Reconstructing biologically coherent cellular profiles from imaging-based spatial transcriptomics

Long Yuan[1,2*], Youyun (Peter) Zheng[3,4], Shuming Zhang[5,6,7], Rameen Beroukhim[8,9,10,11], Atul Deshpande[5,6,12,13*]

[1]Department of Immunology, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[2]Department of Computer Science, Johns Hopkins University Whiting School of Engineering, Baltimore, MD, USA.
[3]Department of Medical Oncology and Cancer Biology, Dana Farber Cancer Institute, Boston, MA, USA.
[4]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA.
[5]Bloomberg-Kimmel Institute for Cancer Immunotherapy, Baltimore, MD, USA.
[6]Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
[7]Convergence Institute, Johns Hopkins University, Baltimore, MD, USA.
[8]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA.
[9]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA.
[10]Broad Institute of MIT and Harvard, Cambridge, MA, USA.
[11]Department of Medicine, Harvard Medical School, Boston, MA, USA.
[12]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA.
[13]Data Science and AI Institute, Johns Hopkins University, Baltimore, MD, USA.

*Corresponding author(s). E-mail(s): lyuan13@jhmi.edu; adeshpande@jhu.edu;

Abstract

In imaging-based spatial transcriptomics, transcript-to-cell assignment shapes downstream biological interpretation including cell typing, ligand-receptor inference, and niche characterization. However, two-dimensional segmentation of volumetric tissue often yields mixed cellular profiles, while cells without detected nuclei are missed entirely, distorting the aforementioned downstream analyses. We present TRACER, which refines cellular representations in imaging-based transcriptomics by leveraging gene–gene coherence and spatial co-localization of transcripts observed directly in the data, without requiring external annotations or reference atlases. TRACER resolves mixed cellular profiles and reconstructs partial cells whose nuclei are not detected, enabling more complete representation of cells within the tissue section. We also introduce coherence-based metrics that quantify transcriptional purity and conflict, enabling platform-agnostic benchmarking of segmentation quality. Across diverse platforms, tissues, and segmentation methodologies, TRACER consistently and reproducibly improves the coherence of cellular profiles and the quality of downstream analyses.

Keywords: Spatial transcriptomics, 3D segmentation correction, segmentation diagnostics, cellular reconstruction

1 Introduction

Imaging-based spatial transcriptomics enable the measurement of gene expression at subcellular resolution within the spatial context of the tissue, offering unprecedented opportunities to study tissue architecture, cell states, and cell–cell interactions in situ [1–3]. Central to these analyses is the accurate assignment of individual transcripts to cells, the fundamental unit for downstream interpretation, including cell-type identification, differential expression analysis, and spatial niche modeling [4, 5]. Errors in imaging-based cell segmentation—typically used for transcript assignment—propagate directly into biological conclusions, making the accuracy of this step a critical determinant of analytical fidelity [4–7].

Despite rapid advances in imaging technologies and computational methods, accurate cell segmentation in spatial transcriptomics remains a challenge. Many pipelines rely on two-dimensional (2D) segmentation derived from nuclear staining, nuclear expansion heuristics, or membrane markers when available [4, 5, 8, 9]. Segmentation methods must contend with variable signal quality, densely packed tissues, and the intrinsic thickness of biological sections. Recent work has highlighted that tissue sectioning itself introduces ambiguity in cellular representation. Tissue sectioning causes contamination by the vertical environment of cells, specifically by partially captured cells with out-of-plane nuclei [10, 11]. Yet most approaches rely on 2D segmentation—derived from nuclear staining, nuclear expansion heuristics, or membrane markers when available—under the assumption that all transcripts within the 2D boundary belong to the same cell[4, 8, 9, 12]. As a result, segmented boundaries may not accurately represent cellular profiles, and subsequent analyses often exhibit mixed transcriptional signatures that are difficult to interpret [13, 14] (Fig. 1A).

At the same time, there is growing recognition that imaging-based spatial transcriptomics data lack a clear gold standard for evaluating segmentation quality [15, 16]. Metrics developed for natural image segmentation focus on geometric overlap or boundary accuracy, which do not readily translate to biological context where cell boundaries are fuzzy, overlapping, and functionally defined [17]. Recent work using matrix factorization of local molecular neighborhoods [18] further demonstrated that segmentation-induced transcript admixture can dominate downstream analyses in imaging-based spatial transcriptomics, underscoring the need for approaches that explicitly address transcript mixing and segmentation uncertainty. As a result, segmentation methods are often assessed indirectly through downstream tasks, such as clustering or marker gene expression, without principled segmentation-agnostic diagnostics to quantify transcriptional coherence or mixing at the cellular level [4, 19].

Here, we introduce an information-theoretic approach for refining cellular representations and evaluating segmentation fidelity in spatial transcriptomics. We present TRACER (Tissue Reconstruction via Associative Clique Extraction and Relation-mapping), a computational framework designed to address the limitations of segmentation in imaging-based spatial transcriptomics by leveraging transcriptional associations directly from the data (Fig. 1B). TRACER uses gene–gene coherence and spatial colocalization to refine cellular representations without requiring external annotations, supervised training, or GPU acceleration. We also provide coherence-based metrics that enable evaluation of transcript-to-cell assignment across platforms and methodologies.

2 Results

2.1 TRACER uses transcriptional coherence to refine transcript assignment

TRACER is a computational framework for refining cellular representations in imaging-based spatial transcriptomics. TRACER leverages gene-gene associations and spatial proximity to address erroneous transcript assignments and to recover fragments of cells whose nuclei lie outside the section plane. The core design of TRACER is motivated by an analogy between natural language processing and spatial transcriptomics, treating cells (or nuclei) as contexts and gene presence as events, enabling robust estimation of gene–gene co-expression independent of transcript counts (Fig. 1B).

Using nuclear segmentation as a reliable context window, TRACER computes a gene–gene normalized pointwise mutual information (NPMI) matrix directly from the data [20–22]. These associations are then used to guide conservative transcript pruning, recovery of partial cell fragments, and constrained hierarchical stitching of fragmented cellular profiles into a three-dimensional (3D) representation (Fig. 1B). The stitching criterion, based on a coherence-change metric (ΔC), selectively merges spatially proximal fragments that are transcriptionally coherent while preventing incompatible merges; we validate this approach in a controlled 3D simulation (Supplementary Fig. 1 ; see Methods).

We distinguish between two stages of TRACER output used throughout this study. TRACER-stitched refers to the intermediate output obtained after NPMI-guided transcript pruning and ΔC-based hierarchical stitching, which yields transcriptionally coherent whole cells while preserving all spatially connected transcript components. This stage primarily resolves transcript mixing and hybrid cellular profiles arising from segmentation errors. TRACER-fine-tuned refers to a subsequent spatial refinement step in which each stitched entity is subjected to a spatial connectivity constraint. Specifically, disconnected transcript components that lie beyond a user-defined distance threshold from the dominant connected component are removed and reassigned through an additional constrained stitching pass. This step eliminates spatially inconsistent outliers while preserving the dominant transcriptional program, yielding the final refined set of whole cells, partial cells, and reconstructed partial cells used for downstream analyses (see Methods).

Unless otherwise stated, results are shown for both TRACER-stitched and TRACER-fine-tuned outputs to illustrate the impact of coherence-based transcript reassignment alone and in combination with spatial fine-tuning.

2.2 NPMI captures lineage-specific gene programs and enables quantitative purity and conflict scoring

We first evaluated whether the NPMI values computed from nuclear contexts captures biologically meaningful gene–gene relationships. In Xenium breast cancer data [23], the NPMI matrix reveals a prominent block-diagonal structure corresponding to known lineage-associated gene programs Fig. 2A). Insets focusing on submatrices of

representative lymphoid (T/NK) and myeloid genes show strong within-lineage positive NPMI values, indicating consistent co-expression across cells Fig. 2B,C).

In contrast, cross-lineage gene pairs typically exhibit negative NPMI, reflecting transcriptional incompatibility rather than simple sparsity Fig. 2A). Gene pairs with NPMI values near zero show little enrichment or depletion of co-occurrence across cells, whereas NA values arise when no co-occurrence is observed in the dataset. These observations demonstrate that NPMI computed from binary presence–absence events recovers biologically meaningful co-expression patterns without requiring expression normalization or reference annotations.

Using this NPMI matrix, we next defined cell-level purity and conflict scores to quantify, respectively, the internal coherence and mutual incompatibility of gene programs within individual cellular profiles (see Methods). Purity and conflict scores, projected onto UMAP embeddings of Xenium breast cancer data, reveal distinct patterns Fig. 2D,F). Consistent with these patterns, both metrics exhibited bimodal distributions—confirmed by two-component Gaussian mixture models (GMM) [24]—separating coherent cells from ambiguous profiles Fig. 2E,G).

To assess the generality of these metrics across platforms and segmentation strategies, we applied the same purity and conflict scoring framework to Xenium lung cancer data with multimodal segmentation [25] and to Multiplexed Error-Robust Fluorescence in situ Hybridization (MERFISH) mouse ileum data segmented with Baysor using a 3D membrane prior [5]. In both datasets, cells with elevated conflict scores and reduced purity were again preferentially localized to narrow connective bands between UMAP clusters, rather than within well-defined cell-type clusters (Supplementary Fig. 2A,B; Supplementary Fig. 3A,B). These regions correspond to transcriptionally intermediate or mixed profiles, consistent with residual transcript sharing or boundary errors, as commonly observed for transitional cell states in low-dimensional embeddings of single-cell transcriptomic data [26–28].

Notably, in contrast to standard nucleus-expanded segmentation, we see fairly unimodal distributions of purity and conflict scores in the Xenium Lung Cancer and Baysor-segmented MERFISH data, as quantified by GMM (Supplementary Fig. 2C,D; Supplementary Fig. 3C,D). This unimodality relative to the standard Xenium segmentation could be attributed to the smaller area of the lung cancer tissue compared to the breast cancer tissue in the standard Xenium, and potential improvements in cell purity due to Baysor-segmentation. However, there are other confounding factors to consider such as the transcript selection and detection characteristics of each dataset. Together, these results indicate that NPMI-derived purity and conflict scores are reliable segmentation-agnostic indicators of transcriptional coherence and mixing across spatial transcriptomics technologies.

2.3 TRACER refines transcript assignment and restores cell-type resolution in the absence of multimodal staining

Based on the ability of the purity and conflict scores to detect transcriptional mixing, we next evaluated whether TRACER improves transcript assignment and downstream cell-type resolution (Fig. 3). We generated a version of the lung cancer dataset with nucleus-expanded cell segmentation. 2D convex hulls of the transcripts assigned to

each cell are non-overlapping for both multimodal segmentation and nucleus-expanded segmentation(Fig. 3A,B), reflecting the strict boundary enforcement of the imaging-based approach. In contrast, TRACER-stitched and TRACER-fine-tuned outputs yield partially overlapping convex hulls (Fig. 3C,D), representing contributions from overlapping cells captured in the tissue section.

UMAP embeddings based on multimodal segmentation or a 5 µm nuclear expansion showed substantial intermixing of lineages (Fig. 3E,F). Correspondingly, dot plots of canonical marker expression showed broad, low-specificity patterns, indicative of hybrid profiles due to transcriptional mixing. Applying TRACER to both segmentation outputs markedly improved manifold structure (Fig. 3G,H; Supplementary Fig. 2E,F): clusters become clearly delineated clusters, and marker expression profiles exhibit higher specificity with reduced cross-lineage contamination. Notably, TRACER applied to nuclear expansion-based segmentation achieved similar or improved cluster separation compared to multimodal segmentation.

TRACER produces deterministic outputs, yielding identical cellular reconstructions across runs (Supplementary Fig. 2G). At the individual cell level, TRACER increased relative purity and decreased relative conflict compared with pre-TRACER assignments (Supplementary Fig. 2H), consistent with reduced transcript mixing after refinement.

2.4 TRACER reveals partial cells arising from three-dimensional tissue geometry

Imaging-based spatial transcriptomics is intrinsically a 3D measurement problem. Most platforms acquire 3D transcript coordinates from tissue sections with non-trivial thickness[29–31]. As a result, 2D segmentation inevitably produces mixed profiles from boundary errors between overlapping adjacent cells. Partial cells with out-of-plane nuclei either contribute transcripts to overlapping in-plane cells—causing further contamination—or are missed entirely, their transcripts remaining unassigned (Fig. 1A).

TRACER addresses these artifacts by separating contaminated profiles and recovering missed partial cells. To formalize this, we define three categories of cells: whole, partial, and reconstructed partial cells. Whole cells are defined as cellular profiles containing a well-focused nucleus within the imaging plane, consistent with inclusion of the major cellular volume. Partial cells are fragments separated by TRACER from whole-cell profiles—representing contamination from overlapping cells with out-of-plane nuclei. Reconstructed partial cells are coherent transcript clusters assembled by TRACER from previously unassigned transcripts—representing potential cells that were missed entirely because they did not overlap with any detected nuclei (Fig. 1A).

To evaluate the biological relevance of these entities, we applied TRACER to the standard Xenium breast cancer dataset. UMAP embeddings from the original segmentation exhibited poorly separated clusters with extensive inter-cluster mixing (Supplementary Fig. 4A, left), consistent with transcript mixing and segmentation errors. TRACER-stitched and TRACER-fine-tuned whole-cell profiles formed compact, well-separated clusters (Supplementary Fig. 4A, middle and right). Projecting original study annotations onto these embeddings revealed that multiple major lineages—including DCIS tumor cells, invasive tumor cells, T cells, myeloid cells, myoepithelial populations, stromal cells, endothelial cells, and perivascular-like cells—were

substantially conflated under the original segmentation (Supplementary Fig. 4B, left). After TRACER refinement, these lineages resolved into discrete clusters corresponding to biologically coherent cell types–using original annotations without re-annotating the cells (Supplementary Fig. 4B, middle and right).

To examine how TRACER resolves transcriptionally ambiguous profiles, we focused on a representative "T cell & tumor hybrid" cell from the original study (Fig. 4A)[23]. After TRACER refinement, transcripts from this single hybrid profile separate into two distinct, biologically coherent entities: one exhibiting a HER2-positive tumor gene program (Fig. 4B,C) and the other exhibiting a canonical $CD8^+$ T cell program (Fig. 4D,E). We observed a similar effect for a cell from the cluster " unlabeled" in the original study (Supplementary Fig. 5A), where TRACER fine-tuning resolved a mixed profile into two coherent entities corresponding to a T cell–like program (Supplementary Fig. 5B,C) and a tumor-like program (Supplementary Fig. 5D,E).

TRACER reconstructs partial cells by embedding unassigned transcripts in 3D physical space and representing them as nodes in a k-nearest-neighbor graph (Fig. 1A). The NPMI prior enforces internal transcriptional coherence within each reconstructed partial cell, allowing structured entities to emerge from the unassigned transcript pool rather than being discarded. UMAP embeddings of the partial and reconstructed partial cells reveal that these entities also form well-defined clusters. Unsupervised Leiden clustering of these partial cells [32] produced clear and biologically interpretable populations (Fig. 4F), each defined by distinct marker programs (Fig. 4G). For example, clusters expressed *LYZ* (macrophages), *LUM* (stromal cells), *CD3E* (T cells), *SLAMF7* (B cells), *ERBB2* (DCIS tumor cells), *PECAM1* (endothelial cells), *ACTA2* (perivascular-like cells), or *KIT* (mast cells).

Spatially, the distribution of partial and reconstructed partial cells closely mirrors that of whole cells (Fig. 4H–J), consistent with vertical continuity of tissue architecture rather than artifactual reassignment. In Fig. 4J (inset), transcripts contributing to reconstructed partial cells were predominantly unassigned under standard segmentation and would have been discarded from downstream analysis. TRACER consolidates these transcripts into coherent reconstructed partial cells whose convex hulls fill the gaps between segmented cells (Fig. 4J, inset), further supporting their biological relevance.

