, , pp. 1–22

doi:

Spartan: activation-aware framework for spatial domain and variable gene discovery

Mohammad Faiz Iqbal Faiz,[1] Elliot Jokl,[2] Rachel Jennings,[2] Karen Piper Hanley,[2] Andrew Sharrocks,[3] Mudassar Iqbal[1*] and Syed Murtuza Baker[1*]

[1]Division of Informatics, Imaging & Data Sciences, School of Health Sciences, University of Manchester, Oxford Road, M13 9PL, Manchester, United Kingdom, [2]Division of Diabetes, Endocrinology & Gastroenterology, School of Medical Sciences, University of Manchester, Oxford Road, M13 9PL, Manchester, United Kingdom and [3]Division of Molecular & Cellular Function, School of Biological Sciences, University of Manchester, Oxford Road, M13 9PL, Manchester, United Kingdom

[*]Corresponding authors. syed.murtuzabaker@manchester.ac.uk, mudassar.iqbal@manchester.ac.uk

Spatial transcriptomics is rapidly advancing toward single-cell-level resolution, revealing complex tissue architectures organized across continuous anatomical gradients. However, accurate identification of spatial domains remains a central computational challenge, as many existing clustering approaches blur anatomical boundaries, merge transitional zones, or fail to resolve localized microstructures. Here we introduce Spartan, an activation-aware multiplex graph framework for high-resolution domain discovery. Spartan integrates spatial topology and Local Spatial Activation (LSA), a neighborhood deviation signal that captures localized transcriptional heterogeneity often attenuated by similarity-based clustering. By jointly modeling cohesion within domains and localized activation structure, Spartan recovers anatomically aligned partitions across spatially resolved transcriptomics technologies including Visium HD, MERFISH, Stereo-seq, and STARmap. We further demonstrate its utility in a high-resolution Visium HD section of developing human esophagus and stomach, where activation-aware graph integration enables precise delineation of complex transitional regions such as the gastroesophageal junction and supports stable multi-scale domain recovery without fragile hyperparameter tuning. Beyond domain identification, Spartan leverages activation-aware structure to detect spatially variable genes associated with localized tissue remodeling. Spartan scales near-linearly with dataset size, providing a robust and interpretable framework for spatial systems-level analysis.

Key words: Spatial transcriptomics, Spatial domains, Spatially variable genes, Local spatial autocorrelation

Introduction

Spatially resolved transcriptomics (SRT) enables gene expression to be measured directly in intact tissue, allowing molecular states to be mapped onto anatomical structures [1, 2, 3]. As SRT technologies increase in resolution and scale, datasets now contain tens of thousands of spatial locations exhibiting fine-grained transcriptional heterogeneity [4]. Accurately resolving spatial organization within such data remains a central computational challenge.

A core task in SRT analysis is spatial domain identification, in which spatial locations are grouped into coherent regions with shared transcriptional programs. Most existing approaches integrate spatial proximity with gene expression similarity to infer such domains, which effectively separates major tissue compartments. However, complex tissues frequently contain localized microenvironments embedded within broader regions, including transitional zones and niche-like structures, where subtle transcriptional shifts define functional specialization. Under these conditions, purely similarity-driven clustering can smooth localized heterogeneity, merge anatomically distinct regions, and overlook fine spatial structure that is biologically meaningful.

To better capture this fine-scale organization, we reasoned that spatial domain discovery should measure not only what each cell or spot expresses and where it resides, but also how its molecular state relates to its immediate neighborhood context. We introduce this additional signal as Local Spatial Activation (LSA), a quantitative measure of neighborhood-relative transcriptional deviation in a low-dimensional latent space. Rather than reinforcing similarity alone, LSA models joint deviation of spatially adjacent spots relative to their shared local transcriptional context. Conceptually, LSA captures localized shifts in molecular programs while preserving spatial coherence. Although related in spirit to local spatial statistics used in computational geography such as local Moran's I [5, 6], LSA is not a statistical test of spatial autocorrelation. Instead, it is formulated as a non-negative, variance-normalized graph signal designed for integration within a multiplex community detection framework.

Motivated by this principle, we developed Spartan (Spatial Activation-Aware Transcriptomic Analysis Network), a framework for spatial domain recovery that integrates three complementary layers within an aggregated graph: spatial topology, LSA, and gene expression connectivity. Unlike existing spatial clustering approaches that encode spatial proximity and

© The Author . Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

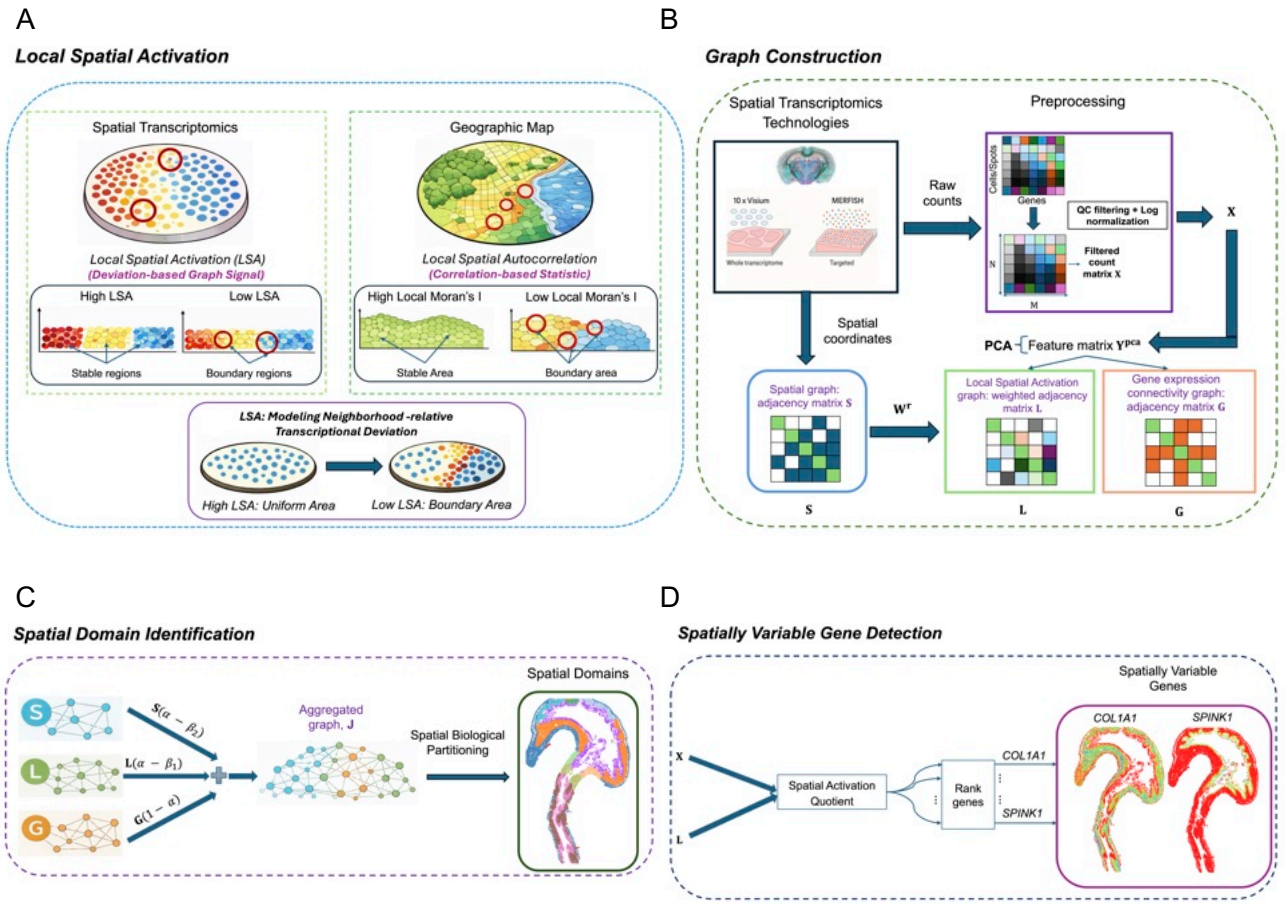

Figure 1. Overview of the Spartan framework for spatial domain identification and spatially variable gene detection. (A) Conceptual comparison between Local Spatial Activation (LSA) in spatial transcriptomics and classical local spatial autocorrelation (Local Moran's I) in geostatistics. Left: In spatial transcriptomics, LSA models neighborhood-relative transcriptional deviation, capturing localized shifts in molecular programs within tissue architecture. High LSA values occur in spatially coherent regions, with lower values at boundaries exhibiting structured transcriptional variation. Right: In geostatistics, Local Moran's I (LISA) measures local spatial autocorrelation, identifying regions where neighboring locations exhibit similar values. Boundary regions exhibit reduced autocorrelation in heterogeneous areas and elevated autocorrelation in spatially coherent regions. Unlike classical statistical autocorrelation measures, LSA is formulated as a constructive graph signal designed for activation-aware community detection in spatial transcriptomics. (B) Graph construction within Spartan. Spartan identifies spatial domains and spatially variable genes (SVGs) through integration of spatial and transcriptomic information. From spatial transcriptomics data, Spartan performs PCA to obtain $\mathbf{Y}^{\mathbf{pca}}$ and constructs three complementary graphs: a spatial graph encoding spot proximity, a Local Spatial Activation (LSA) graph capturing fine-scale spatial variation, and a gene expression connectivity graph encoding local transcriptomic neighborhood structure. (C) These graphs are combined into an aggregated graph (represented by weighted adjacency matrix $\mathbf{J}$), where contributions from the spatial, gene expression connectivity, and LSA graphs are explicitly aggregated. Leiden clustering on $\mathbf{J}$ yields biologically coherent spatial domains. (D) The Spatial Activation Quotient (SAQ), derived from the LSA graph, ranks genes by the strength of their spatially structured activation to identify spatially variable genes (SVGs) (green indicates higher enrichment, and red indicates lower enrichment).

expression similarity within a single homogeneous graph layer, Spartan introduces neighborhood-relative activation as an explicit structural component of the clustering objective. Spartan applies Leiden community detection [7] to this activation-aware graph, enabling efficient and reproducible domain identification while preserving interpretability through explicit layer weighting and transparent parameterization. By incorporating activation as a graph layer, Spartan enables activation-aware clustering while supporting robust scale adaptation across datasets. The aggregated formulation allows Spartan to transition between settings dominated by broad compartmental organization and those requiring sensitivity to fine local heterogeneity.

In high-resolution samples, such as Visium HD sections of the developing human esophagus and stomach, biologically meaningful organization is expressed as thin layers, curved anatomical boundaries, and junctional transition zones, including the gastroesophageal junction (GEJ), where epithelial and stromal programs change gradually rather than forming sharply separated compartments. Resolving such continuous spatial organization illustrates the type of challenge that activation-aware graph integration is designed to address.

We benchmarked Spartan across multiple SRT technologies, including MERFISH, Stereo-seq, STARmap, and osmFISH, spanning diverse resolutions and tissue architectures. Spartan achieved strong and consistent performance across datasets, frequently matching or exceeding recently proposed methods including BANKSY [8], NichePCA [9], and SpatialLeiden [10]. Ablation analyses confirmed the complementary contribution of each graph component, with the multiplex configuration providing the most robust domain recovery across technologies and

experimental settings. Beyond spatial domain identification, Spartan extends naturally to spatially variable gene detection by leveraging LSA to identify genes associated with localized spatial structure, providing a scalable and extensible framework for spatial transcriptomics and future spatial multi-omics integration.

Results

Overview of Spartan

Spartan is a computational framework for the accurate identification of spatial domains and detection of spatially variable genes (SVGs) across diverse SRT platforms. It integrates spatial topology and gene expression structure through three biologically grounded graphs that jointly encode spatial proximity, LSA (conceptual illustration in Fig. 1A), and expression-derived connectivity.

From each SRT sample, Spartan extracts spatial coordinates and a preprocessed count matrix $\mathbf{X}$, performs principal-component analysis to obtain $\mathbf{Y}^{pca}$, and constructs three complementary graphs (Fig. 1B). The spatial graph, represented by adjacency matrix $\mathbf{S}$ (Figs. 1B, 2A), encodes physical neighborhood relationships between spots using either k-nearest-neighbor (KNN) or Delaunay connectivity. A central component of Spartan is the LSA graph, represented by $\mathbf{L}$ (Figs. 1B, 2B). For each pair of neighboring spots i and j, the weight l_{ij} combines their attribute deviations ϕ_i and ϕ_j with a spatial proximity term w^r_{ij}. This formulation quantifies LSA, providing an additional spatial signal that captures neighborhood-relative transcriptional deviation beyond expression similarity alone. The gene expression connectivity graph, represented by adjacency matrix $\mathbf{G}$ (Figs. 1B, 2C), encodes local transcriptomic neighborhood structure in PCA-reduced space.

Spartan integrates these three graphs into an aggregated graph, represented by the weighted adjacency matrix $\mathbf{J}$ (Fig. 1C) that balances spatial proximity, LSA, and expression-derived connectivity. Leiden clustering on $\mathbf{J}$ partitions the tissue into coherent spatial domains that align with known anatomical regions. To further exploit the LSA signal, Spartan computes a Spatial Activation Quotient (SAQ) (Fig. 1D) for each gene by combining $\mathbf{L}$ and $\mathbf{X}$, ranking genes by their spatial activation strength to identify SVGs.

By modeling spatial relationships at multiple levels, including global proximity ($\mathbf{S}$), LSA ($\mathbf{L}$), and transcriptomic similarity ($\mathbf{G}$), Spartan captures biologically coherent tissue organization and spatially patterned gene programs. Its independence from specific SRT technologies and its ability to detect both structured domains and fine-grained local patterns make Spartan a general and extensible framework for spatial omics analysis.

Ablation study of multiplex graph components

Next, we conducted an ablation study to evaluate the contribution of each component in the aggregated graph gr^{mul}. The aggregated graph comprises three elements (Equation 15): (i) the LSA graph gr^{lsa}, represented by the weighted adjacency matrix $\mathbf{L}$, (ii) the gene expression connectivity graph gr^{g}, represented by $\mathbf{G}$, and (iii) the spatial graph gr^{s}, represented by $\mathbf{S}$. We considered three combinations of these graphs that have both structural and expression-based terms: (i) $\mathbf{L}+\mathbf{S}+\mathbf{G}$, (ii) $\mathbf{L}+\mathbf{G}$, and (iii) $\mathbf{S}+\mathbf{G}$. For dual-graph formulations ($\mathbf{L}+\mathbf{G}$ or $\mathbf{S}+\mathbf{G}$), the combined adjacency matrix was defined as $\alpha \cdot (\mathbf{L} \text{ or } \mathbf{S}) + (1-\alpha) \cdot \mathbf{G}$. The β_1 and β_2 parameters were only used in $\mathbf{L}+\mathbf{S}+\mathbf{G}$ configurations. Each combination was evaluated using two spatial connectivity schemes, k-nearest neighbors (KNN) or Delaunay triangulation, resulting in six total configurations. These configurations were evaluated across four datasets representing different SRT technologies. For multi-sample datasets, the multiplex weighting parameter α was selected using a cross-sample hierarchical strategy to avoid sample-specific overfitting (α-selection strategy) (conceptual illustration in Supplementary Fig. 1).

We evaluated all configurations across multiple datasets using the R score and ARI score (Fig. 3A-B). Overall, the ablation study supports the aggregated graph formulation implemented in Spartan. Across eight SRT samples from different technologies, including MERFISH and STARmap, the multiplex configuration ($\mathbf{L}+\mathbf{S}+\mathbf{G}$) performed strongly across all samples. Specifically, the aggregated graph using KNN-based spatial connectivity ranked best in two samples and second best in five, demonstrating strong generalization across datasets. The Delaunay-based configuration also performed competitively, ranking best in one sample, although KNN-based connectivity offers greater stability across all datasets. In contrast, the dual-graph configurations ($\mathbf{L}+\mathbf{G}$ or $\mathbf{S}+\mathbf{G}$) exhibited greater variability across samples and technologies. Notably, these dual configurations displayed an inverse relationship: when $\mathbf{L}+\mathbf{G}$ achieved high performance, $\mathbf{S}+\mathbf{G}$ tended to perform poorly, and vice versa. Together, these findings indicate that the three graph components contribute complementary spatial and transcriptomic information, contributing to consistent clustering performance across diverse SRT datasets.

Validation of Local Spatial Activation (LSA) graph structure

Having established the importance of the multiplex formulation through ablation analysis, we next evaluated whether the Local Spatial Activation (LSA) graph itself captures biologically meaningful spatial organization. We examined the relationship between LSA edge weights and ground truth spatial domains (Supplementary Methods). Because LSA quantifies neighborhood-relative transcriptional activation rather than simple spatial proximity, this analysis tests whether the LSA graph aligns with biologically annotated spatial domains.

