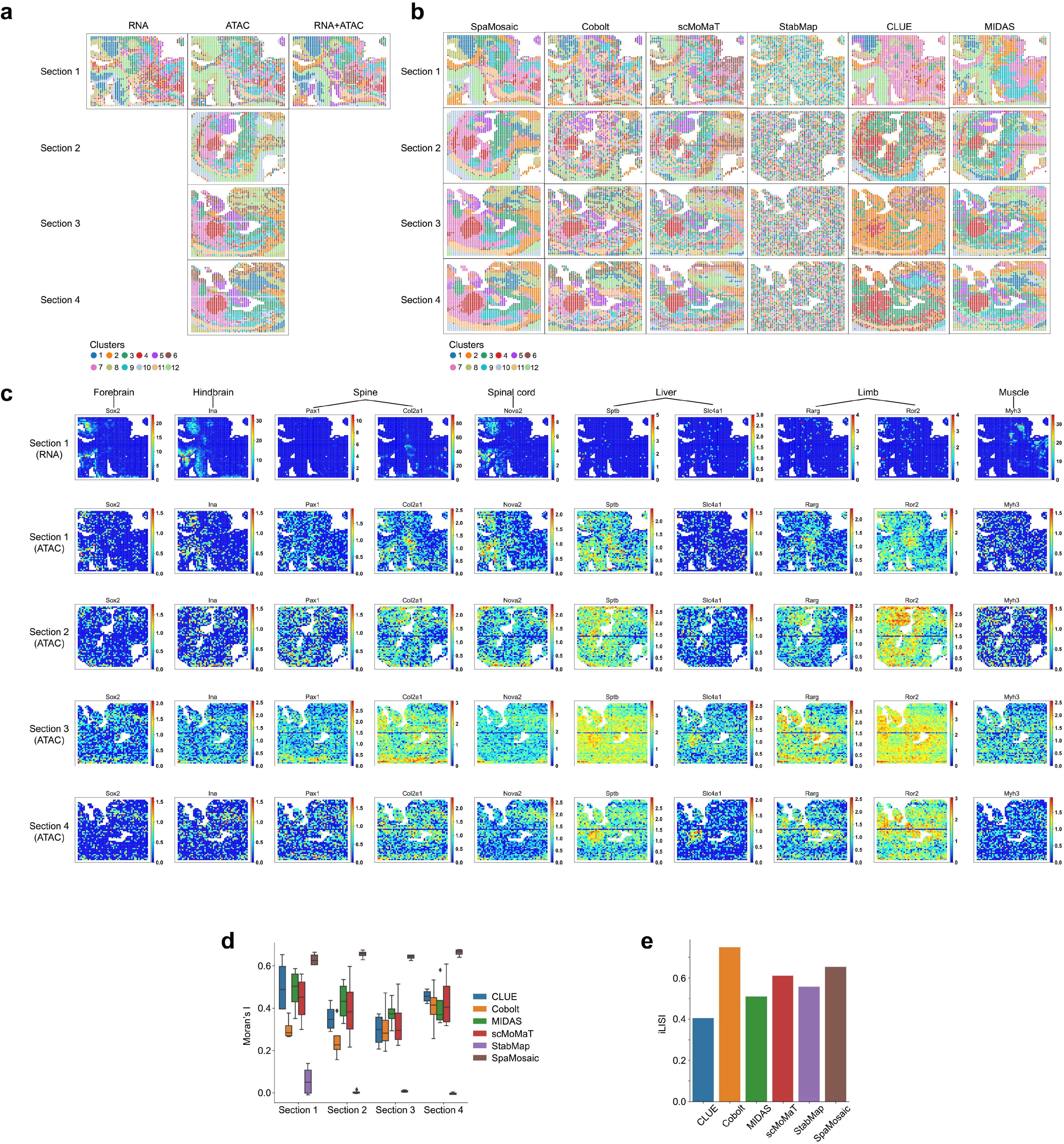
a
Section
RNA
H3K27me3
Section 1
Section 2
Section 3
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6
7
8

