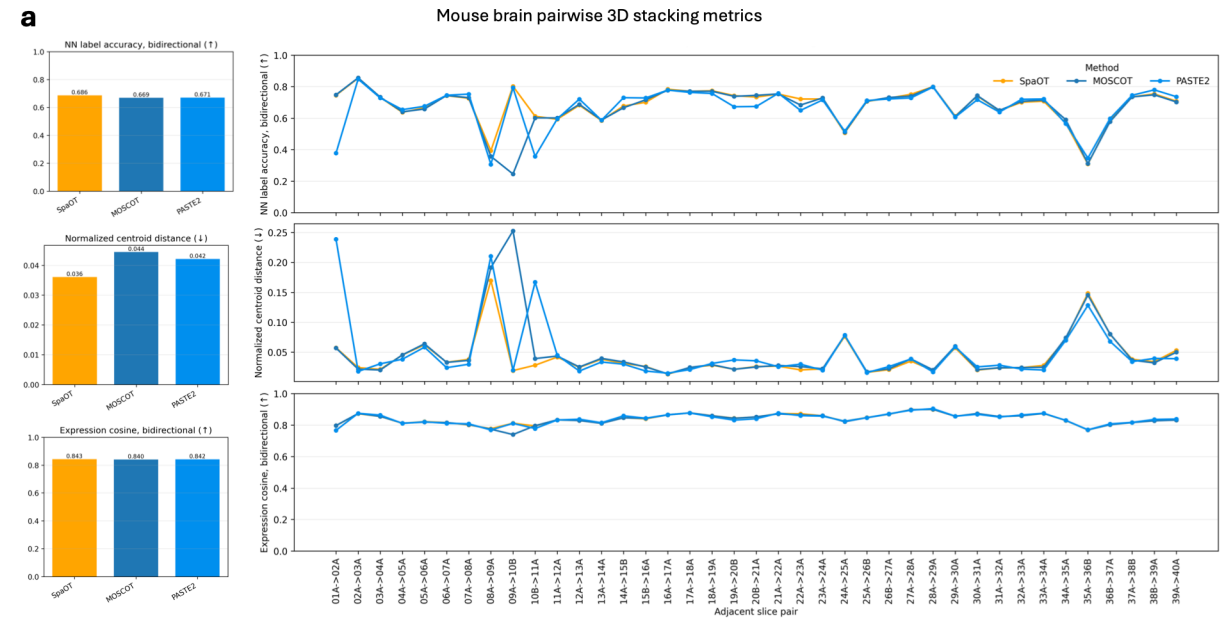
**a**

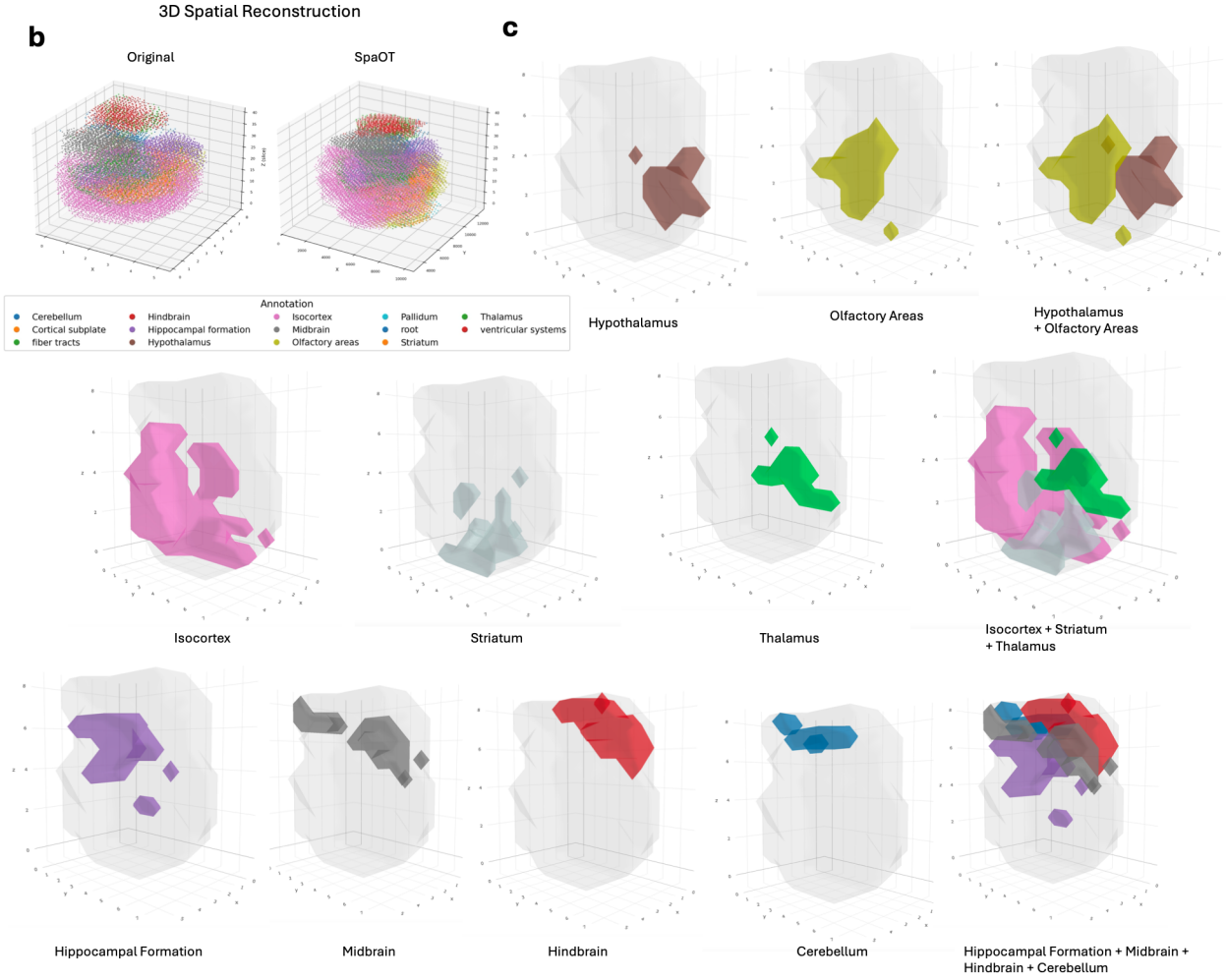
| | Examples in spatial omics | Molecular Profile | Spatial Geometry | KL-divergence | TV-divergence |
|---|---|---|---|---|---|
| | | $\int C(x,y)\,d\gamma(x,y)$ | $\iint L(d_X^r, d_Y^r)\,d\gamma(x,y)d\gamma(x',y')$ | $D_{\phi_1}(\gamma_1^{\otimes 2} \parallel \mu^{\otimes 2}) + D_{\phi_2}(\gamma_2^{\otimes 2} \parallel \nu^{\otimes 2})$ | $\parallel \mu^{\otimes 2} - \gamma_1^{\otimes 2} \parallel_{TV} + \parallel \nu^{\otimes 2} - \gamma_2^{\otimes 2} \parallel_{TV}$ |
| Classic Optimal transport | moscot.time[16] | ✅ | | | |
| Gromov Wasserstein | novoSpaRc* | | ✅ | | |
| Fused Gromov Wasserstein | N/A | ✅ | ✅ | | |
| Fused Gromov Unbalanced Wasserstein | moscot.spatio-temporal[16] | ✅ | ✅ | ✅ | |
| Mass-Constrain Fused Partial Gromov Wasserstein | PASTE2[17] | ✅ | ✅ | | |
| Fused Partial Gromov Wasserstein | SpaOT | ✅ | ✅ | | ✅ |