First, we compared LSA weights for spatial neighbor edges connecting spots within the same spatial domain and those connecting spots across different domains. Intra-domain edges exhibited systematically higher LSA weights than inter-domain edges, indicating that LSA preferentially strengthens connections within spatial domains (median intra-domain edge weights: 0.444; median inter-domain edge weights: 0.357). The difference between these distributions was highly significant according to a Mann–Whitney U test ($p = 3.16 \times 10^{-94}$), with a moderate effect size as measured by the rank-biserial correlation (RBC = 0.246) (Supplementary Fig. 2A–C).

To assess whether LSA weights provide discriminative information about spatial domain structure, receiver operating characteristic (ROC) analysis was performed. In the MERFISH 0.24 sample, LSA edge weights demonstrated measurable ability to distinguish intra-domain from inter-domain relationships (AUC = 0.623) (Supplementary Fig. 2D). Next, permutation analysis confirmed that the observed enrichment of LSA weights within spatial domains exceeded that expected under random domain assignments. The observed difference between median intra-domain and inter-domain LSA weights (≈ 0.0869) was substantially larger than values obtained under random permutations ($p \approx 0.000999$), supporting the conclusion that

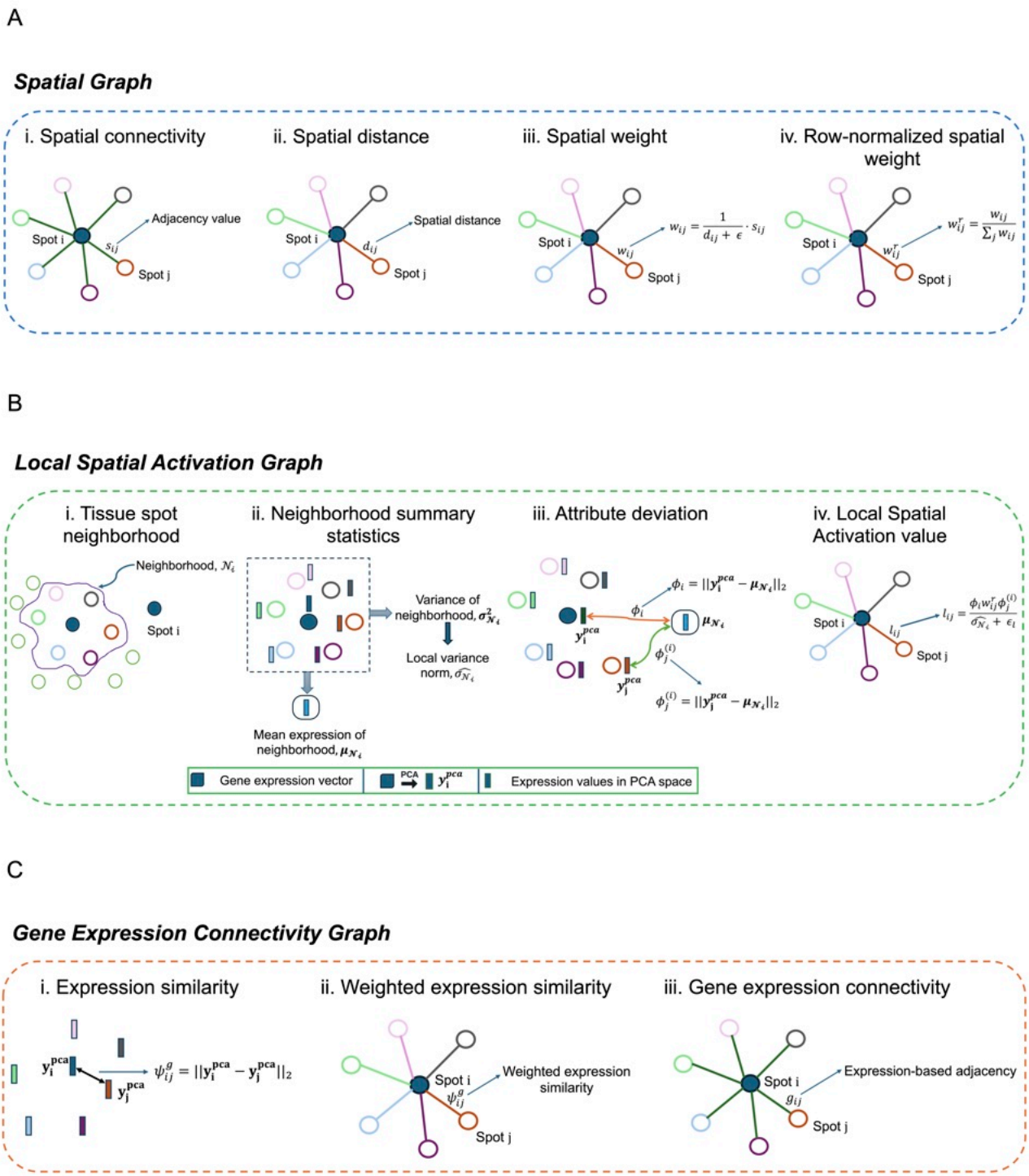

Figure 2. Overview of Spartan's graph construction. (A) Spartan constructs a spatial graph represented by the adjacency matrix $\mathbf{S}$. The spatial weight w_{ij} for each edge is computed from the Euclidean distance d_{ij} between spot i and spot j, together with the adjacency value s_{ij}. The row-normalized weight w^r_{ij} is then obtained from w_{ij}. (B) Spartan constructs the Local Spatial Activation (LSA) graph, represented by the weighted adjacency matrix $\mathbf{L}$. Each edge is assigned a weight l_{ij}, which captures the variance-normalized, spatially weighted joint transcriptional deviation of spots i and j relative to the local neighborhood of spot i. (C) Spartan constructs a gene expression connectivity graph represented by the adjacency matrix $\mathbf{G}$, where each edge is assigned an adjacency value g_{ij} based on KNN relationships in the PCA-reduced expression space.

LSA captures non-random spatial organization in the tissue (Supplementary Fig. 2E).

Similar analysis was done for the Stereo-seq sample E2S3 (Supplementary Fig. 3A-C). Specifically, local LSA magnitudes showed similar discriminative ability for distinguishing boundary from interior regions (AUC = 0.618), consistent with the higher spatial resolution and complex boundary structure characteristic of this technology (Supplementary Fig. 3D).

Spartan reveals biologically coherent domains in the mouse embryo

The Spartan clustering results on the Stereo-seq E9.5E2S2 sample demonstrate its strong effectiveness in identifying biologically meaningful domains (Fig. 3C). Visually, Spartan generates spatially coherent domains with minimal fragmentation, which aligns well with the tissue morphology. This behavior is consistent with the activation-aware graph construction strategy, which leverages LSA in addition to expression similarity and is well suited to early embryo tissues where spatial gradients and sharp regional boundaries coexist.

The contingency table (ground truth vs. predicted clusters) (Fig. 3D) provides a quantitative view of Spartan's mapping accuracy. Most tissues, such as the spinal cord, surface ectoderm, heart, and liver, map cleanly to unique predicted clusters, visible as strong diagonal blocks. The predicted clusters broadly recapitulate the major anatomical compartments annotated [11]: forebrain, heart, liver, spinal cord, mesenchyme, and others (Supplementary Fig. 4A). Some tissues, such as mesenchyme and head mesenchyme, exhibit partial overlap across multiple predicted clusters. This does not necessarily indicate misclassification: mesenchyme is known to be heterogeneous and spread across numerous anatomical regions in E9.5 embryos [11], so splitting into multiple clusters may reflect biological heterogeneity or annotation granularity rather than misclassifications. Thus, Spartan not only recovers the ground truth domains but also reveals finer granularity potentially providing new biological insights.

Overall, these results highlight Spartan's effectiveness not just in matching ground truth but in extending beyond it by resolving complex tissues into finer subdomains.

Spartan demonstrates competitive and statistically supported performance on public SRT datasets

The benchmarking study was divided into two parts. First, Spartan was benchmarked against representative and widely used methods for spatial domain identification using datasets from the SDMBench suite. Second, Spartan was compared against recently proposed methods using multiple datasets, including those from SDMBench as well as the Vizgen MERFISH dataset. Two versions of Spartan were evaluated: Spartan (v1), which employs a Delaunay graph to represent neighborhood connectivity in spatial coordinate space, and Spartan (v2), which employs a KNN graph to represent neighborhood relations.

Benchmarking on SDMBench

Spartan was benchmarked on six datasets from SDMBench: STARmap, STARmap*, MERFISH, BaristaSeq, osmFISH, and Stereo-seq. For competing methods covered by SDMBench, we report the published benchmark results generated under the centralized SDMBench framework using matched cluster counts and standardized evaluation metrics while respecting method-specific preprocessing and architectural settings. Performance was assessed in comparison with five representative and widely used methods: BASS [12], GraphST [13], SpaGCN [14], SCAN-IT [15] and SpaceFlow [16], as well as the standard Leiden clustering algorithm.

The distribution of evaluation metrics (NMI, COM, HOM, and R) across datasets and methods is illustrated (Fig. 3E; Supplementary Figs. 5A, 6A, 7A). Both versions of Spartan attain the highest or near-highest R scores in five of six SDMBench datasets and rank within the top-performing methods across all datasets. In MERFISH, Spartan improves median NMI by approximately +0.05 and HOM by +0.08 relative to the next-best method under the SDMBench protocol. In Stereo-seq, median improvements reach approximately +0.03 in NMI and +0.05 in HOM. In STARmap, STARmap*, and osmFISH, Spartan achieves the highest or near-highest R scores under the centralized SDMBench protocol.

These gains are observed across both sequencing-based and imaging-based platforms under the SDMBench evaluation protocol. The relatively small interquartile ranges indicate consistent performance across datasets, consistent with the stability of the multiplex integration of LSA, neighborhood connectivity, and gene expression connectivity.

The R score heatmap across individual datasets (Fig. 3F; Supplementary Figs. 5B, 6B, 7B) further supports this observation. Spartan consistently achieves high or near-high performance, with particularly strong gains in MERFISH and Stereo-seq.

An aggregated radar plot (Supplementary Fig. 4B) highlights trade-offs across evaluation metrics. Both versions of Spartan display a balanced performance profile across COM and HOM, whereas competing methods often exhibit stronger trade-offs between completeness and homogeneity.

To evaluate whether the observed performance differences reflect systematic cross-sample improvements rather than sampling variability, we performed paired two-sided Wilcoxon signed-rank tests comparing both versions of Spartan (v1 and v2) against each representative method across samples within MERFISH (n=5) and Stereo-seq (n=9). All statistical tests were conducted on per-sample metric values rather than aggregated dataset summaries (not median summaries). P-values were adjusted using the Benjamini–Hochberg procedure within each dataset and metric. Median paired differences and adjusted P-values are summarized in Supplementary Fig. 8.

Across Stereo-seq (n=9), Spartan (v2) demonstrates statistically significant paired improvements relative to several competitors after FDR correction across NMI, HOM, and COM. In MERFISH (n=5), median paired differences remain consistently positive but do not uniformly reach statistical significance after correction, consistent with the smaller number of paired samples. These findings indicate that Spartan's performance gains are reproducible across independent samples and are not attributable to isolated dataset-specific instances.

Benchmarking against recently proposed methods

Three recently published methods for identifying spatial domains were implemented. These included SpatialLeiden [10], NichePCA [9], and BANKSY [8]. Two versions of SpatialLeiden were evaluated: SpatialLeiden (v1), which uses PCA for feature selection, and SpatialLeiden (v2), which employs MULTISPATI PCA [17]. The KNN graph was used to encode the neighborhood connectivity of cells in both versions of SpatialLeiden, consistent with their reported benchmarking configuration.

Five datasets from the SDMBench (MERFISH, STARmap, STARmap*, BaristaSeq, osmFISH) were used to benchmark

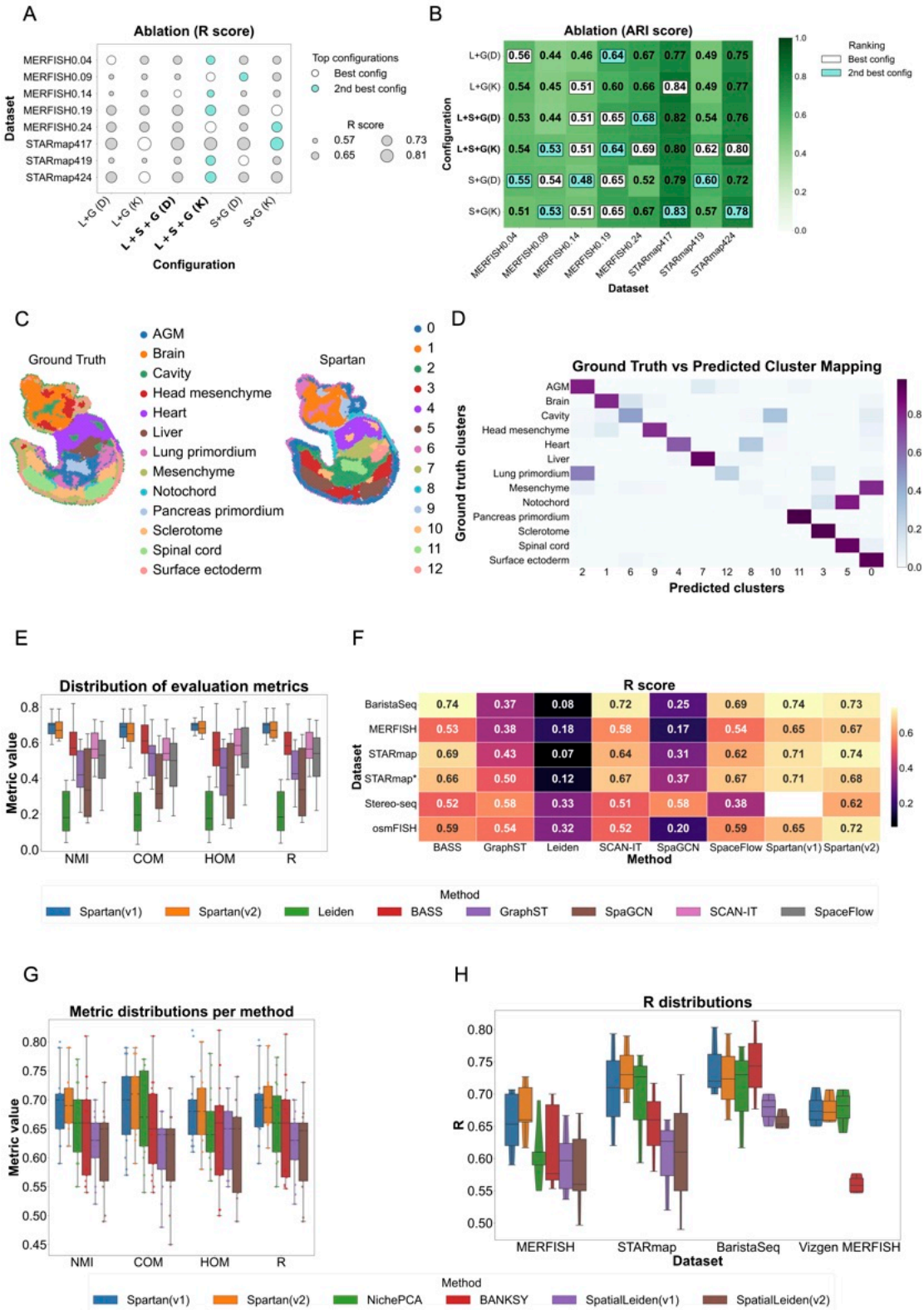

Figure 3. (A–B) Graph ablation. (A) R scores for combinations of LSA (**L**), spatial (**S**), and gene expression connectivity (**G**) graphs using KNN (K) or Delaunay (D) neighborhoods; circle size denotes R, colors indicate best and second-best configurations. (B) Dataset-wise ablation confirms the contribution of each graph component. (C–D) Spatial reconstruction. (C) Example Stereo-seq E9.5E2S2 sample showing ground-truth anatomy (left) and Spartan-derived domains (right). (D) Confusion matrix between predicted clusters and anatomical labels. (E–F) Benchmarking against established methods (SDMBench). (E) Distributions of NMI, COM, HOM, and R using SDMBench-reported results. (F) Dataset-wise R score heatmap. (G–H) Benchmarking against recent methods. (G) Metric distributions for recently published methods evaluated in this study. (H) R score distributions across datasets.

Spartan against these methods. In addition, the Vizgen MERFISH dataset was employed to assess Spartan's scalability and its ability to detect spatial domains in high-resolution samples.

