A Framework for Rigorous Cell Segmentation and Annotation for Xenium Spatial Transcriptomics

ZHIYUAN ZHU[1,2], Helen Huang[2,3], Xiaolu Guo[2], Alexander Hoffmann[2]

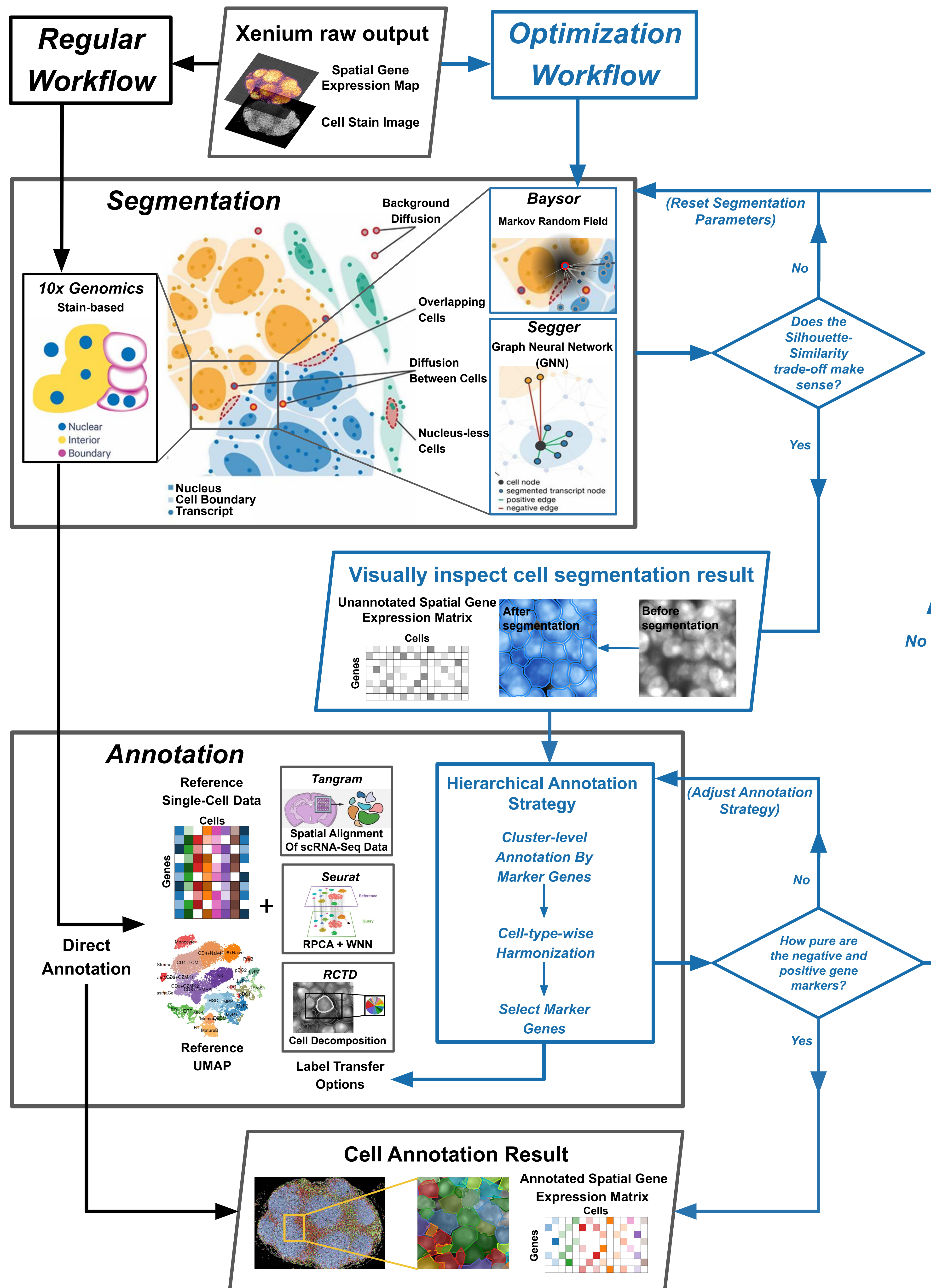[1] B.I.G. Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
[2] Signaling Systems Laboratory, Department of Microbiology, Immunology and Molecular Genetics, UCLA
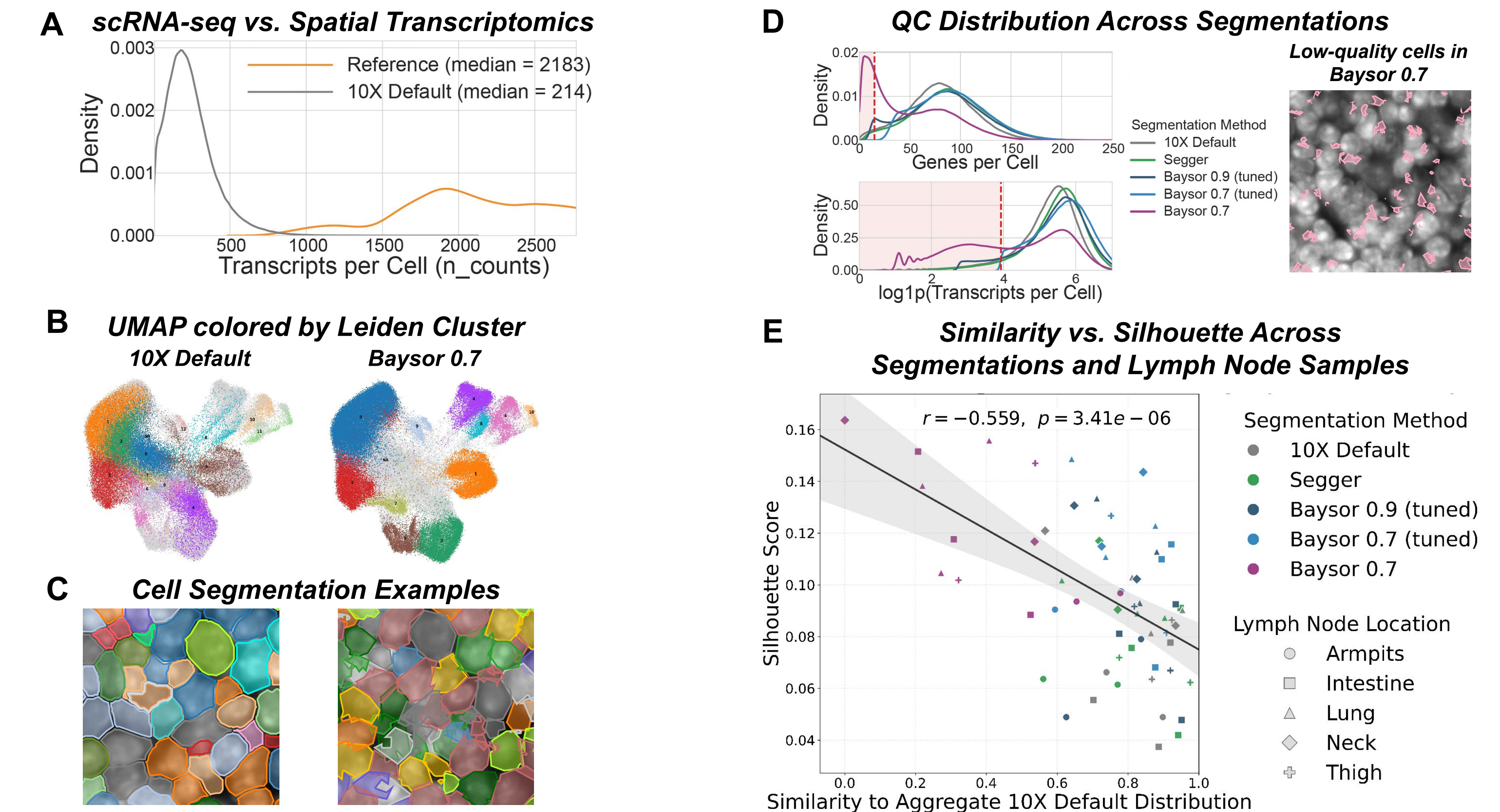[3] Bioinformatics Interdepartmental PhD Program, UCLA