2.5 TRACER resolves hybrid cellular profiles and removes segmentation-derived ligand–receptor artifacts

Because no gold-standard ground truth exists for evaluating segmentation accuracy in imaging-based spatial transcriptomics [15], we assessed TRACER performance indirectly through marker gene specificity—reflecting the primary purpose of segmentation, namely to capture biologically coherent cell types. In the Xenium breast cancer dataset, marker expression profiles show extensive cross-lineage contamination and weak cluster specificity (Supplementary Fig. 5F). In contrast, TRACER outputs yield markedly enhanced marker specificity and clearer lineage-appropriate expression patterns (Supplementary Fig. 5G; Fig. 5A), consistent with reduced transcript mixing and improved cell-type delineation.

Next, we examined two prominent hybrid populations that arise from the segmentation in the original study. The first corresponds to a stromal & T cell hybrid cluster

(Fig. 5B). After TRACER stitching and fine-tuning, transcripts from this hybrid population separate into distinct stromal and T cell clusters (Fig. 5C,D), restoring biologically expected boundaries. A similar pattern emerges in the T cell & tumor hybrid population identified in the original study, where transcriptional resolution separates them into distinct T cell and tumor clusters after TRACER refinement (Supplementary Fig. 4C). These findings illustrate TRACER's ability to resolve mixed-lineage profiles produced by segmentation errors.

To illustrate the impact of segmentation errors on downstream analyses and TRACER's ability to mitigate them, we compare the results of ligand-receptor interaction analysis using the outputs from the original segmentation and TRACER. The transcriptional mixing resulting from the original segmentation leads to the inference of biologically implausible ligand-receptor interactions [33, 34]. For example, in the original segmentation, ductal carcinoma in situ subtype 1 (DCIS-1) cells—previously characterized as low-invasive [23]—spuriously appeared to express immune-restricted ligands and receptors such as *CD80*, *CD8A*, and *CD8B*, result in unlikely DCIS-to-immune cell interactions (Fig. 5E). TRACER refinement substantially reduces these likely artifacts, yielding interaction patterns more consistent with lineage-restricted ligands and receptors and clarifying epithelial–immune communication patterns. These corrections illustrate the downstream effect of TRACER for suppressing spurious ligand–receptor interaction predictions arising from transcript mixing from segmentation-errors.

2.6 Quantitative benchmarking of TRACER relative to 3D membrane–guided segmentation

We leveraged the NPMI-derived purity and conflict scores Fig. 2; Supplementary Figs. 2–3) to benchmark TRACER against existing segmentation approaches across platforms and datasets. To increase robustness to weak associations and technical noise, we extended the original purity and conflict definitions with a symmetric Rectified Linear Unit (ReLU) formulation incorporating a dead zone around zero NPMI values (Methods). We also defined a complementary signal-strength metric, representing the net magnitude of total positive and negative NPMI gene associations within each cell, as well as relative purity and relative conflict scores, computed as the fraction of intracellular gene associations respectively deemed coherent and conflicting. Transcript counts per cell were retained as an orthogonal quality-control measure.

We first evaluated TRACER on the MERFISH mouse ileum dataset introduced in the Baysor study, which uses Cellpose-derived 3D membrane priors for volumetric segmentation [5]. Compared to the original Baysor segmentation, TRACER-stitched whole cells exhibited a marked increase in purity, with further improvement after fine-tuning, accompanied by a reduction in conflict (Fig. 6A; Supplementary Fig. 3I). Transcript counts decreased modestly following TRACER correction, whereas signal strength decreased more substantially (Fig. 6A), indicating that transcripts pruned by TRACER were disproportionately associated with incoherent gene–gene relationships rather than with dominant expression programs. Similar trends were observed in the standard Xenium breast cancer dataset and in multimodal or nucleus-expanded Xenium lung cancer datasets (Supplementary Fig. 4D; Supplementary Fig. 7A,B).

We also benchmarked the fidelity of the whole-cell profiles generated by Baysor, TRACER (Supplementary Fig. 6) using Scrublet, a widely used scRNA-seq doublet detector [35]. Although Scrublet was not developed for imaging-based spatial transcriptomics, segmentation-induced transcript mixing manifests as "doublet-like" profiles in expression space [36]. Baysor-segmented cells exhibited elevated Scrublet scores concentrated in the cluster-bridge regions highlighted by NPMI-derived conflict metrics, consistent with mixed-lineage transcript contamination. Across all expected doublet-rate settings, TRACER-stitched and TRACER-fine-tuned segmentations showed near-zero predicted doublets and uniformly low Scrublet scores, indicating that TRACER restores gene–gene coherence and eliminates mixed-lineage artifacts.

Beyond metric-level comparisons, TRACER also improved downstream structure and interpretability. In the MERFISH ileum dataset, unsupervised Leiden clustering of TRACER-corrected profiles produced more clearly separated UMAP clusters than those obtained from the original Baysor segmentation (Supplementary Fig. 3E), while remaining consistent with the annotated cell types from the original study (Supplementary Fig. 3F). Differential marker gene expression analysis further revealed enhanced signal-to-noise ratios following TRACER stitching and fine-tuning (Fig. 6B–D). In addition, TRACER identified 3,275 partial cells in this dataset (Supplementary Fig. 3G), which formed distinct clusters with coherent marker gene-expression programs (Supplementary Fig. 3H), reinforcing the biological interpretability of TRACER-derived entities.

3 Discussion

Accurate cell segmentation is a foundational requirement for faithful representation of cell types and cell states in imaging-based spatial transcriptomics. However, these segmentation errors arise from applying 2D methods to inherently 3D tissues, leading to hybrid clusters and lineage-mixed profiles. These errors lead to transcript mixing, substantially distorting downstream analyses including cell clustering, marker gene interpretation, cell-type identification, and ligand–receptor interaction analysis.

Recent computational approaches increasingly apply graph neural networks (GNNs) or contrastive learning frameworks to transcript–transcript or cell–cell graphs [9, 37]. While powerful, these approaches typically involve model training and greater computational requirements, with less transparent decision rules. In contrast, TRACER adopts a transparent, deterministic formulation: it computes NPMI representing gene-gene association directly from nucleus-derived contexts in the data and uses these values to guide coherence-based pruning and hierarchical stitching. TRACER requires no neural architectures and thus no model learning. This deterministic structure ensures reproducibility and allows users to inspect and validate every stage of the refinement process. Supplementary Table 1 summarizes the runtime and memory characteristics of TRACER across three benchmark datasets, demonstrating that the full pipeline can be executed efficiently on standard hardware without GPU acceleration.

In addition to TRACER, this work introduces NPMI-derived coherence metrics at the cellular level—purity, conflict, and related quantities—as segmentation-agnostic diagnostics for evaluating transcriptional mixing. These address a critical gap: imaging based segmentation benchmarks cannot capture the fuzzy, overlapping, and functionally defined nature of cell boundaries [15]. Coherence-based metrics provide a platform-independent approach to evaluate assess the biological coherence of cellular profiles, rather than geometric segmentation accuracy. We anticipate that these metrics will be broadly useful for algorithm development, benchmarking, and routine quality control in spatial transcriptomics.

TRACER also reframes partial cells as natural consequences of sectioning a 3D tissue structure, rather than as noise to be discarded. Through modeling, simulation, and empirical analysis, we show that these entities encode biologically meaningful gene programs, spatial organization, and tissue-level structure when treated explicitly. Recent work has independently demonstrated the biological relevance of these unassigned transcript populations [38]. Supplementary Table 2 summarizes how transcripts are reassigned across TRACER stages (pruning, stitching, fine-tuning) into whole cells, partial cells, and reconstructed partial cells across all benchmark datasets. While these entities represent incomplete cellular profiles and should not be used to define fine-grained cell states, they offer valuable information about tissue architecture and spatial niches.

At the same time, important limitations remain. Like many computational methods, TRACER requires parameter tuning, such as the maximum spatial distance used during fine-tuning. Increasing this threshold can merge spatially distant transcripts into overly large partial cells (Supplementary Fig. 7C–F). Although our recommended default thresholds performed consistently across tissues and platforms, future work integrating

adaptive or data-driven distance estimation may further improve robustness. Computing NPMI from noisy input segmentation may yield misleading gene associations, and undersampling of rare genes may create spurious mutual exclusion—an inherent issue with data-driven methods. Furthermore, population-level statistics may not capture rare transitional states, such as cells undergoing epithelial–mesenchymal transition or macrophage repolarization, which may inadvertently be pruned by TRACER despite legitimately expressing markers that are negatively associated at the population level. To mitigate these issues, we computed NPMI from nuclear segmentations, which are less susceptible to boundary errors than whole-cell segmentations, and observed no obvious artifacts. We also adopted a conservative pruning strategy: using a moderate NPMI threshold and ignoring evidence of perfect mutual exclusion between rare gene pairs. Future work could address these limitations by incorporating external data sources such as single-cell RNA-seq atlases for NPMI computation, confidence intervals to account for gene frequency and estimation uncertainty, and trajectory-informed coherence metrics to preserve transitional populations.

Nevertheless, the modular design of TRACER enables independent modification of each stage—co-expression estimation, pruning, stitching, hierarchical assembly, and spatial fine-tuning—without restructuring the overall pipeline. As imaging-based spatial transcriptomics moves toward thicker tissues and true 3D acquisition, coherence-based refinement of cellular profiles will become increasingly critical.

4 Figures

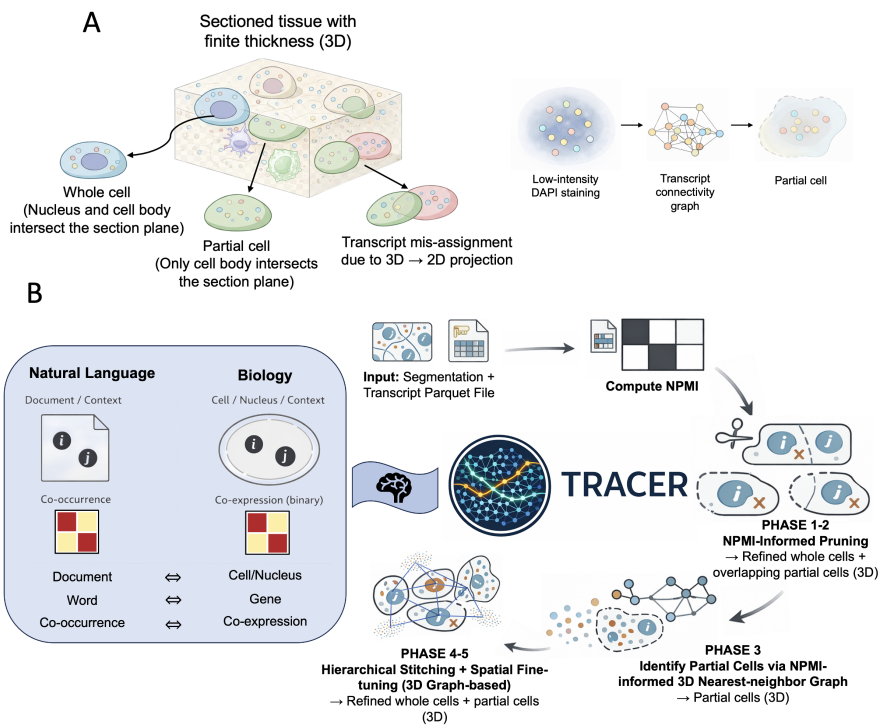


Figure 1. Biological motivation and algorithmic overview of TRACER. (A) 2D-segmentation on non-negligible tissue thickness produces errors in imaging-based spatial transcriptomics . Cells located above or below the imaging plane may intersect the section only partially, producing whole cells (nucleus and cytoplasm captured in-plane) and partial cells (cytoplasmic transcripts captured without an in-plane nucleus). Overlap of cellular profiles along the z-axis causes transcripts from different cells to project into the same two-dimensional footprint, resulting in systematic transcript misassignment. In addition, low-intensity or out-of-plane nuclei can leave clusters of transcripts unassigned under standard segmentation. These transcripts can be represented as a spatial transcript-connectivity graph, allowing coherent partial cellular profiles to be identified. (B) Conceptual framework and workflow of TRACER. TRACER draws an analogy to natural language processing in which nuclei define local context windows and genes are treated as binary events, enabling estimation of a data-driven gene–gene co-expression prior via normalized pointwise mutual information (NPMI). This prior guides: (1) NPMI-informed extraction of transcriptionally coherent whole cells and partial cells (Phases 1–2); (2) identification of partial cellular profiles from spatially structured unassigned transcripts using an NPMI-informed three-dimensional nearest-neighbor graph (Phase 3); and (3) hierarchical stitching and spatial fine-tuning of reconstructed cellular entities using a ΔC coherence-optimization criterion (Phases 4–5). The final output is a refined, three-dimensional–aware cellular representation that accounts for segmentation errors, incomplete cellular profiles, and transcript misassignment.

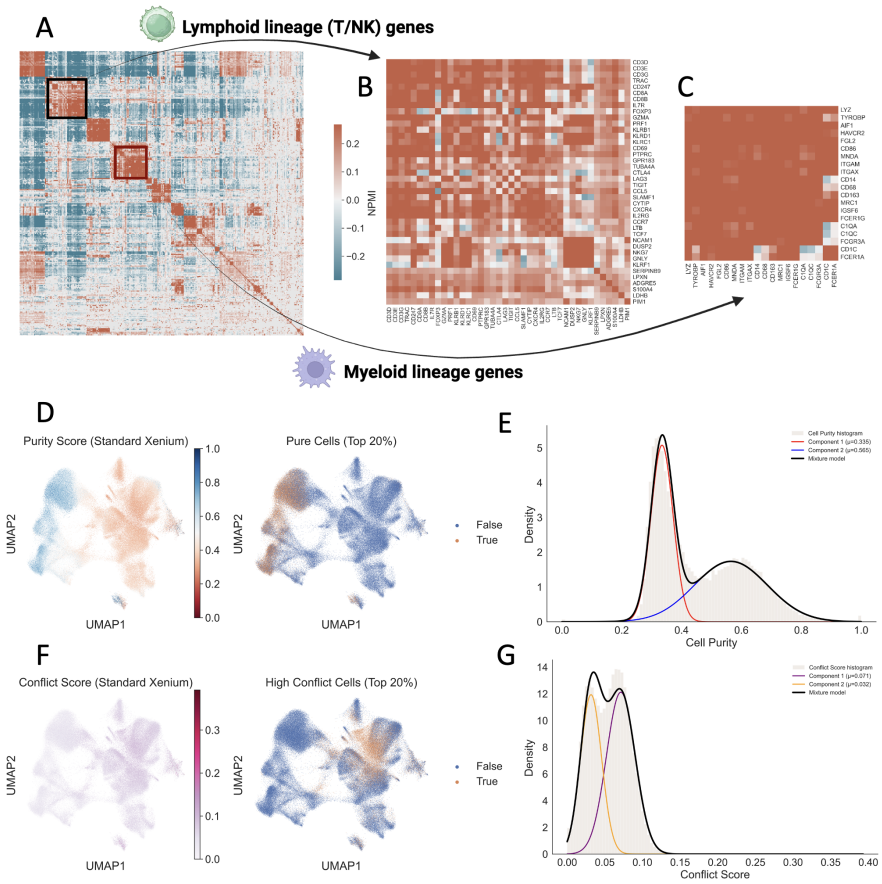


Figure 2. NPMI reveals lineage-resolved gene programs and enables quantitative purity and conflict scoring. (A) Global gene–gene NPMI matrix computed from the standard Xenium breast cancer dataset, showing block-diagonal lineage programs and negative cross-lineage associations. Insets (black boxes) correspond to panels B–C. (B,C) Representative lymphoid (B) and myeloid (C) gene modules with uniformly high within-lineage NPMI. (D) UMAP embedding of whole-cell profiles colored by the NPMI-derived purity score, which quantifies internal gene–gene coherence. Cells in the top 20% of purity scores are indicated, marking transcriptionally coherent entities. (E) Distribution of purity scores across cells with a two-component Gaussian mixture model (GMM), revealing a bimodal structure separating high-purity from low-purity profiles. (F) UMAP embedding colored by the NPMI-derived conflict score, which captures mutually incompatible gene associations within individual cells. Cells in the top 20% of conflict scores localize predominantly to regions connecting distinct UMAP clusters. (G) Distribution of conflict scores and corresponding two-component GMM fit, separating low-conflict and high-conflict states associated with mixed-lineage transcriptional signatures.