To verify faithful baseline implementation, we quantitatively compared our reproduced results against the originally reported values. For SpatialLeiden (v1 and v2 configurations), reproduced median NMI values across MERFISH, STARmap, osmFISH, and BaristaSeq deviated from the originally reported medians by typically 0.01–0.04 (mean absolute deviation ≈0.03), with a single larger deviation of 0.086 in one MERFISH sample. For NichePCA and BANKSY, reproduced median NMI values closely matched published results across Vizgen MERFISH, MERFISH, STARmap, and BaristaSeq (absolute deviation ≤0.04 in all but one case, where BANKSY improved by +0.08 under our implementation). These deviations confirm faithful reproduction of baseline pipelines and indicate that implementation differences are unlikely to account for comparative performance differences.

In several datasets, methods such as NichePCA, BANKSY, and SpatialLeiden exhibit wider spreads and greater variability across datasets. Importantly, Spartan performs well on both COM and HOM simultaneously, whereas competing methods often trade one for the other (Fig. 3G; Supplementary Figs. 9A, 10A, 11A, 12A). This balanced performance is consistent with Spartan's multiplex design, which integrates LSA and neighborhood structure to stabilize performance against dataset-specific challenges.

The R-score distributions across datasets are further examined and both versions of Spartan consistently rank among the top performers, with the strongest gains observed in MERFISH and STARmap (Fig. 3H; Supplementary Figs. 9B-C, 10B-C, 11B-C, 12B-C). These datasets are characterized by fine-grained spatial heterogeneity between neighboring cells and non-uniform cell density. Spartan's incorporation of the LSA component plausibly contributes to improved sensitivity to fine-scale differences by considering attribute deviations of spots or cells relative to their neighborhood.

The ARI scores of Spartan (v1 and v2), NichePCA, SpatialLeiden (v1 and v2), and BANKSY across five datasets are shown in (Fig. 4A-B; Supplementary Fig. 13A-D). Regardless of version, Spartan attains the highest ARI scores in three of five datasets (MERFISH, STARmap, and STARmap*) and ranks second in the remaining two (osmFISH and BaristaSeq).

To quantify cross-sample robustness relative to recently proposed methods, we performed paired two-sided Wilcoxon signed-rank tests comparing Spartan (v1 and v2) against NichePCA and BANKSY across MERFISH and Vizgen MERFISH samples, and against both SpatialLeiden configurations across MERFISH samples. P-values were adjusted using the Benjamini–Hochberg procedure within each dataset and metric. Median paired differences and adjusted P-values are summarized in Supplementary Fig. 14.

Across MERFISH, Spartan (v2) exhibits consistently positive median paired differences across NMI, HOM, and COM relative to competing methods, with several comparisons remaining significant after FDR correction. In Vizgen MERFISH, improvements are directionally consistent, while statistical significance varies across metrics, reflecting the limited number of paired samples. No systematic degradations relative to competing methods were observed. Together, these results indicate that Spartan's performance remains stable across independent samples and is not driven by isolated dataset-specific effects.

Taken together, across both SDMBench and recently proposed methods, Spartan demonstrates stable cross-sample performance and competitive results with statistically supported improvements across multiple technologies in several datasets. The paired statistical analyses confirm that observed gains reflect systematic effects rather than isolated dataset-specific instances. Notably, Spartan (v2), which employs a fixed-degree KNN spatial graph, exhibits the most stable and reproducible improvements across metrics and datasets, suggesting that controlled neighborhood regularization enhances multiplex activation-aware integration in heterogeneous spatial contexts. Under the evaluation protocols applied here, these results suggest that Spartan provides a robust and generalizable framework for spatial domain identification across diverse spatial transcriptomics technologies.

Spartan identifies cell-specific, layer-specific, and region-specific SVGs

In the preoptic and hypothalamic regions of the adult mouse brain sample, obtained from the MERFISH technology, Spartan effectively identifies spatially variable genes (SVGs) that delineate distinct anatomical and cellular domains (Fig. 4C; Supplementary Fig. 15A). The inhibitory neuron marker *Gad1* is detected across nearly all gray matter structures, including the bed nuclei of the stria terminalis (BST), medial preoptic area (MPA), medial preoptic nucleus (MPN), periventricular hypothalamic nucleus (PV), and paraventricular hypothalamic nucleus (PVH), but is absent in the fornix (fx) and paraventricular thalamic nucleus (PVT). This pattern is consistent with the known distribution of GABAergic neurons [18, 19], as the fx and PVT are primarily white matter and glutamatergic, respectively. Conversely, *Mbp* is strongly expressed only in the fx, highlighting oligodendrocyte-rich white matter tracts [20], whereas *Nnat* is restricted to the V3 ventricular zone, marking ependymal cells [21]. The calcium-regulatory gene *Sln* is present in most neuronal and ependymal regions [22] but absent in fx, and the neuropeptide gene *Oxt* is highly specific to the PVH, corresponding to magnocellular neurosecretory neurons [23]. Collectively, these SVGs reveal a coherent molecular anatomy, distinguishing inhibitory neurons, white matter, ependymal cells, and neuroendocrine centers.

In the Stereo-seq developing mouse embryo sample, Spartan captures region- and tissue-specific gene expression patterns that reflect early organogenesis (Fig. 4D; Supplementary Fig. 15B). Cardiac markers such as *Actc1*, *Myh7*, and *Myl7* are detected exclusively in the heart, with *Nppa* showing highly lateralized expression restricted to the left side of the heart, consistent with left–right patterning [24]. Hemoglobin genes (*Hba-a1*, *Hba-a2*) are expressed in the heart, liver, and fluid-filled cavity regions [25], including areas between the sclerotome and lung primordium, highlighting early erythropoietic activity. Other genes, such as *Gm42418*, are confined to cavity regions, while *Cdk8* and *Il31ra* are absent from specific tissues like the surface ectoderm, spinal cord, and brain regions, reflecting selective tissue differentiation. These findings demonstrate Spartan's ability to detect localized embryonic expression patterns consistent with developmental biology [26, 27].

The Vizgen MERFISH mouse brain dataset further highlights Spartan's ability to capture regionally specialized neuronal and glial populations across the adult brain (Fig. 4E; Supplementary Fig. 15C). The excitatory neuron marker *Slc17a7* [28] is strongly expressed in the cerebral cortex, hippocampus, and thalamus, whereas the astrocytic marker *Gfap* is enriched in the hippocampal and hypothalamic regions. *Chat* and *Gpr151* [29] show selective expression in the medial and lateral habenula of the epithalamus, while *Gpr88* is enriched

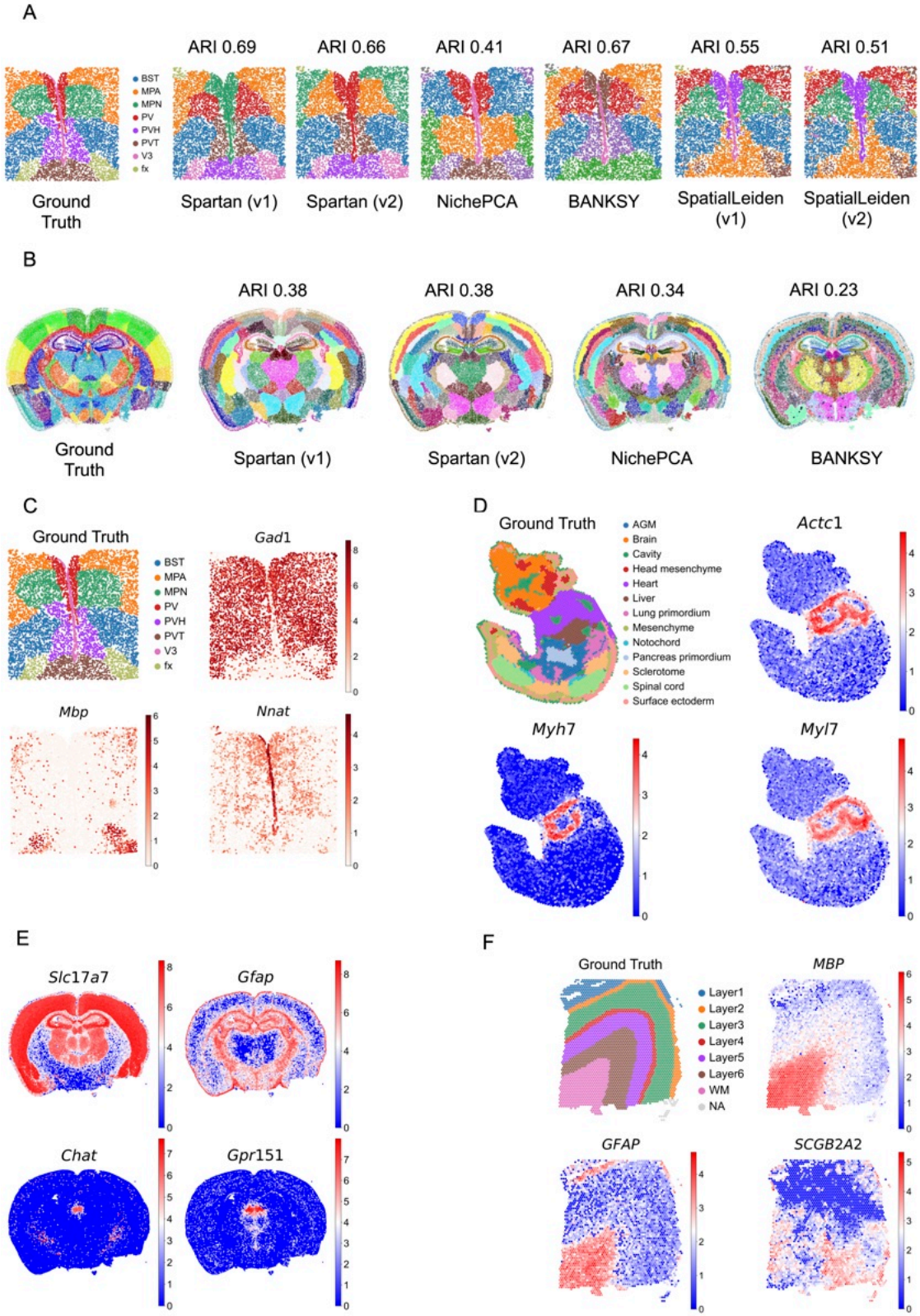

Figure 4. (A–B) Spatial domain identification for MERFISH 0.19 (SDMBench) and Vizgen MERFISH S2R3 samples. Ground-truth annotations (left) are compared with clustering results from Spartan (v1, v2), NichePCA, BANKSY, and SpatialLeiden (v1, v2). ARI values above each map quantify agreement with ground truth. (C) Spatially variable gene (SVG) examples for the MERFISH 0.19 sample. Spatial expression patterns align with anatomical boundaries and support the cluster assignments in (A). (D) Stereo-seq mouse embryo (E9.5E2S2). Ground-truth annotations (left) and spatial expression of representative embryo marker genes (right) show anatomically consistent localization. (E) SVGs for the Vizgen MERFISH mouse brain S2R3 sample, highlighting region- and cell-type–specific expression patterns. (F) 10x Visium human dorsolateral prefrontal cortex (DLPFC) sample 151673. Ground-truth cortical layers (left) and layer-enriched marker genes (right) demonstrate accurate laminar recovery.

in the striatum of the cerebral nuclei, reflecting medium spiny neuron populations [30]. The muscarinic acetylcholine receptor *Chrm1* is present throughout the brain but with lower expression in the thalamus and hypothalamus, highlighting regionally modulated cholinergic signaling [31]. Collectively, these patterns reflect both cortical and subcortical specialization, including excitatory, inhibitory, and modulatory neuronal circuits, as well as astrocytic distribution.

In the 10x Visium human dorsolateral prefrontal cortex (DLPFC) sample, Spartan identifies layer- and white matter-specific SVGs that recapitulate canonical cortical architecture (Fig. 4F; Supplementary Fig. 15D). *MBP* is highly expressed in the white matter (WM), marking oligodendrocyte populations [20], while *GFAP* is enriched in both WM and layer 1, highlighting astrocytic distribution [32] across superficial cortex and subcortical tracts. Layer-specific genes such as *COX6C* and *CST3* exhibit selective expression in combinations of deep and superficial layers (layers 1, 2, 5, and 6), whereas *FBXL16* is detected across all cortical layers except WM, indicating broad gray matter enrichment. *SCGB2A2*, although widely expressed, shows absence in portions of the left upper layers 1–4, reflecting subtle spatial heterogeneity. These findings indicate that Spartan resolves both classical layer-specific markers and fine-scale lateralized expression patterns in the human cortex.

Across all four datasets, Spartan consistently identifies genes with spatial specificity corresponding to known cell types, tissue structures, and functional domains. In the MERFISH mouse hypothalamus sample, the distinction between GABAergic nuclei, white matter tracts, and ependymal zones is clearly captured. In the Stereo-seq mouse embryo sample, tissue-specific expression reflects organogenesis and left–right patterning. In the 10x Visium human DLPFC sample, cortical layer specificity and white matter enrichment are resolved, while in the Vizgen MERFISH mouse brain dataset, regional specialization of neurons and astrocytes is detected across cortical, hippocampal, thalamic, and subcortical structures.

Notably, Spartan also captures subtle spatial heterogeneity and lateralization. Examples include *Nppa* restricted to the left heart, partial absence of *SCGB2A2* in upper cortical layers, and lateralized expression of *COX6C* in DLPFC layers. Such fine-scale patterns illustrate Spartan's sensitivity to both broad regional patterns and localized deviations.

Overall, Spartan's detection of spatially variable genes across multiple organs and developmental stages demonstrates robust spatial pattern detection consistent with known biology. It accurately resolves canonical markers, layer- and region-specific genes, and subtle spatial heterogeneities, making it a powerful platform for mapping molecular architecture in complex tissues.

Spartan preserves transferability of detected SVGs

Next, we evaluated Spartan's ability to preserve the transferability of spatially variable genes (SVGs). A gene identified as spatially variable in one SRT sample should exhibit a similar spatial expression pattern in other samples from the same tissue type [14]. To test this, two different MERFISH samples representing the preoptic region of the mouse hypothalamus were used. The SVGs detected in one sample, *Nnat*, *Oxt*, *Gad1*, *Sln*, and *Mbp*, exhibited similar non-random spatial expression patterns in the other sample (Supplementary Fig. 16A-B). Importantly, SVGs that defined spatial domains in one sample also defined the same spatial domains in the other sample. For instance, *Mbp* showed high expression in the fx region in both samples, while *Sln* was restricted to the V3, MPN, BST, and PVT domains in both. Similarly, the presence of *Oxt* in the PVH region was preserved across samples. These results highlight Spartan's ability to identify reproducible SVGs and accurately capture non-random spatial patterns.

Visium HD analysis: Spartan resolves anatomical layering and functional compartmentalization in the developing esophagus and stomach

To evaluate Spartan in a novel and biologically complex setting, we applied it to an in-house high-resolution Visium HD section of the developing human esophagus and stomach at 10 weeks of gestation. This tissue comprises multiple epithelial cell types, concentric smooth muscle layers, heterogeneous stromal compartments, and vascular niches arranged along curved and continuous anatomical axes. The gastroesophageal junction (GEJ) represents a particularly challenging transitional region in which epithelial identity and stromal programs shift gradually rather than forming sharply demarcated boundaries. This structured resolution of a gradual transition zone suggests that activation-aware integration can partition gradual molecular gradients into spatially coherent domains, while retaining fine-scale substructure that is often lost under purely spatial smoothing.

Spartan resolved spatial domains across the entire specimen that aligned with known anatomical organization, including precise delineation of the GEJ (Fig. 5A). For this analysis, parameters were configured to emphasize the LSA component (**L**), while retaining spatial contiguity (**S**) and expression similarity (**G**), enabling detection of spatially coherent transcriptional programs within fine-grained tissue heterogeneity. Multiple domains were supported by domain-restricted spatially variable genes (SVGs), whereas additional domains reflected anatomically meaningful microstructures captured through local transcriptomic sensitivity. Notably, Spartan identified the GEJ, outer muscularis layer, esophageal mesenchyme compartment, esophageal luminal epithelium, gastric mucosal epithelium, and three quantitatively stratified gastric stromal compartments characterized by graded enrichment of extracellular matrix, glycosaminoglycan, vascular, and immune-associated programs as coherent spatial domains. Compared to NichePCA (Fig. 5B), Spartan preserved anatomically consistent boundaries and spatial continuity along curved tissue structures, consistent with more localized representation of regionally enriched programs.

