

Visualization and Integration of multiple spatial modalities: One Ring to Rule Them All



Trevor D. McKee PhD, CTO

Mission Statement

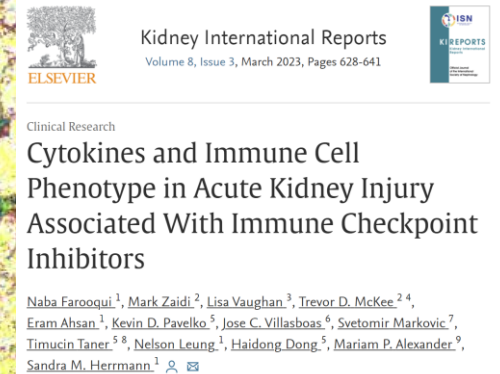
Our mission is to transform the landscape of spatial biology by delivering solutions that demystify complex imaging data.

You stain + We train = Biology explained

Problem: Highly Multiplex Images are Not Human-Readable

Beyond 2-3 colors:
not Human-Readable

	Control	ICI-AIN	ICI-Other	Drug-AIN
CD3	5.4	17.9	7.6	8.9
CD4	0.5	6.8	0.4	3.5
CD8a	0.2	1.6	0.2	0.4
CD11b	4.9	0.2	0.7	0.6
CD11c	0.1	1.7	0.1	0.9
CD14	7.0	1.7	1.9	8.9
CD16	1.0	3.4	3.1	10.8
CD19	0.5	2.1	2.8	5.2
CD20	0.0	0.4	0.0	0.3
CD45	3.2	29.2	5.1	19.0
CD45RO	6.8	8.1	3.0	4.3
CD45RA	0.2	0.5	0.3	0.4
CD68	0.1	0.9	0.1	1.6
FoxP3	5.8	7.6	12.0	17.4
HLA-DR	0.5	7.4	0.6	2.3
GranzymeB	0.2	0.2	0.1	0.1



2020 Mayo Clinic Proceedings,
2022 Kidney International Reports

Data Visualization / Quantification Necessary for Spatial Biology

Our Solutions

Diverse Inputs

Brightfield IHC

Spatial
Proteomics

COMET MxIF
IMC

Spatial
Transcriptomics

STOmics
10x Nanostring

Spatial
Metabolomics

Bruker
MALDI-MSI
DESI-MS

...
(other
spatial
modalities)

Image Converters

Analytical Platforms

VISIO^{PHARM}

HALO



QuPath

scanpy

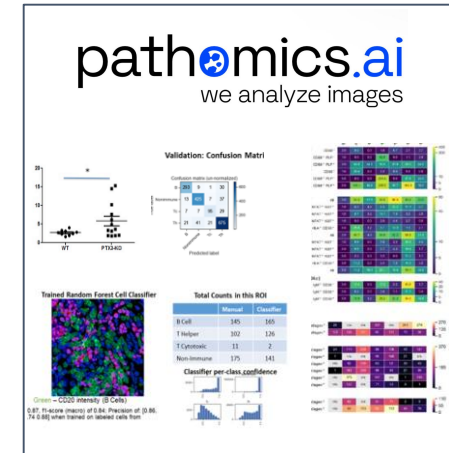
squidpy

MCMICRO

...
(other
platforms)

Tailored Analytics

Downstream Analytics



Standardized Spatial Report
From Data Table to Figures

IMS Integration

(in development)
APIs for Integration

Pathcore

...
(other
platforms)