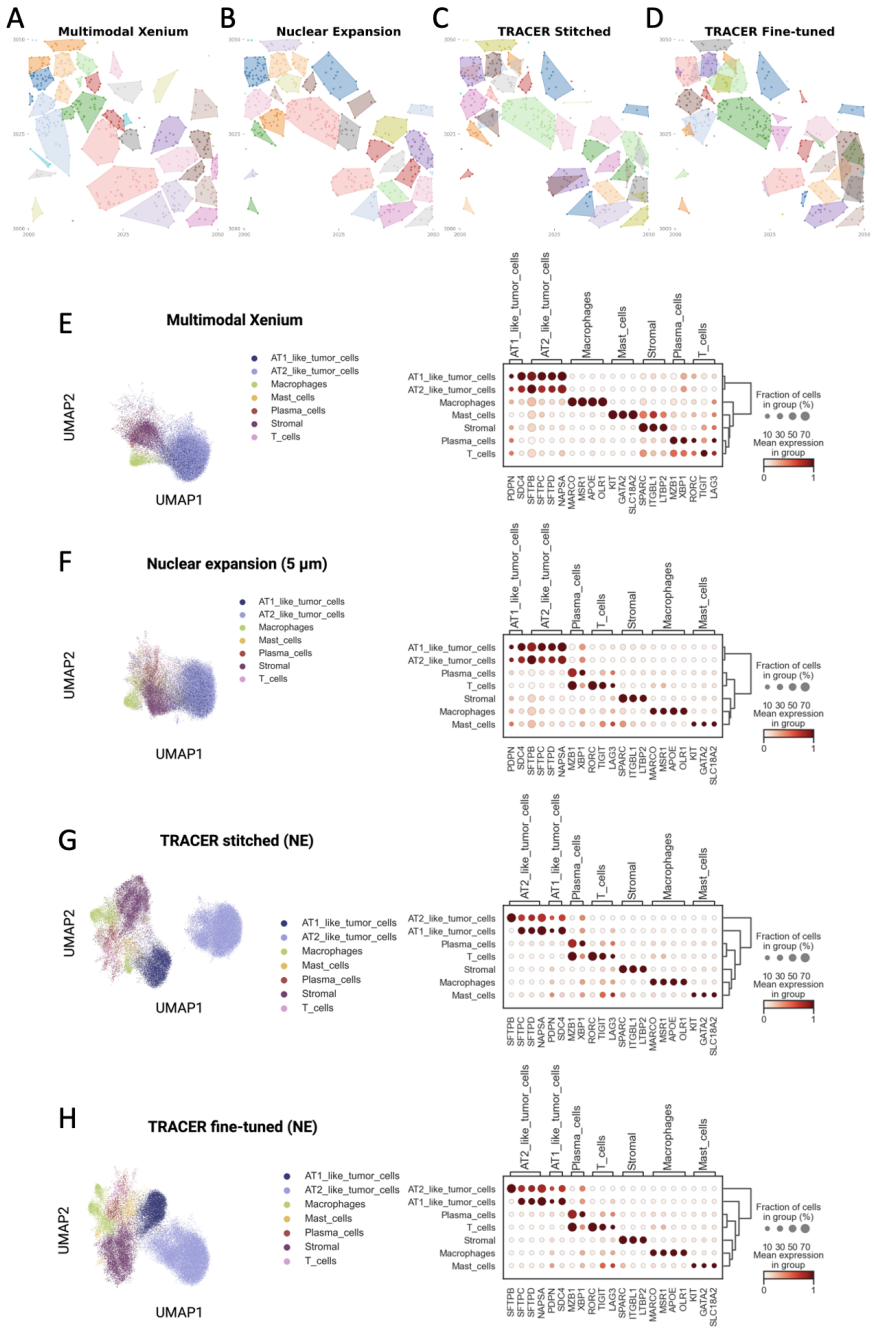

A
Multimodal Xenium
B
Nuclear Expansion
C
TRACER Stitched
D
TRACER Fine-tuned
E
Multimodal Xenium
F
Nuclear expansion (5 µm)
G
TRACER stitched (NE)
H
TRACER fine-tuned (NE)
UMAP1
UMAP2
AT1_like_tumor_cells
AT2_like_tumor_cells
Macrophages
Mast_cells
Plasma_cells
Stromal
T_cells
Fraction of cells in group (%)
10 30 50 70
Mean expression in group
0
1

Figure 3. **TRACER refines transcript assignment and cell-type resolution without multimodal staining.** (A–D) Convex hulls generated from transcripts assigned to cells using (A) multimodal Xenium segmentation; (B) 5-μm nuclear expansion (NE) without multimodal information, both showing strict boundary enforcement. (C) TRACER-stitched refinement applied to the NE baseline; (D) TRACER-fine-tuned refinement applied to the NE baseline, both allowing overlapping cells. Convex hulls are for illustrative purposes only. (E–H) UMAP embeddings and corresponding marker-gene dot plots for each segmentation strategy. (E) Multimodal Xenium segmentation; (F) 5-μm NE without multimodal staining, showing cross-lineage spillover. (G) TRACER-stitched refinement applied to the NE baseline; (H) TRACER-fine-tuned refinement applied to the NE baseline, both exhibiting improved signal-to-noise ratio and clearer lineage separation.

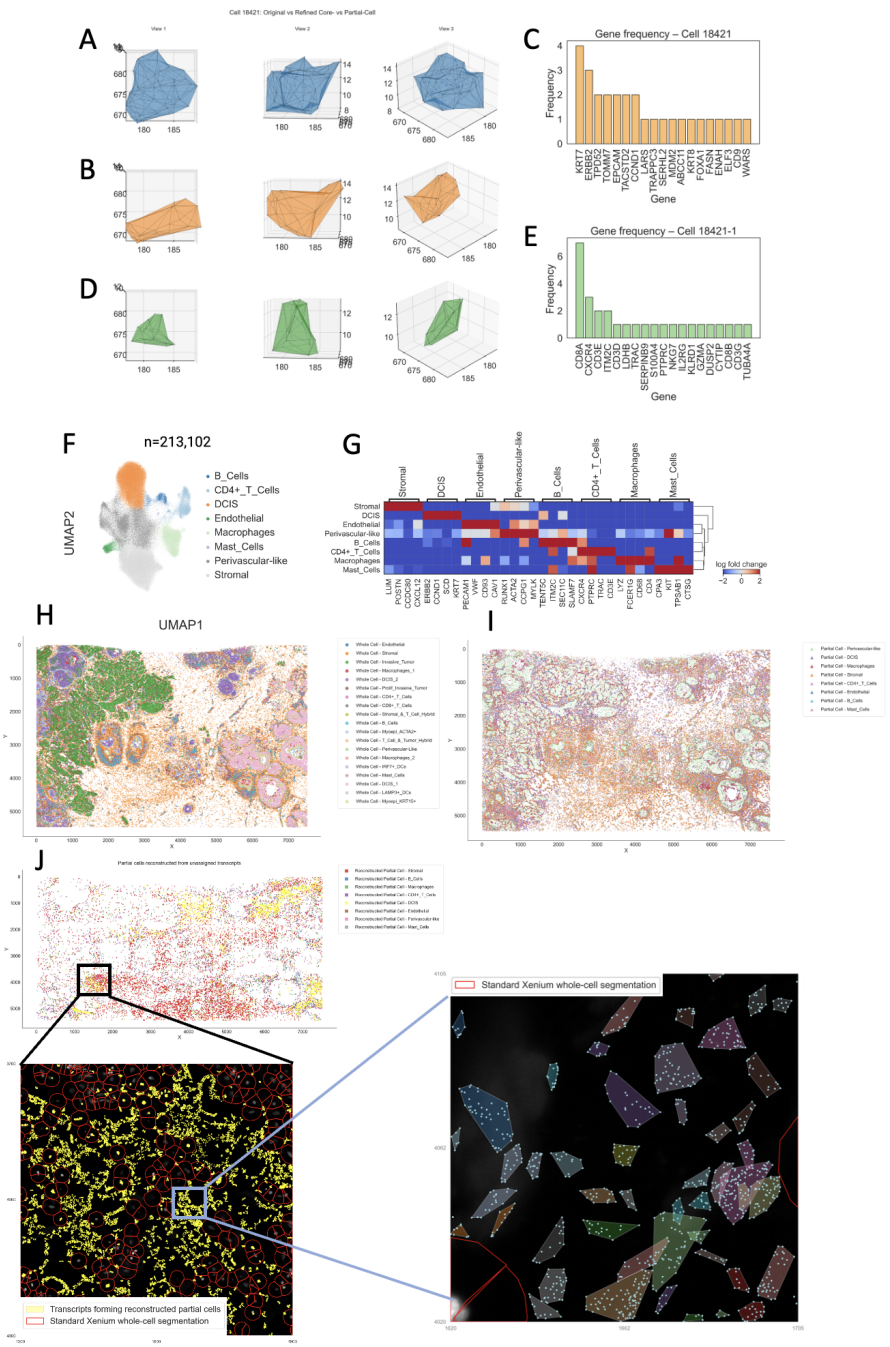

A
B
D
C
Gene frequency – Cell 18421
Frequency
Gene
E
Gene frequency – Cell 18421-1
F
n=213,102
UMAP2
B_Cells
CD4+_T_Cells
DCIS
Endothelial
Macrophages
Mast_Cells
Perivascular-like
Stromal
G
log fold change
H
UMAP1
I
J
Standard Xenium whole-cell segmentation
Transcripts forming reconstructed partial cells
Standard Xenium whole-cell segmentation

Figure 4. **Recovery and characterization of partial cellular profiles with TRACER.** (A) Three-dimensional concave hull constructed from the transcript point cloud originally assigned to a representative T cell–tumor hybrid cell under standard Xenium segmentation. (B,C) TRACER-extracted component corresponding to the tumor-associated subset of transcripts from the hybrid cell: (B) 3D concave hull and (C) transcript frequency profile. (D,E) TRACER-extracted component corresponding to the T cell–associated subset of transcripts: (D) 3D concave hull and (E) transcript frequency profile. (F) UMAP embedding of partial cells identified by TRACER after stitching and spatial fine-tuning. (G) Differential marker-gene expression matrix for Leiden clusters (annotated by marker-gene expression) computed from the partial cell population. (H) Spatial distribution of whole cells in the breast cancer Xenium section, colored by the original cell-type annotations. (I) Spatial distribution of TRACER-derived partial cells originating from segmentation-derived hybrid profiles, colored by Leiden clusters. (J) Spatial distribution of TRACER-derived partial cells reconstructed from previously unassigned transcripts, colored by Leiden clusters. The left inset shows transcripts contributing to reconstructed partial cells (yellow) overlaid on standard Xenium whole-cell segmentation boundaries (red). The right inset shows a convex-hull representation of these reconstructed partial cell transcripts within extracellular regions.

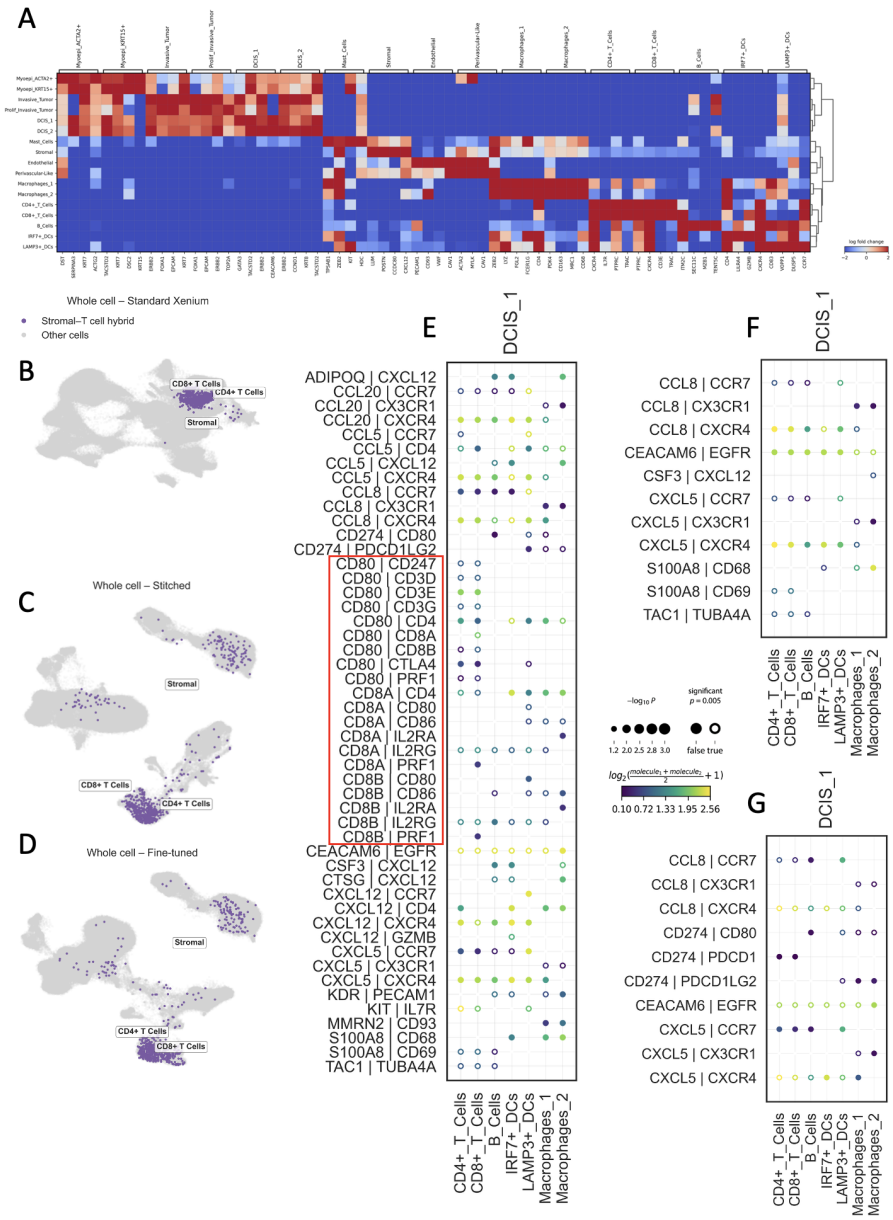

A
Whole cell – Standard Xenium
Stromal–T cell hybrid
Other cells
B
CD8+ T Cells
CD4+ T Cells
Stromal
C
Whole cell – Stitched
Stromal
CD8+ T Cells
CD4+ T Cells
D
Whole cell – Fine-tuned
Stromal
CD4+ T Cells
CD8+ T Cells
E
DCIS_1
ADIPOQ | CXCL12
CCL20 | CCR7
CCL20 | CX3CR1
CCL20 | CXCR4
CCL5 | CCR7
CCL5 | CD4
CCL5 | CXCL12
CCL5 | CXCR4
CCL8 | CCR7
CCL8 | CX3CR1
CCL8 | CXCR4
CD274 | CD80
CD274 | PDCD1LG2
CD80 | CD247
CD80 | CD3D
CD80 | CD3E
CD80 | CD3G
CD80 | CD4
CD80 | CD8A
CD80 | CD8B
CD80 | CTLA4
CD80 | PRF1
CD8A | CD4
CD8A | CD80
CD8A | CD86
CD8A | IL2RA
CD8A | IL2RG
CD8A | PRF1
CD8B | CD80
CD8B | CD86
CD8B | IL2RA
CD8B | IL2RG
CD8B | PRF1
CEACAM6 | EGFR
CSF3 | CXCL12
CTSG | CXCL12
CXCL12 | CCR7
CXCL12 | CD4
CXCL12 | CXCR4
CXCL12 | GZMB
CXCL5 | CCR7
CXCL5 | CX3CR1
CXCL5 | CXCR4
KDR | PECAM1
KIT | IL7R
MMRN2 | CD93
S100A8 | CD68
S100A8 | CD69
TAC1 | TUBA4A
CD4+_T_Cells
CD8+_T_Cells
B_Cells
IRF7+_DCs
LAMP3+_DCs
Macrophages_1
Macrophages_2
F
DCIS_1
CCL8 | CCR7
CCL8 | CX3CR1
CCL8 | CXCR4
CEACAM6 | EGFR
CSF3 | CXCL12
CXCL5 | CCR7
CXCL5 | CX3CR1
CXCL5 | CXCR4
S100A8 | CD68
S100A8 | CD69
TAC1 | TUBA4A
CD4+_T_Cells
CD8+_T_Cells
B_Cells
IRF7+_DCs
LAMP3+_DCs
Macrophages_1
Macrophages_2
−log10 P
1.2 2.0 2.5 2.8 3.0
significant
p = 0.005
false true
0.10 0.72 1.33 1.95 2.56
G
DCIS_1
CCL8 | CCR7
CCL8 | CX3CR1
CCL8 | CXCR4
CD274 | CD80
CD274 | PDCD1
CD274 | PDCD1LG2
CEACAM6 | EGFR
CXCL5 | CCR7
CXCL5 | CX3CR1
CXCL5 | CXCR4
CD4+_T_Cells
CD8+_T_Cells
B_Cells
IRF7+_DCs
LAMP3+_DCs
Macrophages_1
Macrophages_2