Spartan resolved the GEJ as domain 5 (Fig. 5A), defined by the presence of SVGs including *PLA2G2A*, gastric-associated mucins (*MUC5AC*, *MUC5B*), and squamous epithelial-associated markers such as the transcription regulator *ELF3*, and *TACSTD2* [33, 34] (Supplementary Fig. 17A-G). Co-expression of gastric-associated mucins alongside esophageal markers like *TACSTD2* supports a mixed transcriptional identity consistent with a junctional transition zone. Notably, *PLA2G2A* exhibited selective enrichment within this domain, whereas NichePCA distributed the signal across multiple adjacent partitions (Fig. 5A-B), further supporting more localized representation of regionally enriched programs by Spartan. *PLA2G2A* has been reported to function as an antimicrobial effector and modulator of gut microbiota, supporting epithelial barrier associated functions in the gastrointestinal tract [35]. Together, these findings are consistent with activation-aware integration contributing to stable partitioning at transitional interfaces.

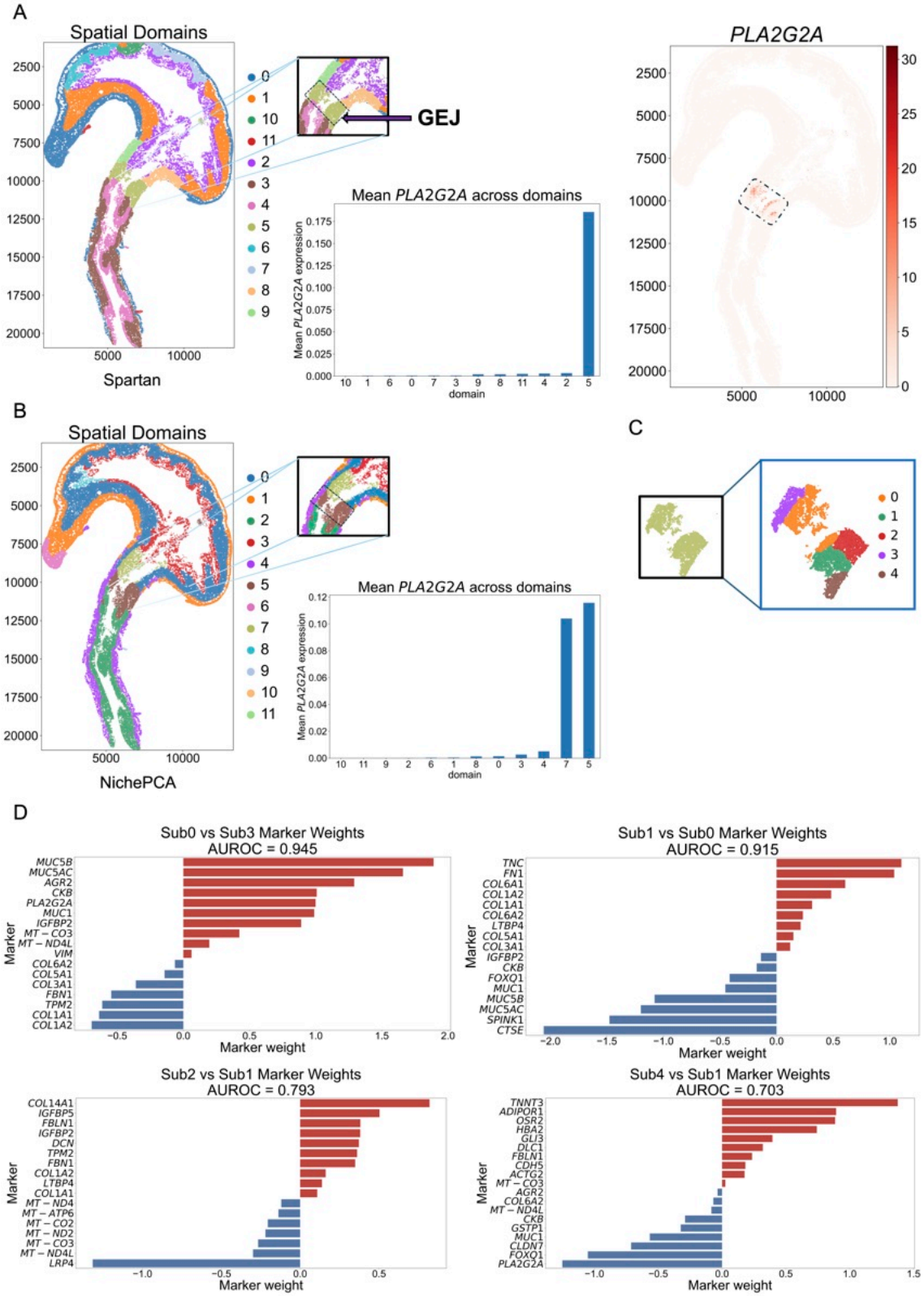

Figure 5. High-resolution domain reconstruction and GEJ subdomain stratification in Visium HD. (A–B) Spatial domains identified by Spartan (A) and NichePCA (B) on a Visium HD section of the developing human esophagus and stomach using 12 domains. Spartan preserves spatially coherent domains in anatomically complex regions, including the gastroesophageal junction (GEJ; Domain 5). Insets highlight the GEJ and spatial expression of *PLA2G2A*, with corresponding mean expression across domains. Domain colors are for visualization only and are not comparable across methods. (C–D) Supervised subdomain analysis within the GEJ. Subdomains were identified by reapplying Spartan to Domain 5 only. Marker weights correspond to logistic-regression coefficients from pairwise subdomain classification, and AUROC values (0.703–0.945) quantify subdomain separability.

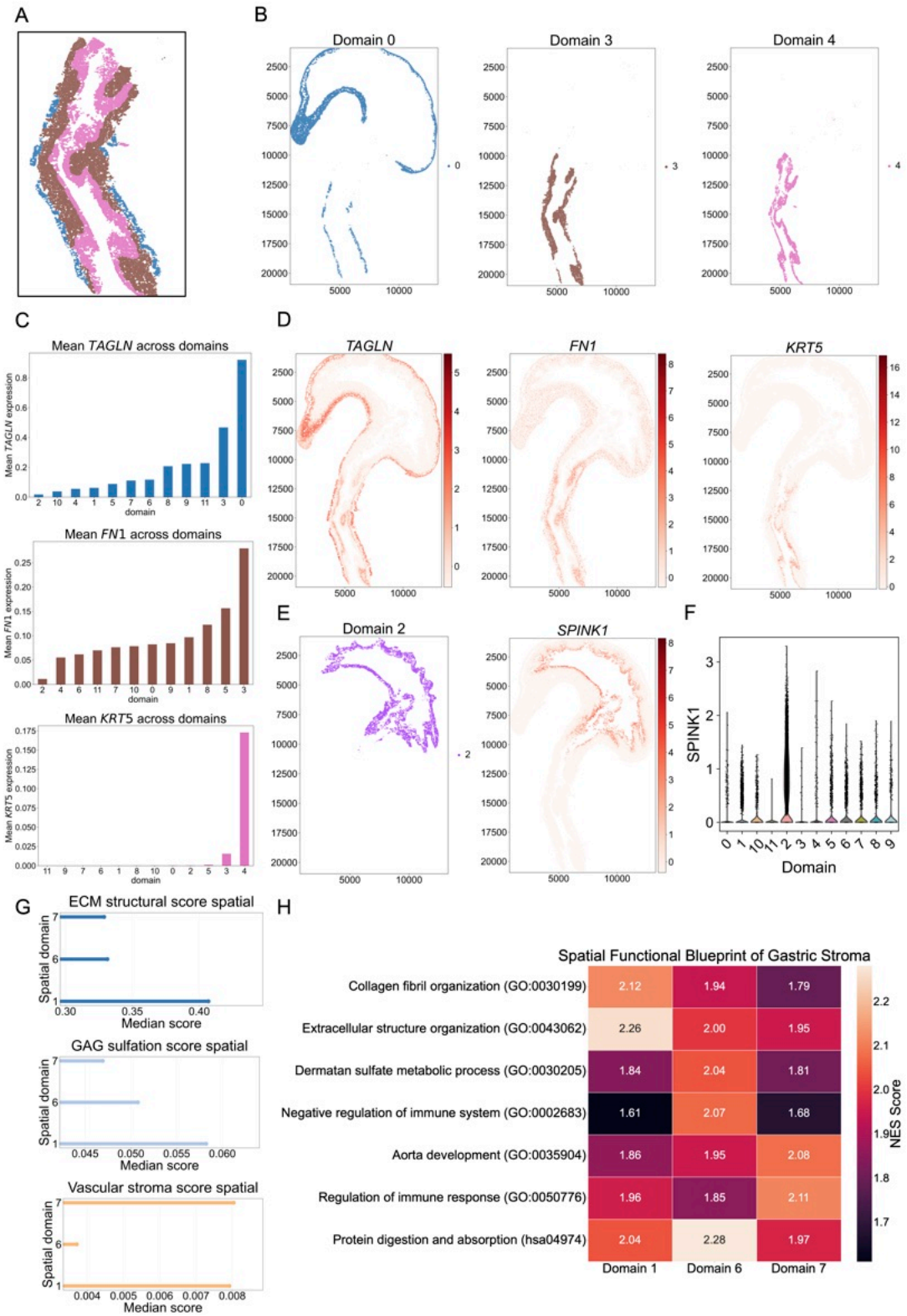

Figure 6. Reconstruction of esophageal laminar architecture and stromal functional programs using Visium HD. (A) Spartan resolves the concentric laminar organization of the developing esophageal wall, delineating epithelial, stromal and muscular compartments. (B) Spatial domains corresponding to outer muscularis (Domain 0), stromal interface (Domain 3), and luminal epithelium (Domain 4). (C) Mean expression of canonical markers across domains. *TAGLN* is enriched in Domain 0, *FN1* in Domain 3, and *KRT5* in Domain 4. (D) Spatial expression maps of *TAGLN*, *FN1*, and *KRT5* align with domain boundaries. (E) Gastric mucosal epithelium (Domain 2) with localized *SPINK1* expression. (F) Violin plot confirming *SPINK1* enrichment in Domain 2. (G) Median spatially smoothed module scores across stromal domains (Domains 1, 6, 7) reveal continuous functional gradients across adjacent compartments. (H) Gene set enrichment heatmap summarizing domain-associated programs: Domain 1 shows extracellular matrix enrichment, Domain 6 sulfated proteoglycan and immune-regulatory pathways, and Domain 7 vascular-associated signaling.

To further characterize spatial structure within the GEJ, we recursively applied Spartan within domain 5 to identify subdomains capturing localized transcriptional programs (Fig. 5C). Adjacent subdomains were differentiated using pairwise differential expression (DE) analysis, and discriminative marker sets were evaluated using logistic regression, yielding AUROC scores ranging from 0.703 to 0.945, indicating robust separability.

Subdomain 0 (Sub0) represented a secretory gastric epithelial-like compartment oriented towards the lumen, reflected by mucin-associated expression patterns (Supplementary Fig. 17B-C). Pairwise comparison against a left-side stromal subdomain (Sub3) yielded high discriminative performance (AUROC=0.945), consistent with distinct epithelial and collagen-rich stromal molecular states (Fig. 5D). On the right side of the GEJ, a remodeling-associated stromal compartment (Sub1) was distinguished by extracellular matrix reorganization programs (AUROC=0.915). Additional right-side subdomains captured deeper stromal niches with separable signaling and infrastructure associated transcriptional profiles (AUROC range 0.703–0.793)(Fig. 5D), illustrating the ability of Spartan to resolve nested spatial microenvironments within a continuous transition zone. This analysis revealed asymmetry across the junction, with a coherent mucosal and stromal interface on the left side, and a multi-layered arrangement with distinct matrix remodeling and signaling associated compartments on the right side.

Beyond the GEJ, Spartan resolved the concentric multilaminar architecture of the fetal esophagus (Fig. 6A), identifying an outer muscularis layer (domain 0), an intermediate fibroblast-rich mesenchymal stroma (domain 3), and an inner luminal epithelium (domain 4) (Fig. 6B). In contrast, NichePCA grouped the luminal epithelium and intermediate stroma into a single domain. The muscularis externa (domain 0) formed a coherent contractile scaffold defined by smooth muscle associated SVGs, anchored by *TAGLN* [36] (Fig. 6B-D; Supplementary Fig. 18A-E). Quantitative analysis confirmed maximal enrichment within domain 0 and reduced enrichment at the GEJ, consistent with distinct molecular prioritization at the junction relative to surrounding muscle (Fig. 6C; Supplementary Fig. 18F-G). Domain 3 was supported by extracellular matrix associated stromal programs anchored by *FN1*, consistent with a fibroblast-rich supportive compartment (Fig. 6B-D; Supplementary Fig. 19). Domain 4 was supported by epithelial stratification associated expression patterns, including *KRT5*, a canonical basal epithelial marker in stratified epithelia [37] (Fig. 6B-D; Supplementary Fig. 20).

Spartan also identified the gastric mucosal epithelium (domain 2), characterized by a spatially coherent secretory program and epithelial identity (Fig. 6E). This compartment was supported by enrichment of a set of gastric epithelial marker genes including the gene encoding the mucosal protective protein *SPINK1* [38, 39] (Fig. 6E-F; Supplementary Fig. 21).

Surrounding this epithelium, Spartan resolved the gastric stroma into three specialized spatial domains, capturing regional heterogeneity that NichePCA grouped as a single domain. Domain 1 represents the structural framework of the gastric corpus, exhibiting the highest median extracellular matrix (ECM) and glycosaminoglycan (GAG) sulfation scores (Fig. 6G-H). Domain 6 represents the antral reservoir, showing peak enrichment for protein digestion and the dermatan sulfate metabolic process, while also contributing to immune regulation (Fig. 6G-H). Domain 7 is associated with deeper submucosal layers and is defined by programs linked to vascular development and immune response regulation (Fig. 6G-H).

Beyond coarse domain recovery, Spartan identifies localized stromal microstructures that provide a mechanistic justification for sub-compartmentalization. Domain 10, a spatially restricted niche embedded within gastric stroma, exhibits strong enrichment for erythroid-associated genes such as *SLC4A1* and *HBA2* (Supplementary Fig. 22A-D). The corresponding functional enrichment patterns highlight oxygen transport and heme-related biology (Supplementary Fig. 22E) supporting the interpretation that this region represents a blood-associated microenvironment rather than a generic fibroblast compartment. Stromal regions are frequently grouped into broad partitions by neighborhood-based approaches, whereas activation-aware integration supports separation of smaller but biologically coherent niches.

Similarly, the analysis of Domains 8 and 9 supports Spartan's ability to distinguish closely adjacent stromal subregions on either side of the lumen at the entrance to the stomach following the gastroesophageal junction (Supplementary Fig. 23A). While both domains belong to a broadly stromal context, their differential signatures suggest that they encode distinct microenvironments rather than noisy fragmentation (Supplementary Fig. 23B). Domain 8 shows strong enrichment of extracellular matrix-associated markers, including fibrillar collagens and matrix remodeling genes, exemplified by *FBLN1*, consistent with a structural stromal compartment (Supplementary Fig. 23B and Supplementary Fig. 23C). In contrast, Domain 9 is characterized by reduced levels of these markers but moderately enriched vascular-associated transcripts such as *PECAM1*, *CLDN5*, *PLVAP*, and *EGFL7*, supporting a vessel-associated stromal niche (Supplementary Fig. 23B and Supplementary Fig. 23D). Logistic regression based discrimination provides a quantitative measure of separability between these domains, further strengthening the argument that Spartan is resolving functional microstructure rather than producing arbitrary clusters (Supplementary Fig. 24). Together, these results demonstrate that Spartan can resolve spatially coherent anatomical layers and functionally distinct stromal compartments in a complex developing tissue.

To further assess whether the gene programs underlying the spatial domains identified by Spartan reflect meaningful spatial structure, we evaluated the spatial properties of SAQ-ranked genes within this dataset.

Validation of SAQ-based SVGs in Visium HD sample

To validate the spatial relevance of SAQ-based gene ranking, we evaluated whether genes with high SAQ scores exhibit stronger spatial autocorrelation than background genes (Supplementary Methods; Supplementary Fig. 25). Moran's I statistics were computed for all genes using the spatial neighbor graph. The mean Moran's I for top-ranked genes was 0.3048, whereas candidate background genes exhibited a mean Moran's I of 0.0153, indicating stronger spatial autocorrelation among SAQ-ranked genes.

The distribution of Moran's I values for top-ranked SAQ genes was compared with that of candidate background genes using a Mann–Whitney U test (Supplementary Fig. 25), yielding a highly significant difference ($p \approx 1.2 \times 10^{-19}$). Enrichment of significantly spatially autocorrelated genes among the top SAQ genes was further evaluated using Fisher's exact test (p-value $= 5.68 \times 10^{-7}$, odds ratio $= \infty$) (Supplementary Fig. 25).