a
RNA
ATAC
RNA+ATAC
Section 1
Section 2
Section 3
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6
Annotation for SpaMosaic's clusters
● Cortex ● Cp ● VI ● Cortex ● Ccg ● Ls
c
Section 1 and Section 2
RNA
Syt1
Mef2c
Pgm2l1
Map1b
C1ql3
Adcy1
Nrgn
Napb
Pde10a
Gng7
Foxp1
Rgs9
Bcl11b
Smpd3
Atp2b1
Arpp21
Slc1a2
Atp1a2
Camk2a
Sparcl1
Malat1
Chd3
Slc7a11
Igf2
Meg3
Snhg11
Atp2b4
Dgkg
Gria1
Ttc3
Ahi1
Cacna2d2
Zbtb20
Ccnd2
Sox4
Msi2
Nnat
Gm973
Syne2
Nfib
Mbp
Plp1
Mobp
Mal
Mag
Qk
Cnp
Scd2
3
2
1
0
−1
−2
−3
1
2
4
6
3
5
d
Section 3
Adcy5 (Cp)
Pde10a (Cp)
Msi2 (VI)
Mef2c (Cortex)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
1
2
3
4
e
Moran's I
0.0
0.2
0.4
0.6
0.8
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
f
iLISI
0
0.2
0.4
0.6
0.8
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
g
RNA
ATAC
Section 1
Section2-RNA
Section2-ATAC
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6
h
● Section1-RNA
● Section2-RNA
● Section1-ATAC
● Section2-ATAC
i
FOSCTTM
0
0.1
0.2
0.3
0.4
0.5
MS
0
0.02
0.04
0.06
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

a
Section
RNA
ATAC
Section 1
Section 2
Section 3