**b**

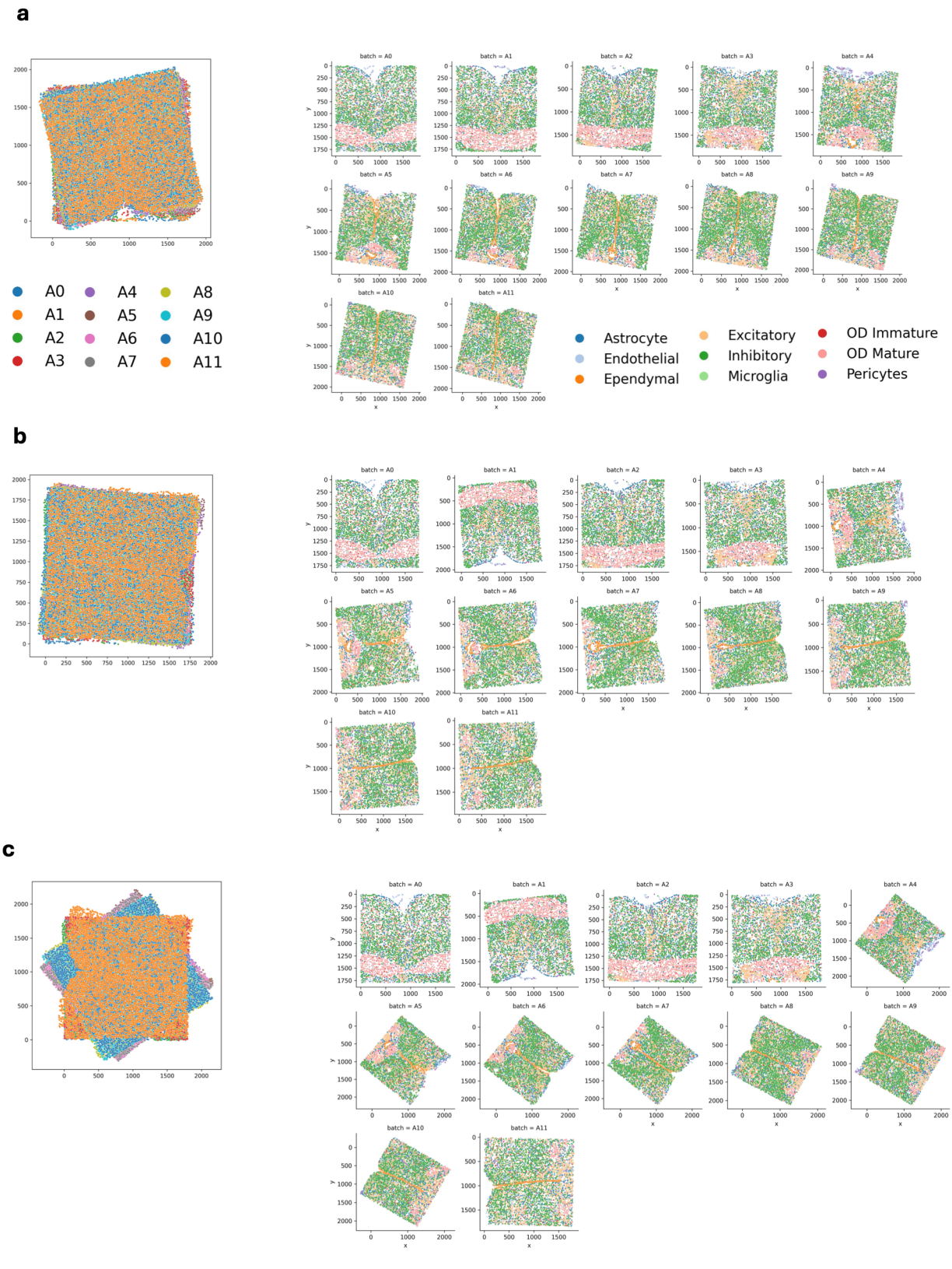
| Four properties of a **metric** | Formulation | FUGW | FPGW(SpaOT) |
|---|---|---|---|
| Non-negativity | $d(x,y) \geq 0$ | ✅ | ✅ |
| Identity of indiscernibles | $d(x,y) \geq 0 \leftrightarrow x \cong y$ | Not established | ✅ |
| Symmetry | $d(x,y) = d(y,x)$ | ✅ | ✅ |
| Triangle inequality | $d(x,z) \leq d(x,y) + d(y,z)$ | Not established | Relaxed ✅, when $q = 1$ |

**Supplementary Fig. 1 | Ablation analysis and metric comparison of the proposed SpaOT framework. a,** Comparison of optimal transport (OT) formulations based on the incorporated information and regularization of the transport objective. Classical OT uses only molecular profiles (e.g., function ***moscot.time***), Gromov Wasserstein (GW) uses only spatial geometry (e.g., **novoSpaRc**), and Fused GW jointly integrates molecular profiles and spatial geometry. Fused Unbalanced GW (FUGW) further introduces Kullback-Leibler (KL) divergence relaxation (e.g., function ***moscot.spatiotemporal***). Mass-constrained Fused Partial GW enforces a hard mass constraint without additional divergence regularization (e.g., **PASTE2**). In contrast, the proposed FPGW replaces the KL-divergence with Total Variance (TV)-divergence regularization while preserving both molecular and spatial information (e.g., **SpaOT**). The mathematical formulations of the molecular profile, spatial geometry, KL-divergence, and TV-divergence terms are summarized in the table, and their complete definitions are provided in Eq. (7) and Eq. (8) in the Methods section. **b,** Comparison of whether FUGW and FPGW (SpaOT) meet the metric properties. Here $x, y, z \in \mathcal{X}$ are all elements of the sample space $\mathcal{X}$. Compared to FUGW, FPGW satisfies non-negativity, identity of indiscernibles, symmetry, and a relaxed triangle inequality when $q = 1$ (**Theorem 1** in the Methods section), providing a more principled distance metric in spatial omics research. * denotes: Moriel, N., Senel, E., Friedman, N. *et al.* NovoSpaRc: flexible spatial reconstruction of single-cell gene expression with optimal transport. *Nat Protoc* **16**, 4177–4200 (2021).