Finally, spatial expression maps of the top six SAQ-ranked genes illustrate localized expression patterns consistent with spatial structure (Supplementary Fig. 26).

Discussion and conclusion

Spatial transcriptomics is rapidly transitioning from coarse tissue maps to high-resolution profiling, where tens of thousands of spatial locations capture subtle biological structure along continuous anatomical axes. This shift exposes a computational limitation: many clustering approaches separate broad compartments but blur tissue boundaries, merge transitional zones, or miss localized microstructures that are biologically meaningful, particularly in developmental tissue where programs vary gradually yet change sharply at interfaces. Spartan is designed to address this challenge through an aggregated graph formulation that integrates spatial topology, Local Spatial Activation (LSA), an activation-aware component designed to capture localized transcriptomic deviation relative to neighborhood context, complementing spatial topology and gene expression connectivity.

Conceptually, LSA draws inspiration from local spatial statistics in that it evaluates neighborhood context; however, unlike classical autocorrelation measures, LSA is formulated as a non-negative graph signal in PCA-reduced space, enabling direct integration into weighted community detection frameworks. Within transcriptionally coherent domains, LSA values tend to be low because neighboring locations exhibit minimal deviation from their local context. At spatial transitions or nested microstructures, LSA increases when joint deviation from the neighborhood exceeds background variance. When integrated with gene expression connectivity, this activation-aware signal contributes additional contrast at localized transitions, influencing community detection toward partitions that better respect anatomical interfaces.

Spartan remains computationally practical as datasets increase in resolution and gene panel size (Supplementary Figs. 27–28). Runtime scales approximately linearly with the number of spatial locations, and high-resolution datasets such as Visium HD are processed within minutes on CPU. Efficiency derives from sparse local graphs and linear-time community detection, avoiding dense pairwise computations or model training. These properties position Spartan as a computationally efficient and scalable framework that supports high-resolution datasets without reliance on model training or specialized hardware.

A practical advantage of Spartan is that it does not require ground-truth labels. Ground truth was used primarily for benchmarking and ablation studies. In unlabeled datasets, Spartan can be applied by varying the Leiden resolution to obtain partitions at the desired level of biological granularity. Optional stability-based pruning metrics may further constrain the search space but should be viewed as workflow guidance rather than definitive unsupervised proof of a universally optimal partition.

Spartan was designed for robustness to parameterization. Clustering performance varies smoothly across a broad range of multiplex weight allocations (β_1, β_2) without sharp optima (Supplementary Fig. 29), indicating that improvements are not driven by fragile tuning. For benchmarking, a single dataset-level α was selected using a cross-sample consensus procedure and subsequently fixed, avoiding per-sample optimization (Supplementary Fig. 30); performance remains stable across a wide α range. The smooth dependence on multiplex weights clarifies the role of LSA: its contribution can be increased without destabilizing clustering, which is particularly advantageous in high-resolution datasets with sharp spatial transitions, such as Visium HD and Vizgen MERFISH, where increased activation weighting enhances sensitivity to localized transitions while preserving spatial coherence. In this framework, LSA provides a structured signal that highlights localized transcriptional transitions within spatial neighborhoods.

We demonstrate Spartan's biological relevance using a cell-level Visium HD section of developing human esophagus and stomach at 10 weeks of gestation, containing layered smooth muscle, fibroblast-rich stroma, glandular epithelium, vascular niches, and the complex gastroesophageal junction (GEJ). Spartan resolves spatial domains across the specimen and separates the GEJ as a distinct transitional territory consistent with its mixed molecular identity. Recursive application within the GEJ reveals asymmetrically organized subdomains, supporting the interpretation that activation-aware graph integration enhances sensitivity to nested microenvironments in continuous tissue regions relative to expression-only or topology-only integration.

To assess robustness to clustering granularity, we varied the Leiden resolution for the Visium HD sample to obtain cluster counts spanning coarse to fine partitioning ($R = 7$–20) while fixing all other hyperparameters (Supplementary Figs. 31-32). Spartan exhibits stable behavior across this range. At low resolution ($R = 7$), it collapses into spatially contiguous macro-domains reflecting global tissue organization while preserving the GEJ as a coherent transition region. At higher resolution ($R = 16$–20), it introduces local subdomains through hierarchical refinement without disrupting anatomical continuity or fragmenting the GEJ. In contrast, NichePCA shows reduced anatomical alignment at low resolution and increasingly fragmented or stripe-like partitions at higher cluster counts.

Finally, Spartan extends beyond domain discovery to spatially variable gene detection. By leveraging activation-aware structure, Spartan identifies genes associated with localized transitions and domain-restricted programs, supporting interpretability and enabling mechanistic follow-up. Spartan therefore functions not only as a clustering method but also as a framework for investigating how spatial organization relates to localized transcriptional programs.

In conclusion, Spartan introduces an activation-aware multiplex graph framework that integrates spatial topology, Local Spatial Activation, and gene expression connectivity. The method is robust to parameter variation while permitting adaptive weighting in high-resolution settings where microdomain boundaries are critical. By linking Local Spatial Activation to principles of local spatial autocorrelation in geostatistics, Spartan provides an interpretable and extensible framework for spatial systems-level analysis, with potential applications in cell–cell communication, developmental tissue remodeling, and spatial multi-omics integration.

Methods

Data preprocessing

Raw SRT data were processed using standard workflows described previously [40, 41]. Gene expression matrices were library-size normalized using `Scanpy`'s *pp.normalize_total* and log-transformed using *pp.log1p* [40]. The resulting filtered log-normalized matrix is denoted as $\mathbf{X} \in \mathbb{R}^{N \times M}$, where N represents the number of spots (or cells) and M the number of genes. For sequencing-based technologies such as 10x Visium,

Visium HD, and Stereo-seq, highly variable genes (HVGs) were selected using `Scanpy`'s *pp.highly_variable_genes* functionality [40]. Throughout this work, each row $\mathbf{X}_i$ is referred to as a spot; the formulation applies equally when $\mathbf{X}_i$ represents a cell (e.g., MERFISH).

Principal component analysis (PCA) was applied to $\mathbf{X}$ to obtain a lower-dimensional representation of spots in principal-component space, $\mathbf{Y^{pca}} \in \mathbb{R}^{N\times P}$, where $P \ll M$ is the number of principal components. The feature matrix is represented as

$$\mathbf{Y^{pca}} = \begin{bmatrix} y^{pca}_{1,1} & y^{pca}_{1,2} & \cdots & y^{pca}_{1,P} \\ \vdots & \vdots & \ddots & \vdots \\ y^{pca}_{N,1} & y^{pca}_{N,2} & \cdots & y^{pca}_{N,P} \end{bmatrix}. \tag{1}$$

Equivalently, $\mathbf{Y^{pca}}$ can be expressed in terms of row vectors:

$$\mathbf{Y^{pca}} = \begin{bmatrix} \mathbf{y^{pca}_1} \\ \vdots \\ \mathbf{y^{pca}_N} \end{bmatrix}, \quad \mathbf{y^{pca}_i} \in \mathbb{R}^{1\times P}. \tag{2}$$

Spartan is also implemented within the `SpatialData` framework [42], enabling standardized multimodal I/O and interoperability across spatial omics data types.

Spatial graph

The spatial graph is expressed as $gr^{\mathrm{s}} = (V_{\mathrm{s}}, E_{\mathrm{s}})$ where $V_{\mathrm{s}} = \{1, 2, \cdots, N\}$ and E_{s} is represented in the form of a binary adjacency matrix, $\mathbf{S} \in \mathbb{R}^{N\times N}$. Spatial neighbor graphs were constructed using `Squidpy` [41]. Each element of the matrix $\mathbf{S}$, s_{ij}, represents:

$$s_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are neighbors based on spatial proximity} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

The spatial proximity is determined by constructing either: (i) a k-nearest neighbor (KNN) graph based on Euclidean distance in 2D space using spatial coordinates, where for each spot i its k nearest spatial neighbors are identified. The KNN construction yields a directed adjacency by default; we optionally symmetrized it depending on the `Squidpy` coordinate setting. Alternatively, (ii) a Delaunay graph is constructed, in which spot i shares edges with neighboring spots defined by the Delaunay triangulation of the spatial coordinates. The Delaunay construction yields an undirected planar graph. Accordingly, $\mathbf{S}$ is generally asymmetric under the directed KNN construction and symmetric under the Delaunay construction. The matrix $\mathbf{S}$ can be represented in terms of row vectors:

$$\mathbf{S} = \begin{bmatrix} \mathbf{s_1} \\ \vdots \\ \mathbf{s_N} \end{bmatrix}, \quad \mathbf{s_i} \in \mathbb{R}^{1\times N} \tag{4}$$

Each vector $\mathbf{s_i} = (s_{i1}, \cdots, s_{iN})$ represents the spatial neighborhood of spot i where each s_{ij} is defined according to Eq. 3. Each row $\mathbf{s_i}$ contains nonzero entries corresponding to the spatial neighbors of spot i (with out-degree K_i in the directed KNN case), and the matrix is stored in sparse format.

Row-normalized spatial weights matrix

The spatial weights matrix, $\mathbf{W} \in \mathbb{R}^{N\times N}$, is constructed where each element of the matrix w_{ij} represents the measure of spatial proximity between spot i and spot j. The spatial weight w_{ij} is calculated as:

$$w_{ij} = \frac{1}{d_{ij}} \cdot s_{ij} \tag{5}$$

where d_{ij} is the Euclidean distance between spatial coordinates of spots i and j. To ensure numerical stability, distances were offset by $\epsilon = 10^{-3}$, i.e., $w_{ij} = \frac{1}{d_{ij}+\epsilon} s_{ij}$. We row-normalized $\mathbf{W}$ to obtain $\mathbf{W^r} = (w^r_{ij})$, where $w^r_{ij} = \frac{w_{ij}}{\sum_{j=1}^{N} w_{ij}}$. The matrix $\mathbf{W^r}$ is stored as a sparse matrix. Distances d_{ij} were computed only for graph-connected pairs (i, j) using the neighbor graph construction utilities in `Squidpy` [41]. The $\mathbf{S}$ provides a binary connectivity structure based on raw spatial proximity, while the row-normalized weight matrix $\mathbf{W^r}$ refines this by accounting for distance. We use the row-normalized matrix $\mathbf{W^r}$ to construct the matrix $\mathbf{L}$ in the next section.

Local Spatial Activation (LSA) graph

The LSA graph is expressed as $gr^{\mathrm{lsa}} = (V_{\mathrm{lsa}}, E_{\mathrm{lsa}})$ where $V_{\mathrm{lsa}} = \{1, 2, \cdots, N\}$ and E_{lsa} is represented as the weighted adjacency matrix, $\mathbf{L} \in \mathbb{R}^{N\times N}$. $\mathbf{L}$ captures neighborhood-relative transcriptional deviation between spatially adjacent spots. Each element l_{ij} gives the LSA weight associated with the spatial edge from spot i to spot j.

We divide the calculation of the LSA weight l_{ij} for a spatial edge $i \to j$ into three components:

Summary statistics of spot i

We calculate two summary statistics: the mean and variance of the closed neighborhood $\mathcal{N}_i$ of spot i. The closed neighborhood $\mathcal{N}_i$ consists of spot i and its K_i spatial neighbors, as determined by the matrix $\mathbf{S}$.

We calculate the mean feature vector of $\mathcal{N}_i$ as:

$$\mu_{\mathcal{N}_{\mathrm{i}}} = \frac{\mathbf{y^{pca}_i} + \sum_{j=1}^{K_i} \mathbf{y^{pca}_j}}{K_i + 1}, \quad \mathbf{y^{pca}_i}, \quad \mathbf{y^{pca}_j} \in \mathbf{Y^{pca}} \tag{6}$$

where $\mathbf{y^{pca}_i}$ and $\mathbf{y^{pca}_j}$ are the feature vectors of spot i and spot j, respectively.

We calculate the local variance and the local variance norm of the neighborhood $\mathcal{N}_i$ as follows.

First, define the deviation vectors:

$$\mathbf{dv_i} = \mathbf{y^{pca}_i} - \mu_{\mathcal{N}_{\mathrm{i}}}, \quad \mathbf{dv^{(i)}_j} = \mathbf{y^{pca}_j} - \mu_{\mathcal{N}_{\mathrm{i}}}, \tag{7}$$

where $\mathbf{dv_i}$ and $\mathbf{dv^{(i)}_j}$ denote deviations from the neighborhood mean.

The neighborhood variance vector is then computed using elementwise products:

$${\sigma_{\mathcal{N}_{\mathrm{i}}}}^2 = \frac{\mathbf{dv_i} \odot \mathbf{dv_i} + \sum_{j=1}^{K_i} \mathbf{dv^{(i)}_j} \odot \mathbf{dv^{(i)}_j}}{K_i + 1}, \quad {\sigma_{\mathcal{N}_{\mathrm{i}}}}^2 \in \mathbb{R}^P \tag{8}$$

where $\odot$ denotes elementwise (Hadamard) multiplication.

The local variance norm of the neighborhood $\mathcal{N}_i$ is defined as

$$\hat{\sigma_{\mathcal{N}_i}} = \|{\sigma_{\mathcal{N}_{\mathrm{i}}}}^2\|_2 \tag{9}$$

The mean feature vector, $\mu_{\mathcal{N}_{\mathrm{i}}} \in \mathbb{R}^P$, summarizes the transcriptional context of the surrounding tissue (neighborhood) and provides a smooth representation of the local expression trends. The local variance norm of the neighborhood $\mathcal{N}_i$, $\hat{\sigma_{\mathcal{N}_i}}$, is defined as the Euclidean norm ($\|\cdot\|_2$) of the neighborhood variance vector ${\sigma_{\mathcal{N}_{\mathrm{i}}}}^2$. This ℓ_2 aggregation provides an isotropic summary of multivariate variance across principal components and is invariant to orthogonal rotations of the PCA coordinate system.

Attribute deviation

We measure the attribute deviation of spot i from its surroundings, $\mathcal{N}_i$:

$$\phi_i = \|\mathbf{dv_i}\|_2, \tag{10}$$

The quantity ϕ_i is the attribute deviation norm, measuring how much spot i deviates from its neighborhood in PCA-reduced expression space. Lower values indicate that spot i is transcriptionally similar to its neighborhood mean, whereas higher values indicate stronger local deviation.

Similarly, we calculate the attribute deviation of neighbor j of spot i, from $\mathcal{N}_i$ as:

$$\phi_j^{(i)} = \|\mathbf{dv_j^{(i)}}\|_2, \tag{11}$$

Note that $\phi_j^{(i)}$ is computed with respect to $\mu_{\mathcal{N}_i}$, making $\mathbf{L}$ asymmetric by construction.

Local Spatial Activation (LSA)

The LSA weight associated with the spatial edge $i \rightarrow j$ is defined as:

$$l_{ij} = \frac{\phi_i \, w_{ij}^r \, \phi_j^{(i)}}{\sigma_{\hat{\mathcal{N}}_i} + \epsilon_l}, \quad l_{ij} \geq 0 \tag{12}$$

Here, l_{ij} consists of three key components: the quantities ϕ_i and $\phi_j^{(i)}$ measure the magnitude of transcriptional deviation of spots i and j relative to the local neighborhood of spot i, while w_{ij}^r denotes the row-standardized spatial proximity weight between spots i and j. The small quantity ϵ_l ensures numerical stability. The local variance norm $\sigma_{\hat{\mathcal{N}}_i}$ normalizes LSA values relative to local neighborhood variability. LSA is formulated as a non-negative activation signal on the spatial neighbor graph rather than as a statistical test of spatial autocorrelation. Thus, l_{ij} encodes a variance-normalized, spatially weighted joint transcriptional deviation relative to the local neighborhood context of spot i within gr^{lsa}. The resulting matrix $\mathbf{L}$ is therefore an asymmetric weighted adjacency matrix and serves as a key component integrated into the aggregated graph (Eq. 15), where it is balanced against gene expression connectivity and spatial proximity.

