

Integrated Spatial Transcriptomics

Neighborhood Comparison

HuBMAP Hackathon

Team iST

Team Overview

- Team Name: iST
- Team Members:
 - Chenchen Zhu, Stanford (Snyder Lab)
 - Stephanie Reinert, WUSTL (Jain Lab)
 - Asmita Lagwankar, WUSTL (Jain Lab)

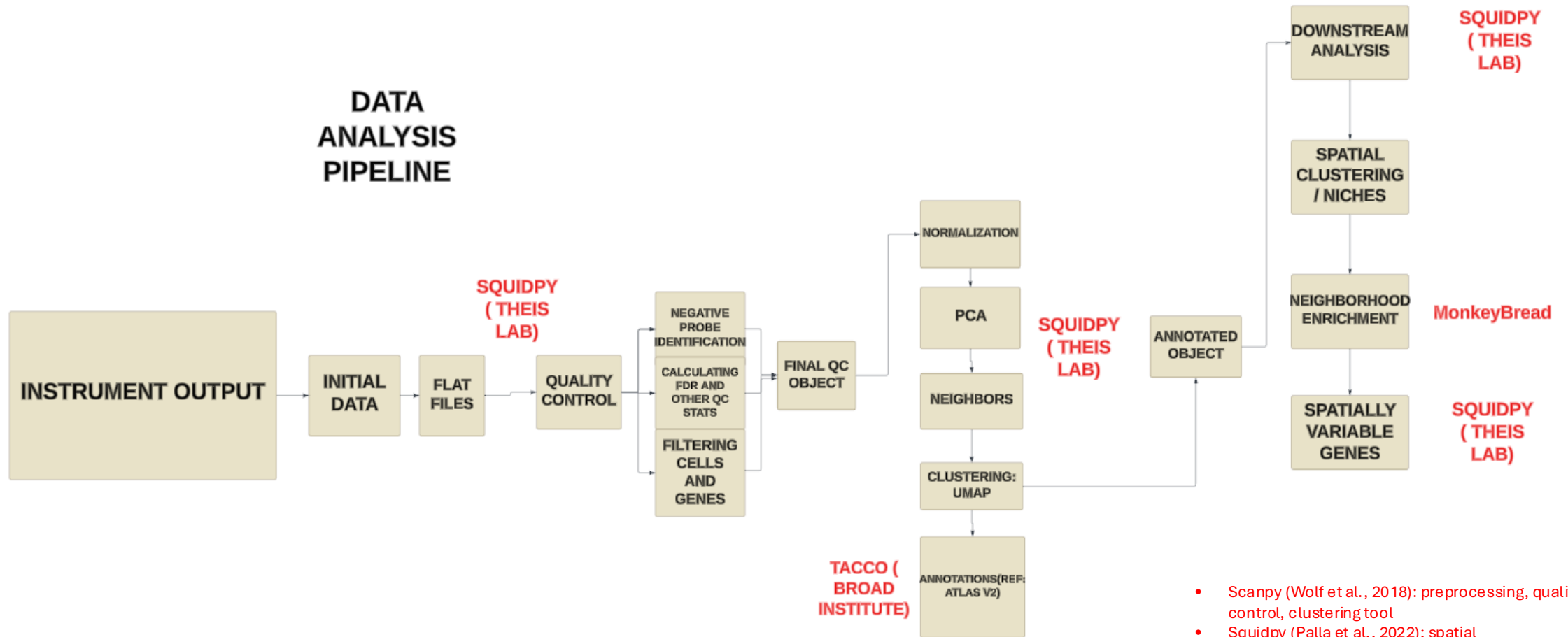
Problem and Goal

- There is a lack of existing pipelines to analyze imaging-based spatial transcriptomics within the HuBMAP community
- There are many packages that are solely for sequencing-based spatial transcriptomics technologies (Visium)
- Goal:
 - Perform quality control of imaging-based spatial transcriptomics datasets for kidney, intestine, and lung
 - Perform cell type annotation of ST datasets
 - Conduct neighborhood analysis