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6

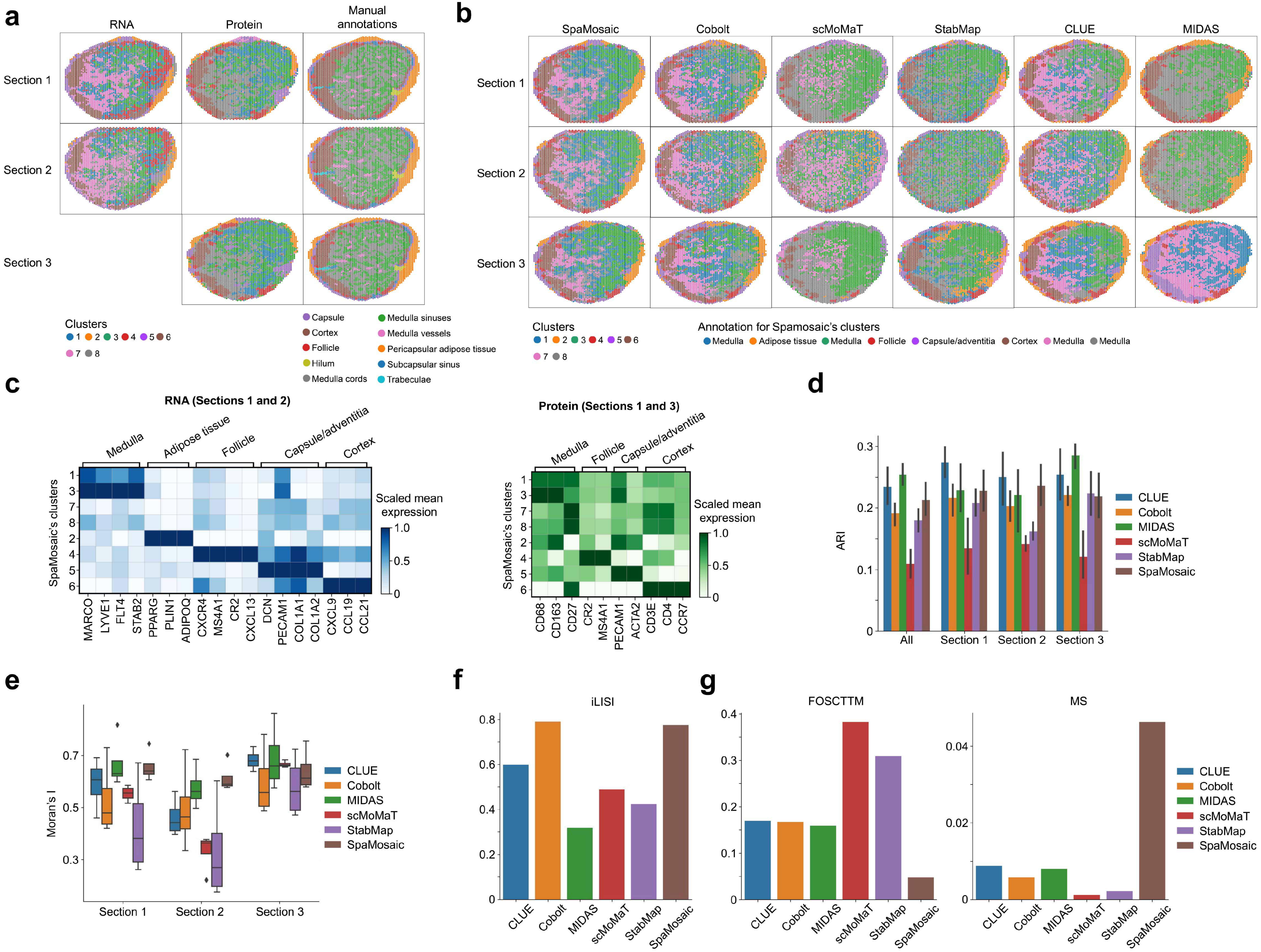
a
RNA
H3K27ac
RNA+H3K27ac
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
1 2 3 4 5 6
Annotation for SpaMosaic's clusters
Ls/cortex Cortex Cortex Cp Ccg Vl
c
Section 1 and Section 2
RNA
Syt1
Pgm2l1
Map1b
Atp1b1
Adcy1
Slc24a2
Mef2c
Napb
Mef2c
C1ql3
Syt1
Grin2b
Nrxn1
Camk2a
Pgm2l1
Tenm2
Mbp
Plp1
Mobp
Mal
Mag
Cldn11
Cnp
Tspan2
Pde10a
Gng7
Foxp1
Bcl11b
Rgs9
Smpd3
Atp2b1
Arpp21
Meg3
Atp1a2
Slc1a2
Dgkg
Camk2a
Gm26924
Sparcl1
Zic1
Zbtb20
Ccnd2
Nnat
Sox4
Msi2
Syne2
Ptma
Btg1
3 2 1 0 −1 −2 −3
2 3 5 4 1 6
d
Section 3
Mef2c (Cortex)
Satb2 (Cortex)
Tspan2 (Ccg)
Msi2 (Vl)
Pde10a (Cp)
e
Moran's I
0.0 0.2 0.4 0.6 0.8
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
f
iLISI
0 0.2 0.4 0.6 0.8
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
g
RNA
H3K27ac
Section 1
Section 2-RNA
Section 3-H3K27ac
Clusters
1 2 3 4 5 6
Clusters
1 2 3 4 5 6
h
Section 1-RNA
Section 2-RNA
Section 1-H3K27ac
Section 2-H3K27ac
i
FOSCTTM
0 0.1 0.2 0.3 0.4 0.5
MS
0 0.01 0.02 0.03 0.04
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