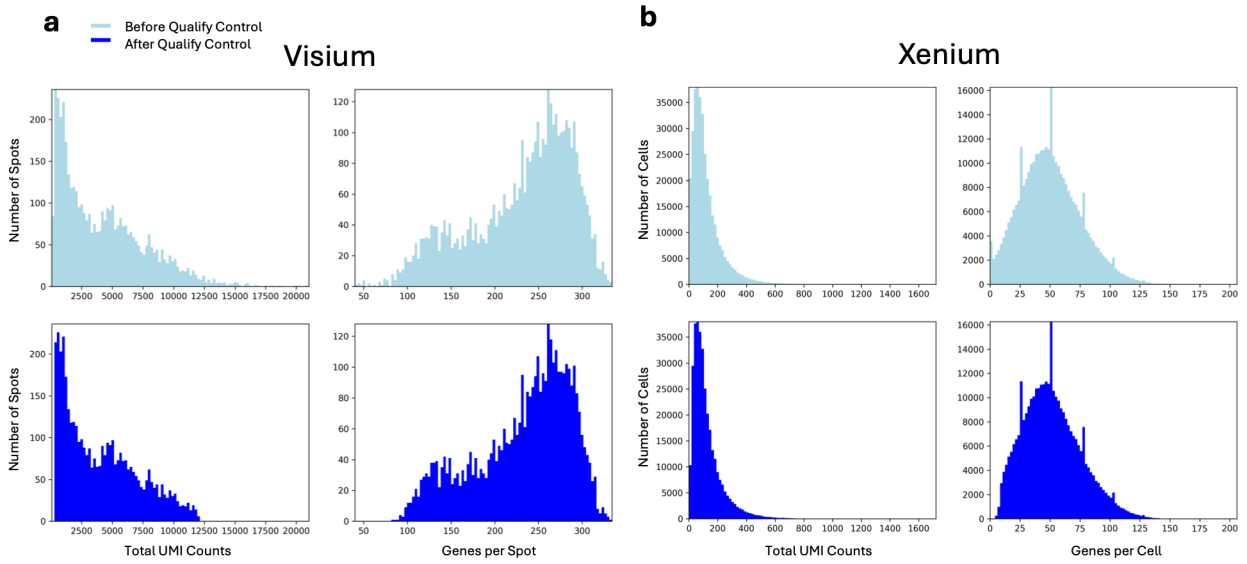

a
Mouse brain pairwise 3D stacking metrics
NN label accuracy, bidirectional (↑)
Normalized centroid distance (↓)
Expression cosine, bidirectional (↑)
SpaOT
MOSCOT
PASTE2
Method
Adjacent slice pair


b
3D Spatial Reconstruction
Original
SpaOT
Annotation
Cerebellum
Cortical subplate
fiber tracts
Hindbrain
Hippocampal formation
Hypothalamus
Isocortex
Midbrain
Olfactory areas
Pallidum
root
Striatum
Thalamus
ventricular systems
c
Hypothalamus
Olfactory Areas
Hypothalamus
+ Olfactory Areas
Isocortex
Striatum
Thalamus
Isocortex + Striatum
+ Thalamus
Hippocampal Formation
Midbrain
Hindbrain
Cerebellum
Hippocampal Formation + Midbrain +
Hindbrain + Cerebellum

