Characterizing Cellular Heterogeneity and Spatial Organization in Lung Adenocarcinoma with Visium HD

Sameen S. Hossain[1] and Dr. Katarzyna M. Tyc[2,3]

[1]College of Humanities and Sciences, School of Life Sciences and Sustainability, Virginia Commonwealth University

[2]Department of Biostatistics, School of Public Health, Virginia Commonwealth University, [3]Massey Bioinformatics Shared Resource Core

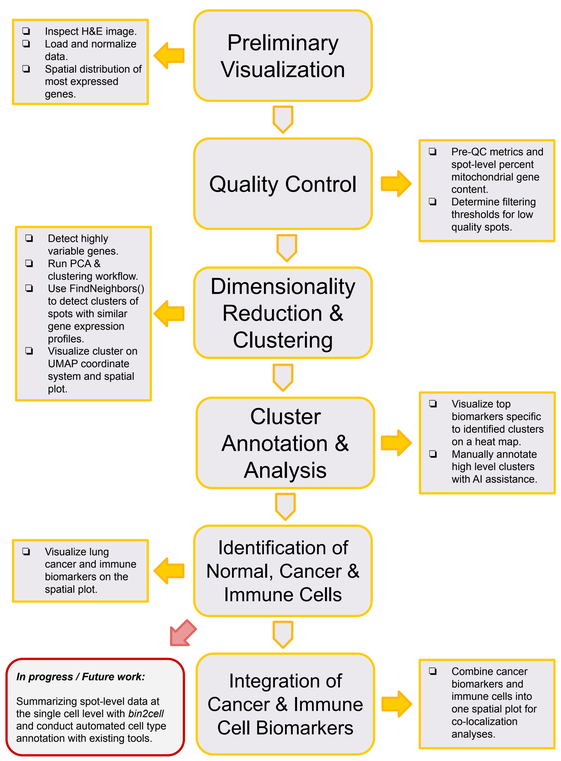
Introduction

Spatial transcriptomics is an emerging technology that enables the profiling of gene expression while preserving the spatial context of cells within intact tissue sections. This approach offers unprecedented insights into tissue architecture and cellular neighborhoods, which are critical for understanding disease mechanisms and identifying clinically relevant biomarkers.

Visium HD, the latest high-resolution spatial transcriptomics platform developed by 10x Genomics, allows for detailed transcriptome mapping across diverse tissue types. Analyzing such data requires robust computational tools. Seurat, an R package originally developed by the Satija Lab for single-cell RNA-seq analysis, has been adapted to handle spatial data, including Visium HD. Best practices in spatial transcriptomics include quality control, normalization, dimensionality reduction, and clustering—tasks supported by a growing suite of analytical tools.

In this study, we establish a streamlined, best-practice workflow for analyzing high-resolution Visium HD data using the Seurat framework. We apply this pipeline to a human lung adenocarcinoma dataset, the most prevalent form of non-small cell lung cancer (NSCLC) in the United States. Our goal is to map tumor and immune cell populations within the tumor microenvironment and highlight the utility of spatial methods for cancer research and precision medicine.

Methodology



Results

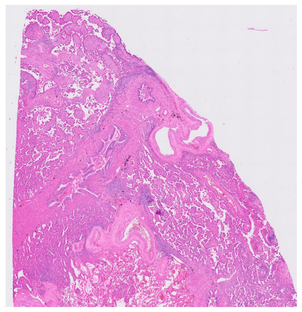


Figure 1. Hematoxylin and eosin (H&E)-stained image of human lung cancer tissue. Visium HD transcriptomic data were obtained from the 10x Genomics Data Portal.

Figure 2. Spot-level spatial expression of *NAPSA*, a canonical gene marker of Tumor Cells with retained AT2 identity. Though enriched in Cluster 0, *NAPSA* is expressed across multiple clusters (see Fig. 6).

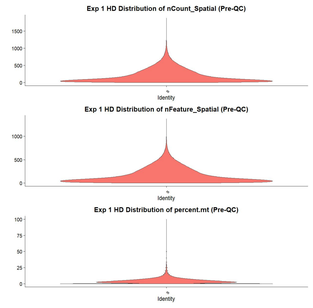


Figure 3. Pre-QC Metrics. 448,109 total spots were analyzed. Median UMI count per spot: 145. Median unique gene count per spot: 122. Median percent mitochondrial content: 3.43%.

