

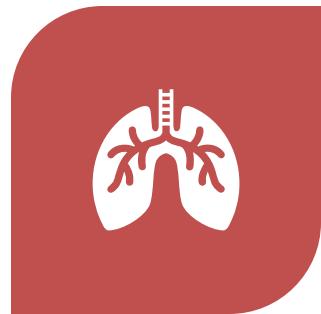
Spatial Transcriptomics Analysis of Lung Cancer Samples to Identify the Impact of Chronic Stress

Visium Pipeline Workflow & Key Findings

Ritupam Sarma

March 2025

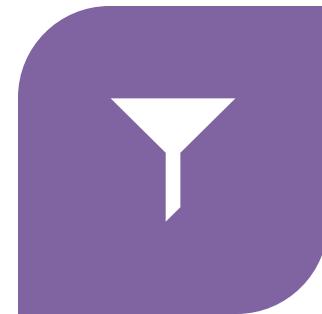
Introduction



BACKGROUND ON LUNG CANCER &
STRESS RESPONSE



OBJECTIVE: COMPARE GENE
EXPRESSION IN LUNG CANCER
SAMPLES SUBJECTED TO DIFFERENT
STRESS LEVELS.



METHODOLOGY: SPATIAL
TRANSCRIPTOMICS USING 10X
GENOMICS VISIUM & SEURAT PIPELINE

Background

Stress in this project includes anything that has a chronic effect on the amygdala activity. It includes conditions like anxiety and depression.

20 10x Visium samples.

Stress Levels determined by
Questionnaires filled out by patients, and
Blood glucocorticoid levels.

Loss of one high stress sample due to detachment.

Tumor Classification

Types of Tumors detected:

NSCLC

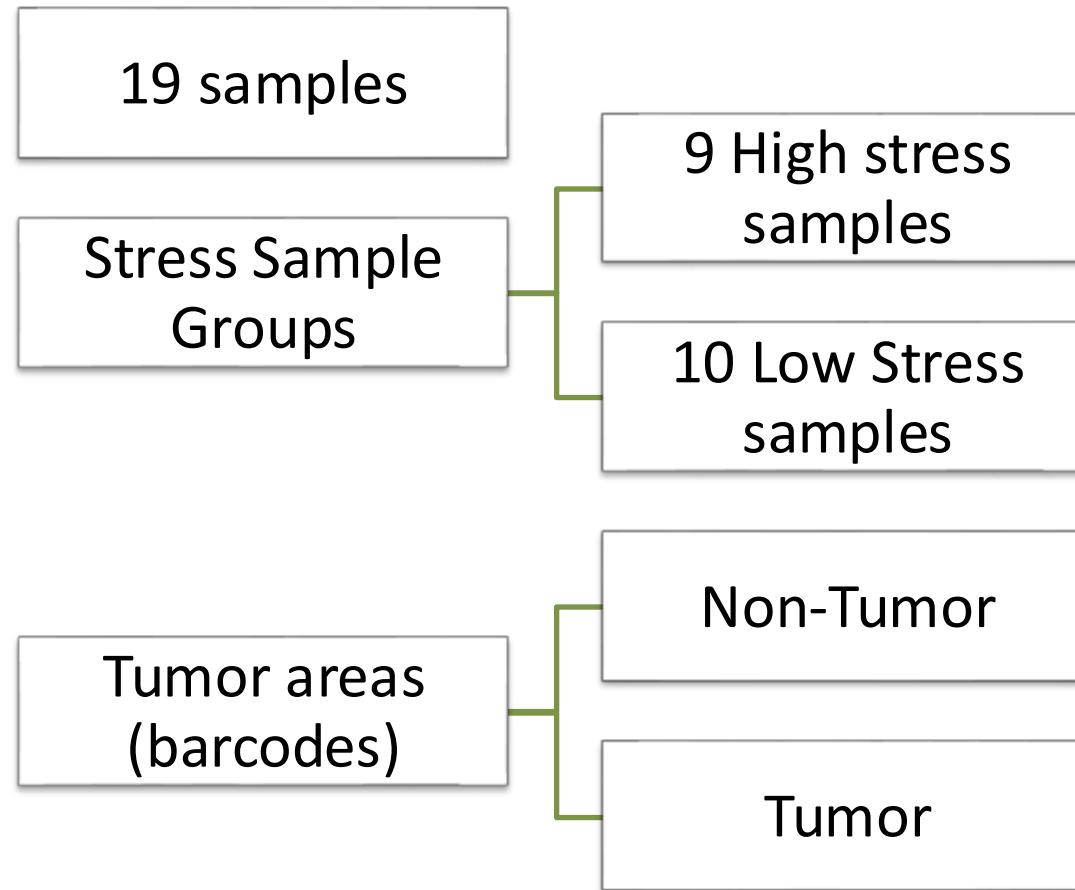
AC

PEC



For the initial analysis, all pathology annotations including Tumor cells is classified as TUMOR and everything else (non-tumor area, healthy tissue, etc.) is classified as NON-TUMOR.