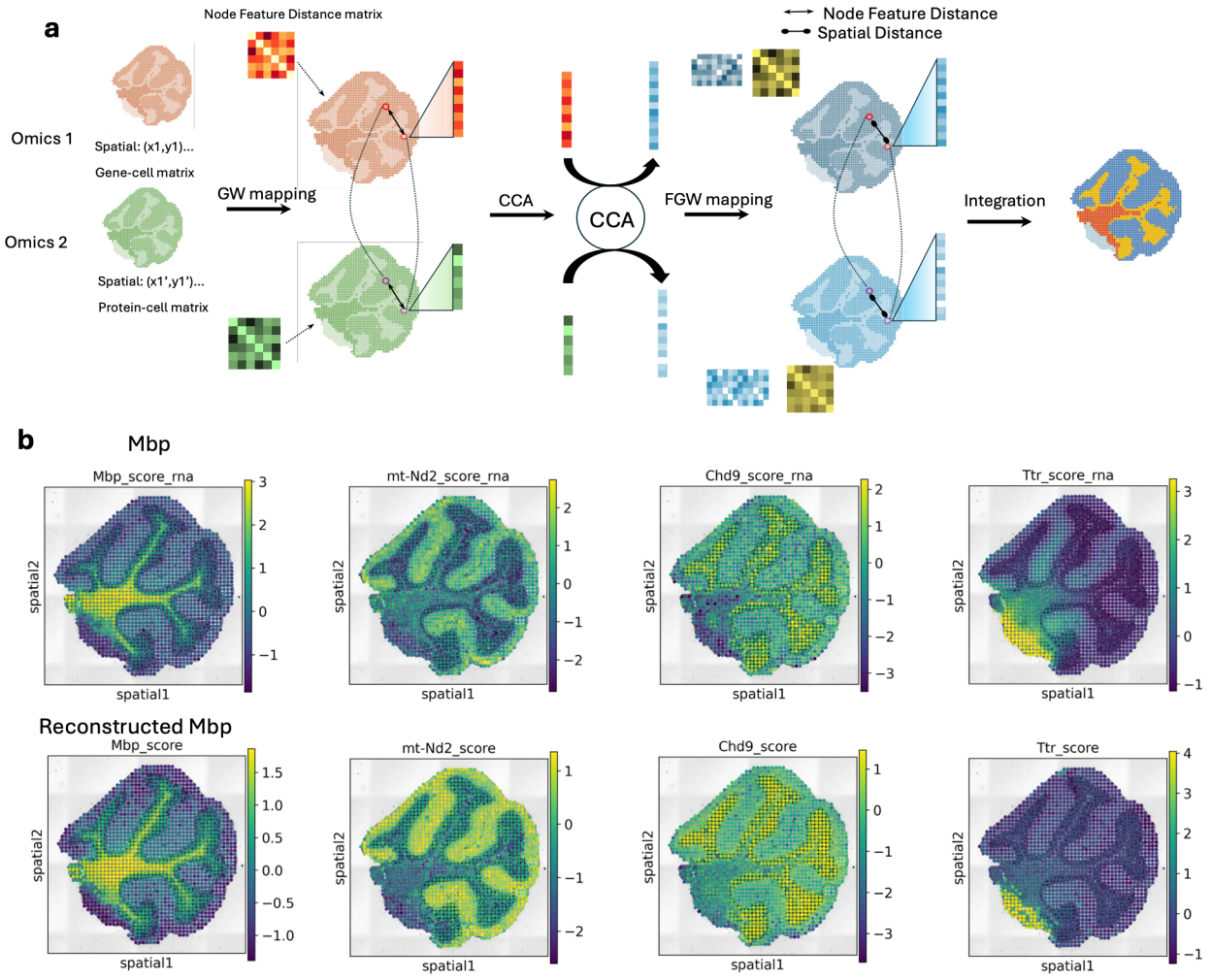
**Supplementary Fig. 2 | SpaOT performs 3D reconstruction of serial mouse brain sections.** **a,** Quantitative evaluation of 3D reconstruction performance on 40 consecutive coronal half-mouse brain sections. Adjacent sections were aligned sequentially using affine transformations derived from optimal transport plans and assembled into a 3D volume. Reconstruction quality was assessed using nearest-neighbor (NN) label accuracy, normalized centroid distance, and expression cosine similarity. Bars represent the mean performance across reconstructed sections for SpaOT and the benchmark methods (**left**), while line plots show the performance for individual brain sections (**right**). **b,** The 3D stacking visualization of the original mouse brain samples and the SpaOT-aligned samples. **c,** Three-dimensional reconstruction of major anatomical regions using Allen Brain Atlas (ABA) neuroanatomical annotations. Colored meshes represent reconstructed brain structures, including the hypothalamus, olfactory areas, isocortex, striatum, thalamus, hippocampal formation, midbrain, hindbrain, and cerebellum. Composite visualizations show the spatial relationships among reconstructed regions.


a
A0
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
A11
batch = A0
batch = A1
batch = A2
batch = A3
batch = A4
batch = A5
batch = A6
batch = A7
batch = A8
batch = A9
batch = A10
batch = A11
Astrocyte
Endothelial
Ependymal
Excitatory
Inhibitory
Microglia
OD Immature
OD Mature
Pericytes
b
batch = A0
batch = A1
batch = A2
batch = A3
batch = A4
batch = A5
batch = A6
batch = A7
batch = A8
batch = A9
batch = A10
batch = A11
c
batch = A0
batch = A1
batch = A2
batch = A3
batch = A4
batch = A5
batch = A6
batch = A7
batch = A8
batch = A9
batch = A10
batch = A11

**Supplementary Fig. 3 | Benchmarking 3D reconstruction of serial MERFISH data. a,** Consecutive alignment of 12 MERFISH-generated mouse hypothalamic preoptic area slices using SpaOT. The left panel shows the reconstructed stack of all slices, colored by slice identity (A0–A11). Individual pairwise alignments between adjacent slices are shown on the right. Cells are colored according to annotated cell types. **b,** Consecutive alignment results generated by moscot on the same dataset. **c,** Consecutive alignment results generated by PASTE2 on the same dataset.