Figure 5. TRACER resolves hybrid cellular profiles and removes segmentation-derived ligand–receptor artifacts. (A) Differential marker-gene expression matrix for Leiden clusters derived from standard Xenium segmentation after TRACER fine-tuning. Sharpened lineage-specific profiles and reduced cross-lineage signal are observed. (B) UMAP embedding of cells derived from the standard Xenium segmentation, colored by original study annotations. A cluster annotated as "Stromal–T cell hybrid" forms a contiguous manifold and contains cells whose expression profiles include both stromal- and T-cell–associated markers. (C,D) UMAP embeddings of cells after TRACER refinement of the standard Xenium baseline. In TRACER-stitched (C) and TRACER–fine-tuned (D) embeddings, cells originally grouped within the hybrid manifold occupy two distinct regions corresponding to stromal-associated and T cell–associated expression profiles. No external annotations or marker-based constraints were used during TRACER processing; annotations are shown for visualization only. (E–G) Ligand–receptor permutation analysis (FDR < 0.05) between DCIS-1 and immune cell groups. Using standard Xenium segmentation (E), DCIS-1 cells appear to express immune-restricted ligands (e.g., CD80, CD8A, CD8B), producing unlikely interactions. These interactions are not observed when using TRACER-stitched (F) or TRACER-fine-tuned (G) transcript assignments, in which DCIS-1 retains only epithelial-associated ligand–receptor signals.

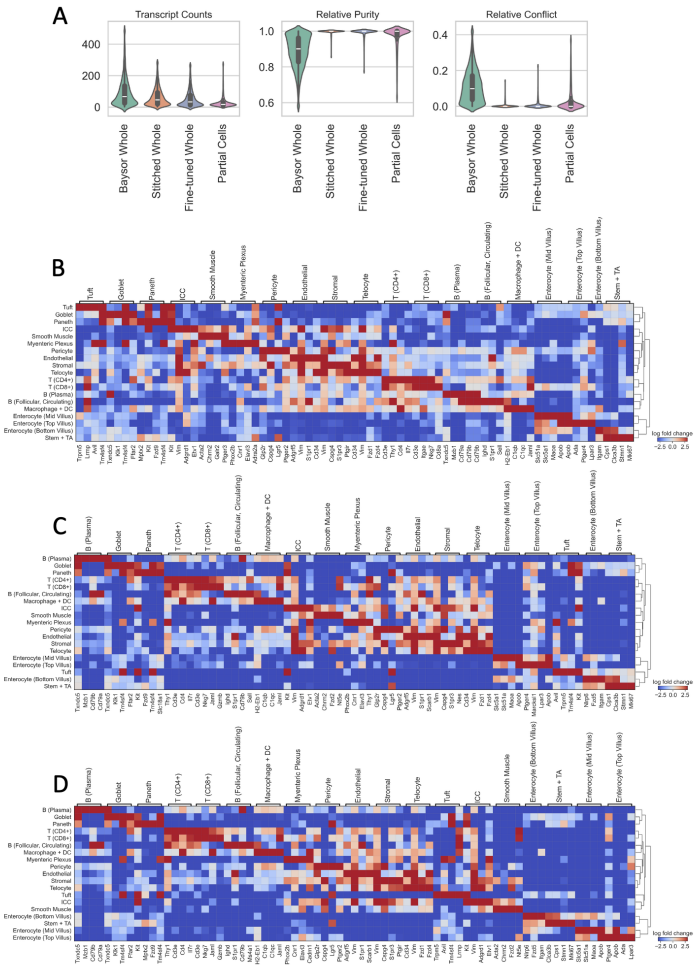
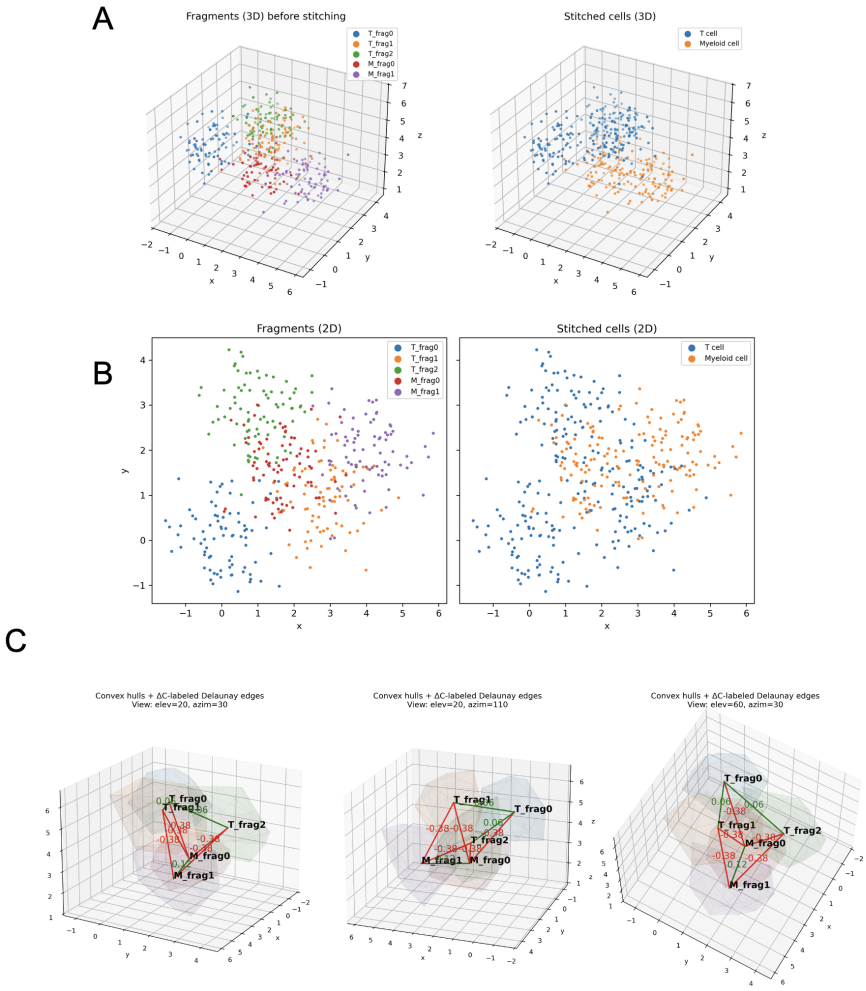
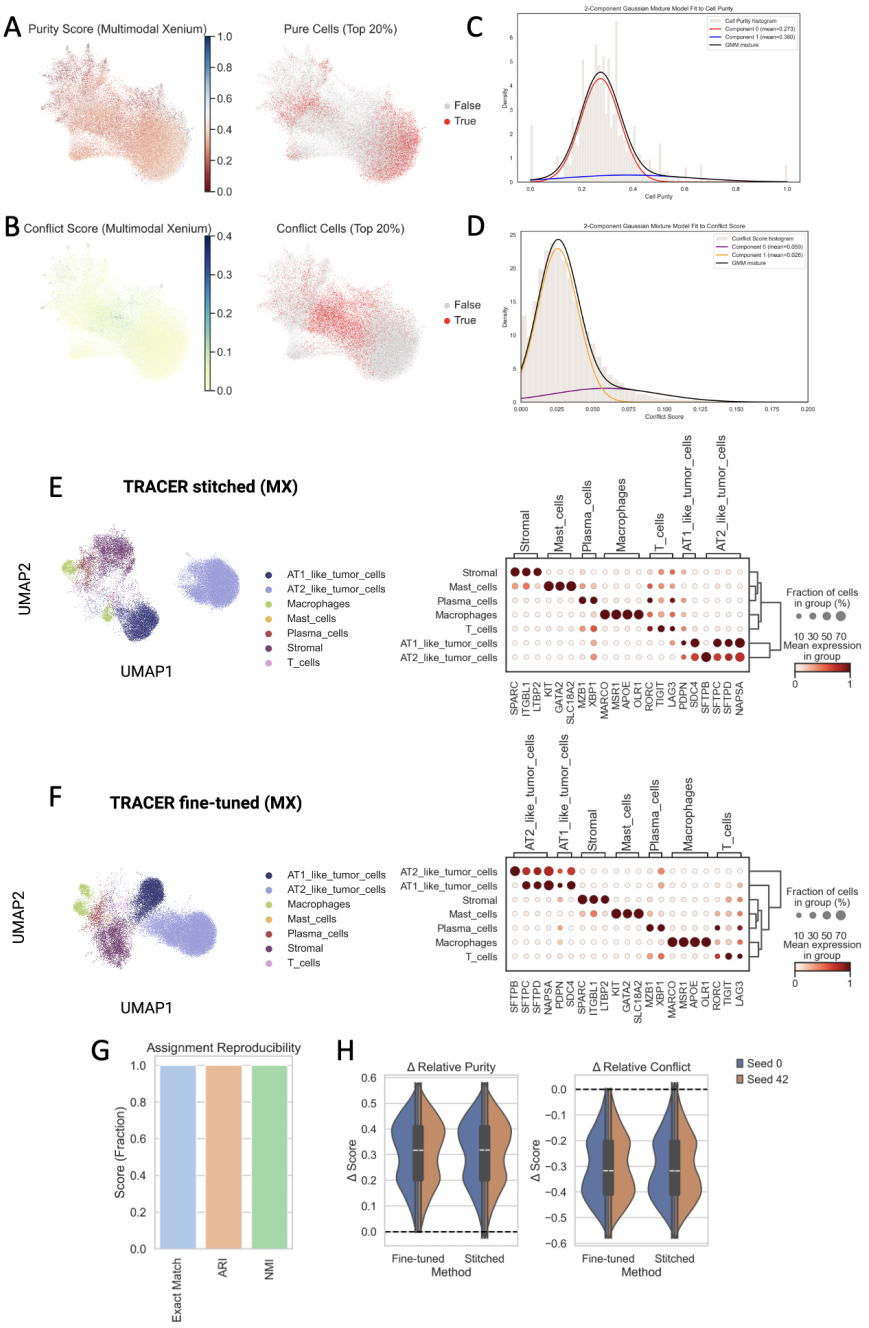


Figure 6. Quantitative characterization of TRACER output relative to 3D membrane–guided segmentation. (A) Summary distributions of transcript counts, NPMI-derived relative purity, and relative conflict for cellular entities obtained using different segmentation strategies: Baysor with Cellpose-based 3D membrane priors, TRACER-stitched whole cells, TRACER-fine-tuned whole cells, and TRACER-extracted partial cells. Relative purity reflects the internal agreement of gene–gene associations, whereas relative conflict reflects the presence of mutually incompatible gene pairs. TRACER-derived cells exhibit consistently higher relative purity and lower relative conflict compared with the Baysor segmentation. (B) Differential marker-gene expression matrix for cell types annotated in the Baysor-segmented mouse ileum dataset, representing the transcriptional organization obtained from a 3D membrane–guided reference segmentation. (C,D) Differential marker-gene expression matrices for cell groups identified using TRACER-stitched (C) and TRACER-fine-tuned (D) outputs. Despite no cell-type annotations or marker genes were used during TRACER refinement, TRACER output exhibit improved signal-to-noise ratio relative to Baysor. Annotations are shown for visualization only.