Stress and Tumor Groups



Key Data Stats

Number of reads

- Range of reads per spot: between **19,144 (low PEC)** and **30,002 (high PEC)**.
- Average: **25,581**.

Number of genes detected

The median number of genes detected per spot varies widely (**426 to 5,650**).

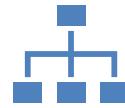
Workflow Overview



1. Data Pre-processing



2. Normalization & Feature Selection



3. PCA & Clustering



4. UMAP Visualization



5. Differential Gene Expression Analysis



6. Marker Gene Identification

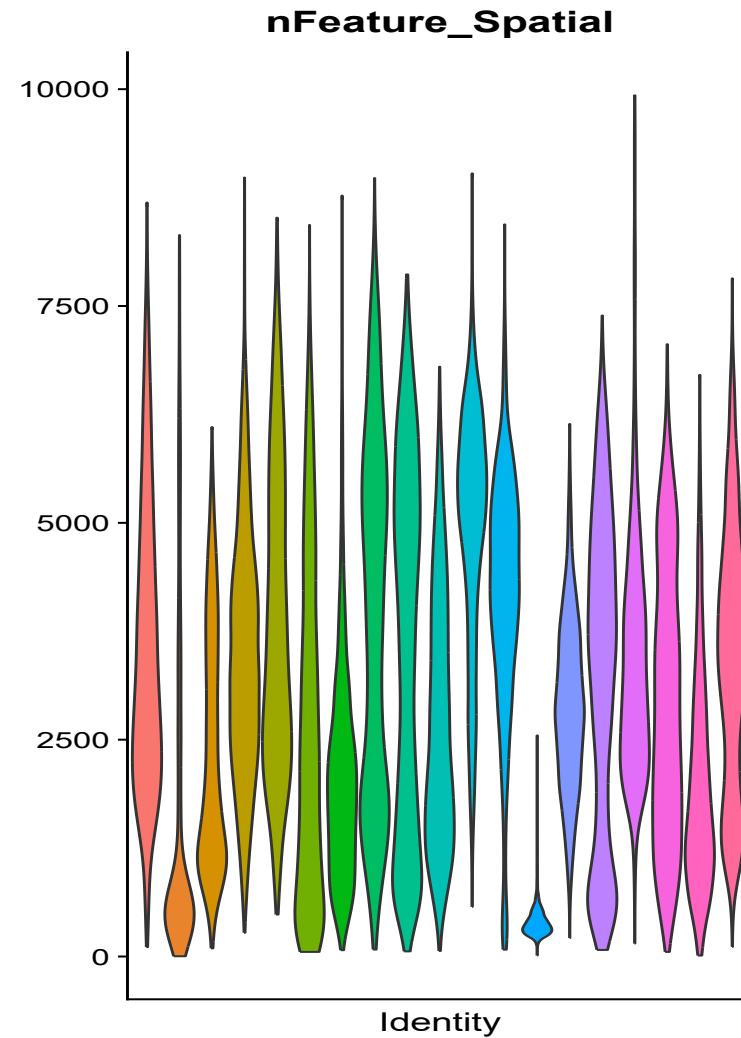
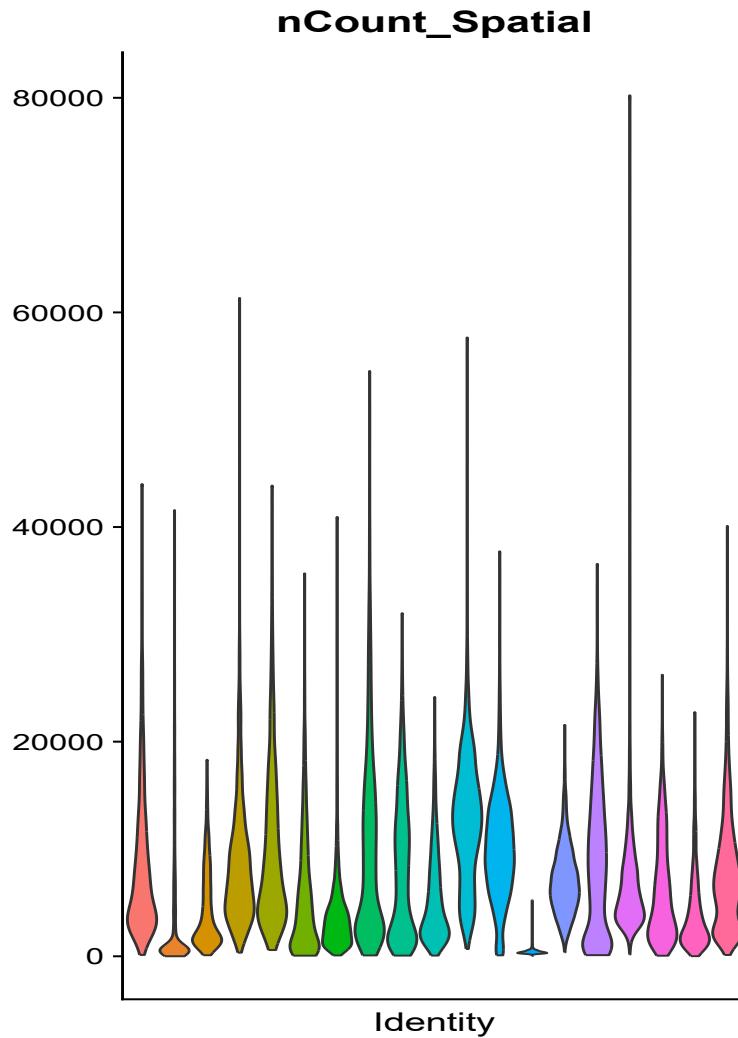
Data Preprocessing