Gene expression connectivity graph

We construct a gene expression connectivity graph in the form of a KNN graph in the PCA space, gr^{g}. The gr^{g} is expressed as $(V_{\text{g}}, E_{\text{g}})$ where $V_{\text{g}} = \{1, 2, \cdots, N\}$ and E_{g} is represented by the binary adjacency matrix, $\mathbf{G} \in \mathbb{R}^{N \times N}$. The gr^{g} encodes the neighborhood relationship between two spots based on their transcriptomic similarity. The transcriptomic distance in PCA space between spots i and j, ψ_{ij}^g, was computed as:

$$\psi_{ij}^g = \|\mathbf{y_i^{pca}} - \mathbf{y_j^{pca}}\|_2, \quad \mathbf{y_i^{pca}}, \quad \mathbf{y_j^{pca}} \in \mathbf{Y^{pca}} \tag{13}$$

The adjacency matrix $\mathbf{G}$ encodes directed KNN neighborhood connectivity. Each entry of $\mathbf{G}$, denoted g_{ij}, is defined as:

$$g_{ij} = \begin{cases} 1, & \text{if } j \text{ is among the } k\text{-nearest neighbors of } i \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Here, ψ_{ij}^g is used to identify the k nearest neighbors for each spot in PCA space. The adjacency matrix $\mathbf{G}$ forms the gene expression connectivity component of the aggregated graph (Eq. 15), complementing LSA and raw spatial proximity.

Aggregated graph

The aggregated graph gr^{mul} is a directed graph that combines three components: gr^{s}, gr^{lsa}, and gr^{g}. The aggregated graph, gr^{mul}, is expressed as $(V_{\text{mul}}, E_{\text{mul}})$ where $V_{\text{mul}} = \{1, 2, \cdots, N\}$ and E_{mul} is represented by an asymmetric weighted adjacency matrix $\mathbf{J} \in \mathbb{R}^{N \times N}$ and is formed as:

$$\mathbf{J} = (\alpha - \beta_1) \times \mathbf{L} + (1 - \alpha) \times \mathbf{G} + (\alpha - \beta_2) \times \mathbf{S} \tag{15}$$

where $\alpha \in (0, 1)$ and $\beta_1 + \beta_2 = 0.5$ with $\beta_1, \beta_2 \leq \alpha$. In this formulation, the matrix $\mathbf{J}$ balances three complementary sources of structure: $\mathbf{L}$ that encodes LSA information of all spots, the gene expression connectivity matrix $\mathbf{G}$, and the raw spatial adjacency matrix $\mathbf{S}$. An entry j_{ij} of the matrix $\mathbf{J}$ is high when spot j belongs to the neighborhood $\mathcal{N}_i$ (the neighborhood of spot i) and is supported either by high transcriptomic similarity or strong joint deviation relative to the neighborhood $\mathcal{N}_i$. The parameter α regulates the overall trade-off between spatial and expression-derived connectivity. At the same time, the correction terms β_1 and β_2 ensure that the combined weight of the structural terms $\mathbf{L}$ and $\mathbf{S}$ equals $2\alpha - 0.5$, thereby controlling the total structural contribution relative to the expression-based connectivity term $\mathbf{G}$. We set $\beta_1 + \beta_2 = 0.5$ so that α controls the overall structural–transcriptomic trade-off while β_1 and β_2 set the relative allocation between $\mathbf{L}$ and $\mathbf{S}$. With β_1 and β_2 fixed within each experimental configuration, α is the sole free parameter governing this trade-off. This yields a balanced integration of spatial and expression-derived information. No additional per-layer rescaling was applied. Since $s_{ij}, g_{ij} \in \{0, 1\}$ and $l_{ij} \geq 0$, $\mathbf{J}$ is non-negative, satisfying the requirements of modularity-based community detection.

Spatial biological partitioning

We use the Leiden clustering algorithm to partition the aggregated graph $gr^{\text{mul}} = (V_{\text{mul}}, E_{\text{mul}})$. We use the RBConfigurationVertexPartition quality function (`leidenalg`), which supports directed graphs. The objective $Q(\mathbf{J}, \gamma)$ uses the weighted adjacency matrix $\mathbf{J}$ and resolution parameter γ. It partitions V_{mul} into R non-overlapping clusters $\{C_1, C_2, \cdots, C_R\}$. The quality function $Q(\mathbf{J}, \gamma)$ maximizes the total edge weight within clusters relative to a configuration-model null. In this case, we refer to these clusters as spatial domains.

In practice, the resolution parameter γ controls the granularity of spatial domains. Together with the layer parameters $(\alpha - \beta_1, 1 - \alpha, \alpha - \beta_2)$, it allows flexible adjustment across SRT datasets of varying spatial scales and technologies. This approach allows the algorithm to leverage the connectivity information encoded by $\mathbf{S}$ and $\mathbf{G}$, which provide neighborhood guidance, and the matrix $\mathbf{L}$, which introduces neighborhood-relative activation structure not captured by spatial proximity or expression-derived connectivity alone. Modularity optimization on $\mathbf{J}$ therefore favors partitions that concentrate activation-weighted and connectivity-supported interactions within clusters while penalizing spurious cross-domain links. This formulation enables the detection of spatial domains that reflect both spatial cohesion and localized transcriptional heterogeneity.

α-selection strategy

We formulated a hierarchical α-selection strategy for determining the multiplex weighting parameter α of the matrix $\mathbf{J}$. This α-selection procedure was used only in benchmarking and ablation on annotated datasets, and was not used in unlabeled biological analyses. Annotations were used only to set dataset-level hyperparameters for evaluation.

For a given multi-sample SRT dataset $D = \{\mathrm{Sm}_1, \cdots, \mathrm{Sm}_T\}$, where Sm_i denotes the ith sample and T is the total number of samples, the α-selection strategy consists of four stages:

1. **Range selection.** For each sample Sm_i in D, we defined the search ranges of the multiplex weighting parameter $\alpha \in [\alpha_l, \alpha_r]$ and the resolution parameter $\gamma \in [\gamma_l, \gamma_r]$. For every candidate $\alpha_i \in [\alpha_l, \alpha_r]$, we performed Leiden clustering across all γ values, yielding a set of clustering assignments $C^{\alpha_i} = \{C^{\alpha_i}_{\gamma_l}, \cdots, C^{\alpha_i}_{\gamma_r}\}$.
2. **Cluster-count alignment.** For benchmarking datasets with curated annotations, we retained clustering assignments whose number of clusters matched the annotated domain count, forming $C^{(1)\alpha_i} \subset C^{\alpha_i}$. This cluster-count alignment follows standard SDMBench benchmarking conventions. For each $C^{(1)\alpha_i}_j \in C^{(1)\alpha_i}$, we computed the evaluation metrics ARI, NMI, HOM, COM, together with nLSAS.
3. **Spatial-stability prior.** We further refined $C^{(1)\alpha_i}$ by retaining only clustering assignments whose nLSAS values lay within the central stability band (40th–70th percentiles by default). The nLSAS score is fully unsupervised and derived solely from the dataset-specific LSA graph and the predicted partition structure. This percentile-based filtering removes spatially fragmented partitions (low nLSAS) and over-smoothed partitions (high nLSAS), thereby favoring configurations that balance activation coherence with domain resolution.

   The resulting pruned set $C^{(2)\alpha_i} \subset C^{(1)\alpha_i} \subset C^{\alpha_i}$ was used to compute the median of each evaluation metric across all $C^{(2)\alpha_i}_j \in C^{(2)\alpha_i}$.

   For each sample Sm_i, we thus obtained a collection of candidate α values and their corresponding median performance summaries, stored in a table $\mathcal{T}_{\mathrm{Sm}_i}$, where each row represents one α_i and its associated summaries. Each table $\mathcal{T}_{\mathrm{Sm}_i}$ was ranked by median NMI, and only the top K candidates were retained to mitigate the influence of outliers (default K=50). The resulting table $\mathcal{T}'_{\mathrm{Sm}_i}$ therefore contains high-performing structural weight configurations for that sample. Although ranking was performed using median NMI for stability, final performance evaluation was conducted using multiple metrics, as described in the benchmarking Results section Spartan demonstrates competitive and statistically supported performance on public SRT datasets.

   Repeating Steps 1–3 for all samples in D yields a collection of tables $\mathcal{T} = \{\mathcal{T}'_{\mathrm{Sm}_1}, \mathcal{T}'_{\mathrm{Sm}_2}, \cdots, \mathcal{T}'_{\mathrm{Sm}_T}\}$.
4. **Cross-sample consensus.** Let $\mathbb{S}_{\alpha_{\mathrm{Sm}_i}}$ denote the set of α values contained in $\mathcal{T}'_{\mathrm{Sm}_i} \in \mathcal{T}$. We computed the intersection of all such sets to obtain a consensus set of α values, $\mathbb{S}_{\alpha_{\mathrm{cm}}}$, thereby enforcing cross-sample generalizability by restricting selection to α values that were consistently stable across samples (when the intersection was non-empty). In cases where the consensus intersection is empty, the union of top-K candidates is used with cross-sample averaging.

   For each $\alpha_{\mathrm{cm}_i} \in \mathbb{S}_{\alpha_{\mathrm{cm}}}$, we extracted its median performance summaries from all $\mathcal{T}'_{\mathrm{Sm}_i}$ tables and calculated the mean of the median NMI values across samples. The α_{cm_i} yielding the highest mean of the per-sample median NMI values under this criterion was designated as α_{best}.

   The resulting α_{best} represents a dataset-level structural weighting parameter for the matrix $\mathbf{J}$. Once selected, α_{best} was fixed for all samples within that dataset and not re-optimized per sample or per resolution.

Parameter selection

Ablation study

The k value for the KNN graph was set to 10 for constructing the spatial graph and to 15 for constructing the gene expression connectivity graph. The parameter α, which balances spatial and transcriptomic contributions, was varied between 0.5 and 0.9. Two additional parameters, β_1 and β_2, controlled the relative weighting of the LSA and spatial neighborhood connectivity components; these were fixed at 0.26 and 0.24, respectively, to emphasize spatial connectivity slightly more strongly. The resolution parameter γ for the Leiden algorithm was varied between 0.5 and 0.9. For all datasets, the number of ground truth clusters was provided as input for evaluation.

Benchmarking on SDMBench

The benchmarking parameters for Spartan were configured as follows. In feature selection and graph construction, the number of principal components (P) was fixed at 30 for all datasets. A total of 3,000 highly variable genes (HVGs) were selected, following [10], with HVGs required only for the Stereo-seq dataset. The k value for constructing the gr^s KNN graph was set to 10, and the gr^g KNN graph was constructed using 15 nearest neighbors. Except for the Stereo-seq dataset, the number of immediate neighbors was set to 4 and the ring size was set to 1.

Parameter ranges were defined as described in Section α-selection strategy. The weight parameter α was bounded between 0.5 and 0.9 across datasets. The parameters β_1 and β_2 were fixed at 0.26 and 0.24, respectively. The resolution parameter γ was varied between 0.5 and 0.9 for most datasets but extended to the range [0.5, 1.2] for osmFISH and Stereo-seq. The percentile thresholds for nLSAS were fixed at the 40 and 70 percentiles for all datasets, except for STARmap where they were fixed at 40 and 80.

Benchmarking against recently proposed methods

We used the same benchmarking pipeline for Spartan to evaluate its performance on five SDMBench datasets as discussed above. To evaluate NichePCA and BANKSY on SDMBench datasets, we adopted a similar benchmarking pipeline as detailed in [9]. Specifically, for each dataset, the number of nearest neighbors was varied between 5 and 20, and the value yielding the highest mean NMI across samples within that dataset was selected. For SpatialLeiden, we chose their best-performing configuration as reported in their SDMBench benchmarking, namely the configuration that uses a KNN graph to represent spatial connectivity. We adopted their technology-specific layer weight ratios as used in their SDMBench benchmarking.

For the Vizgen MERFISH dataset, we employed a different benchmarking workflow similar to that used in NichePCA. Specifically, we restricted the maximum number of clusters to 60, consistent with the evaluation protocol used in NichePCA. The range selection step was similar to Section Benchmarking on SDMBench, but here α was varied between 0.77 and 0.87, γ was varied between 2.0 and 4.0 for the Delaunay configuration and between 1.0 and 2.0 for the KNN configuration. The parameters β_1 and β_2 were fixed at 0.20 and 0.30, reflecting the higher spatial resolution characteristic of imaging-based datasets. For a given configuration (KNN/Delaunay), we selected a single dataset-level α by, for each candidate α, taking the maximum NMI achieved over the predefined resolution sweep (subject to the cluster cap), and then selecting the α that maximized this value consistently across samples. Similarly, for NichePCA and BANKSY, we varied the number of nearest neighbors between

5 and 29 and the γ between 1.0 and 2.0 for NichePCA, 2.0 and 4.0 for BANKSY, and selected the dataset-level k yielding the highest NMI across samples under the same resolution sweep. This applied dataset-level parameter selection across methods.

For all benchmarking comparisons, parameter ranges for competing methods were explored within the bounds recommended by their respective publications, and configurations were selected according to their reported benchmarking strategies. No additional dataset-specific supervision was introduced beyond what was described in the respective original studies. This ensures that cross-method comparisons reflect methodological design rather than asymmetric tuning procedures.

Spatial Activation Quotient

To identify genes with non-random spatial expression patterns [spatially variable genes (SVGs)] , we introduce the *Spatial Activation Quotient* (SAQ), denoted by $\mathbf{v^{saq}} \in \mathbb{R}^M$. This metric ranks genes according to the degree to which their expression profiles exhibit localized spatial structure rather than random variation.

We begin with the filtered gene expression matrix $\mathbf{X} \in \mathbb{R}^{N\times M}$, where N is the number of spots and M is the number of genes. Each gene vector is first column-wise mean-centered to remove global expression offsets:

$$\mathbf{X}^{'} = \mathbf{X} - \mu^{\mathbf{g}}, \quad \mu_j^g = \frac{1}{N}\sum_{i=1}^{N} X_{ij}, \quad j = 1, 2, \cdots, M \tag{16}$$

where $\mathbf{X}^{'} \in \mathbb{R}^{N\times M}$ is the mean-centered expression matrix and $\mu^{\mathbf{g}} \in \mathbb{R}^M$ is the gene-wise mean expression vector.

Next, we propagate the mean-centered expression through the LSA graph to obtain an LSA-weighted neighborhood representation of the expression for each gene:

$$\mathbf{X^{sm}} = \mathbf{L}\mathbf{X}^{'}, \quad \mathbf{X^{sm}} \in \mathbb{R}^{N\times M} \tag{17}$$

Here, $\mathbf{X}^{sm}$ represents an LSA-weighted neighborhood aggregation (with weights given by $\mathbf{L}$). $\mathbf{L}$ is the weighted adjacency matrix of the LSA graph. The product $\mathbf{L}\mathbf{X}^{'}$ represents a spatially weighted smoothing of expression values, such that each spot's value reflects its local neighborhood context. For SAQ scoring, $\mathbf{L}$ is treated as a fixed spot–spot operator constructed from PCA-reduced spot embeddings, whereas gene scores are computed in the original mean-centered gene expression space.

We then quantify the alignment between the original expression profile of a gene and its spatially smoothed profile. For gene j, this is given by:

$$v_j^{sp} = \sum_{i=1}^{N} X_{ij}^{'} X_{ij}^{sm}, \quad j = 1, 2, \cdots, M \tag{18}$$

Equivalently, if $\mathbf{x}_j' \in \mathbb{R}^N$ denotes the mean-centered expression vector of gene j across all spots (i.e., the jth column of $\mathbf{X}'$), then

$$v_j^{sp} = \mathbf{x}_j'^{\top}\mathbf{L}\mathbf{x}_j', \tag{19}$$

highlighting that SAQ corresponds to a graph-weighted quadratic form measuring spatial coherence of gene j. This quadratic-form interpretation also motivates the permutation-based null model used for SVG identification, described below.

To normalize for gene-specific variability and expression magnitude, we compute the squared ℓ_2 norm of each gene's mean-centered profile:

$$v_j^{norm} = \sum_{i=1}^{N} {X_{ij}^{'}}^2 = \|\mathbf{x}_j^{'}\|_2^2, \quad j = 1, 2, \cdots, M \tag{20}$$

Finally, the spatial activation quotient is defined as the ratio:

$$v_j^{saq} = \frac{v_j^{sp}}{v_j^{norm}}, \quad \mathbf{v^{saq}} \in \mathbb{R}^M. \tag{21}$$

This normalization rescales the graph-weighted quadratic form by total gene variance, yielding a dimensionless measure of relative spatial coherence across genes.

Spatially variable gene identification

We computed the Spatial Activation Quotient (SAQ) for each gene (M total) using the weighted adjacency matrix $\mathbf{L}$, which represents the LSA graph of a given sample. The SAQ can be interpreted as a normalized graph quadratic form, analogous to a Rayleigh quotient [43], providing a measure of how strongly a gene's expression pattern aligns with the spatial neighborhood structure encoded by $\mathbf{L}$. Genes with higher SAQ values exhibit stronger non-random spatial organization and are therefore candidates for spatially variable genes (SVGs).