a
Section
RNA
H3K27ac
Section 1
Section 2
Section 3
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section1
Section 2
Section 3
Clusters
1
2
3
4
5
6


a
RNA
ATAC
RNA+ATAC
Section 1
Section 2
Section 3
Section 4
Clusters
1 2 3 4 5 6 7 8 9 10 11 12
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Section 4
Clusters
1 2 3 4 5 6 7 8 9 10 11 12
c
Forebrain
Hindbrain
Spine
Spinal cord
Liver
Limb
Muscle
Sox2
Ina
Pax1
Col2a1
Nova2
Sptb
Slc4a1
Rarg
Ror2
Myh3
Section 1 (RNA)
Section 1 (ATAC)
Section 2 (ATAC)
Section 3 (ATAC)
Section 4 (ATAC)
d
Moran's I
0.0
0.2
0.4
0.6
Section 1
Section 2
Section 3
Section 4
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
e
iLISI
0.0
0.2
0.4
0.6
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

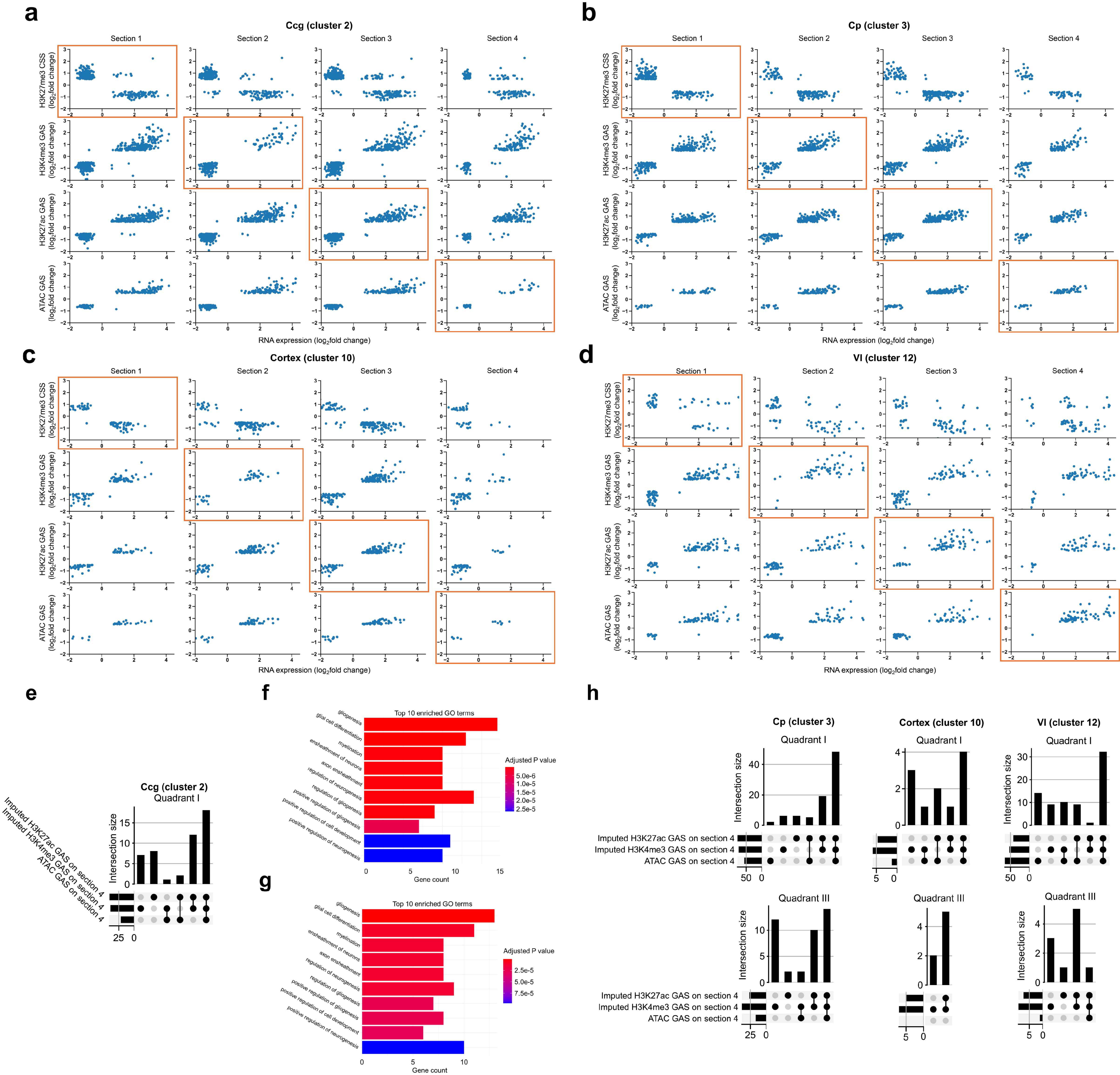
a
Section
ATAC
Section 1
Section 2
Section 3
Section 4
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section 1
Section 2
Section 3
Section 4
Clusters
1
2
3
4
5
6
7
8
9
10
11
12