UCLA Institute for Quantitative & Computational Biosciences

signaling systems
www.signalingsystems.ucla.edu

Workflow of Cell Segmentation and Annotation for Xenium

Regular Workflow

Xenium raw output
Spatial Gene Expression Map
Cell Stain Image

Optimization Workflow

Segmentation

10x Genomics Stain-based
- Nuclear
- Interior
- Boundary

- Background Diffusion
- Overlapping Cells
- Diffusion Between Cells
- Nucleus-less Cells

- Nucleus
- Cell Boundary
- Transcript

Baysor Markov Random Field

Segger Graph Neural Network (GNN)
- cell node
- segmented transcript node
- positive edge
- negative edge

(Reset Segmentation Parameters)

Does the Silhouette-Similarity trade-off make sense? — No / Yes

Visually inspect cell segmentation result

Unannotated Spatial Gene Expression Matrix
Cells / Genes

After segmentation / Before segmentation

No

Annotation

Reference Single-Cell Data
Cells / Genes

Tangram Spatial Alignment Of scRNA-Seq Data

Seurat RPCA + WNN

RCTD Cell Decomposition

Reference UMAP

Direct Annotation

Label Transfer Options

Hierarchical Annotation Strategy

Cluster-level Annotation By Marker Genes
↓
Cell-type-wise Harmonization
↓
Select Marker Genes

(Adjust Annotation Strategy)

How pure are the negative and positive gene markers? — No / Yes

Cell Annotation Result

Annotated Spatial Gene Expression Matrix
Cells / Genes

Trade-off between Cell Clusterability & Mis-segmentation

A scRNA-seq vs. Spatial Transcriptomics
— Reference (median = 2183)
— 10X Default (median = 214)
Density / Transcripts per Cell (n_counts)

B UMAP colored by Leiden Cluster
10X Default / Baysor 0.7

C Cell Segmentation Examples

D QC Distribution Across Segmentations
Density / Genes per Cell
Density / log1p(Transcripts per Cell)
Segmentation Method: 10X Default, Segger, Baysor 0.9 (tuned), Baysor 0.7 (tuned), Baysor 0.7
Low-quality cells in Baysor 0.7

E Similarity vs. Silhouette Across Segmentations and Lymph Node Samples
$r = -0.559, \ p = 3.41e - 06$
Silhouette Score / Similarity to Aggregate 10X Default Distribution
Segmentation Method: 10X Default, Segger, Baysor 0.9 (tuned), Baysor 0.7 (tuned), Baysor 0.7
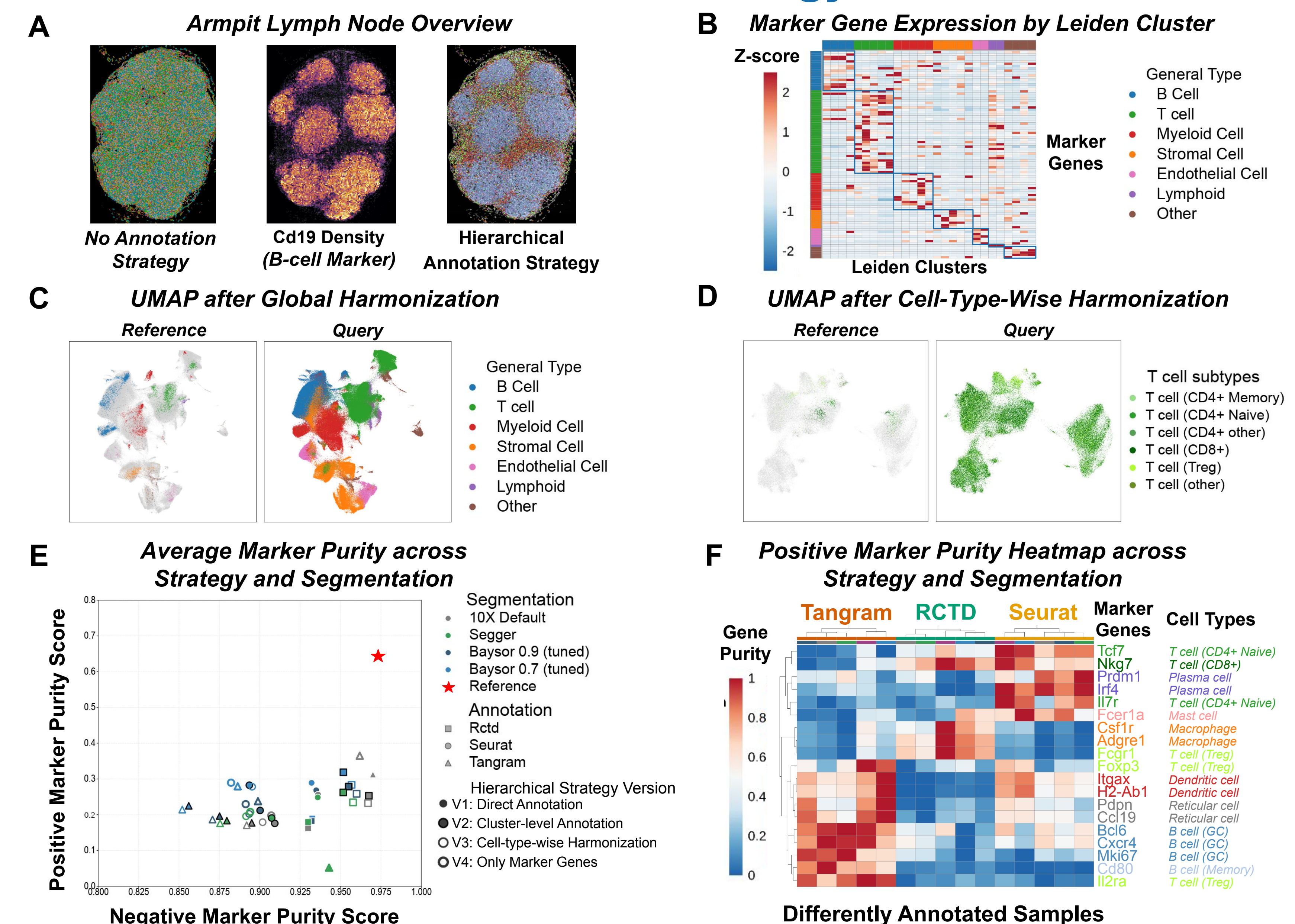Lymph Node Location: Armpits, Intestine, Lung, Neck, Thigh