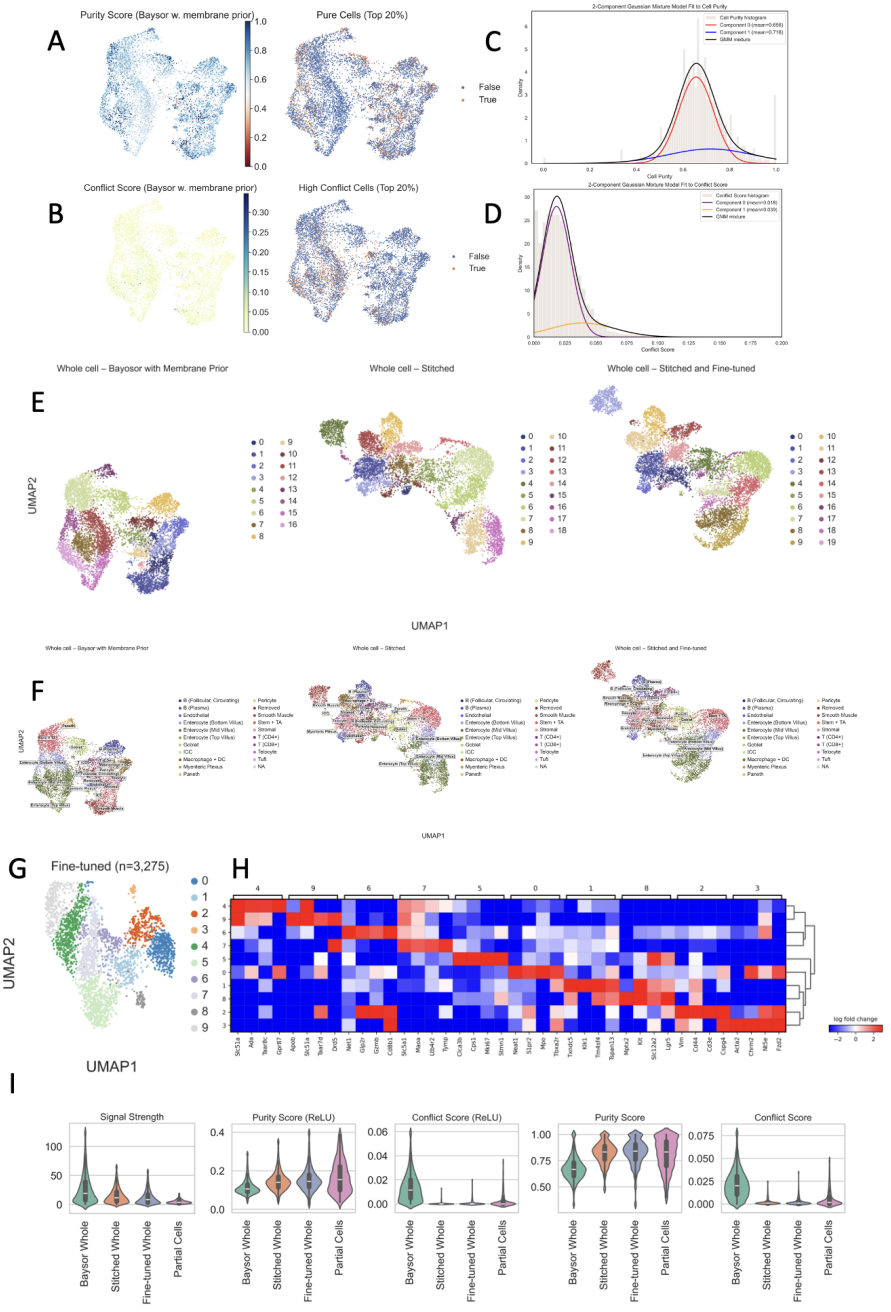
5 Supplementary Materials



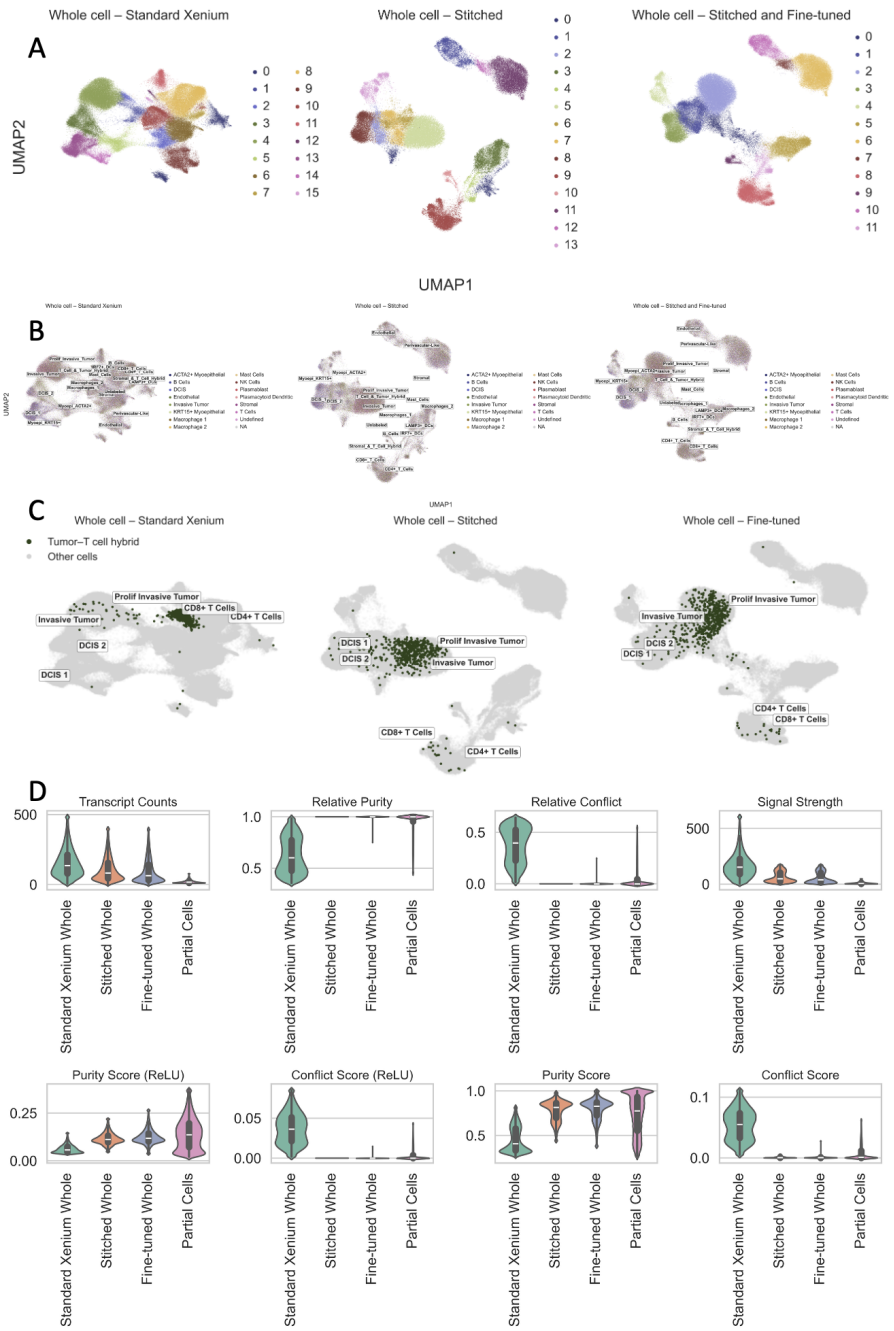
Supplementary Figure 1. **ΔC-guided stitching resolves fragmented cellular profiles in a simulated 3D system.** (A) Simulated three-dimensional transcript point clouds representing five fragments (T_frag0, T_frag1, T_frag2, M_frag0, M_frag1) derived from two underlying cells (one T cell and one myeloid cell). (B) Two-dimensional projections of the same fragments, showing substantial spatial overlap that would confound merging strategies based solely on 2D segmentation strategies. (C) ΔC-based coherence-guided stitching applied to all fragment pairs. Candidate merges are evaluated by the change in transcriptional coherence (ΔC), enabling selective merging of fragments originating from the same cell while rejecting merges between transcriptionally incompatible fragments. Edge colors denote ΔC values for each candidate pair, illustrating coherent (positive ΔC) versus incoherent (negative ΔC) associations.


A
Purity Score (Multimodal Xenium)
Pure Cells (Top 20%)
False
True
B
Conflict Score (Multimodal Xenium)
Conflict Cells (Top 20%)
False
True
C
D
E
TRACER stitched (MX)
UMAP2
UMAP1
AT1_like_tumor_cells
AT2_like_tumor_cells
Macrophages
Mast_cells
Plasma_cells
Stromal
T_cells
Fraction of cells in group (%)
10 30 50 70
Mean expression in group
F
TRACER fine-tuned (MX)
UMAP2
UMAP1
G
Assignment Reproducibility
Score (Fraction)
Exact Match
ARI
NMI
H
Δ Relative Purity
Δ Relative Conflict
Δ Score
Fine-tuned
Stitched
Method
Seed 0
Seed 42

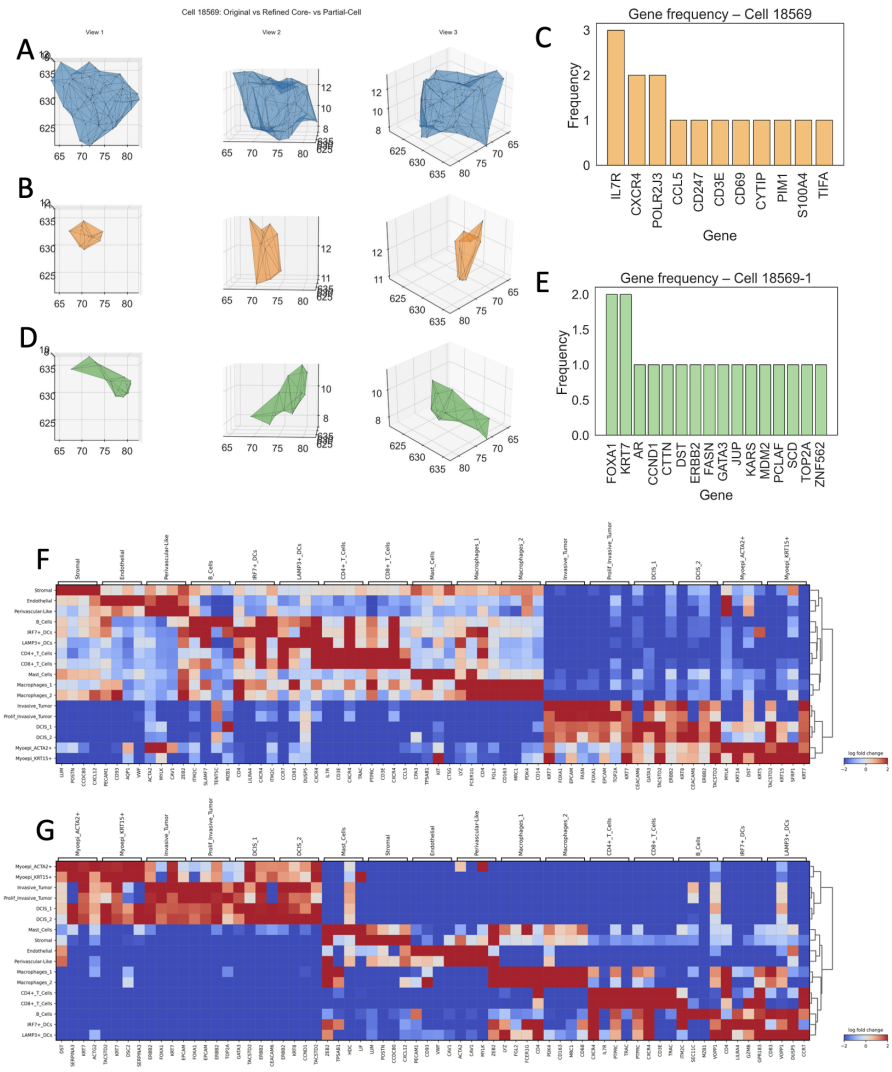
Supplementary Figure 2. **TRACER refinement and purity/conflict profiling in multimodal Xenium lung cancer data.** (A) UMAP embedding of whole-cell profiles from a multimodal Xenium (MX) lung cancer dataset colored by the NPMI-derived purity score. Cells within the top 20th percentile of the purity distribution are highlighted. (B) UMAP embedding colored by the NPMI-derived conflict score, with the top 20th percentile of high-conflict cells highlighted. (C,D) Distributions of purity (C) and conflict (D) scores fitted using two-component Gaussian mixture models (GMMs). (E,F) UMAP embeddings and corresponding marker-gene dot plots for TRACER refinements applied to the multimodal Xenium segmentation. (E) TRACER-stitched refinement. (F) TRACER-fine-tuned refinement. (G) Assignment reproducibility across two independent TRACER runs with different random seeds (0 and 42), quantified using exact-match fraction, adjusted Rand index (ARI), and normalized mutual information (NMI). TRACER assignments are identical across runs (exact-match fraction = 1.0). (H) Per-cell changes in relative purity and relative conflict across the two runs, with refinement-associated shifts toward higher purity and lower conflict at the per-cell level.


A
Purity Score (Baysor w. membrane prior)
Pure Cells (Top 20%)
False
True
B
Conflict Score (Baysor w. membrane prior)
High Conflict Cells (Top 20%)
False
True
C
2-Component Gaussian Mixture Model Fit to Cell Purity
Cell Purity Histogram
GMM mixture
Cell Purity
Density
D
2-Component Gaussian Mixture Model Fit to Conflict Score
Conflict Score Histogram
GMM mixture
Conflict Score
Density
E
Whole cell – Bayosor with Membrane Prior
Whole cell – Stitched
Whole cell – Stitched and Fine-tuned
UMAP2
UMAP1
F
Whole cell – Bayosor with Membrane Prior
Whole cell – Stitched
Whole cell – Stitched and Fine-tuned
UMAP2
UMAP1
G
Fine-tuned (n=3,275)
UMAP2
UMAP1
H
log fold change
I
Signal Strength
Purity Score (ReLU)
Conflict Score (ReLU)
Purity Score
Conflict Score
Baysor Whole
Stitched Whole
Fine-tuned Whole
Partial Cells

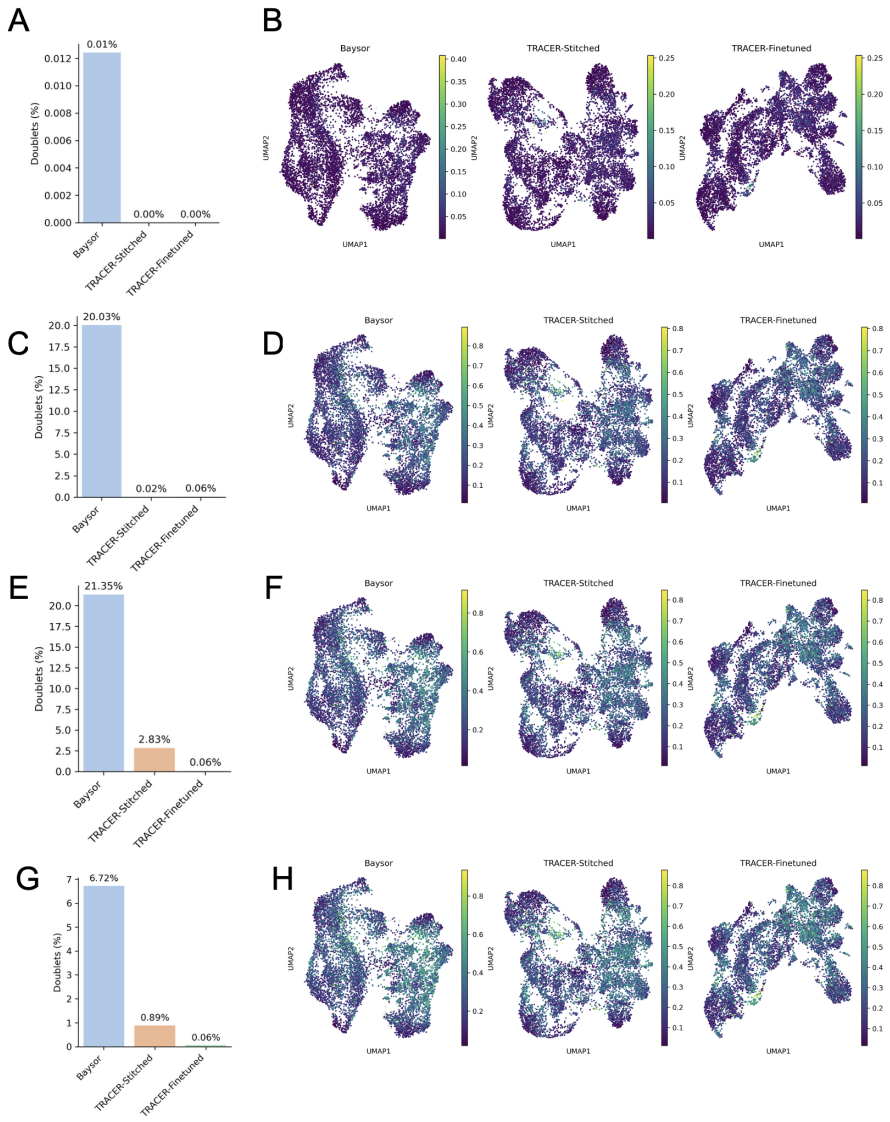
Supplementary Figure 3. Purity/conflict scoring, TRACER refinement, and coherence metrics for MERFISH mouse ileum segmentation with membrane-informed 3D priors. (A) UMAP embedding of whole-cell profiles obtained from Baysor segmentation incorporating Cellpose-derived three-dimensional membrane priors, colored by NPMI-derived purity scores. Cells in the top 20th percentile of purity are highlighted. (B) UMAP embedding colored by NPMI-derived conflict scores, with high-conflict cells (top 20th percentile) highlighted. (C,D) Distributions of purity (C) and conflict (D) scores fitted with two-component Gaussian mixture models (GMMs). (E) Leiden clustering overlaid on UMAP embeddings for Baysor + membrane priors (left), TRACER-stitched whole cells (middle), and TRACER–fine-tuned whole cells (right). (F) Cell-type annotations from the original study overlaid on the corresponding embeddings in (E). (G) UMAP embedding of TRACER-identified partial cells after stitching and spatial fine-tuning, colored by Leiden cluster identity. (H) Differential marker-gene expression matrix for Leiden clusters shown in (G). (I) Summary metrics computed per cellular entity across segmentation strategies, including signal strength, NPMI-based purity (ReLU formulations) and conflict (ReLU formulations). Metrics are shown for membrane-informed Baysor cells, TRACER-stitched whole cells, TRACER–fine-tuned whole cells, and TRACER-extracted partial cells, highlighting increased purity and reduced conflict in TRACER-derived cellular entities relative to the Baysor.