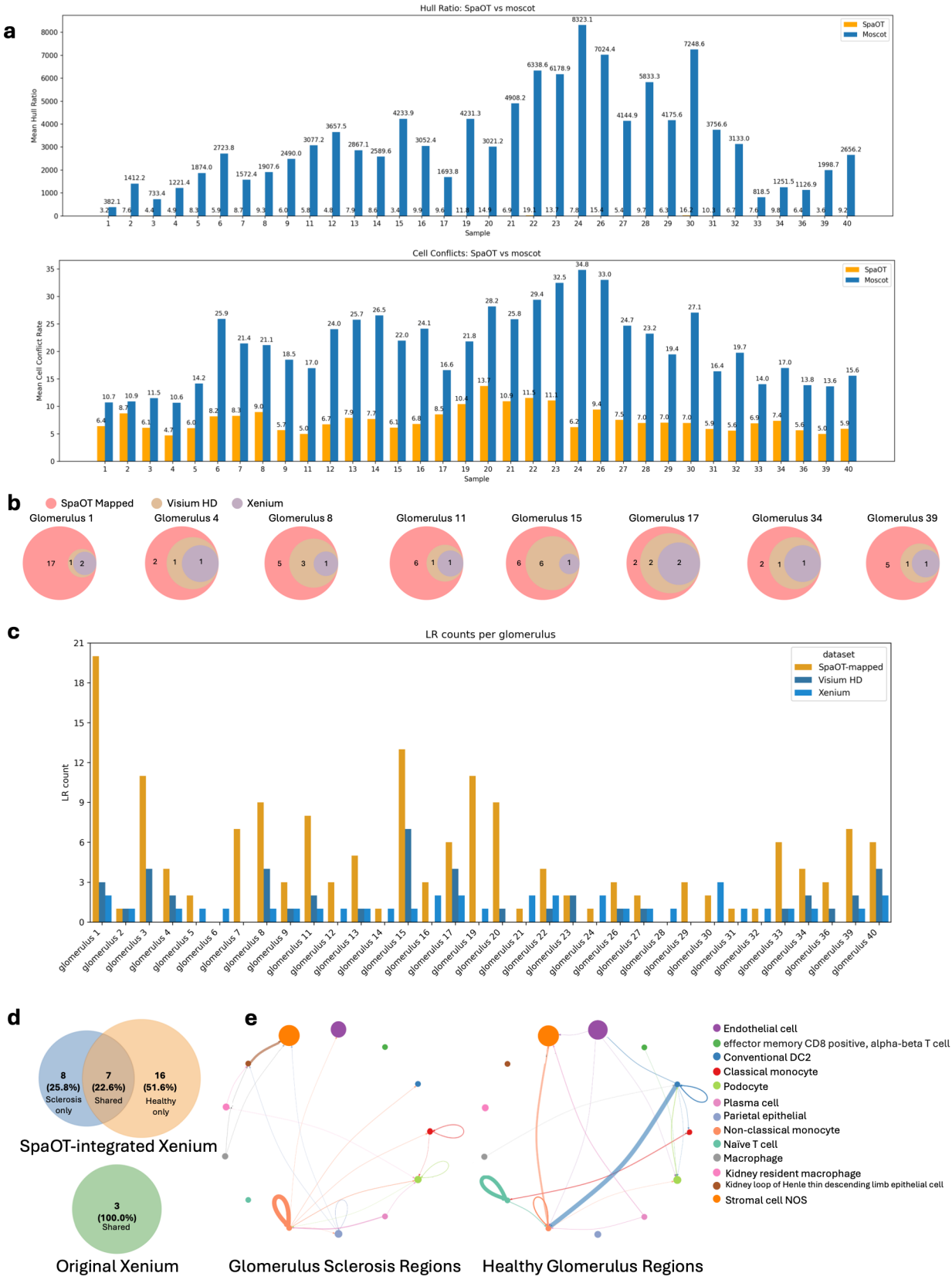



**Supplementary Fig. 4 | Distribution before and after Quality-Control (QC) assessment of Visium and Xenium datasets. a,** Distribution of transcript counts and detected genes per spot in the CRC Visium dataset. Histograms show the total UMI counts per spot (**left column**) and the number of detected genes per spot (**right column**) before QC (**top row**) and after QC (**bottom row**). **b,** Distribution of transcript counts and detected genes per cell in the Xenium dataset. Histograms display total transcript counts per cell (**left column**) and the number of detected genes per cell (**right column**) before QC (**top row**) and after QC (**bottom row**).