To assess statistical significance, we performed permutation-based calibration under a spatial-randomness null model. Specifically, for each permutation we randomly permuted the spot indices within each mean-centered gene vector $\mathbf{x}_j'$, disrupting its alignment with the fixed graph operator $\mathbf{L}$ while preserving the marginal expression distribution. For each permuted instance, the SAQ statistic was recomputed to generate a gene-wise null distribution reflecting the range of values expected under spatial randomness.

Rather than estimating empirical tail probabilities, we estimated the null mean and variance of the SAQ from permutation scores and computed a standardized z-score for each observed value. One-sided p-values were obtained from the upper tail of the standard normal distribution, reflecting enrichment for positive spatial coherence under $\mathbf{L}$.

Finally, we applied False Discovery Rate (FDR) correction across all genes using the Benjamini–Hochberg procedure [44]. Genes passing the threshold (FDR < 0.05) were considered significantly spatially variable genes (SVGs).

Evaluation metrics

We evaluated spatial domain identification using the following metrics:

Adjusted Rand Index

The Adjusted Rand Index (ARI) measures the agreement between a predicted clustering and a ground truth partition while correcting for random label assignments. Given two clustering assignments: $U = \{U_1, U_2, \cdots, U_{\mathcal{R}}\}$ being the ground truth partition and $V = \{V_1, V_2, \cdots, V_{\mathcal{C}}\}$ being the predicted clustering of Spartan, of N spots. Let n_{ij} denote the number of spots simultaneously in cluster U_i and V_j.

Let $a_i = \sum_j n_{ij}$ and $b_j = \sum_i n_{ij}$ represent the cluster sizes in U and V respectively. The ARI is defined as:

$$\text{ARI} = \frac{\sum_{i,j}\binom{n_{ij}}{2} - \frac{\sum_i\binom{a_i}{2}\sum_j\binom{b_j}{2}}{\binom{N}{2}}}{\frac{1}{2}[\sum_i\binom{a_i}{2} + \sum_j\binom{b_j}{2}] - \frac{\sum_i\binom{a_i}{2}\sum_j\binom{b_j}{2}}{\binom{N}{2}}} \tag{22}$$

The numerator measures the excess agreement between the two clusterings relative to that expected by random chance,

while the denominator normalizes the score so that random labelings yield an expected value of zero and identical clusterings yield a value of one. The ARI is invariant to label permutations and robust to differences in cluster sizes. In this work, we use the ARI as implemented in scikit-learn [45].

Normalized Mutual Information

The Normalized Mutual Information (NMI) quantifies the mutual dependence between a predicted clustering and a ground truth labeling. Given two clustering assignments: $U = \{U_1, U_2, \cdots, U_{\mathcal{R}}\}$ being the ground truth partition and $V = \{V_1, V_2, \cdots, V_{\mathcal{C}}\}$ being the predicted clustering of Spartan, of N spots.

Let $n_{ij} = |U_i \cap V_j|$, $a_i = \sum_j n_{ij}$ and $b_j = \sum_i n_{ij}$. Define the probability distributions $\mathcal{P}_U(i) = \frac{a_i}{N}$, $\mathcal{P}_V(j) = \frac{b_j}{N}$, and joint-probability $\mathcal{P}_{UV}(i,j) = \frac{n_{ij}}{N}$. The mutual information is:

$$MI(U,V) = \sum_{i,j} \mathcal{P}_{UV}(i,j) \log \frac{\mathcal{P}_{UV}(i,j)}{\mathcal{P}_U(i)\mathcal{P}_V(j)} \tag{23}$$

The entropies of the two clustering assignments are:

$$H(U) = -\sum_i \mathcal{P}_U(i) \log \mathcal{P}_U(i) \tag{24}$$

$$H(V) = -\sum_j \mathcal{P}_V(j) \log \mathcal{P}_V(j) \tag{25}$$

The NMI is defined as:

$$\text{NMI(U,V)} = \frac{MI(U,V)}{\frac{1}{2}(H(U) + H(V))} \tag{26}$$

The NMI(U,V) $\in$ [0,1]. A value of 1 indicates perfect agreement, and a value near 0 indicates essentially independent assignments. In our experiments, we compute NMI using scikit-learn [45].

Homogeneity and Completeness

Homogeneity (HOM) and Completeness (COM) are information-theoretic metrics that assess the quality of clustering with respect to known true cluster labels. Homogeneity (HOM) measures whether each predicted cluster contains only spots of a single ground truth cluster, while completeness assesses whether all spots of a given ground truth cluster are assigned to the same predicted cluster.

Formally, given ground truth labels $U = \{U_1, U_2, \cdots, U_{\mathcal{R}}\}$ and predicted clusters $V = \{V_1, V_2, \cdots, V_{\mathcal{C}}\}$, the conditional entropies are defined as:

$$H(U|V) = -\sum_{i,j} \frac{n_{ij}}{N} \log \frac{n_{ij}}{b_j} \tag{27}$$

$$H(V|U) = -\sum_{i,j} \frac{n_{ij}}{N} \log \frac{n_{ij}}{a_i} \tag{28}$$

where $n_{ij} = |U_i \cap V_j|$, $a_i = \sum_j n_{ij}$, $b_j = \sum_i n_{ij}$, and N is the total number of spots.

The entropy of true and predicted labels are $H(U) = -\sum_i \frac{a_i}{N} \log \frac{a_i}{N}$ and $H(V) = -\sum_j \frac{b_j}{N} \log \frac{b_j}{N}$. HOM and COM are then defined as:

$$\text{HOM} = 1 - \frac{H(U|V)}{H(U)} \tag{29}$$

$$\text{COM} = 1 - \frac{H(V|U)}{H(V)} \tag{30}$$

We compute both scores using scikit-learn's implementation [45].

R score

While individual clustering metrics such as NMI, homogeneity (HOM), and completeness (COM) each capture distinct aspects of clustering quality, they can vary in sensitivity depending on cluster purity or fragmentation. To obtain a balanced, single-value summary of clustering performance across these complementary information-theoretic measures, we define the R score [46] as their arithmetic mean:

$$\text{R} = \frac{\text{NMI} + \text{HOM} + \text{COM}}{3} \tag{31}$$

A higher R score indicates simultaneously high mutual information, cluster purity, and label completeness, thus providing a robust overall indicator of alignment between predicted and ground truth structures.

Normalized Local Spatial Activation Score (nLSAS)

The normalized Local Spatial Activation Score (nLSAS) quantifies the distribution of LSA-derived activation mass across clusters. It measures how activation mass derived from the LSA graph is distributed among cluster-specific regions, thereby reflecting structural consistency and fragmentation characteristics of a clustering solution.

Let $V = \{V_1, V_2, \cdots, V_{\mathcal{C}}\}$ denote the set of clusters, where $\mathcal{C}$ is the total number of clusters. For each spot $i \in \{1, 2, \cdots, N\}$, we define its LSA as:

$$L_i = \sum_{j=1}^{N} l_{ij}, \quad l_{ij} \in \mathbf{L} \tag{32}$$

where l_{ij} represents the LSA of spot i with respect to spot j, derived from the LSA matrix $\mathbf{L}$. Since $\mathbf{L}$ is asymmetric, L_i corresponds to the row-sum (outgoing activation mass) of spot i.

For each cluster $V_k \in V$, the cluster-level activation mass is computed as:

$$L_k = \sum_{i \in V_k} L_i. \tag{33}$$

We then determine the maximum and minimum activation mass across all clusters:

$$L_{\max} = \max_{k \in \{1,2,\cdots,\mathcal{C}\}} \{L_k\}, \tag{34}$$

$$L_{\min} = \min_{k \in \{1,2,\cdots,\mathcal{C}\}} \{L_k\}. \tag{35}$$

To ensure comparability across clusters and datasets, we normalize the cluster activation masses to the range $[0, 1]$ as:

$$\tilde{L}_k = \frac{L_k - L_{\min}}{L_{\max} - L_{\min} + 10^{-10}}, \tag{36}$$

where a small constant (10^{-10}) avoids division by zero when $L_{\max} = L_{\min}$ and ensures numerical stability.

Finally, the normalized Local Spatial Activation Score (nLSAS) is defined as the average normalized activation across all clusters:

$$\text{nLSAS} = \frac{1}{\mathcal{C}} \sum_{k=1}^{\mathcal{C}} \tilde{L}_k. \tag{37}$$

nLSAS summarizes how LSA-derived activation mass is distributed across clusters and provides a dataset-normalized measure of structural heterogeneity. nLSAS was used as an unsupervised stability prior during α-selection.

Datasets

To assess Spartan's performance in spatial domain identification and spatially variable gene (SVG) detection, we have evaluated Spartan on a diverse group of SRT datasets, coming from many different technologies and tissue types. Seven of those datasets have been used for spatial domain identification and four of them for SVG detection. Across the manuscript, we use the term sample to describe a single slice or a tissue section of the SRT. We use dataset to describe a set of SRT samples. The description of each dataset are given below:

Stereo-seq

The Stereo-seq dataset is sequencing-based SRT dataset obtained by high-resolution technology that spans the full transcriptome. The dataset consists of nine samples from mouse embryo depicting organogenesis [11]. The samples of the dataset [47] have the following number of spots: 5,913 (E9.5E1S1), 5,292 (E9.5E2S1), 4,356 (E9.5E2S2), 5,059 (E9.5E2S3), 5,797 (E9.5E2S4), 18,408 (E10.5E1S1), 18,647 (E10.5E1S2), 18,670 (E10.5E1S3), and 8,494 (E10.5E2S1). The corresponding number of genes: 25,568 (E9.5E1S1), 23,756 (E9.5E2S1), 24,107 (E9.5E2S2), 24,238 (E9.5E2S3), 23,398 (E9.5E2S4), 25,201 (E10.5E1S1), 25,244 (E10.5E1S2), 25,647 (E10.5E1S3), and 22,385 (E10.5E2S1).

BaristaSeq

The BaristaSeq dataset is a SRT dataset obtained by barcoded BarisTaSeq technology [48] which allows faster image acquisition. The dataset [47] consists of three samples from the mouse primary cortex. The number of cells in the samples: 1,525 (Slice 1), 2,042 (Slice 2), and 1690 (Slice 3). The size of the gene panel is 79.

MERFISH

The MERFISH dataset is SRT dataset obtained from imaging-based MERFISH technology. The dataset represents the preoptic region of the mouse hypothalamus [49]. The dataset [47] consists of five samples with number of cells: 5,488 (Bregma-0.04), 5,557 (Bregma-0.09), 5,926 (Bregma-0.14), 5,803 (Bregma-0.19), and 5,543 (Bregma-0.24). The gene panel size is 155. The regional annotations of the dataset is published in [12].

osmFISH

The osmFISH dataset is an imaging-based SRT dataset obtained by nonbarcoded osmFISH technology [50]. The dataset [47] consists of one sample representing the mouse somatosensory cortex. The gene panel is made up of 33 marker genes. The sample consists of 4,839 cells. The regional annotations of the dataset is published in [51].

STARmap

The STARmap dataset is an imaging-based SRT dataset obtained by in-situ RNA-sequencing STARmap technology [52]. The dataset [47] consists of three samples from the high cognitive region of the brain, the medial prefrontal cortex. The samples consist of the following number of cells: 1,049 (BZ5), 1,053 (BZ9), and 1,088 (BZ14). The gene panel size is 166. All samples are annotated to four distinct layer structures [12].

*STARmap**

The STARmap* is imaging-based SRT dataset, version of the STARmap dataset but has 1020 genes. The dataset [47] consist of one sample with 1207 cells. The dataset represents the visual cortex with regions enriched with excitatory, non-neuronal, and inhibitory cell types, expertly annotated [52].

Vizgen MERFISH

The Vizgen MERFISH is imaging-based SRT dataset obtained by the MERFISH technology. The dataset contains a gene panel of 483 genes [53]. This includes canonical cell type markers, nonsensory G-protein coupled receptors (GPCRs), and Receptor Tyrosine Kinases (RTKs). Cell type markers are mapped using Vizgen's own pipeline. The dataset consists of four samples derived from the coronal section of the mouse brain. The number of cells in the samples are:78,329 (S1R1), 88884 (S1R2), 83,546 (S2R1), and 85,958 (S2R3). The samples are regionally annotated by [9].

10x Visium

The 10x Visium is spot sequencing-based SRT dataset. The dataset consists of 12 samples from the human dorsolateral pre-frontal cortex (DLPFC) region [54]. Each sample spans six neuronal layers and the white matter. The number of spots in the samples [47]: 4,226 (151507), 4,384 (151508), 4,789 (151509), 4,634 (151510), 3,661 (151669), 3,498 (151670), 4,110 (151671), 4,015 (151672), 3,639 (151673), 3,673 (151674), 3,592 (151675), and 3,460 (151676). The total number of genes in the dataset is 33,538. The dataset is expertly annotated with six neuronal layers and the white matter [55].

Visium HD dataset

High-resolution spatial transcriptomic profiling was conducted on a 10-week post-conception (wpc) human fetal gastroesophageal specimen using the Visium HD Spatial Gene Expression platform (10x Genomics). Developmental staging was determined by foot length (8 mm) in combination with ultrasonographic assessment.

Human tissue collection and processing

Human embryonic material was obtained from elective terminations of pregnancy under ethical approval from the North West Research Ethics Committee (23/NW/0039), with written informed consent from participants and in accordance with the Human Tissue Authority Codes of Practice.

The esophagus, stomach, and proximal duodenum were dissected and fixed within 1 h post-collection in 4% paraformaldehyde under RNase-free conditions. Tissue was processed and paraffin-embedded using standard protocols. Sagittal sections (5 µm thickness) were obtained, with haematoxylin and eosin (H&E) staining performed on every eighth section for anatomical reference.

Sections selected for spatial transcriptomic profiling were processed following the Visium HD FFPE Tissue Preparation

Handbook (CG000684) and the Visium HD Spatial Gene Expression Reagent Kits User Guide (CG000685). Sections were deparaffinised, destained, de-crosslinked, and permeabilised prior to overnight probe hybridisation. Following probe ligation and washing, probe transfer was performed using the 10x Genomics CytAssist instrument. Libraries were prepared using the Visium HD Spatial Gene Expression Reagent Kit according to the manufacturer's protocol and sequenced on an Illumina NovaSeq 6000 platform (SP flow cell, 100-cycle configuration).

Data preprocessing and cell-level binning

Raw sequencing data were processed using the Space Ranger pipeline (10x Genomics, version 3.1.2) with alignment to the GRCh38 reference genome (release 2020-A).

To achieve cell-level resolution, histological images were analysed using QuPath (ersion 0.5.1-arm64). Nuclei were detected based on intensity and morphological parameters derived from H&E staining, followed by watershed-based expansion of nuclear boundaries to approximate whole-cell areas. Segmentation parameters included intensity threshold of 0.1, background radius 5 μm, and a minimum cell area of 10 μm^2.

Custom cell-specific bins were generated by assigning Visium HD barcodes to segmented cell polygons based on centroid overlap between barcode spatial coordinates and cell boundaries. For each segmented cell, gene-wise UMI counts from overlapping barcodes were aggregated to generate a gene-by-cell count matrix.

Cells were filtered to retain those with $\geq$20 UMIs and a maximum cell area of 2000 μm^2.

The final preprocessed dataset comprised 84,732 segmented cells and 18,085 detected genes following quality control filtering.

Data availability

The authors analyzed eight publicly available and one in-house SRT datasets. The data were acquired from the following websites or accession numbers: (1) Stereo-seq, BaristaSeq, MERFISH, osmFISH, STARmap, STARmap*, 10x Visium datasets with regional annotations were downloaded from `http://sdmbench.drai.cn/`; (2) Vizgen MERFISH dataset was downloaded from `https://console.cloud.google.com/storage/browser/public-datasets-vizgen-merfish;tab=objects?prefix=&forceOnObjectsSortingFiltering=false`; (3) The Visium HD dataset generated in this study is available in ArrayExpress under accession number E-MTAB-16671.

Code availability

An open-source implementation of Spartan can be downloaded from `https://github.com/MohammadFaizIqbalFaiz/spartan-st`, where code and scripts for reproducing the results and figures are provided.

Acknowledgments

The authors thank all the contributors of the public data which was used in this study. Mohammad Faiz Iqbal Faiz was funded by BBSRC DTP award (BB/T008725/1).