Overall Approach

- Generation of Xenium, MERFISH, and CosMx datasets for different tissues (part of Supplementary Admin)
- Established a Workspace Notebook on HuBMAP HIVE for QC
- Compare QC metrics for ST assays
- Implement and evaluate cell type annotations based on reference
- Comparison of different neighborhood analysis softwares
- Established a Workspace Notebook on HuBMAP HIVE for neighborhood analysis
- Comparison of neighborhood analysis across tissue types in Xenium data

Implementation



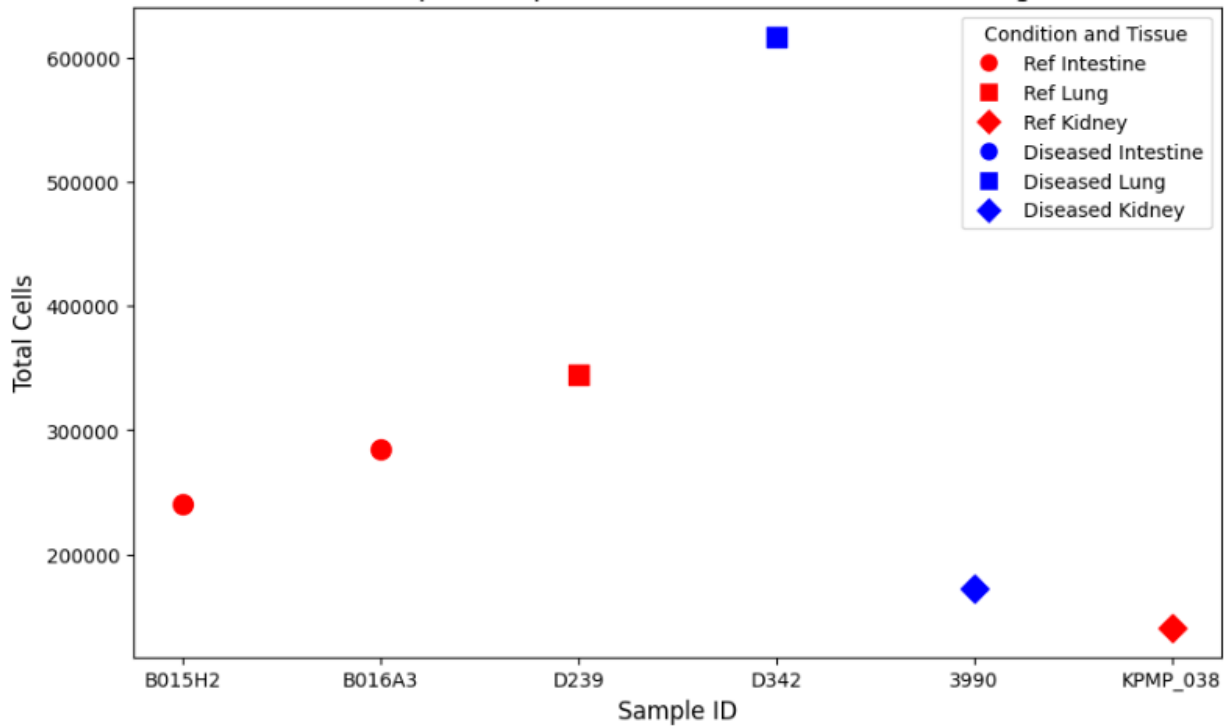
- Scanpy (Wolf et al., 2018): preprocessing, quality control, clustering tool
- Squidpy (Palla et al., 2022): spatial transcriptomics visualization tool
- TACCO (Mages et al., 2023): cell label transfer / annotation tool
- Monkeybread (Bernstein, 2023, preprint): neighborhood analysis tool