Quality Control: Removed spots with zero reads.

Normalization: Log-normalization and scaling using Seurat

Feature Selection: Top 2000 variable features using VST

Violin Plots per sample



Dimensionality Reduction

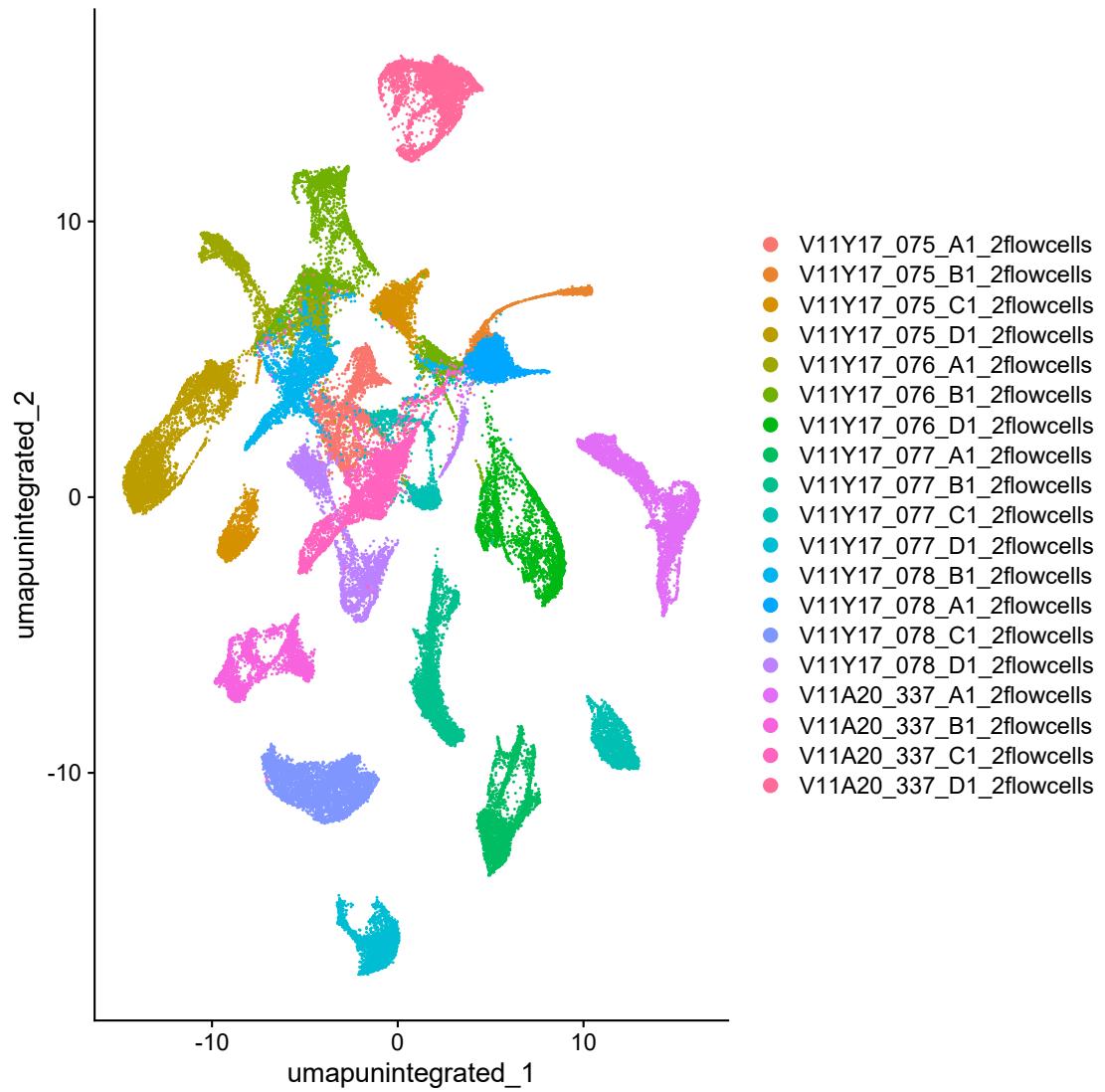
Principal Component Analysis