References

1. Vivien Marx. Method of the year: spatially resolved transcriptomics. *Nature Methods*, 18(1):9–14, 2021.
2. Michaela Asp, Joseph Bergenstråhle, and Joakim Lundeberg. Spatially resolved transcriptomes—next generation tools for tissue exploration. *BioEssays*, 42(10):1900221, 2020.
3. Tsai-Ying Chen, Li You, Jose Angelito U Hardillo, and Miao-Ping Chien. Spatial transcriptomic technologies. *Cells*, 12(16):2042, August 2023.
4. Lukas Valihrach, Daniel Zucha, Pavel Abaffy, and Mikael Kubista. A practical guide to spatial transcriptomics. *Mol. Aspects Med.*, 97(101276):101276, June 2024.
5. P. A. P. Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37(1-2):17–23, 1950.
6. Luc Anselin. Local indicators of spatial association—LISA. *Geogr. Anal.*, 27(2):93–115, April 1995.
7. V A Traag, L Waltman, and N J van Eck. From louvain to leiden: guaranteeing well-connected communities. *Sci. Rep.*, 9(1):5233, March 2019.
8. Vipul Singhal, Nigel Chou, Joseph Lee, Yifei Yue, Jinyue Liu, Wan Kee Chock, Li Lin, Yun-Ching Chang, Erica Mei Ling Teo, Jonathan Aow, Hwee Kuan Lee, Kok Hao Chen, and Shyam Prabhakar. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Genet.*, 56(3):431–441, March 2024.
9. Darius P Schaub, Behnam Yousefi, Nico Kaiser, Robin Khatri, Victor G Puelles, Christian F Krebs, Ulf Panzer, and Stefan Bonn. PCA-based spatial domain identification with state-of-the-art performance. *Bioinformatics*, 41(1), December 2024.
10. Niklas Müller-Bötticher, Shashwat Sahay, Roland Eils, and Naveed Ishaque. SpatialLeiden: spatially aware leiden clustering. *Genome Biol.*, 26(1):24, February 2025.
11. Ao Chen, Sha Liao, Mengnan Cheng, Kailong Ma, Liang Wu, Yiwei Lai, Xiaojie Qiu, Jin Yang, Jiangshan Xu, Shijie Hao, Xin Wang, Huifang Lu, Xi Chen, Xing Liu, Xin Huang, Zhao Li, Yan Hong, Yujia Jiang, Jian Peng, Shuai Liu, Mengzhe Shen, Chuanyu Liu, Quanshui Li, Yue Yuan, Xiaoyu Wei, Huiwen Zheng, Weimin Feng, Zhifeng Wang, Yang Liu, Zhaohui Wang, Yunzhi Yang, Haitao Xiang, Lei Han, Baoming Qin, Pengcheng Guo, Guangyao Lai, Pura Muñoz-Cánoves, Patrick H Maxwell, Jean Paul Thiery, Qing-Feng Wu, Fuxiang Zhao, Bichao Chen, Mei Li, Xi Dai, Shuai Wang, Haoyan Kuang, Junhou Hui, Liqun Wang, Ji-Feng Fei, Ou Wang, Xiaofeng Wei, Haorong Lu, Bo Wang, Shiping Liu, Ying Gu, Ming Ni, Wenwei Zhang, Feng Mu, Ye Yin, Huanming Yang, Michael Lisby, Richard J Cornall, Jan Mulder, Mathias Uhlén, Miguel A Esteban, Yuxiang Li, Longqi Liu, Xun Xu, and Jian Wang. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 185(10):1777–1792.e21, May 2022.
12. Zheng Li and Xiang Zhou. BASS: multi-scale and multi-sample analysis enables accurate cell type clustering and spatial domain detection in spatial transcriptomic studies. *Genome Biol.*, 23(1):168, August 2022.
13. Yahui Long, Kok Siong Ang, Mengwei Li, Kian Long Kelvin Chong, Raman Sethi, Chengwei Zhong, Hang Xu, Zhiwei Ong, Karishma Sachaphibulkij, Ao Chen, Li Zeng, Huazhu Fu, Min Wu, Lina Hsiu Kim Lim, Longqi Liu, and Jinmiao Chen. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat. Commun.*, 14(1):1155, March 2023.
14. Jian Hu, Xiangjie Li, Kyle Coleman, Amelia Schroeder, Nan Ma, David J Irwin, Edward B Lee, Russell T Shinohara, and Mingyao Li. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods*, 18(11):1342–1351, November 2021.

15. Zixuan Cang, Xinyi Ning, Annika Nie, Min Xu, and Jing Zhang. Scan-it: Domain segmentation of spatial transcriptomics images by graph neural network. In *Proceedings of the 32nd British Machine Vision Conference (BMVC 2021)*, 2021. Paper1139.
16. Honglei Ren, Benjamin L Walker, Zixuan Cang, and Qing Nie. Identifying multicellular spatiotemporal organization of cells with SpaceFlow. *Nat. Commun.*, 13(1):4076, July 2022.
17. Stéphane Dray, Sonia Saïd, and Françis Débias. Spatial ordination of vegetation data using a generalization of wartenberg's multivariate spatial correlation. *J. Veg. Sci.*, 19(1):45–56, February 2008.
18. Pauline Yahr. Sex difference and response to testosterone in gabaergic cells of the medial preoptic nucleus and ventral bed nuclei of the stria terminalis in gerbils. *Horm. Behav.*, 59(4):473–476, April 2011.
19. Mathias Uhlén, Linn Fagerberg, M. Hallström, Björn and others many˙ Gad1 – brain tissue expression (hypothalamus). The Human Protein Atlas, version 19, 2024. `https://v19.proteinatlas.org/ENSG00000128683GAD1/brain/Hypothalamus` (accessed 13 Nov. 2025).
20. Robert Gould and Scott Brady. Identifying mRNAs residing in myelinating oligodendrocyte processes as a basis for understanding internode autonomy. *Life (Basel)*, 13(4):945, April 2023.
21. Jessica L Braun, Mia S Geromella, Sophie I Hamstra, and Val A Fajardo. Neuronatin regulates whole-body metabolism: is thermogenesis involved? *FASEB Bioadv.*, 2(10):579–586, October 2020.
22. Nishadh Rathod, Jessi J Bak, Joseph O Primeau, M'lynn E Fisher, Lennane Michel Espinoza-Fonseca, Mary Joanne Lemieux, and Howard S Young. Nothing regular about the regulins: Distinct functional properties of SERCA transmembrane peptide regulatory subunits. *Int. J. Mol. Sci.*, 22(16):8891, August 2021.
23. Seoyoung Son, Steffy B Manjila, Kyra T Newmaster, Yuan-Ting Wu, Daniel J Vanselow, Matt Ciarletta, Todd E Anthony, Keith C Cheng, and Yongsoo Kim. Whole-brain wiring diagram of oxytocin system in adult mice. *J. Neurosci.*, 42(25):5021–5033, June 2022.
24. Nathaly Anto Michel, Senka Ljubojevic-Holzer, Heiko Bugger, and Andreas Zirlik. Cellular heterogeneity of the heart. *Front. Cardiovasc. Med.*, 9:868466, April 2022.
25. Deborah Hay, Jim R Hughes, Christian Babbs, James O J Davies, Bryony J Graham, Lars Hanssen, Mira T Kassouf, A Marieke Marieke Oudelaar, Jacqueline A Sharpe, Maria C Suciu, Jelena Telenius, Ruth Williams, Christina Rode, Pik-Shan Li, Len A Pennacchio, Jacqueline A Sloane-Stanley, Helena Ayyub, Sue Butler, Tatjana Sauka-Spengler, Richard J Gibbons, Andrew J H Smith, William G Wood, and Douglas R Higgs. Genetic dissection of the α-globin super-enhancer in vivo. *Nat. Genet.*, 48(8):895–903, August 2016.
26. Qingquan Zhang, Daniel Carlin, Fugui Zhu, Paola Cattaneo, Trey Ideker, Sylvia M Evans, Joshua Bloomekatz, and Neil C Chi. Unveiling complexity and multipotentiality of early heart fields. *Circ. Res.*, 129(4):474–487, August 2021.
27. Tianyi Sun, Alexander Grassam-Rowe, Zhaoli Pu, Yangpeng Li, Huiying Ren, Yanru An, Xinyu Guo, Wei Hu, Ying Liu, Yuqing Zheng, Zhu Liu, Kun Kou, Xianhong Ou, Tangting Chen, Xuehui Fan, Yangyang Liu, Shu Tu, Yu He, Yue Ren, Ao Chen, Zhouchun Shang, Zhidao Xia, Lucile Miquerol, Nicola Smart, Henggui Zhang, Xiaoqiu Tan, Weinian Shou, and Ming Lei. Dbh+ catecholaminergic cardiomyocytes contribute to the structure and function of the cardiac conduction system in murine heart. *Nat. Commun.*, 14(1):7801, November 2023.
28. Xianchao Du, Jiashuo Li, Minghui Li, Xinxin Yang, Zhipeng Qi, Bin Xu, Wei Liu, Zhaofa Xu, and Yu Deng. Research progress on the role of type I vesicular glutamate transporter (VGLUT1) in nervous system diseases. *Cell Biosci.*, 10(1):26, March 2020.
29. Beatriz Antolin-Fontes, Kun Li, Jessica L Ables, Michael H Riad, Andreas Görlich, Maya Williams, Cuidong Wang, Sylvia M Lipford, Maria Dao, Jianxi Liu, Henrik Molina, Nathaniel Heintz, Paul J Kenny, and Ines Ibañez-Tallon. The habenular g-protein-coupled receptor 151 regulates synaptic plasticity and nicotine intake. *Proc. Natl. Acad. Sci. U. S. A.*, 117(10):5502–5509, March 2020.
30. Renaud Massart, Jean-Philippe Guilloux, Virginie Mignon, Pierre Sokoloff, and Jorge Diaz. Striatal GPR88 expression is confined to the whole projection neuron population and is regulated by dopaminergic and glutamatergic afferents. *Eur. J. Neurosci.*, 30(3):397–414, August 2009.
31. Mohammad Golam Sabbir. Cholinergic receptor muscarinic 1 co-localized with mitochondria in cultured dorsal root ganglion neurons, and its deletion disrupted mitochondrial ultrastructure in peripheral neurons: Implications in alzheimer's disease. *J. Alzheimers. Dis.*, 98(1):247–264, 2024.
32. Shelley L Forrest, Jordan Hanxi Kim, Daniel R Crockford, Katharine Huynh, Rosie Cheong, Samantha Knott, Madison A Kane, Lars M Ittner, Glenda M Halliday, and Jillian J Kril. Distribution patterns of astrocyte populations in the human cortex. *Neurochem. Res.*, 48(4):1222–1232, April 2023.
33. F Wang, X Liu, P Yang, L Guo, C Liu, H Li, S Long, Y Shen, and H Wan. Loss of TACSTD2 contributed to squamous cell carcinoma progression through attenuating TAp63-dependent apoptosis. *Cell Death Dis.*, 5(3):e1133, March 2014.
34. The Human Protein Atlas. The human protein atlas: TACSTD2 expression in esophagus, 2026.
35. Yoshitaka Taketomi, Yoshimi Miki, and Makoto Murakami. Old but new: Group IIA phospholipase A2 as a modulator of gut microbiota. *Metabolites*, 12(4):352, April 2022.
36. Kiyomi Tsuji-Tamura, Saori Morino-Koga, Shingo Suzuki, and Minetaro Ogawa. TAGLN, a canonical marker of smooth muscle cells, is present in endothelial cells and involved in angiogenesis. *J. Cell Sci.*, 134(15), July 2021.
37. Hongming Huang, Kai Sen Tan, Suizi Zhou, Tian Yuan, Jing Liu, Hsiao Hui Ong, Qianmin Chen, Junxiao Gao, Minghong Xu, Zhenchao Zhu, Qianhui Qiu, and De Yun Wang. p63+krt5+ basal cells are increased in the squamous metaplastic epithelium of patients with radiation-induced chronic rhinosinusitis. *Radiat. Oncol.*, 15(1):222, September 2020.
38. The Human Protein Atlas. The human protein atlas: SPINK1 expression in stomach, 2026.
39. Yong H. Sheng, Garrett Z. Ng, Kim M. Summers, Alison L. Every, Gareth Price, Sumaira Z. Hasnain, Philip Sutton, and Michael A. McGuckin. Influence of the muc1 cell surface mucin on gastric mucosal gene expression profiles in response to helicobacter pylori infection in mice. *Frontiers in Cellular and Infection Microbiology*, Volume 10 - 2020, 2020.

40. F Alexander Wolf, Philipp Angerer, and Fabian J Theis. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.*, 19(1), December 2018.
41. Giovanni Palla, Hannah Spitzer, Michal Klein, David Fischer, Anna Christina Schaar, Louis Benedikt Kuemmerle, Sergei Rybakov, Ignacio L Ibarra, Olle Holmberg, Isaac Virshup, Mohammad Lotfollahi, Sabrina Richter, and Fabian J Theis. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods*, 19(2):171–178, February 2022.
42. Luca Marconato, Giovanni Palla, Kevin A Yamauchi, Isaac Virshup, Elyas Heidari, Tim Treis, Wouter-Michiel Vierdag, Marcella Toth, Sonja Stockhaus, Rahul B Shrestha, Benjamin Rombaut, Lotte Pollaris, Laurens Lehner, Harald Vöhringer, Ilia Kats, Yvan Saeys, Sinem K Saka, Wolfgang Huber, Moritz Gerstung, Josh Moore, Fabian J Theis, and Oliver Stegle. SpatialData: an open and universal data framework for spatial omics. *Nat. Methods*, 22(1):58–62, January 2025.
43. Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, England, June 2012.
44. Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.*, 57(1):289–300, January 1995.
45. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Michaël Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
46. Zhiyuan Yuan, Fangyuan Zhao, Senlin Lin, Yu Zhao, Jianhua Yao, Yan Cui, Xiao-Yong Zhang, and Yi Zhao. Benchmarking spatial clustering methods with spatially resolved transcriptomics data. *Nat. Methods*, 21(4):712–722, April 2024.
47. Zhiyuan Yuan, Fangyuan Zhao, Senlin Lin, Yu Zhao, Jianhua Yao, Yan Cui, Xiao-Yong Zhang, and Yi Zhao. Sdmbench: A spatial domain benchmark for spatial transcriptomics data. `http://sdmbench.drai.cn/`, 2024. Accessed: 2025-11-13.
48. Xiaoyin Chen, Yu-Chi Sun, George M Church, Je Hyuk Lee, and Anthony M Zador. Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res.*, 46(4):e22–e22, February 2018.
49. Jeffrey R Moffitt, Dhananjay Bambah-Mukku, Stephen W Eichhorn, Eric Vaughn, Karthik Shekhar, Julio D Perez, Nimrod D Rubinstein, Junjie Hao, Aviv Regev, Catherine Dulac, and Xiaowei Zhuang. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, 362(6416):eaau5324, November 2018.
50. Simone Codeluppi, Lars E Borm, Amit Zeisel, Gioele La Manno, Josina A van Lunteren, Camilla I Svensson, and Sten Linnarsson. Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods*, 15(11):932–935, November 2018.
51. Simone Codeluppi et al. Data and code availability: Expression data (loom file) with osmfish data. `http://linnarssonlab.org/osmFISH/availability/`, 2018. Linnarsson Lab, Accessed: 2025-11-13.
52. Xiao Wang, William E Allen, Matthew A Wright, Emily L Sylwestrak, Nikolay Samusik, Sam Vesuna, Kathryn Evans, Cindy Liu, Charu Ramakrishnan, Jia Liu, Garry P Nolan, Felice-Alessio Bava, and Karl Deisseroth. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*, 361(6400):eaat5691, July 2018.
53. Inc. Vizgen. Merfish mouse brain receptor map — mouse brain map dataset. `https://info.vizgen.com/mouse-brain-map?submissionGuid=a66ccb7f-87cf-4c55-83b9-5a2b6c0c12b9`, 2025. Accessed: 2025-11-13.
54. Kristen R Maynard, Leonardo Collado-Torres, Lukas M Weber, Cedric Uytingco, Brianna K Barry, Stephen R Williams, Joseph L Catallini, 2nd, Matthew N Tran, Zachary Besich, Madhavi Tippani, Jennifer Chew, Yifeng Yin, Joel E Kleinman, Thomas M Hyde, Nikhil Rao, Stephanie C Hicks, Keri Martinowich, and Andrew E Jaffe. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.*, 24(3):425–436, March 2021.
55. KristenR. Maynard, Leonardo Collado-Torres, LukasM. Weber, Cedric Uytingco, BriannaK. Barry, StephenR. Williams, JosephL. II Catallini, MatthewN. Tran, Zachary Besich, Madhavi Tippani, Jennifer Chew, Yifeng Yin, JoelE. Kleinman, ThomasM. Hyde, Nikhil Rao, StephanieC. Hicks, Keri Martinowich, and AndrewE. Jaffe. Spatial transcriptomics data (10×visium) for the human dorsolateral prefrontal cortex hosted in spatiallibd. `http://research.libd.org/spatialLIBD/`, 2021. Accessed: 2025-11-13.