Results

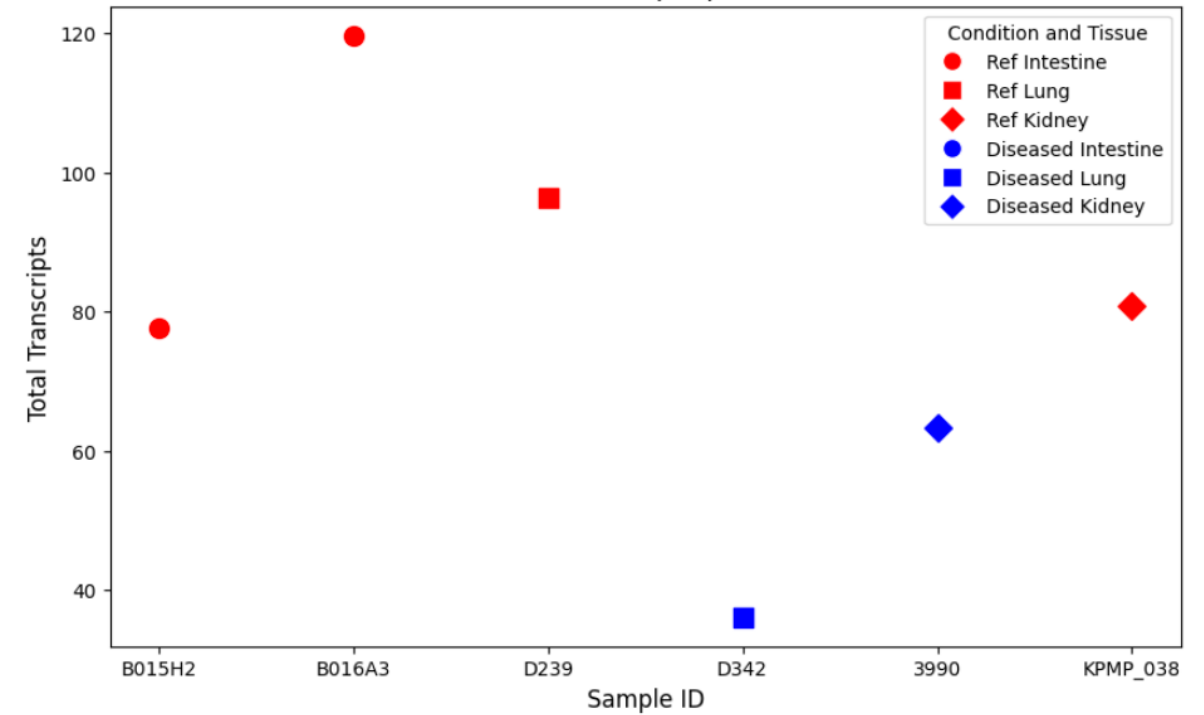
- Required Inputs:
 - Cell metadata (cells.csv)
 - Cell by gene matrix (cell_feature_matrix.h5) read into an anndata object (adata)
 - Reference annotation object with single cell RNA-seq data in adata
- Benefit to the community: end-to-end pipeline for the analysis of image-based spatial transcriptomic data including QC, cell type annotation, and neighborhood analysis
- Example data: three Xenium samples one from each organ run using the [300 KIL Panel]
- Status: Ready to go