A
Whole cell – Standard Xenium
Whole cell – Stitched
Whole cell – Stitched and Fine-tuned
UMAP2
UMAP1
B
C
Whole cell – Standard Xenium
Whole cell – Stitched
Whole cell – Fine-tuned
Tumor–T cell hybrid
Other cells
Prolif Invasive Tumor
Invasive Tumor
CD8+ T Cells
CD4+ T Cells
DCIS 1
DCIS 2
D
Transcript Counts
Relative Purity
Relative Conflict
Signal Strength
Purity Score (ReLU)
Conflict Score (ReLU)
Purity Score
Conflict Score
Standard Xenium Whole
Stitched Whole
Fine-tuned Whole
Partial Cells

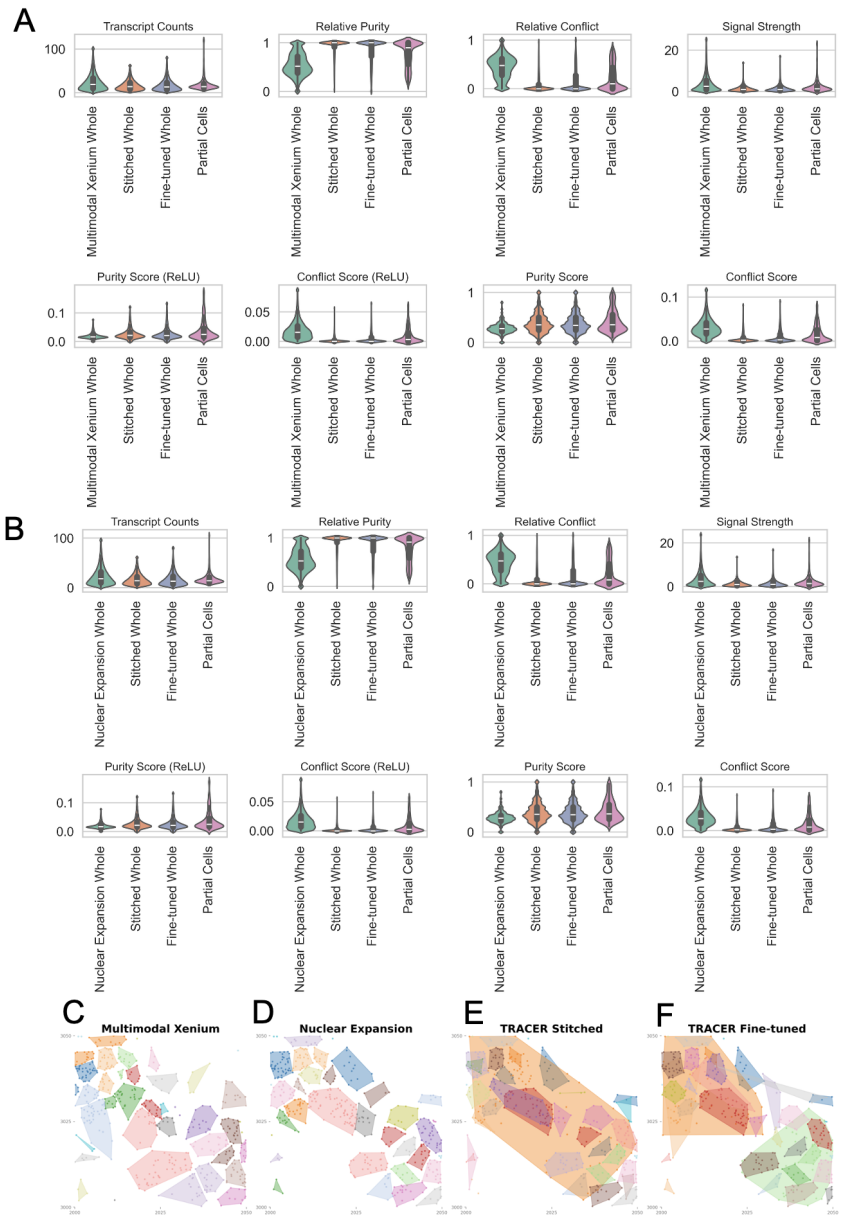
Supplementary Figure 4. **TRACER-induced changes in clustering structure, cell-type separation, and coherence metrics in standard Xenium breast cancer data.** (A) Leiden clustering overlaid on UMAP embeddings for the standard Xenium segmentation (left), TRACER-stitched whole cells (middle), and TRACER–fine-tuned whole cells (right). (B) Cell-type annotations from the original study overlaid on the corresponding embeddings in (A). These labels were not used in TRACER and are shown for visualization only. (C) Distribution of cells originally annotated as a tumor–T cell hybrid population. Under the original segmentation (left), these cells occupy a shared manifold. After TRACER stitching (middle) and subsequent fine-tuning (right), the same cells separate into distinct tumor and T cell manifolds. (D) Summary metrics computed per cellular entity across segmentation strategies, including transcript counts, NPMI-derived purity (ReLU formulations), relative purity, relative conflict, signal strength, and conflict (ReLU formulations). Purity metrics quantify internal gene–gene coherence; conflict metrics quantify mutually incompatible associations; signal strength reflects total coherence magnitude (sum of |NPMI| values). Comparisons are shown for standard Xenium whole-cell segmentation, TRACER-stitched whole cells, TRACER–fine-tuned whole cells, and TRACER-extracted partial cells, highlighting increased purity and reduced conflict after TRACER refinement.



Supplementary Figure 5. Resolution of mixed-lineage cellular profiles and coherence-guided decomposition of ambiguous cells. (A) Three-dimensional concave hull constructed from the transcript point cloud originally assigned to an unlabeled cell under the standard Xenium segmentation. This cell displayed promiscuous marker gene expression, preventing confident assignment in the original study. Three orthogonal views of the reconstructed concave hull are shown. (B,C) TRACER-derived component corresponding to a T cell, represented as a 3D concave hull (B) together with its gene frequency profile (C). (D,E) TRACER-derived component corresponding to a tumor cell, represented as a 3D concave hull (D) and its gene frequency profile (E). (F) Differential marker gene expression matrix for cell-type annotations defined by the original study applied to the standard Xenium segmentation, showing mixed-lineage signal across annotated groups. (G) Differential marker gene expression matrix for the same published cell-type annotations applied to TRACER-stitched correction, showing more distinct lineage-associated marker expression. Cell-type labels are derived entirely from the original peer-reviewed dataset and were not used during TRACER refinement.



Supplementary Figure 6. Scrublet-based expression doublet scores across segmentation strategies. (A–B) Scrublet applied with an expected doublet rate of 0.02. (A) Predicted doublet percentages for Baysor, TRACER-stitched, and TRACER-fine-tuned whole-cell profiles. (B) UMAP embeddings colored by Scrublet doublet scores. (C–D) Scrublet predictions using an expected doublet rate of 0.20. (E–F) Scrublet predictions using an expected doublet rate of 0.25. (G–H) Scrublet predictions using an expected doublet rate of 0.30. Across expected doublet priors, TRACER-derived cellular profiles exhibit consistently near-zero predicted doublet fractions and remain stable across Scrublet settings, whereas Baysor-derived profiles show substantially higher predicted doublet rates and stronger sensitivity to the assumed doublet prior.


A
Transcript Counts
Relative Purity
Relative Conflict
Signal Strength
Purity Score (ReLU)
Conflict Score (ReLU)
Purity Score
Conflict Score
Multimodal Xenium Whole
Stitched Whole
Fine-tuned Whole
Partial Cells
B
Nuclear Expansion Whole
C
Multimodal Xenium
D
Nuclear Expansion
E
TRACER Stitched
F
TRACER Fine-tuned

Supplementary Figure 7. **Summary coherence metrics and convex-hull visualizations for lung cancer multimodal Xenium across segmentation strategies.** (A) Summary metrics computed per cellular entity for multimodal Xenium whole-cell segmentation, TRACER-stitched whole cells, TRACER-fine-tuned whole cells, and TRACER-extracted partial cells. Metrics include transcript counts, NPMI-derived purity (ReLU formulation), relative purity, relative conflict, signal strength, and conflict (ReLU formulation). Purity metrics quantify internal gene–gene coherence; conflict metrics quantify mutually incompatible gene associations; signal strength reflects the total magnitude of NPMI contributions (sum of |NPMI| values). After TRACER refinement, purity increases and conflict decreases compared with multimodal segmentation. (B) Summary metrics computed per cellular entity for 5-μm nuclear-expansion segmentation, TRACER-stitched whole cells, TRACER-fine-tuned whole cells, and TRACER-extracted partial cells. After TRACER refinement, purity increases and conflict decreases compared with 5-μm nuclear-expansion segmentation. (C–F) Convex hulls generated from transcript assignments under each segmentation method. (C) Multimodal Xenium segmentation. (D) 5-μm nuclear-expansion segmentation without multimodal information. (E) TRACER-stitched refinement applied to the nuclear-expansion baseline using relaxed spatial constraints (15 μm). (F) TRACER-fine-tuned refinement applied to the nuclear-expansion baseline using relaxed spatial constraints (15 μm).

Supplementary Table 1. **TRACER runtime and memory usage on three benchmark datasets.** Wall-clock time was measured via SLURM (`sacct`); per-stage runtimes were obtained from internal logs. All jobs used 17 allocated CPU cores and 96 GB requested memory.

Dataset	Platform	# transcripts	Alloc-CPUs	Req-Mem	MaxRSS (GiB)	Stage 1 (min)	Stage 2 (min)	Stage 3 (min)	Stage 4 (min)	Stage 5 (min)	Total (min)
Breast cancer Xenium	Xenium	28,059,774	17	96G	26.9	2.30	84.40	8.20	72.82	142.30	311.1 (5:11 h)
Lung cancer Xenium	Xenium	1,436,900	17	96G	2.45	0.30	0.10	1.00	0.50	4.10	6.3 (0:06 h)
Mouse ileum MERFISH	MERFISH	819,665	17	96G	4.57	0.06	0.01	0.13	0.09	1.12	1.75 (0:01 h)

Supplementary Table 2. **Transcript assignment across TRACER stages on three benchmark datasets.** For each dataset and TRACER stage, the table reports the number of transcripts assigned to whole cells, partial cells, reconstructed partial cells, or remaining unassigned.

Dataset	min_comp_size	Stage	Unassigned	Whole cell	Partial cell	Reconstructed partial cell	Total transcripts
Breast cancer Xenium	25	Original	1,988,154	26,071,620	0	0	28,059,774
Breast cancer Xenium	25	Stage 1	5,921,403	18,313,439	3,824,932	0	28,059,774
Breast cancer Xenium	25	Stage 2	5,299,316	18,313,439	3,824,932	622,087	28,059,774
Breast cancer Xenium	25	Stage 3 (TRACER stitched)	5,299,316	18,326,600	3,816,187	617,671	28,059,774
Breast cancer Xenium	25	Stage 4	5,299,316	16,119,790	6,022,997	617,671	28,059,774
Breast cancer Xenium	25	Stage 5 (TRACER finetuned)	5,299,316	16,307,533	5,921,089	531,836	28,059,774
Lung cancer Xenium	10	Original	1,822	1,435,078	0	0	1,436,900
Lung cancer Xenium	10	Stage 1	99,128	1,113,094	224,678	0	1,436,900
Lung cancer Xenium	10	Stage 2	98,949	1,113,094	224,678	179	1,436,900
Lung cancer Xenium	10	Stage 3 (TRACER stitched)	98,949	1,200,105	137,678	168	1,436,900
Lung cancer Xenium	10	Stage 4	98,949	1,088,743	249,040	168	1,436,900
Lung cancer Xenium	10	Stage 5 (TRACER finetuned)	98,949	1,225,256	112,673	22	1,436,900
Mouse ileum MERFISH	10	Original	0	819,665	0	0	819,665
Mouse ileum MERFISH	10	Stage 1	129,604	509,098	180,963	0	819,665
Mouse ileum MERFISH	10	Stage 2	129,604	509,098	180,963	0	819,665
Mouse ileum MERFISH	10	Stage 3 (TRACER stitched)	129,604	537,832	152,229	0	819,665
Mouse ileum MERFISH	10	Stage 4	129,604	436,311	253,750	0	819,665
Mouse ileum MERFISH	10	Stage 5 (TRACER finetuned)	129,604	445,028	245,033	0	819,665

Note: The mouse ileum MERFISH dataset was preprocessed and filtered by the original study prior to TRACER analysis. Both Xenium datasets (breast cancer and lung cancer) correspond to raw, unpreprocessed transcript-level data.

6 Methods

6.1 Data acquisition and preprocessing

All spatial transcriptomics data were obtained from previously published work, including data from Xenium (standard and multimodal configurations) and MERFISH platforms. The breast cancer Xenium dataset with nucleus expanded (NE) segmentation and associated cell-type annotations dataset were obtained from the original study [23]. The lung cancer Xenium dataset with multimodal segmentation was obtained from a previously published study [25]. We resegmented this dataset with NE segmentation using Xenium Ranger to obtained the dataset used in Fig. 3. Xenium transcripts were filtered using a quality value threshold of QV > 30 prior to analysis. MERFISH data were obtained from the Baysor study demonstrating three-dimensional cell segmentation using Cellpose-based priors [5]. From this dataset, transcripts with nuclear probability greater than 0.9, as reported by the Baysor output, were considered nuclear and used for downstream analysis.

All analyses were performed independently for each tissue section unless otherwise stated.

6.2 Translating normalized pointwise mutual information (NPMI) to spatial transcriptomics

6.2.1 Conceptual framework

The coherence framework underlying this work is built on the translation of normalized pointwise mutual information (NPMI), originally developed in natural language processing, to spatial transcriptomics under a cell-as-context / gene-as-word paradigm.

In natural language processing, a document defines a unit of co-occurrence, and words are binary events indicating presence or absence within that document. Importantly, PMI and NPMI are defined over event occurrence, not word frequency. This formulation is robust to document length variation and focuses on semantic co-occurrence patterns.

We adopted this framework directly for biological data by defining:

- Context (document): a segmented nucleus
- Word (token): a gene
- Event i: gene i is present in the nucleus
- Event j: gene j is present in the nucleus
- Event (i, j): both genes i and j are present in the same nucleus

This formulation explicitly avoids dependence on transcript counts, which are subject to technical noise and segmentation artifacts, and instead emphasizes co-expression programs across cells, cross-lineage incompatibilities, and doublet or merged-cell signatures.

6.2.2 Nucleus-centered binary gene representation

Each segmented nucleus was treated as a context window, and a gene was considered present within a nucleus if and only if at least $N_{\min}$ transcripts were detected inside its boundary. In this study, we used $N_{\min} = 2$. This yielded a binary presence–absence matrix

$$M \in \{0,1\}^{C \times G}, \tag{1}$$

where C denotes the number of nuclei and G the number of genes.

No normalization, scaling, or transcript count weighting was applied beyond platform-specific quality filtering. This representation intentionally discards transcript abundance information in favor of robust detection of gene co-occurrence patterns across cells.