Elbow Plot used to determine
the number of PCs

Dimensionality Reduction and
clustering

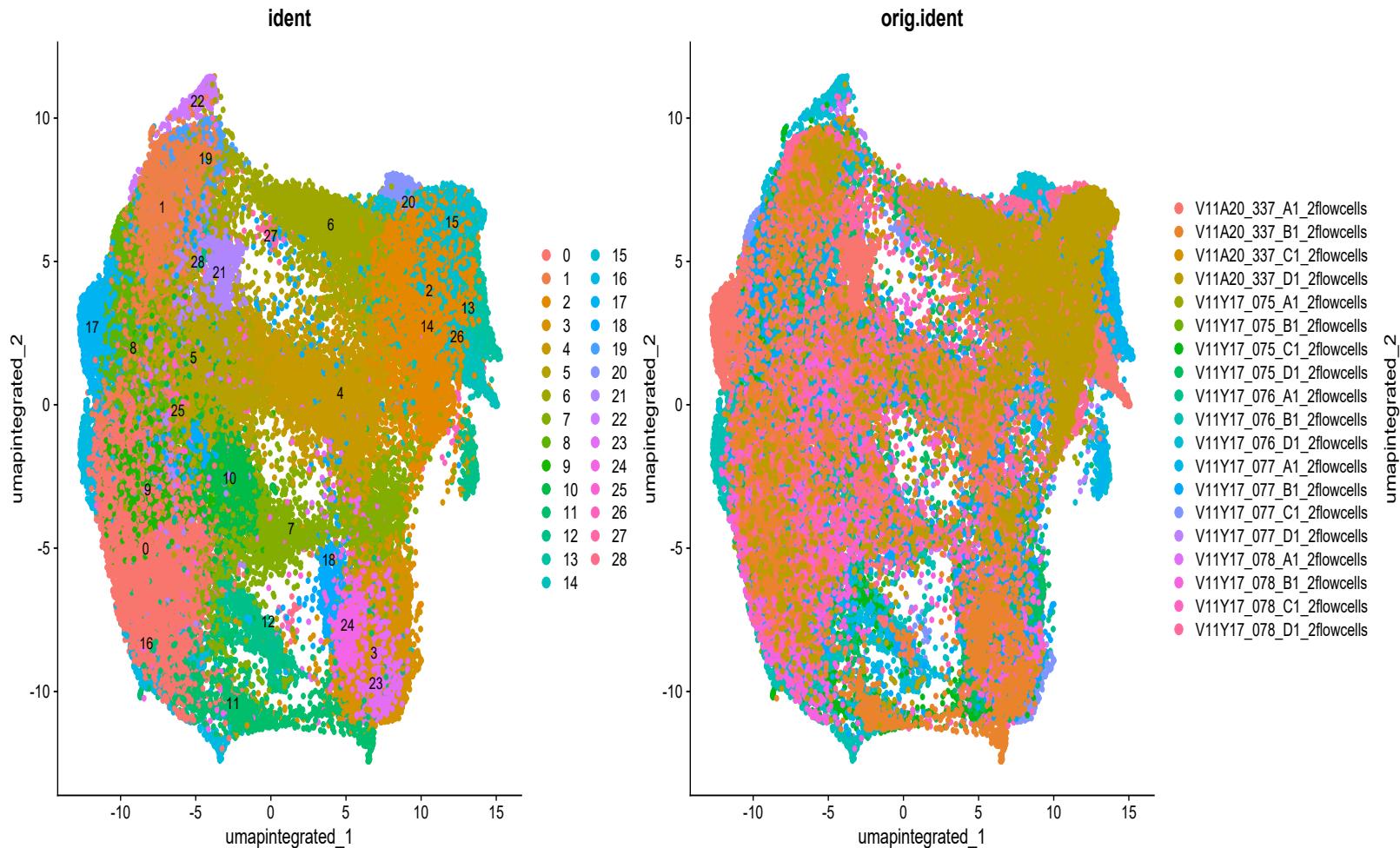
Unintegrated UMAP

Clustering without
batch correction



Integrated UMAP

Used **Harmony** for batch correction:
reduced batch effects, and
Allowed biologically meaningful clustering.