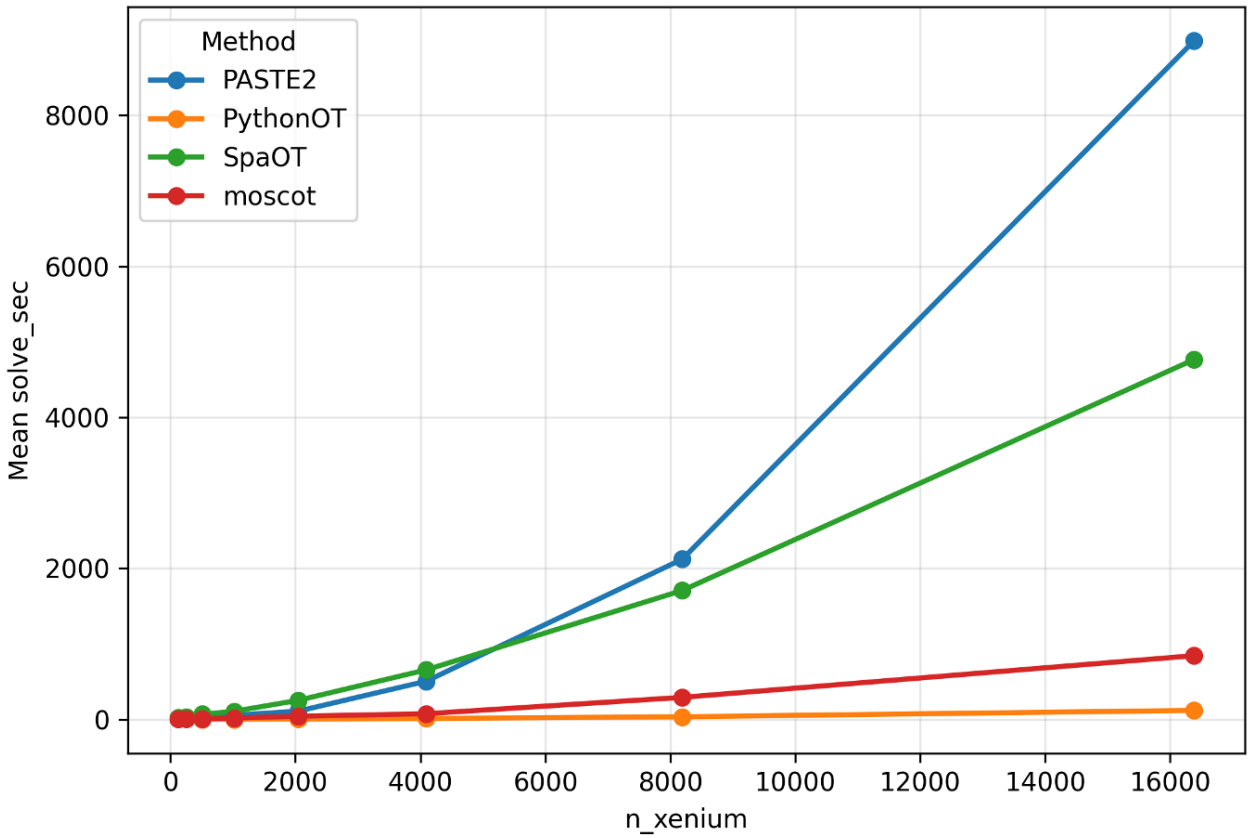



**Supplementary Fig. 5 | Performance of SpaOT cross-modality integration in mouse cerebellum datasets. a,** Schematic overview of the SpaOT framework for cross-modality alignment. Spatial transcriptomics (MAGIC-seq) and spatial proteomics (PLATO) data are shown as an example. Because molecular features from different modalities are not directly comparable, SpaOT first constructs intra-modality node-feature distance matrices and performs an initial alignment using only the structural term derived from within-modality relationships. The resulting transport plan establishes cross-modality correspondences, which are subsequently used for Canonical Correlation Analysis (CCA) to identify shared latent dimensions. These shared embeddings are incorporated as node features for a second-stage SpaOT alignment, in which spatial coordinates define the structural term. The final integrated representation preserves spatial organization while capturing biological associations across modalities, including gene-protein and gene-metabolite relationships. **b,** Examples of molecular feature reconstruction and cross-modal signal enhancement after integration. Top row, spatial expression patterns of representative genes (*Mbp*, *mt-Nd2*, *Chd9*, and *Ttr*) measured in the spatial transcriptomics dataset. Bottom row, reconstructed multimodal scores obtained by integrating corresponding transcriptomic, proteomic, and metabolomic signals. For each molecular group, the combined score was calculated from features reported to be co-localized in the original study: *Mbp* - CLD11 - 845.68 m/z, *mt-Nd2* - GRM1 - 787.55 m/z, *Chd9* - FUBP1 - 909.57 m/z, and *Ttr* - CLIC6 - 811.56 m/z. Compared with gene expression alone, the integrated multimodal