We selected nuclei rather than whole-cell segments as the primary context unit because nuclear segmentation is generally more reliable across platforms and tissues, whereas whole-cell boundaries are often inferred indirectly and may introduce systematic spatial uncertainty—particularly in standard Xenium datasets where cell outlines are extrapolated from nuclear signal. Using nuclei as contexts therefore provides a conservative and platform-agnostic basis for computing gene–gene associations.

6.2.3 Probability estimation

Let C denote the total number of nuclei. For each gene pair (i, j), we estimated marginal and joint probabilities as

$$P(i) = \frac{1}{C} \sum_{c=1}^{C} \mathbb{I}\left(M_{c,i} = 1\right), \tag{2}$$

$$P(j) = \frac{1}{C} \sum_{c=1}^{C} \mathbb{I}\left(M_{c,j} = 1\right), \tag{3}$$

$$P(i,j) = \frac{1}{C} \sum_{c=1}^{C} \mathbb{I}\left(M_{c,i} = 1 \wedge M_{c,j} = 1\right). \tag{4}$$

Each gene pair induces a complete four-part partition of the event space:

$$P(i,j) \quad \text{both genes present}, \tag{5}$$

$$P(i,\neg j) \quad \text{gene } i \text{ only}, \tag{6}$$

$$P(\neg i,j) \quad \text{gene } j \text{ only}, \tag{7}$$

$$P(\neg i,\neg j) \quad \text{neither gene present}. \tag{8}$$

By construction, these probabilities sum to one for every gene pair, ensuring a valid probabilistic interpretation.

6.2.4 NPMI computation

Pointwise mutual information (PMI) was computed as

$$\mathrm{PMI}(i,j) = \log_2 \left(\frac{P(i,j)}{P(i)P(j)} \right). \tag{9}$$

Because PMI is known to be biased toward rare events, particularly when $P(i,j)$ is small, we used normalized PMI (NPMI):

$$\mathrm{NPMI}(i,j) = \frac{\mathrm{PMI}(i,j)}{-\log P(i,j)} \in [-1,1]. \tag{10}$$

NPMI values of 0 indicate statistical independence, positive values indicate enrichment above independence, and negative values indicate depletion or mutual exclusivity. This normalization enables meaningful comparison across gene pairs with widely varying marginal frequencies.

The full gene–gene NPMI matrix was computed in a reference-free, data-driven manner using nuclei selected to represent the central distribution of transcript counts for Xenium datasets (excluding the lowest and highest 20th percentiles) and all nuclei for MERFISH datasets.

6.3 Associative clique extraction and transcript consistency modeling

6.3.1 Distinguishing whole cells and partial cells

In tissue-based spatial transcriptomics, sectioning of non-negligible thickness produces captures cells to varying degrees that may correspond either to a majority of the cellular contents being captured in the section or only a partial capture truncated by sectioning geometry. To explicitly account for this variation, we distinguished between whole cells and partial cells based on combined imaging and transcriptomic criteria.

Whole cells were defined as segmented regions exhibiting a strong, in-focus nuclear DAPI signal within the analyzed imaging plane, consistent with the presence of the nucleus and the majority of the cell body. In contrast, partial cells were defined as segmented regions containing cytoplasmic signal or transcripts but lacking a clearly in-plane nucleus, consistent with cells whose nuclei reside in adjacent sections along the z-axis. Partial cells arise frequently at tissue boundaries, within densely packed regions, or in areas with reduced nuclear signal, and may contribute spurious transcripts to neighboring cells.

Rather than discarding partial profiles, we treated them as explicit intermediate entities and incorporated them into downstream modeling in a conservative, coherence-driven manner.

6.3.2 NPMI-based conservative transcript pruning

To resolve transcript assignment ambiguity while preserving biologically coherent gene programs, we implemented a two-pass conservative pruning strategy based on gene–gene support derived from NPMI.

For each segmented cell, the set of expressed genes was evaluated using the precomputed gene–gene NPMI matrix computed from all nuclei in the dataset. For a given cell, each gene was assessed based on its aggregate mutual support with other genes expressed within the same cell, defined by pairwise NPMI values.

In the first pruning pass, genes failing a predefined NPMI consistency criterion were removed from the original cell and reassigned to a corresponding partial-cell identifier. This step preserved the core, internally consistent gene set of each cell while isolating weakly supported or potentially spurious transcripts without discarding them.

In the second pruning pass, the same NPMI-based evaluation was applied to these partial gene sets. Genes that again failed the consistency criterion were conservatively reassigned as unassigned rather than being forced into any cellular entity.

This two-stage procedure minimized over-pruning while progressively separating transcriptionally coherent gene programs from inconsistent associations. Importantly, partial cells were retained as explicit entities, allowing recovery of biologically meaningful but incomplete expression programs that would otherwise be lost.

6.4 Reconstruction of partial cells from unassigned transcripts

A subset of transcripts could not be confidently assigned to segmented cells, either because they were unassigned in the original segmentation or remained unassigned after two-pass pruning. Such unassigned transcripts may arise from extracellular regions that nonetheless reflect coherent cellular expression programs—for example, when a cell body lies within the analyzed plane but the nucleus is located in an adjacent section along the z-axis, or when nuclear signal is insufficient to pass automated segmentation thresholds.

To capture these structures, we introduced the concept of reconstructed partial cells, defined as spatially localized groups of transcripts lacking a prior segmentation mask but exhibiting strong internal spatial connectivity and transcriptional coherence consistent with a single cell.

Partial cell identification was performed using transcripts remaining unassigned after the two-pass pruning procedure, together with transcripts that were unassigned in the initial segmentation. These transcripts were embedded in three-dimensional physical space using their spatial coordinates and represented as nodes in a k-nearest neighbor (kNN) graph. Edges were formed between transcripts that were mutual nearest neighbors and lay within a fixed Euclidean distance threshold, enforcing local spatial connectivity. Connected components of this graph were extracted to yield candidate partial cell regions.

Components containing fewer than a minimum number of transcripts were discarded to reduce spurious aggregation. For the remaining components, transcriptional coherence was evaluated using the same NPMI framework applied to segmented cells. Within each component, genes were iteratively pruned using the aforementioned greedy

NPMI-based strategy, removing transcripts whose genes lacked sufficient mutual support with the dominant gene co-expression program of the component. Components retaining internally consistent gene programs after pruning were labeled as partial cells, while transcripts removed during this step were conservatively discarded.

Partial cells were treated as distinct cellular entities in downstream analyses and assigned unique identifiers, while maintaining explicit provenance from the unassigned transcript pool. This approach does not assume that unassigned transcripts are artifactual; instead, it enables recovery of transcriptionally coherent cellular profiles missed by nucleus-centered segmentation while preserving conservative handling of ambiguous signals.

6.5 Hierarchical stitching of cellular entities using ΔC coherence optimization

6.5.1 Entity representation and spatial constraints

To reconcile fragmented cellular entities arising from segmentation artifacts and partial profiles, we performed constrained hierarchical stitching based on transcriptional coherence. Each entity—whole cell, partial cell, or reconstructed partial cell—was represented by its centroid in physical space and by the set of genes associated with its transcripts.

Candidate merges were restricted to spatially adjacent entities, defined using a Delaunay triangulation of entity centroids (computed in three dimensions unless otherwise stated). This constraint limited stitching to biologically plausible spatial neighborhoods.

6.5.2 Coherence score based on gene–gene NPMI

For an entity with gene set $G = \{g_1, \ldots, g_k\}$, we defined a coherence score

$$C(G) = \text{purity}(G) - \text{conflict}(G), \tag{11}$$

computed from the gene–gene NPMI matrix.

For all gene pairs (g_i, g_j) with $i < j$, and finite NPMI values w_{ij}, purity and conflict were defined as

$$\text{purity}(G) = \frac{1}{|P|} \sum_{(i,j)\in P} \mathbb{I}\left(w_{ij} > \theta\right), \tag{12}$$

$$\text{conflict}(G) = \frac{1}{K} \sum_{(i,j)\in N} \left(-w_{ij}\right), \tag{13}$$

where P is the set of all observed gene pairs with finite NPMI, $N \subseteq P$ denotes pairs with negative NPMI values, $\mathbb{I}(\cdot)$ is the indicator function, $\theta = 0.05$ accounts for biological noise, and $K = \binom{k}{2}$ is the total number of possible gene pairs. If fewer than two genes were present or no valid gene pairs were observed, coherence was defined as zero.

Purity is normalized by the number of observed pairs $|P|$ because it measures the fraction of gene pairs exhibiting strong positive co-expression. Conflict, in contrast, quantifies the total incompatible gene–gene association pressure, and is therefore normalized by the total number of possible gene pairs $K = \binom{k}{2}$. This choice ensures that conflict increases both with the number and magnitude of contradictory associations, providing a stable and size-consistent indicator of mixed-lineage profiles.

Mathematical properties of the purity and conflict score

To justify the statistical behavior and interpretability of the NPMI-derived purity and conflict functions, we establish several basic properties. Proofs are provided in Supplementary Note 1.

Remark 1 (Boundedness of purity and conflict)

Let $G = \{g_1, \ldots, g_k\}$ be a gene set expressed in a cell, let $w_{ij} \in [-1, 1]$ denote gene–gene NPMI values, let $P = \{(i, j) : w_{ij} \text{ finite}\}$, $N = \{(i, j) : w_{ij} < 0\}$, $K = \binom{k}{2}$ be the total possible gene pairs, and let $\theta \in [0, 1)$ be a small noise tolerance (e.g., 0.05). Then

$$0 \leq \text{purity}(G) = \frac{1}{|P|} \sum_{(i,j)\in P} \mathbb{I}\left(w_{ij} > \theta\right) \leq 1, \tag{14}$$

$$0 \leq \text{conflict}(G) = \frac{1}{K} \sum_{(i,j)\in N} (-w_{ij}) \leq 1. \tag{15}$$

Remark 2 (Coherence score is bounded and well-posed)

Define $C(G) = \text{purity}(G) - \text{conflict}(G)$. Then $-1 \leq C(G) \leq 1$.

Remark 3 (Thresholding at θ removes stochastic noise under independence)

Let w_{ij} be empirical NPMI estimated from Bernoulli co-occurrence in sparse spatial transcriptomics. Assume independence: $w_{ij} \approx 0$ with sampling fluctuations on the order of $\mathcal{O}(n^{-1/2})$. Then

$$\mathbb{E}[\mathbb{I}(w_{ij} > \theta)] \to 0 \quad \text{for any fixed } \theta > 0, \tag{16}$$

so $\mathbb{E}[\text{purity}(G)] \approx 0$ for independent gene pairs. The threshold θ is a noise floor ensuring that purity only reflects genuinely consistent positive associations, not stochastic sampling variance.

Remark 4 (Monotonicity)

Let G be fixed and consider a single gene pair (a, b). If w_{ab} increases and crosses the threshold θ, then purity(G) weakly increases and conflict(G) does not increase, therefore $C(G)$ increases. If w_{ab} decreases further into the negative region, then conflict(G) strictly increases, purity(G) does not increase, and $C(G)$ strictly decreases.

6.5.3 ReLU-based coherence formulation

To reduce sensitivity to weak associations and emphasize stronger evidence, we additionally implemented a symmetric ReLU-based coherence formulation. For each gene–gene NPMI value w, we applied

$$\mathrm{ReLU}_\tau(w) = \begin{cases} w - \tau, & w > \tau, \\ w + \tau, & w < -\tau, \\ 0, & |w| \le \tau, \end{cases} \tag{17}$$

where τ defines a dead zone around zero. Similar to θ, we used $\tau = 0.05$ to account for biological noise.

Purity and conflict were then defined as

$$\mathrm{purity}_{\mathrm{ReLU}}(G) = \frac{1}{K} \sum_{i<j} \max\left(\mathrm{ReLU}_\tau(w_{ij}), 0\right), \tag{18}$$

$$\mathrm{conflict}_{\mathrm{ReLU}}(G) = \frac{1}{K} \sum_{i<j} \max\left(-\mathrm{ReLU}_\tau(w_{ij}), 0\right), \tag{19}$$

where $K = \binom{k}{2}$. The coherence score was again defined as $C_{\mathrm{ReLU}}(G) = \mathrm{purity}_{\mathrm{ReLU}}(G) - \mathrm{conflict}_{\mathrm{ReLU}}(G)$.

Mathematical properties of the ReLU-based coherence formulation

To justify the statistical behavior and interpretability of the ReLU-based coherence formulation, we establish several basic properties. Proofs are provided in Supplementary Note 1.

Remark R1 (ReLU coherence preserves boundedness and monotonicity)

Since $w_{ij} \in [-1, 1]$, the transformed values satisfy $\tilde{w}_{ij} \in [-1, 1]$. The ReLU transformation preserves sign ordering: increasing w_{ij} increases $\tilde{w}_{ij}$, making w_{ij} more negative decreases $\tilde{w}_{ij}$. Therefore, $0 \le \mathrm{purity}_{\mathrm{ReLU}}(G), \mathrm{conflict}_{\mathrm{ReLU}}(G) \le 1$, and $C_{\mathrm{ReLU}}(G)$ remains bounded in $[-1, 1]$. Values with $|w_{ij}| \le \tau$ (dead zone) contribute exactly zero to both purity and conflict.

6.5.4 ΔC-based merge criterion with simplicity penalty

For two candidate entities U and V with gene sets G_U and G_V, we evaluated the change in coherence upon merging as

$$\Delta C = C(G_U \cup G_V) - \max\left(C(G_U), C(G_V)\right). \tag{20}$$

To discourage trivial merges driven solely by small gene sets, we optionally applied a simplicity penalty:

$$\Delta C = \left(C(G_U \cup G_V) - \frac{1}{|G_U| + |G_V|} \right) - \max \left(C(G_U) - \frac{1}{|G_U|},\ C(G_V) - \frac{1}{|G_V|} \right). \tag{21}$$

This penalty promoted absorption of smaller, fragmented entities into larger coherent ones while preventing over-merging of independent structures. Merges were accepted only when $\Delta C \geq 0$.

Mathematical properties of the ΔC merge criterion

To justify the statistical behavior and interpretability of the ΔC-based merge, we establish several basic properties. Proofs are provided in Supplementary Note 1.

Remark 5 (ΔC favors biologically coherent unions)

Let U and V be two cell fragments. Then $\Delta C \geq 0$ only when the union does not introduce disproportionately many negative NPMI pairs. In particular, if both G_U and G_V are internally coherent but cross-pairs are conflicting, then $\Delta C < 0$ and stitching is rejected; if cross-pairs reinforce the dominant gene program, $\Delta C > 0$ and stitching is accepted.

6.5.5 Constrained hierarchical stitching

Hierarchical merging was performed greedily using a union–find structure, prioritizing candidate merges with the largest positive ΔC. To preserve biological interpretability, merges between two entities that both contained a whole cell were disallowed.

After convergence, each stitched cluster was assigned a representative label, prioritizing whole cells, followed by partial cells, and then reconstructed partial cells. Spatial coherence was then enforced by splitting any stitched entity whose transcripts formed multiple disconnected spatial components, retaining the largest component under the original label.

As an optional post-processing step, transcripts remaining unassigned could be reassigned to the nearest partial cell, reconstructed partial cell, or optionally whole cell within a user-defined spatial distance threshold, enabling recovery of locally coherent transcripts without altering core assignments.

6.6 Coherence-derived metrics for segmentation evaluation

To enable quantitative evaluation of transcriptional coherence at the cellular level, we derived a set of metrics from the gene–gene NPMI framework. For a cellular entity with gene set G, purity and conflict scores were computed as defined above, using either the raw NPMI formulation or the ReLU-based formulation with dead-zone threshold τ.