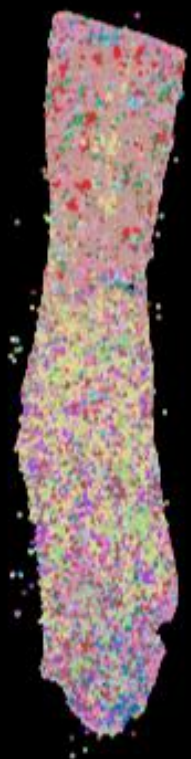
QC Plots

Total Cells



Total Transcripts per Cell

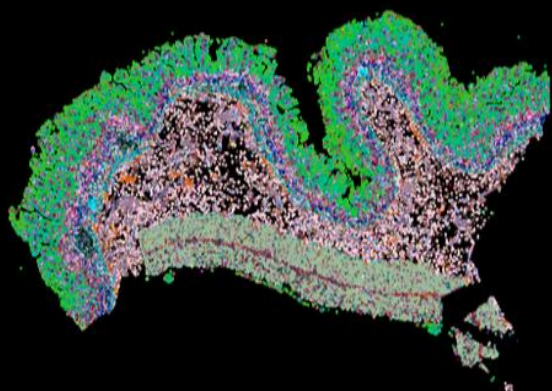




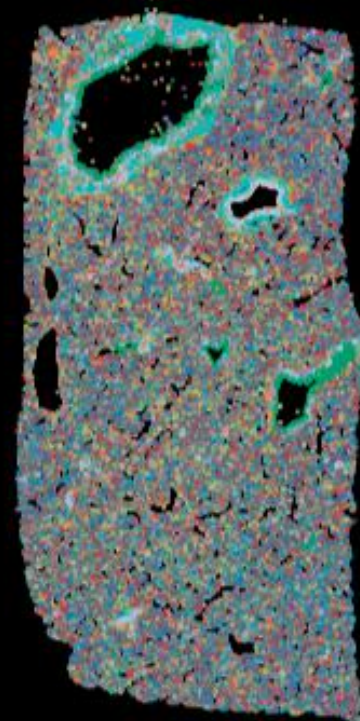
Kidney Tissue

- | | |
|------------|-----------|
| ● ATL | ● Myeloid |
| ● Ad | ● NEU |
| ● CNT | ● PC |
| ● DCT | ● PEC |
| ● DTL | ● POD |
| ● EC | ● PT |
| ● FIB | ● PapE |
| ● IC | ● TAL |
| ● Lymphoid | ● VSM/P |

Intestine Tissue



- | | | |
|-----------------------------------|-------------------------------|--------------------------------|
| ● Adipocytes | ● EnteroendocrineUn 1 | ● Myofibroblasts/SM 3 |
| ● B Cells | ● Epithelial | ● Myofibroblasts/SM DES High |
| ● Best4+ Enterocytes | ● Exocrine | ● NEUROG3high |
| ● CD4 | ● Glia | ● NK |
| ● CD8 | ● Goblet | ● Neurons |
| ● Crypt Fibroblasts 1 WNT2B+ | ● I Cells | ● Paneth |
| ● Crypt Fibroblasts 2 | ● ICC | ● Pericytes |
| ● Crypt Fibroblasts 3 RSPO3+ | ● ILC | ● Plasma |
| ● CyclingImmune | ● Immature Enterocytes | ● S Cells |
| ● CyclingTA 1 | ● Immature Goblet | ● Secretory Specialized MUC5B+ |
| ● CyclingTA 2 | ● K Cells | ● Secretory Specialized MUC6+ |
| ● D Cells | ● L Cells | ● Stem |
| ● DC | ● Lymphatic Endothelial Cells | ● T Cells |
| ● Endothelial-CD36+ Microvascular | ● Mast | ● TA1 |
| ● Endothelial-Venules | ● Mo Cells | ● TA2 |
| ● Enterochromaffin | ● Mono_Macrophages | ● Tuft |
| ● Enterocytes | ● Myofibroblasts/SM 1 | ● Unknown |
| ● EnteroendocrineUn | ● Myofibroblasts/SM 2 | ● Villus Fibroblasts WNT5B+ |