Spatial Visualization

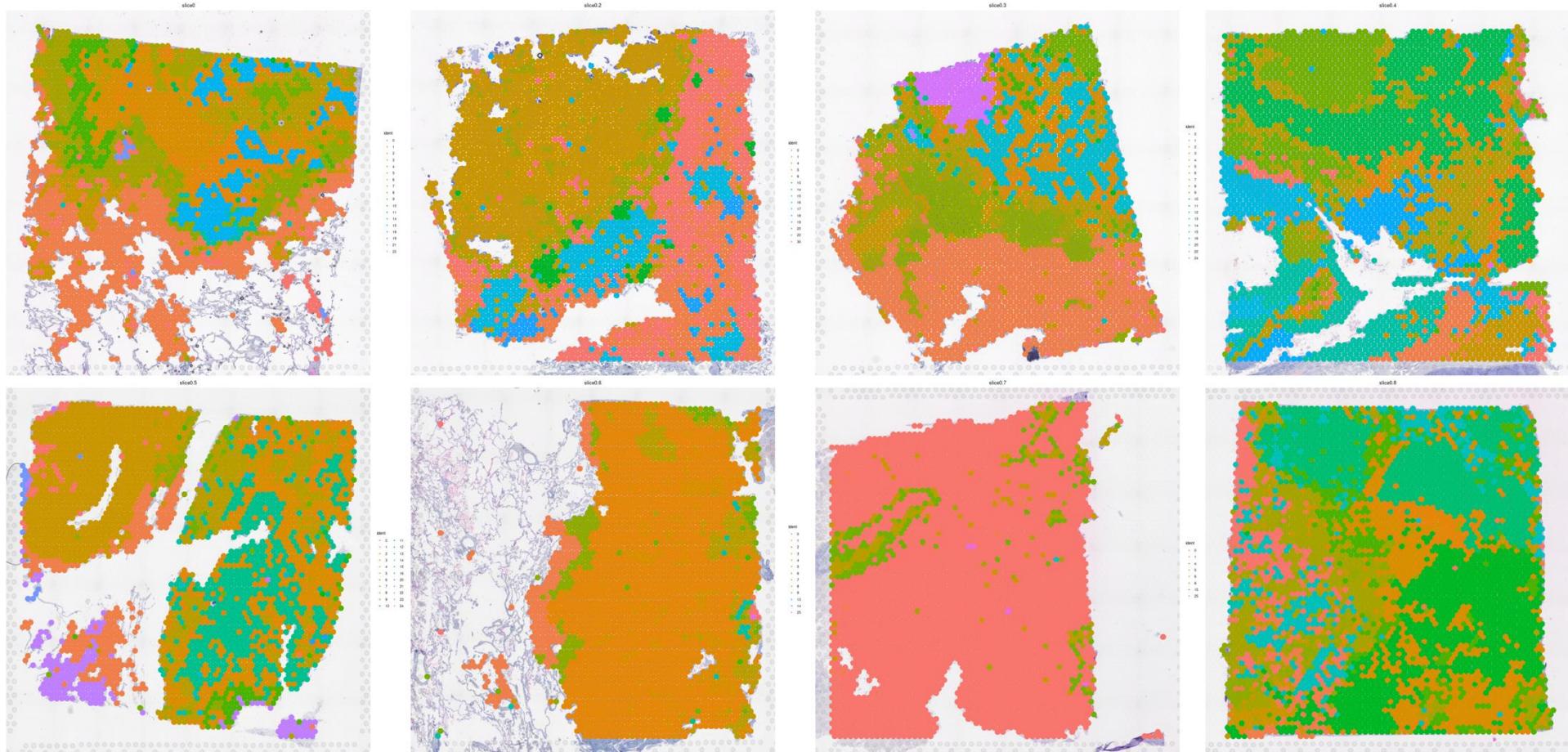


SpatialDimPlot of clusters