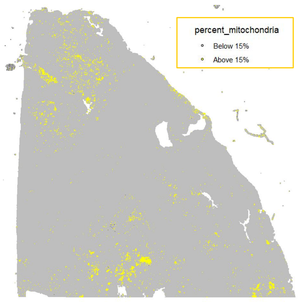


Figure 4. Distribution of percent mitochondrial gene content per spot. Spots with over 15% mt gene content are removed (yellow; 2.6% of total spots). Gray spots are kept for all downstream analyses (436,463 out of the original dataset).

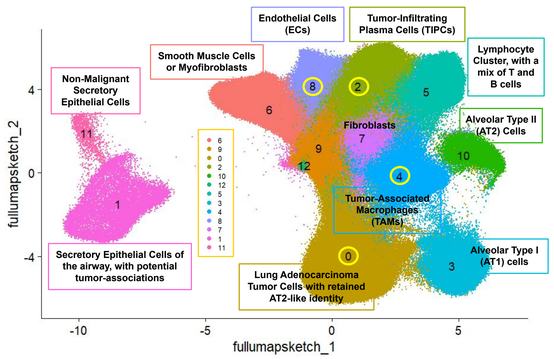


Figure 5. Dimensionality reduction of Visium HD gene expression data. Roughly 50K spots were selected using SketchData() to map the full dataset. High-level communities were identified using FindNeighbors(), and resulting clusters are visualized in UMAP space. Thirteen major cell classes are highlighted. Clusters of interest are circled in yellow.

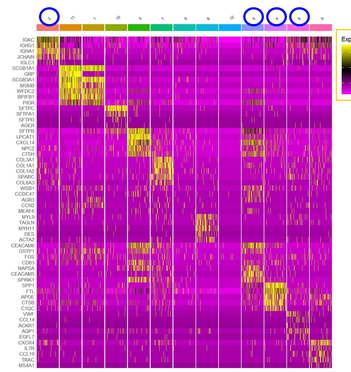


Figure 6. Heatmap of top marker genes for each of the 13 clusters, identified using FindAllMarkers(). Clusters of interest are circled in blue.

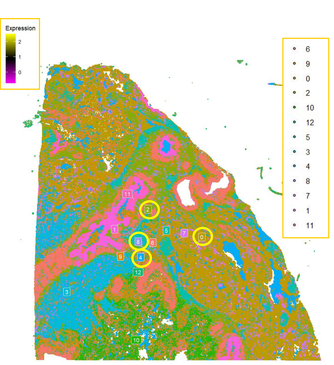


Figure 7. Spatial distribution of identified clusters. Distinct tissue architecture emerges across the section. Clusters of interest are circled in yellow.

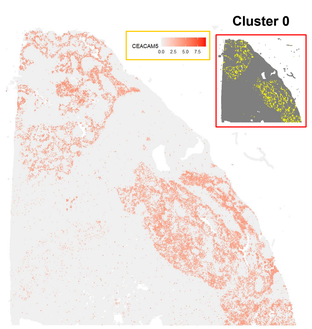


Figure 8. Cluster 0: Lung Adenocarcinoma Tumor Cells with retained AT2-like identity (yellow, inset). The *CEACAM5* gene marks their distinct transcriptional profile (red).

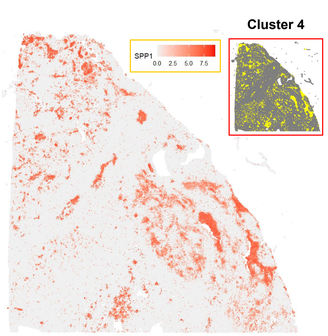


Figure 9. Cluster 4: SPP1⁺ tumor-associated macrophages (TAMs) (yellow, inset). The *SPP1* gene is known for producing Osteopontin (red).