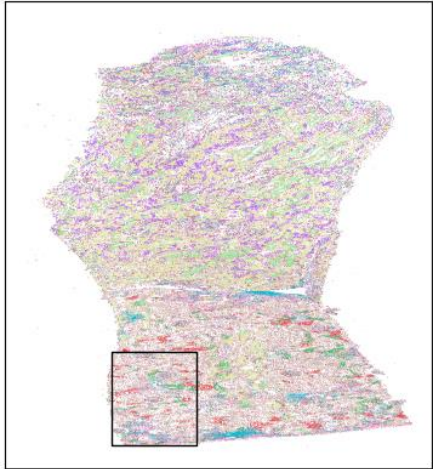


Lung Tissue

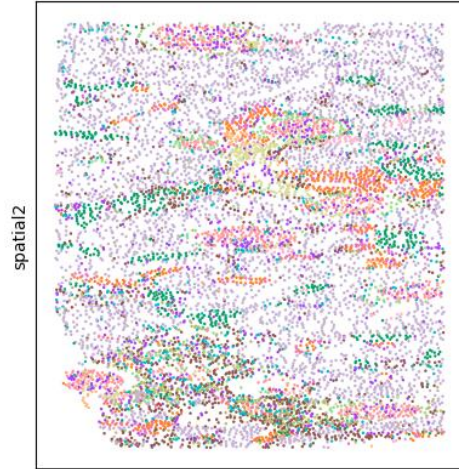
- | |
|------------------|
| ● Endo.Capillary |
| ● Endo.Main |
| ● Epi.Airway |
| ● Epi.Alveolar |
| ● Epi.Glandular |
| ● Imm.Granuloid |
| ● Imm.Lymphoid |
| ● Imm.Myeloid |
| ● Mes.Alveolar |
| ● Mes.Aw.Vasc |
| ● Neural |

Zoom Into Areas of Interest

v2.subclass.l1

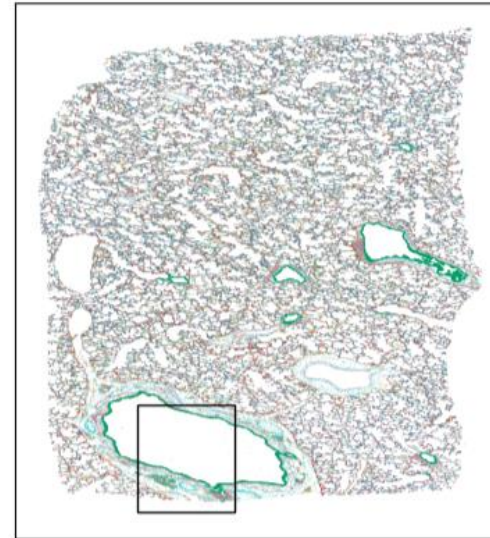


v2.subclass.l1

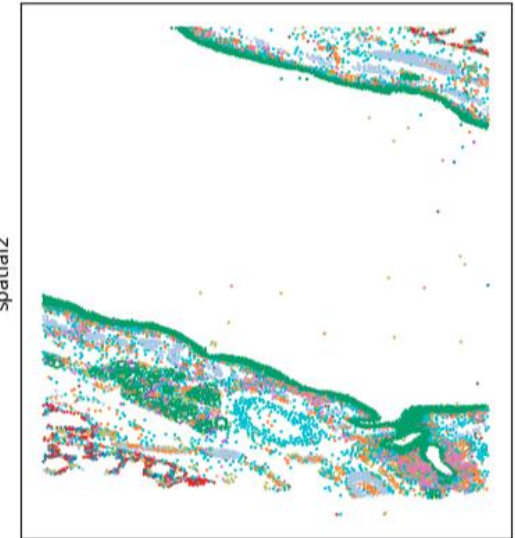


- Ad
- CNT
- DCT
- DTL
- EC
- FIB
- IC
- Lymphoid
- Myeloid
- NEU
- PC
- PEC
- POD
- PT
- PapE
- TAL
- VSM/P