on Visium slides

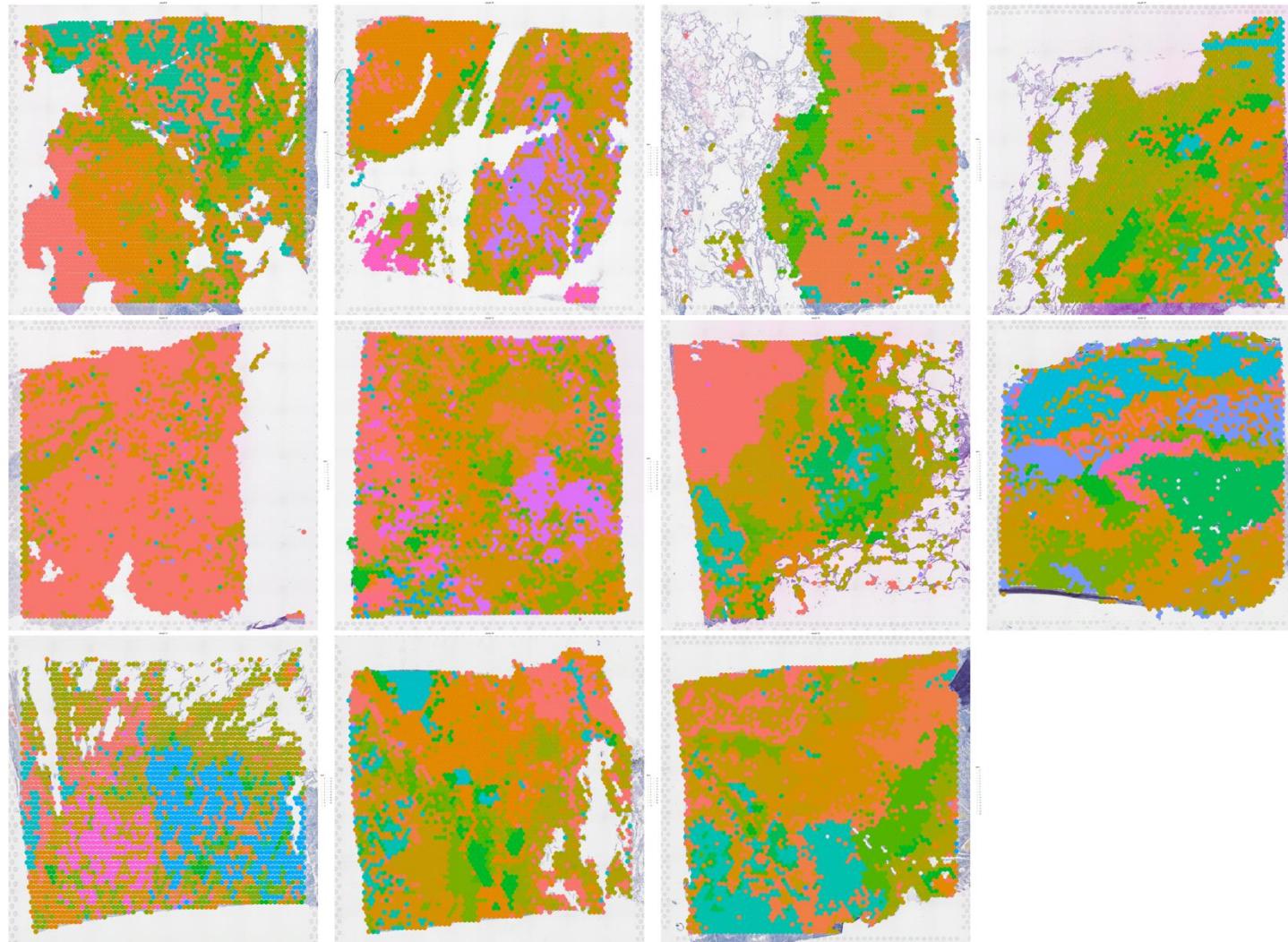


Overlay expression for
spatial viewing.