We further defined the signal strength of a cellular entity as

$$S(G) = \sum_{(i,j) \in P} |w_{ij}|\,, \tag{22}$$

where w_{ij} denotes the NPMI value for gene pair (g_i, g_j). Signal strength captures the total magnitude of coherent and incoherent gene–gene associations within a cell.

To normalize coherence contributions, we defined relative purity and relative conflict as

$$\text{relative purity}(G) = \frac{\sum \max(w_{ij}, 0)}{S(G)}, \tag{23}$$

$$\text{relative conflict}(G) = \frac{\sum \max(-w_{ij}, 0)}{S(G)}, \tag{24}$$

such that the two quantities sum to one when $S(G) > 0$ and at least one valid gene pair is present.

Transcript count per cellular entity was recorded as the total number of transcripts assigned to that entity after segmentation or refinement and used as a quality control measure. All metrics were computed independently for each cellular entity.

6.7 Single-cell analysis

Single-cell preprocessing and downstream analysis were performed using Scanpy [39]. For the breast cancer dataset, partial and reconstructed partial cells with fewer than 10 transcripts from at least 5 unique genes were removed; all other datasets included all cells. Raw counts were retained, followed by total-count normalization and log-transformation. PCA was computed on all genes to avoid variance-based gene selection bias. A k-nearest neighbor graph ($k = 15$) was constructed, and unsupervised clustering was performed using the Leiden algorithm (resolution = 1.0) with fixed parameters across datasets to ensure consistent interpretation of UMAP manifolds. UMAP embeddings were computed for visualization, and differential marker genes were displayed using Scanpy's matrix plot function.

6.8 Scrublet-based expression doublet detection

Doublet detection was performed using the Scanpy wrapper for Scrublet [35] to benchmark segmentation-induced mixed-lineage profiles in an orthogonal expression-only framework. For each segmentation strategy (Baysor, TRACER-stitched, TRACER-fine-tuned), Scrublet was applied to the corresponding whole-cell expression matrix using default parameters to ensure comparability across methods. We evaluated three expected doublet rates (0.02, 0.2, 0.25, and 0.3) to assess robustness of doublet-score behavior across common parameter settings.

6.9 Ligand–receptor interaction analysis

Ligand–receptor (LR) interactions were inferred using Squidpy [33] with the OmniPath ligand–receptor database [34]. For each segmentation strategy (standard Xenium, TRACER-stitched, TRACER-fine-tuned), LR analysis was performed on the same AnnData object structure using identical parameters to ensure comparability.

Permutation-based testing was performed with $n_{\text{perms}} = 100$ and threshold = 0 to avoid filtering out ligands or receptors in sparsely expressed clusters. P-values

were corrected for multiple testing using the Benjamini–Hochberg procedure, and interactions with false discovery rate (FDR) < 0.05 were considered significant.

6.10 Computational performance and scalability of TRACER

On a dual-socket Intel Xeon Platinum 8580 node (120 cores, 503 GiB RAM), TRACER processed a 28 million–transcript Xenium breast cancer dataset in approximately 5.2 hours using 17 CPU cores, with a peak memory footprint of 27 GiB (Supplementary Table 1). For smaller Xenium and MERFISH datasets (1.4 million and 0.8 million transcripts, respectively), end-to-end runtime was under 7 minutes and under 2 minutes. Across datasets, the most time-consuming components were the global coherence step (Stage 4) and final stitching (Stage 5), reflecting the cost of constructing and updating large spatial graphs.

Declarations

Data availability: All datasets analyzed in this study are publicly available. The multimodal Xenium lung cancer dataset (case: TSU-20, section: 1) can be downloaded from https://kero.hgc.jp/Ad-SpatialAnalysis_2024.html. The MERFISH mouse ileum dataset is available on Dryad (https://datadryad.org/dataset/doi:10.5061/dryad.jm63xsjb2). The standard Xenium breast cancer dataset can be accessed via https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast. The curated dataset used for benchmarking is available at Zenodo (DOI: 10.5281/zenodo.18892839)

Code availability: The code for TRACER is publicly available in the GitHub repository https://github.com/imlong4real/TRACER under the Apache License 2.0.

Competing interests: The authors declare no competing interests.

Acknowledgements: The authors would like to thank the Breakthrough Cancer Glioblastoma and Data Science TeamLab members, Alexander Ling, Nathalie Agar, Kenny Yu, Charles Couturier, Jennifer Gantchev, Wesley Tansey, and Christopher Tosh for their insightful feedback and assistance.

Funding: This work was supported by NIH/NCI grants U24CA284156 and 5U01CA253403-03; the Maryland Cigarette Restitution Fund Research Grant to the Johns Hopkins Medical Institutions (FY25); and Break Through Cancer.

References

[1] Gulati, G.S., D'Silva, J.P., Liu, Y., Wang, L., Newman, A.M.: Profiling cell identity and tissue architecture with single-cell and spatial transcriptomics. Nature Reviews Molecular Cell Biology **26**(1), 11–31 (2025) https://doi.org/10.1038/s41580-024-00768-2

[2] Rao, A., Barkley, D., França, G.S., Yanai, I.: Exploring tissue architecture using spatial transcriptomics. Nature **596**(7871), 211–220 (2021) https://doi.org/10.1038/s41586-021-03634-9

[3] Williams, C.G., Lee, H.J., Asatsuma, T., Vento-Tormo, R., Haque, A.: An introduction to spatial transcriptomics for biomedical research. Genome Medicine **14**(1), 68 (2022) https://doi.org/10.1186/s13073-022-01075-1

[4] Chen, H., Li, D., Bar-Joseph, Z.: Scs: cell segmentation for high-resolution spatial transcriptomics. Nature Methods **20**(8), 1237–1243 (2023) https://doi.org/10.1038/s41592-023-01939-3

[5] Petukhov, V., Xu, R.J., Soldatov, R.A., Cadinu, P., Khodosevich, K., Moffitt, J.R., Kharchenko, P.V.: Cell segmentation in imaging-based spatial transcriptomics. Nature Biotechnology **40**(3), 345–354 (2022) https://doi.org/10.1038/s41587-021-01044-w

[6] Ozirmak Lermi, N., Molina Ayala, M., Hernandez, S., Lu, W., Khan, K., Serrano, A., Lubo, I., Hamana, L., Tomczak, K., Barnes, S., *et al.*: Comparison of imaging based single-cell resolution spatial transcriptomics profiling platforms using ffpe tumor samples. Nature Communications **16**, 8499 (2025) https://doi.org/10.1038/s41467-025-63414-1

[7] Wang, H., Huang, R., Nelson, J., Gao, C., Tran, M., Yeaton, A., Krishna, S., *et al.*: Systematic benchmarking of imaging spatial transcriptomics platforms in ffpe tissues. Nature Communications **16**, 10215 (2025) https://doi.org/10.1038/s41467-025-64990-y

[8] Fu, X., Lin, Y., Lin, D.M., Mechtersheimer, D., Wang, C., Ameen, F., Ghazanfar, S., Patrick, E., Kim, J., Yang, J.Y.H.: Bidcell: Biologically-informed self-supervised learning for segmentation of subcellular spatial transcriptomics data. Nature Communications **15**, 509 (2024) https://doi.org/10.1038/s41467-023-44560-w

[9] Jin, K., Zhang, Z., Zhang, K., Viggiani, F., Callahan, C., Tang, J., Aronow, B.J., Shu, J.: Bering: joint cell segmentation and annotation for spatial transcriptomics with transferred graph embeddings. Nature Communications **16**, 6618 (2025) https://doi.org/10.1038/s41467-025-60898-9

[10] Tiesmeyer, S., Müller-Bötticher, N., Malt, A., Ma, L., Marco-Salas, S., Kiessling, P., Horn, P., Guillot, A., Kuemmerle, L.B., Tacke, F., al.: Identifying 3d signal

overlaps in spatial transcriptomics data with ovrlpy. Nature Biotechnology (2026) https://doi.org/10.1038/s41587-026-03004-8

[11] Yapp, C., Nirmal, A.J., Zhou, F., Wong, A.Y.H., Tefft, J.B., Lu, Y.D., Shang, Z., Maliga, Z., Llopis, P.M., Murphy, G.F., Lian, C.G., Danuser, G., Santagata, S., Sorger, P.K.: Highly multiplexed 3d profiling of cell states and immune niches in human tumors. Nature Methods **22**(10), 2180–2193 (2025) https://doi.org/10.1038/s41592-025-02824-x

[12] Wu, L., Beechem, J.M., Danaher, P.: Using transcripts to refine image based cell segmentation with fastreseg. Scientific Reports **15**, 30508 (2025) https://doi.org/10.1038/s41598-025-08733-5

[13] Fang, S., Chen, B., Zhang, Y., Sun, H., Liu, L., Liu, S., Li, Y., Xu, X.: Computational approaches and challenges in spatial transcriptomics. Genomics, Proteomics & Bioinformatics **21**(1), 24–47 (2023) https://doi.org/10.1016/j.gpb.2022.10.001

[14] Song, L., Wang, L., He, Z., Cui, X., Peng, C., Xu, J., Yong, Z., Liu, Y., Fei, J.-F.: Improving spatial transcriptomics with membrane-based boundary definition and enhanced single-cell resolution. Small Methods **9**(5), 2401056 (2025) https://doi.org/10.1002/smtd.202401056

[15] Plummer, J.T., Dezem, F.S., Cook, D.P., Park, J., Zhang, L., Liu, Y., *et al.*: Standardized metrics for assessment and reproducibility of imaging-based spatial transcriptomics datasets. Nature Biotechnology (2025) https://doi.org/10.1038/s41587-025-02811-9

[16] Ren, P., Zhang, R., Wang, Y., Zhang, P., Luo, C., Wang, S., Li, X., *et al.*: Systematic benchmarking of high-throughput subcellular spatial transcriptomics platforms across human tumors. Nature Communications **16**, 9232 (2025) https://doi.org/10.1038/s41467-025-64292-3

[17] Ding, R., Celikay, K., Ni, M., Hou, Y., Zhou, Y., Rohr, K.: Deep learning-based quality control using subcellular rna spatial distribution patterns for cell segmentation in spatial transcriptomics data. Small Methods **10**(1), 00885 (2026) https://doi.org/10.1002/smtd.202500885

[18] Mitchel, J., Gao, T., Petukhov, V., Cole, E., Kharchenko, P.V.: Impact and correction of segmentation errors in spatial transcriptomics. Nature Genetics **58**(2), 434–444 (2026) https://doi.org/10.1038/s41588-025-02497-4

[19] Li, Y., Lac, L., Liu, Q., Hu, P.: St-cellseg: Cell segmentation for imaging-based spatial transcriptomics using multi-scale manifold learning. PLoS Computational Biology **20**(6), 1012254 (2024) https://doi.org/10.1371/journal.pcbi.1012254

[20] Bouma, G.: Normalized (pointwise) mutual information in collocation extraction. In: Proceedings of the Biennial GSCL Conference (2009)

[21] Church, K.W., Hanks, P.: Word association norms, mutual information, and lexicography. In: Proceedings of the 27th Annual Meeting on Association for Computational Linguistics, pp. 76–83 (1989). https://doi.org/10.3115/981623.981633

[22] Levy, O., Goldberg, Y.: Neural word embedding as implicit matrix factorization. In: Advances in Neural Information Processing Systems (2014)

[23] Janesick, A., Shelansky, R., Gottscho, A.D., Wagner, F., Williams, S.R., Rouault, M., Beliakoff, G., Morrison, C.A., Oliveira, M.F., Sicherman, J.T., *et al.*: High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. Nature Communications **14**, 8353 (2023) https://doi.org/10.1038/s41467-023-43458-x

[24] Reynolds, D.: Gaussian mixture models. In: Encyclopedia of Biometrics, pp. 659–663. Springer, ??? (2009). https://doi.org/10.1007/978-0-387-73003-5_196

[25] Takano, Y., Suzuki, J., Nomura, K., Fujii, G., Zenkoh, J., Kawai, H., *et al.*: Spatially resolved gene expression profiling of tumor microenvironment reveals key steps of lung adenocarcinoma development. Nature Communications **15**, 10637 (2024) https://doi.org/10.1038/s41467-024-54671-7

[26] Saelens, W., Cannoodt, R., Todorov, H., Saeys, Y.: A comparison of single-cell trajectory inference methods. Nature Biotechnology **37**(5), 547–554 (2019) https://doi.org/10.1038/s41587-019-0071-9

[27] Trapnell, C., Cacchiarelli, D., Grimsby, J., *et al.*: The dynamics and regulators of cell fate decisions revealed by pseudotemporal ordering of single cells. Nature Biotechnology **32**(4), 381–386 (2014) https://doi.org/10.1038/nbt.2859

[28] Wagner, A., Regev, A., Yosef, N.: Revealing the vectors of cellular identity with single-cell genomics. Nature Biotechnology **34**(11), 1145–1160 (2016) https://doi.org/10.1038/nbt.3711

[29] Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., Cai, L.: Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. Nature **568**(7751), 235–239 (2019) https://doi.org/10.1038/s41586-019-1049-y

[30] Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., Zhuang, X.: Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science **362**(6416) (2018) https://doi.org/10.1126/science.aau5324

[31] Si, Y., Lee, J.S., Jun, G., Kang, H.M., Lee, J.H.: Spatial omics enters the microscopic realm: opportunities and challenges. Trends in Genetics **41**(9), 774–787 (2025) https://doi.org/10.1016/j.tig.2025.05.002

[32] Traag, V.A., Waltman, L., Eck, N.J.: From louvain to leiden: guaranteeing well-connected communities. Scientific Reports **9**, 5233 (2019) https://doi.org/10.1038/s41598-019-41695-z

[33] Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., *et al.*: Squidpy: a scalable framework for spatial omics analysis. Nature Methods **19**(2), 171–178 (2022) https://doi.org/10.1038/s41592-021-01358-2

[34] Türei, D., Schaul, J., Palacio-Escat, N., Bohár, B., Bai, Y., Ceccarelli, F., Çevrim, E., Daley, M., Darcan, M., Dimitrov, D., *et al.*: Omnipath: integrated knowledge-base for multi-omics analysis. Nucleic Acids Research **54**(D1), 652–660 (2025) https://doi.org/10.1093/nar/gkaf1126

[35] Wolock, S.L., Lopez, R., Klein, A.M.: Scrublet: Computational identification of cell doublets in single-cell transcriptomic data. Cell Systems **8**(4), 281–2919 (2019) https://doi.org/10.1016/j.cels.2018.11.005

[36] Xu, R.J., Zhang, H., Cadinu, P., Nicol, P.B., Herrmann, U.S., Lee, T., Geistlinger, L., Irizarry, R.A., Moffitt, J.R.: An image-based transcriptomics atlas reveals the regional and microbiota-dependent molecular, cellular, and spatial structure of the murine gut. Cell Host & Microbe (2026) https://doi.org/10.1016/j.chom.2026.01.018

[37] Heidari, E., Moorman, A., Unyi, D., Pasnuri, N., Rukhovich, G., Calafato, D., Mathioudaki, A., Chan, J.M., Nawy, T., Gerstung, M., Pe'er, D., Stegle, O.: Segger: Fast and accurate cell segmentation of imaging-based spatial transcriptomics data (2025) https://doi.org/10.1101/2025.03.14.643160

[38] Salas, S.M., Dammann, M., Rubens, R.K., Drummer, F., Halle, L., Becker, S., Theis, F.J.: Exploration of rna outside segmented cells in spatial transcriptomics reveals extrasomatic rna organization (2025) https://doi.org/10.64898/2025.12.07.692889

[39] Wolf, F.A., Angerer, P., Theis, F.J.: Scanpy: large-scale single-cell gene expression data analysis. Genome Biology **19**(1), 15 (2018) https://doi.org/10.1186/s13059-017-1382-0