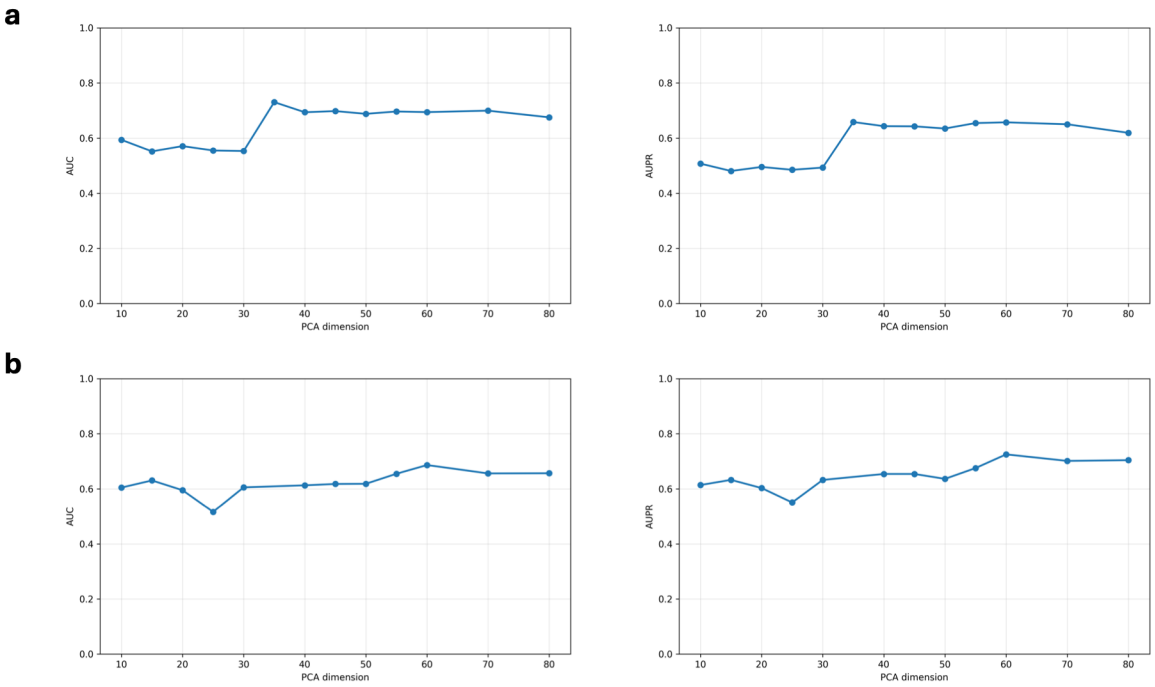
scores exhibited clearer spatial boundaries and more distinct enrichment patterns within the corresponding anatomical regions.


a
Hull Ratio: SpaOT vs moscot
Mean Hull Ratio
Sample
SpaOT
Moscot
Cell Conflicts: SpaOT vs moscot
Mean Cell Conflict Rate
b
SpaOT Mapped
Visium HD
Xenium
Glomerulus 1
Glomerulus 4
Glomerulus 8
Glomerulus 11
Glomerulus 15
Glomerulus 17
Glomerulus 34
Glomerulus 39
c
LR counts per glomerulus
LR count
dataset
SpaOT-mapped
Visium HD
Xenium
d
8 (25.8%) Sclerosis only
7 (22.6%) Shared
16 (51.6%) Healthy only
SpaOT-integrated Xenium
3 (100.0%) Shared
Original Xenium
e
Glomerulus Sclerosis Regions
Healthy Glomerulus Regions
Endothelial cell
effector memory CD8 positive, alpha-beta T cell
Conventional DC2
Classical monocyte
Podocyte
Plasma cell
Parietal epithelial
Non-classical monocyte
Naïve T cell
Macrophage
Kidney resident macrophage
Kidney loop of Henle thin descending limb epithelial cell
Stromal cell NOS

**Supplementary Fig. 6 | SpaOT-enhanced integration improves detection of spatial cellular communication in glomerular regions. a.** Quantitative comparison of hull ratio (**left**) and cell conflict (**right**) for SpaOT and moscot across all glomerulus regions. **b,** Representative examples of ligand-receptor (LR) signaling detected within glomerular regions using COMMOT. Each panel shows a glomerulus analyzed using Visium HD data alone, Xenium data alone, or SpaOT-integrated data. **c,** Barplot of detected LR interactions across 40 glomerular regions using COMMOT. **d,** Comparison of CellChat-inferred signaling interactions (LR pairs) between the original and SpaOT-integrated Xenium datasets. Venn diagram shows overlap between LR pairs detected in the SpaOT-integrated dataset (**top**), another Venn diagram shows 3 overlapped LR pairs detected both in healthy and sclerotic glomeruli of the Original Xenium dataset (**bottom**). **e,** Cell-cell communication analysis within healthy and sclerotic glomerular regions using CellChat on SpaOT-integrated Xenium data.




**Supplementary Fig. 7 | Runtime scalability on CRC research.** Comparison of optimal transport-based approaches including PASTE2, PythonOT, SpaOT, and moscot for Xenium-to-Visium mapping on the colorectal cancer (CRC) dataset[26]. The x-axis shows the number of Xenium cells used in the integration, which is randomly sampled from 128, 256, 512, 1024, 2048, 4096, 8192, and 16384 from the original Xenium dataset, respectively. The mapped Visium dataset was held constant at 4296 cells. The y-axis shows the mean runtime (seconds) on a computer node equipped with an AMD EPYC 7742 64-Core Processor CPU and 72GB RAM.




**Supplementary Fig. 8 | Sensitivity analysis to PCA dimensionality in Breast Cancer stratification research.** The X-axis is the PCA dimensions for SpaOT input, and the results are evaluated by a logistic regression classifier with the criteria of Area Under the Receiver Operating Characteristic Curve (AUROC) and Area Under the Precision-Recall Curve (AUPR). **a,** Training on METABRIC and testing on Jackson cohort. **b,** Training on Jackson and testing on METABRIC cohort. Left column, AUROC; right column, AUPR. Classification performance is shown across different numbers of principal components used as model input, indicating that the classifier is robust to the choice of PCA dimension.