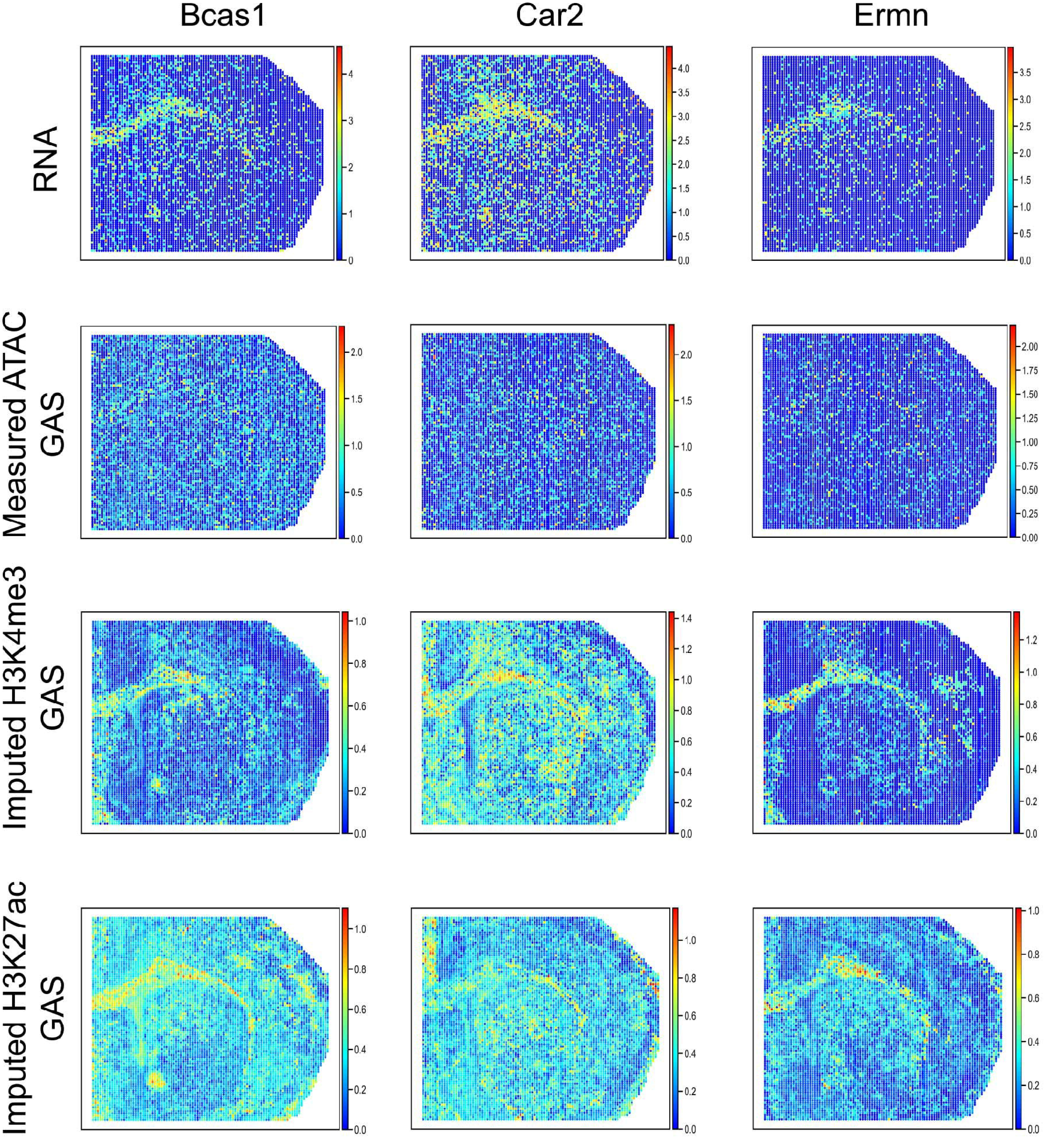
a
Section
RNA
Protein
● Section 1 ● Section 2 ● Section 3
c
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
● Section 1
● Section 2
● Section 3
Clusters
● 1 ● 6
● 2 ● 7
● 3
● 4
● 5
Manual annotation
● Connective & epithelial tissue
● Germinal center
● Lymphoid follicle
● Tonsillar parenchyma
b
Section 1
Section 2
Section 3
d
RNA
Protein
Section 1
Section 2-RNA
Section 2-Protein
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6 ● 7
Clusters
● 1 ● 2 ● 3 ● 4 ● 5 ● 6 ● 7
e
UMAP2
UMAP1
● Section 1-RNA
● Section 2-RNA
UMAP2
UMAP1
● Section 1-Protein
● Section 2-Protein