v2.subclass.l1

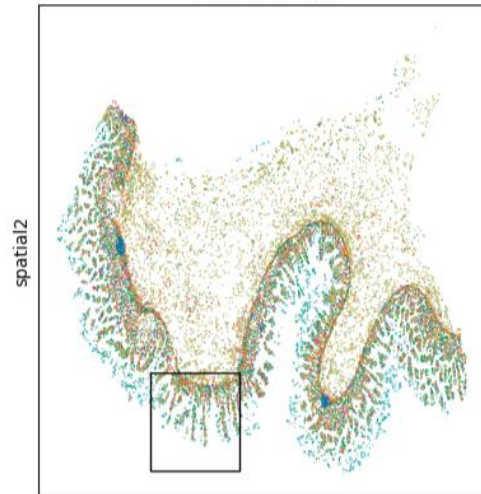


v2.subclass.l1

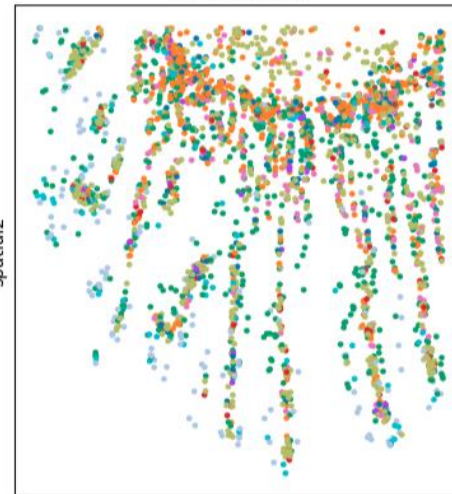


- Endo.Capillary
- Endo.Main
- Epi.Airway
- Epi.Alveolar
- Epi.Glandular
- Imm.Granuloid
- Imm.Lymphoid
- Imm.Myeloid
- Mes.Alveolar
- Mes.Aw.Vasc
- Neural

v2.subclass.l1



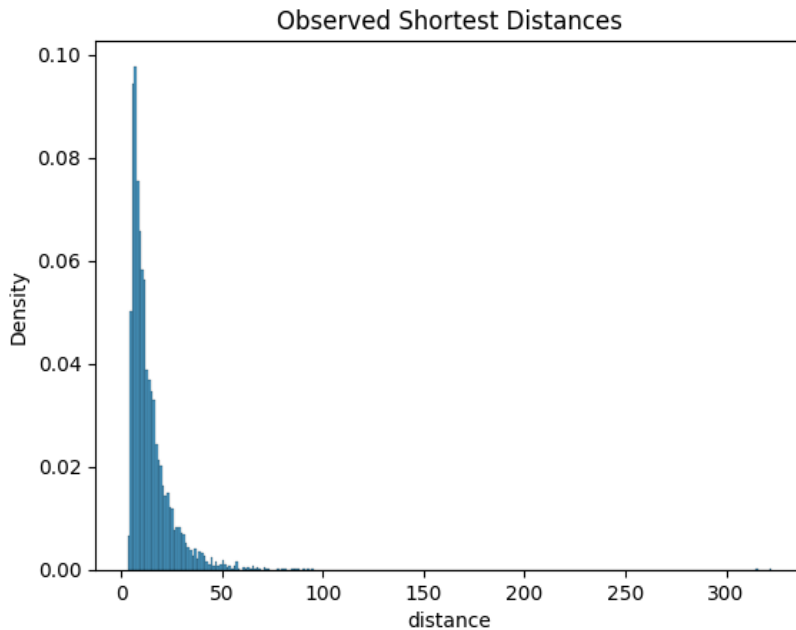
v2.subclass.l1



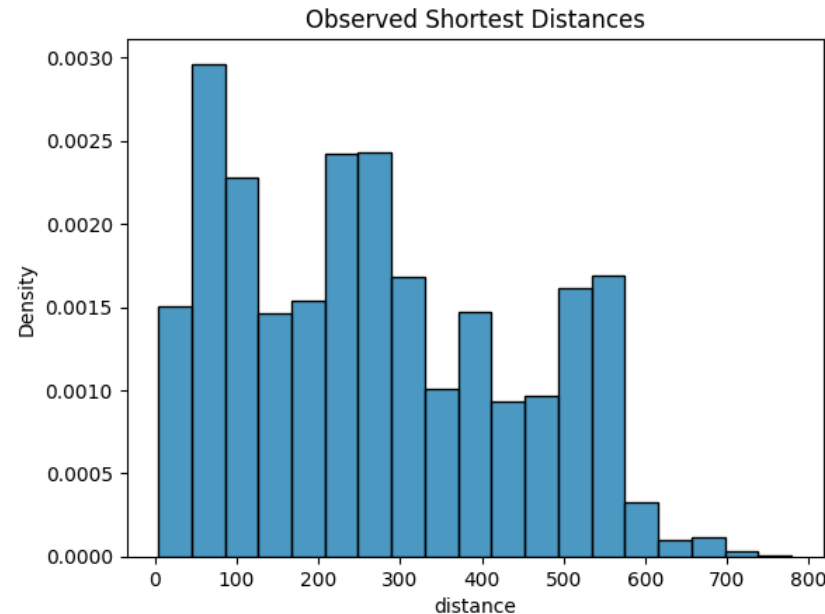
- B Cells
- CD4
- CD8
- CyclingImmune
- DC
- ILC
- Mast
- Mono_Macrophages
- NK
- T Cells