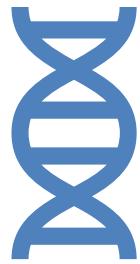
Spatial Plot (8/19)



Spatial Plot (11/19)



Differential Expression Analysis



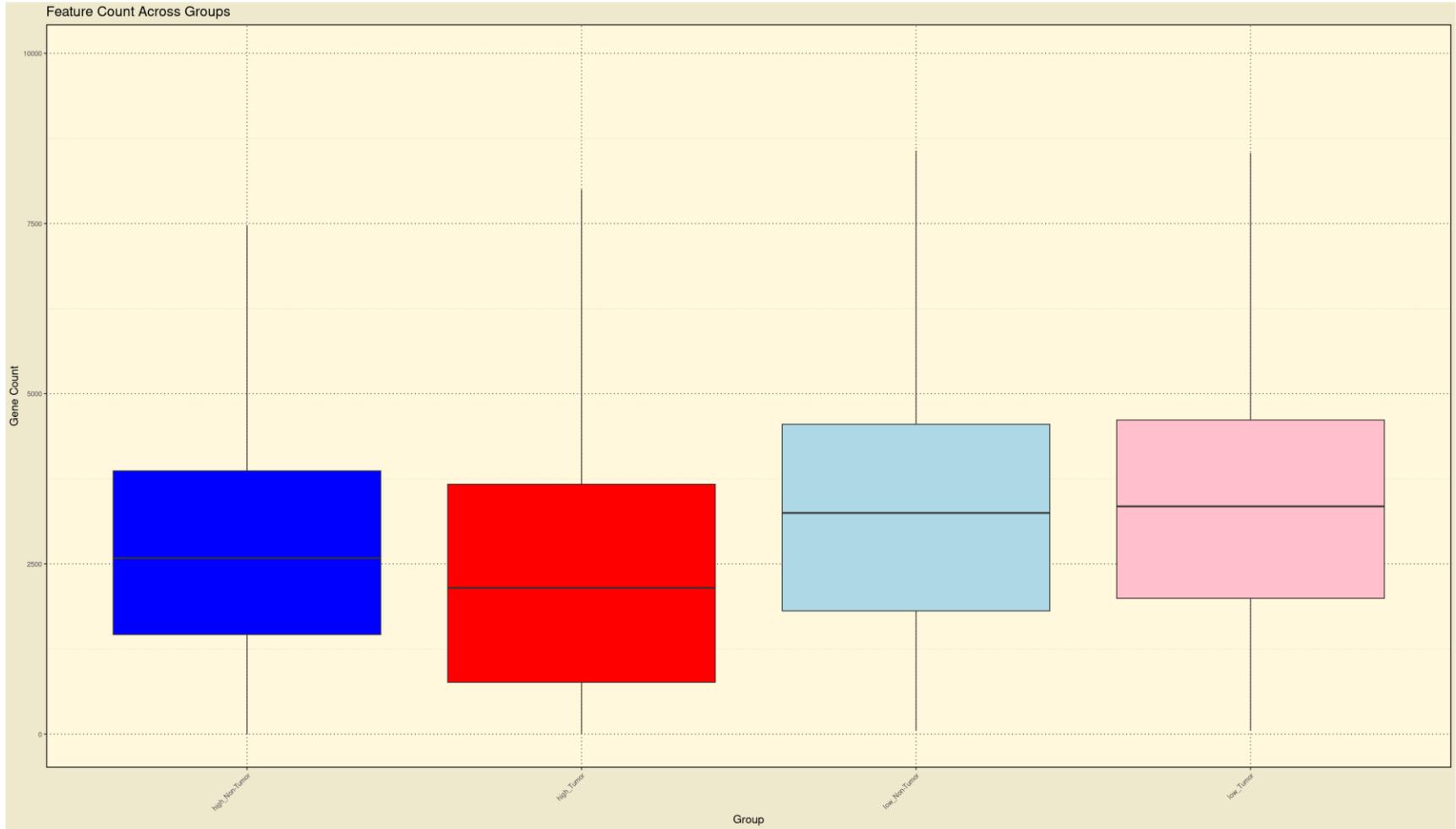
Identified differentially expressed genes (DEGs)
between high and low stress groups, and between
tumor and non-tumor barcodes.