a
RNA
Protein
Manual annotations
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6
7
8
Capsule
Cortex
Follicle
Hilum
Medulla cords
Medulla sinuses
Medulla vessels
Pericapsular adipose tissue
Subcapsular sinus
Trabeculae
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6
7
8
Annotation for Spamosaic's clusters
Medulla
Adipose tissue
Medulla
Follicle
Capsule/adventitia
Cortex
Medulla
Medulla
c
RNA (Sections 1 and 2)
Medulla
Adipose tissue
Follicle
Capsule/adventitia
Cortex
SpaMosaic's clusters
1
3
7
8
2
4
5
6
MARCO
LYVE1
FLT4
STAB2
PPARG
PLIN1
ADIPOQ
CXCR4
MS4A1
CR2
CXCL13
DCN
PECAM1
COL1A1
COL1A2
CXCL9
CCL19
CCL21
Scaled mean expression
1.0
0.5
0
Protein (Sections 1 and 3)
Medulla
Follicle
Capsule/adventitia
Cortex
SpaMosaic's clusters
1
3
7
8
2
4
5
6
CD68
CD163
CD27
CR2
MS4A1
PECAM1
ACTA2
CD3E
CD4
CCR7
Scaled mean expression
1.0
0.5
0
d
ARI
0.3
0.2
0.1
0
All
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
e
Moran's I
0.7
0.5
0.3
Section 1
Section 2
Section 3
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
f
iLISI
0.8
0.6
0.4
0.2
0
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
g
FOSCTTM
0.4
0.3
0.2
0.1
0
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
MS
0.04
0.02
0
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic
CLUE
Cobolt
MIDAS
scMoMaT
StabMap
SpaMosaic

a
Section
RNA
Protein
Section 1
Section 2
Section 3
c
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
MIDAS
Section
Section 1
Section 2
Section 3
Clusters
1
2
3
4
5
6
7
8
Manual annotations
Capsule
Cortex
Follicle
Hilum
Medulla cords
Medulla sinuses
Medulla vessels
Pericapsular adipose tissue
Subcapsular sinus
Trabeculae
b
Section 1
Section 2
Section 3
d
RNA
Protein
Section 1
Section2-RNA
Section2-Protein
Clusters
1
2
3
4
5
6
7
8
Clusters
1
2
3
4
5
6
7
8
e
UMAP2
UMAP1
Section 1-RNA
Section 2-RNA
UMAP2
UMAP1
Section 1-Protein
Section 2-Protein

a
Section
RNA
Section 1
Section 2
Section 3
Section 4
b
SpaMosaic
Cobolt
scMoMaT
StabMap
CLUE
Section
Section 1
Section 2
Section 3
Section 4
Clusters
1
2
3
4
5
6
7
8
9
10
11
12