(A) Image-based spatial transcriptomics exhibits significant **data sparsity**, capturing only 1-10% of the transcripts captured from scRNA-seq data from the same tissue.
(B-C) To overcome the limitations of stain-based methods in complex tissues, transcript-distribution-aware tools (Baysor and Segger) improved cell separability in the latent space, evident from UMAP projection colored by Leiden clustering.
(C-D) Critically, we found that naively optimizing for separability leads to over-segmentation, characterized by numerous low-quality cell segments with near-zero gene counts.
(E) We propose a framework to navigate this trade-off. By plotting silhouette score (clusterability) against a Wasserstein distance (approximate mis-segmentation), we inform parameter tuning to balance clusterability and mis-segmentation.

Hierarchical Annotation Strategy and Evaluation

A Armpit Lymph Node Overview
No Annotation Strategy / *Cd19 Density (B-cell Marker)* / *Hierarchical Annotation Strategy*

B Marker Gene Expression by Leiden Cluster
Z-score / Leiden Clusters
General Type / Marker Genes:
- B Cell
- T cell
- Myeloid Cell
- Stromal Cell
- Endothelial Cell
- Lymphoid
- Other

C UMAP after Global Harmonization
Reference / Query
General Type: B Cell, T cell, Myeloid Cell, Stromal Cell, Endothelial Cell, Lymphoid, Other

D UMAP after Cell-Type-Wise Harmonization
Reference / Query
T cell subtypes: T cell (CD4+ Memory), T cell (CD4+ Naive), T cell (CD4+ other), T cell (CD8+), T cell (Treg), T cell (other)

E Average Marker Purity across Strategy and Segmentation
Positive Marker Purity Score / Negative Marker Purity Score
Segmentation: 10X Default, Segger, Baysor 0.9 (tuned), Baysor 0.7 (tuned), Reference
Annotation: Rctd, Seurat, Tangram
Hierarchical Strategy Version: V1: Direct Annotation, V2: Cluster-level Annotation, V3: Cell-type-wise Harmonization, V4: Only Marker Genes

F Positive Marker Purity Heatmap across Strategy and Segmentation
Tangram / RCTD / Seurat
Gene Purity
Marker Genes / Cell Types / Differently Annotated Samples

(A) Direct application of label transfer tools fails to identify known biological structures (e.g., CD19-rich B Cell Germinal Centers), prompting our development of a Hierarchical Annotation Strategy for improved accuracy.
(B-D) Our strategy constrains the annotation process by starting with a broad, marker-based classification, followed by cell-type-wise data harmonization to prevent lineage mis-assignment and mitigate batch effects.
(E) Evaluation using a Marker Purity metric demonstrates our hierarchical approach outperforms direct annotation, achieving superior positive and negative purity scores.
(F) Critically, we found that different annotation tools exhibit strong marker-specific biases, which can influence final cell-type assignments and highlights the importance of careful cross-validation.

Acknowledgements