Findings are grouped into four categories:

- High stress, non-tumor
- Low stress, non-tumor
- High stress, tumor
- Low stress, tumor

Differential Gene Expression across the 4 Groups



Marker Gene Analysis

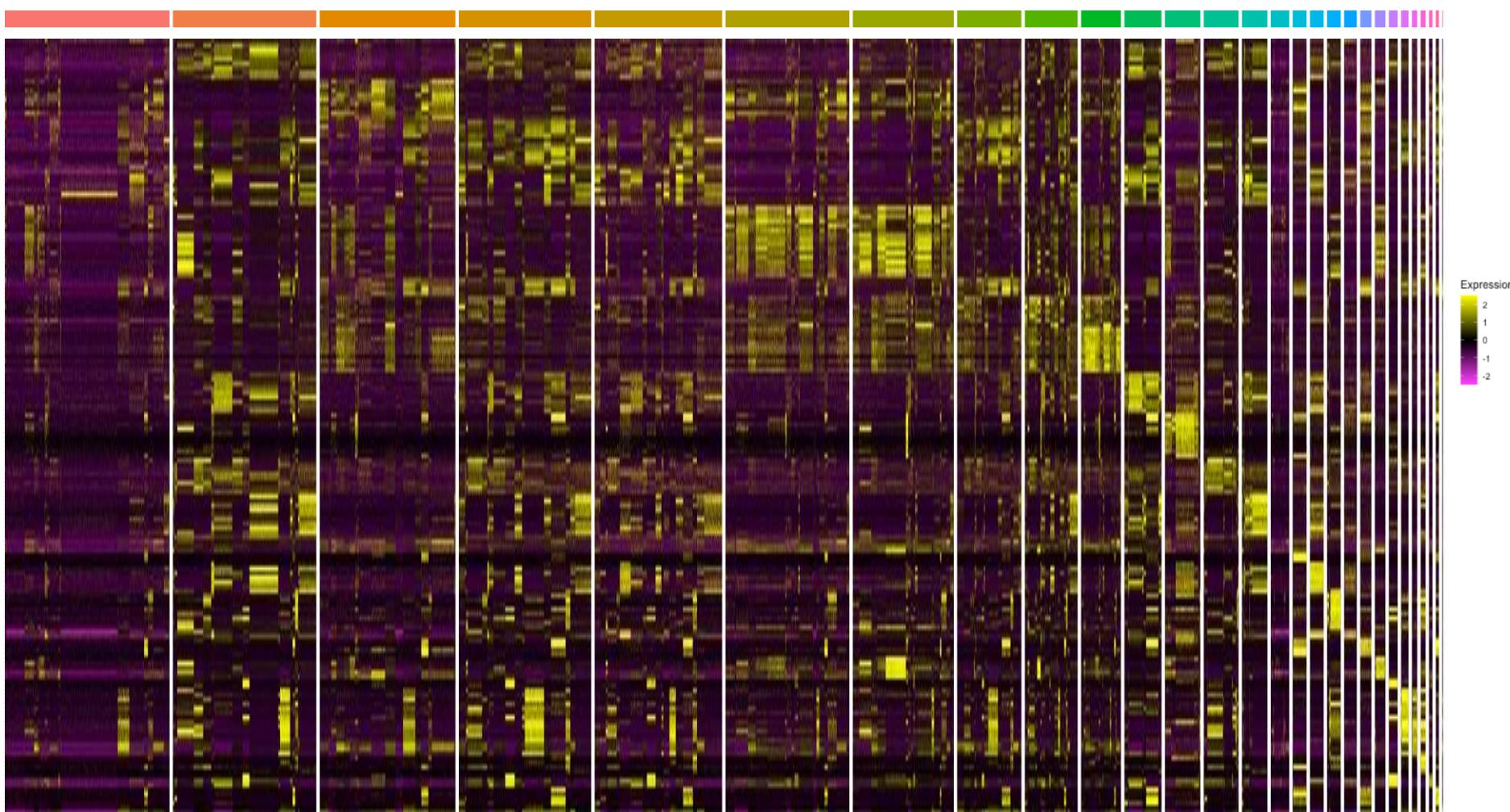


Identified top 10 marker genes for each cluster.



Visualizing expression via marker heatmap.

Top 10 Marker Heatmap per Cluster



Conclusion & Future Directions



Key Findings:

There is a marked difference between gene expression in high-stress and low-stress samples.



Future Work:

- Further validation of key marker genes.
- Manually annotating the clusters.
- Overlaying key marker genes on spatial plot.
- Integration with other omics data.