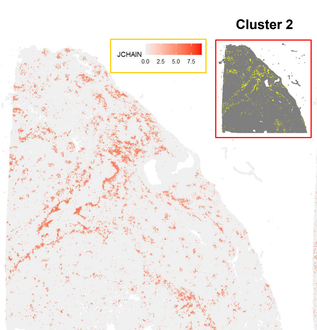


Figure 10. Cluster 2: Tumor-Infiltrating Plasma Cells, or TIPCs (yellow, inset). The *JCHAIN* gene marks this antibody-producing subtype (red).

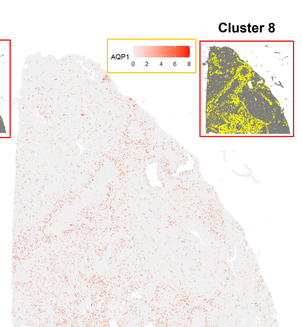


Figure 11. Cluster 8: Endothelial Cells (ECs) involved in vascular activation or inflammation (yellow, inset). The *AQP1* gene marks this activated endothelial subtype (red).

Future Directions

Further analysis of cell-cell communication will enhance understanding of tumor-immune crosstalk and therapeutic targets. We plan to apply single-cell segmentation for precise biomarker mapping and study immune cell topography alongside copy number alteration (CNA)-defined tumor subclones. These efforts aim to clarify the roles of TAMs and TIPCs in lung adenocarcinoma progression and support development of targeted therapies.

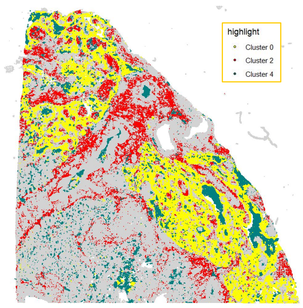
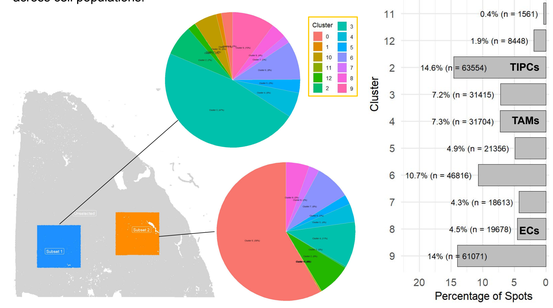


Figure 12. *Top Right:* Spatial Overlay of Clusters 0 (Lung Adenocarcinoma Tumor Cells), Cluster 4 (Tumor Associated Macrophages, TAMs) and Cluster 2 (Tumor Infiltrating Plasma Cells, TIPCs). Regions of spatial proximity and overlap between tumor and immune cells highlight potential sites of tumor–immune interaction.

Bottom Left: Two regions were selected to examine local cellular composition. Pie charts indicate the percentage distribution of individual clusters within each region. Cell type profiles vary significantly between regions, reflecting spatial heterogeneity in the tumor microenvironment.

Bottom Right: Bar plot showing the relative spot coverage of the 13 clusters identified in the dataset (see Fig. 5). Each bar represents the proportion of total spots assigned to a given cluster, highlighting differences in abundance across cell populations.



Conclusion

This study leveraged high-resolution Visium HD spatial transcriptomics to characterize the cellular architecture of human lung adenocarcinoma tissue. We identified thirteen distinct clusters representing tumor cells, immune infiltrates, and stromal components.

Malignant clusters showed retention of alveolar type 2 (AT2)-like features, marked by NAPSA and CEACAM5 expression, while tumor-associated macrophages (SPP1⁺ TAMs) and tumor-infiltrating plasma cells (TIPCs) were spatially localized within the tumor microenvironment. Endothelial cells exhibiting vascular activation were also identified, suggesting ongoing angiogenic and inflammatory processes.

Our integration of spatial gene expression with cell-type annotation reveals complex tissue organization and highlights molecular markers defining functionally relevant subpopulations.

Acknowledgements

Thank you to the VCU Honors Summer Undergraduate Research Program (HSURP) for sponsoring this research.

References

Satija Lab. (2024). *Analysis, visualization, and integration of Visium HD spatial datasets with Seurat.* https://satijalab.org/seurat/articles/visiumhd_analysis_vignette