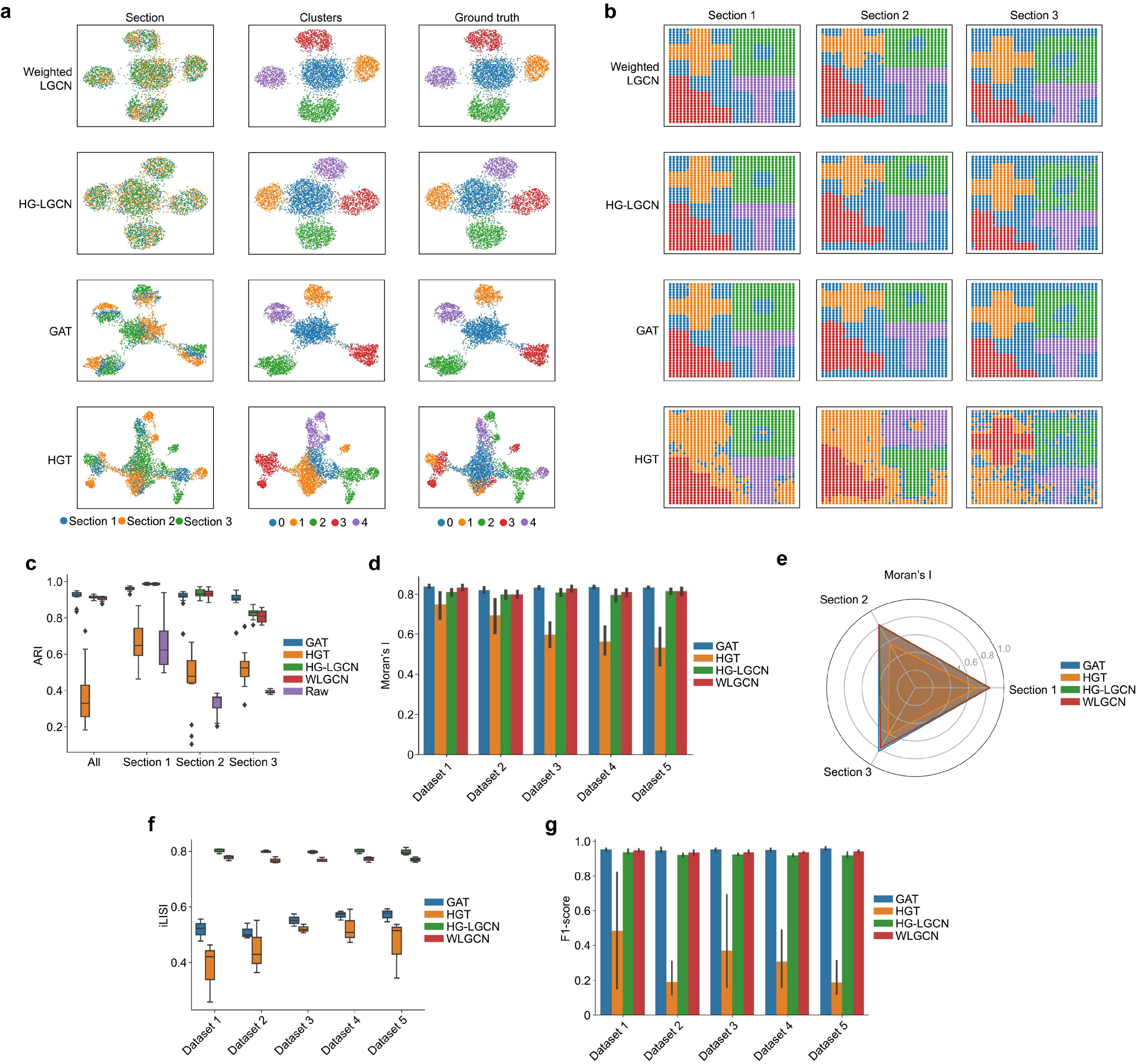

a
Ccg (cluster 2)
b
Cp (cluster 3)
c
Cortex (cluster 10)
d
VI (cluster 12)
Section 1
Section 2
Section 3
Section 4
H3K27me3 CSS (log$_2$fold change)
H3K4me3 GAS (log$_2$fold change)
H3K27ac GAS (log$_2$fold change)
ATAC GAS (log$_2$fold change)
RNA expression (log$_2$fold change)
e
Ccg (cluster 2)
Quadrant I
Intersection size
Imputed H3K27ac GAS on section 4
Imputed H3K4me3 GAS on section 4
ATAC GAS on section 4
f
Top 10 enriched GO terms
gliogenesis
glial cell differentiation
myelination
ensheathment of neurons
axon ensheathment
regulation of neurogenesis
regulation of gliogenesis
positive regulation of gliogenesis
positive regulation of cell development
positive regulation of neurogenesis
Adjusted P value
5.0e-6
1.0e-5
1.5e-5
2.0e-5
2.5e-5
Gene count
g
Top 10 enriched GO terms
gliogenesis
glial cell differentiation
myelination
ensheathment of neurons
axon ensheathment
regulation of neurogenesis
regulation of gliogenesis
positive regulation of gliogenesis
positive regulation of cell development
positive regulation of neurogenesis
Adjusted P value
2.5e-5
5.0e-5
7.5e-5
Gene count
h
Cp (cluster 3)
Cortex (cluster 10)
VI (cluster 12)
Quadrant I
Quadrant III
Intersection size
Imputed H3K27ac GAS on section 4
Imputed H3K4me3 GAS on section 4
ATAC GAS on section 4