Biological Insights – Shortest Distances in Kidney Cells

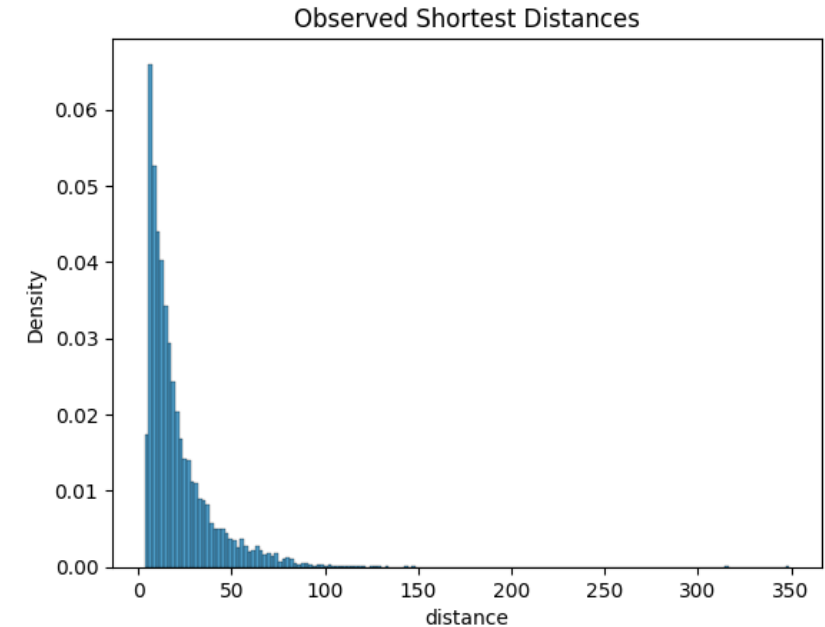
Lymphoid and Fibroblast



Podocytes and PapE

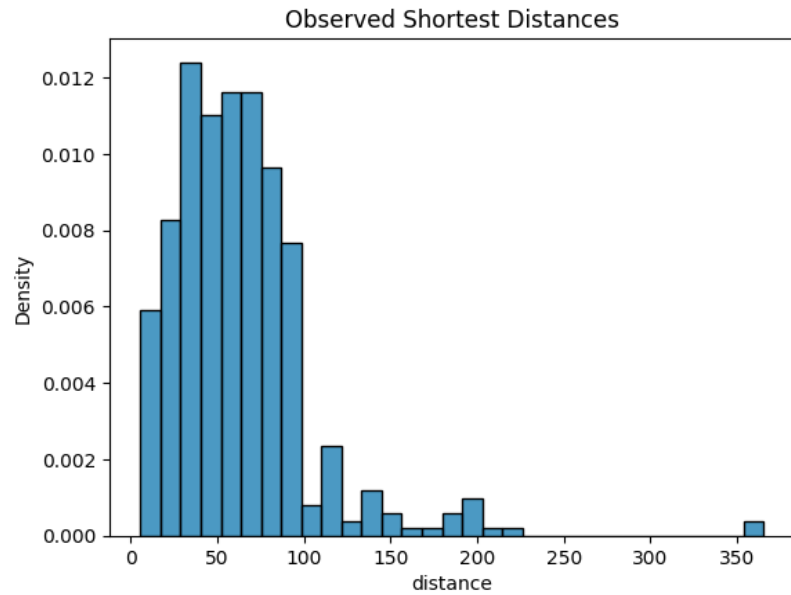


Lymphoid and Myeloid

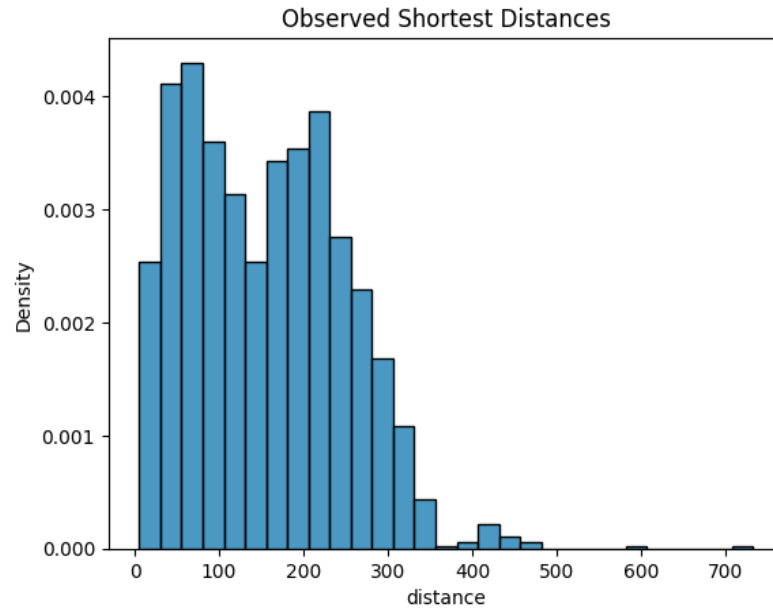


Biological Insights – Shortest Distances Among Similar Cells Across Organs

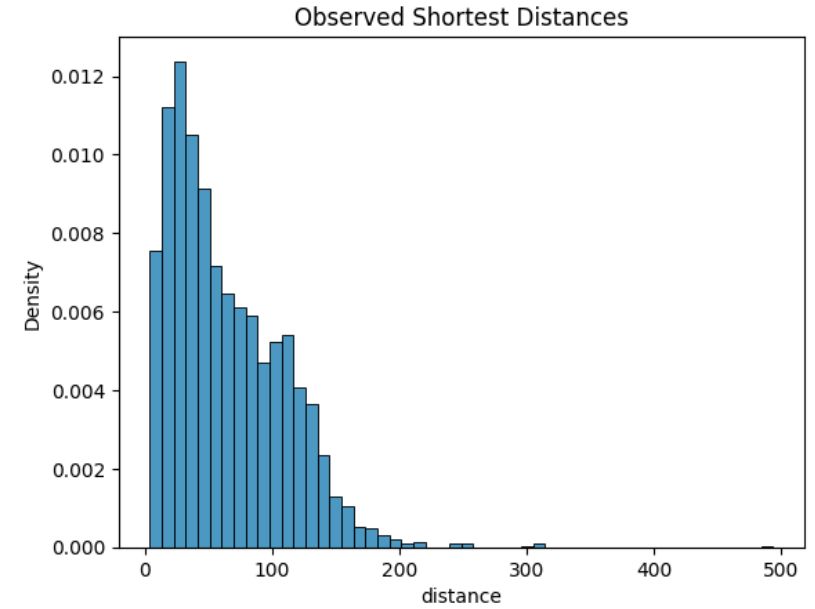
Kidney B and NK Cells



Lung B and NK Cells



Intestine B and NK Cells



Challenges and Conclusions

- Summary: We have established an end-to-end pipeline for the analysis of image-based spatial transcriptomics data.
- During the process, we evaluated several packages for analyzing spatial data. Many of the packages did not work on our datasets as they were optimized for sequencing-based spatial transcriptomics.
- Overall experience: We have a better understanding of HuBMAP Workspaces and how to use them. We also prototyped our pipeline.
- Feedback for Future Hackathons
 - Increased interaction between teams and among team members would be good, as many teams are tackling the same problems
 - Collaborative notebooks would be helpful