Bcas1
Car2
Ermn
RNA
Measured ATAC GAS
Imputed H3K4me3 GAS
Imputed H3K27ac GAS
0
1
2
3
4
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.5
1.0
1.5
2.0
0.0
0.5
1.0
1.5
2.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0

a
Section 4, Cp (cluster 3)
Imputed H3K27me3 CSS (log2fold change)
Imputed H3K4me3 GAS (log2fold change)
Imputed H3K27ac GAS (log2fold change)
ATAC GAS (log2fold change)
I
II
III
IV
b
GO enrichment
Locomotory behavior
Regulation of amine transport
Amine transport
Regulation of catecholamine secretion
Catecholamine secretion
Regulation of dopamine secretion
Dopamine secretion
Response to cocaine
Regulation of nervous system process
Gamma-aminobutyric acid transport
Catecholamine transport
Regulation of neuron differentiation
Positive regulation of ion transport
Adenylate cyclase-modulating G protein-coupled receptor signaling pathway
Response to dopamine
Cellular response to dopamine
Cellular response to monoamine stimulus
Cellular response to catecholamine stimulus
H3K27me3-IV (133 genes)
H3K4me3-I (177 genes)
H3K27ac-I (238 genes)
GeneRatio
0.02
0.04
0.06
0.08
0.10
p.adjust
0.002
0.001
c
Section 4, Cortex (cluster 10)
Imputed H3K27me3 CSS (log2fold change)
Imputed H3K4me3 GAS (log2fold change)
Imputed H3K27ac GAS (log2fold change)
ATAC GAS (log2fold change)
d
GO enrichment
Telencephalon regionalization
Amacrine cell differentiation
Forebrain regionalization
Corpus callosum development
Telencephalon development
Forebrain development
Pallium development
Pattern specification process
Axonogenesis
Cerebral cortex development
regionalization
H3K27me3-IV (27 genes)
H3K4me3-I (52 genes)
H3K27ac-I (55 genes)
GeneRatio
0.075
0.100
0.125
0.150
p.adjust
0.04
0.03
0.02
0.01

a
Mean squared error
P=8e-7
P=2e-10
P=1e-7
P=2e-9
0.15
0.10
0.05
0
H3K27me3 (Section 1)
H3K4me3 (Section 2)
H3K27ac (Section 3)
ATAC (Section 4)
Pearson correlation coefficient
P=1e-7
P=6e-11
P=6e-9
P=7e-8
0.6
0.4
0.2
0
Single-modality
Four modalities
b
H3K27me3 (section 1)
H3K4me3 (section 2)
H3K27ac (section 3)
ATAC (section 4)
n_genes_by_counts
total_counts
pct_counts_mt
value

a
Mouse postnatal brain (RNA→ATAC)
Mouse embryonic brain (RNA → ATAC)
0
0.2
0.4
0.6
0.8
AUROC
Spot-spot PCC
Peak-peak PCC
Spot-spot CMD
Peak-peak CMD
MIDAS
MultiVI
BABEL
SpaMosaic
b
KNN-smoothed ground truth
Mouse postnatal brain (RNA → ATAC)
Mouse embryonic brain (RNA → ATAC)
0
0.2
0.4
0.6
0.8
1.0
AUROC
Spot-spot PCC
Peak-peak PCC
Spot-spot CMD
Peak-peak CMD
MIDAS
MultiVI
BABEL
SpaMosaic
c
Lymph node (RNA → ADT)
Tonsil (RNA → ADT)
0
0.2
0.4
0.6
0.8
1.0
Spot-spot PCC
Protein-protein PCC
Spot-spot CMD
Protein-protein CMD
MIDAS
TotalVI
BABEL
SpaMosaic


a
Section
Clusters
Ground truth
Weighted LGCN
HG-LGCN
GAT
HGT
Section 1 Section 2 Section 3
0 1 2 3 4
0 1 2 3 4
b
Section 1
Section 2
Section 3
Weighted LGCN
HG-LGCN
GAT
HGT
c
ARI
1.0
0.8
0.6
0.4
0.2
All
Section 1
Section 2
Section 3
GAT
HGT
HG-LGCN
WLGCN
Raw
d
Moran's I
0.8
0.6
0.4
0.2
0
Dataset 1
Dataset 2
Dataset 3
Dataset 4
Dataset 5
GAT
HGT
HG-LGCN
WLGCN
e
Moran's I
Section 2
Section 1
Section 3
0.6
0.8
1.0
GAT
HGT
HG-LGCN
WLGCN
f
iLISI
0.8
0.6
0.4
Dataset 1
Dataset 2
Dataset 3
Dataset 4
Dataset 5
GAT
HGT
HG-LGCN
WLGCN
g
F1-score
1.0
0.8
0.6
0.4
0.2
0
Dataset 1
Dataset 2
Dataset 3
Dataset 4
Dataset 5
GAT
HGT
HG-LGCN
WLGCN