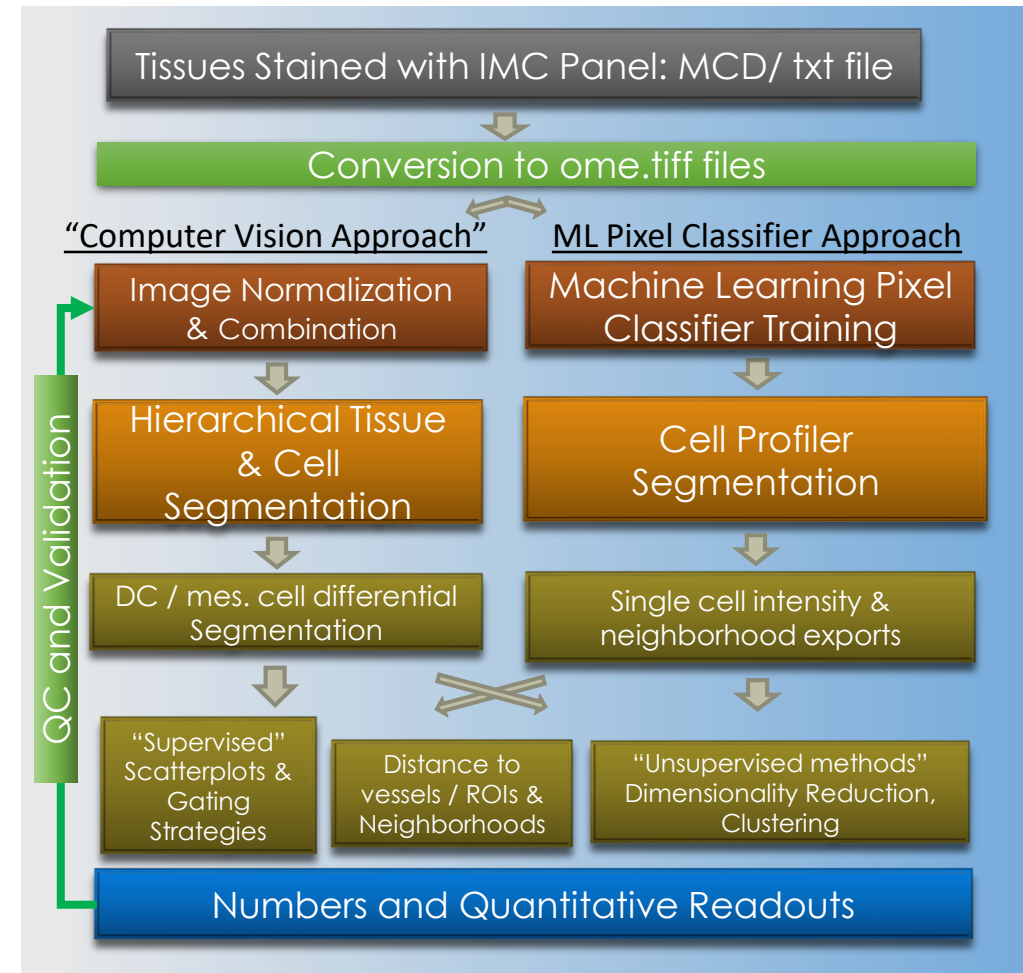
pathomics.ai

Develop / Deploy Custom Pipelines

- Provision of **Services**, **Downstream Analytics** and **Products** to Deliver Interpretable Results from Complex Spatial Biology Data to Academic, Industry, CRO clients info@pathomics.ai

pathomics.ai Generalized Multiplex Analysis Pipeline

- Quantitative analysis requires several steps, incorporated in an end-to-end pipeline for analysis
- An open-source converter (<https://github.com/STTARR/imc-converter>) can convert IMC files to multi-channel OME-Tiff standard (STOmics ome.tiff converter, others in pipeline)
- Choice of “Computer vision” or machine learning pixel approach can be used for segmentation.
- Classification can follow either “supervised” gating strategies or “unsupervised” dimensionality reduction & clustering approaches.
- Critical to success is a robust QC and validation strategy to ensure accuracy



Inputs

Cell Segmentation

Classification

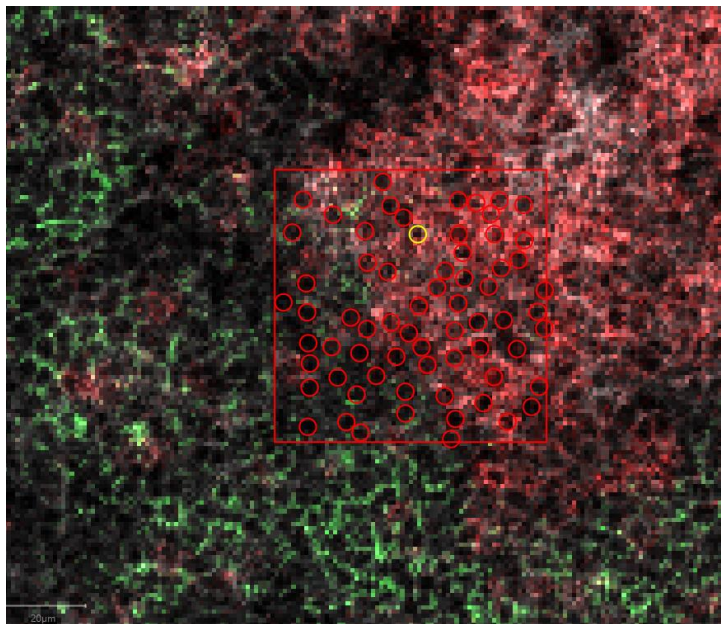
Quality Control

To Supervise or Not To Supervise?

- It Depends: Are you a Pathologist or a Single Cell Biologist?
- What is the scientific / biological question?
- Cell Segmentation, Modality, Tissue all impact – validation needed
- You can't beat physics! – diffusion of transcripts, oblique cell cuts
- It may be useful to get creative, or use combined approaches

How to segment without nuclei? (or: What's the worst you can do?)

Manual Annotations (Membrane)

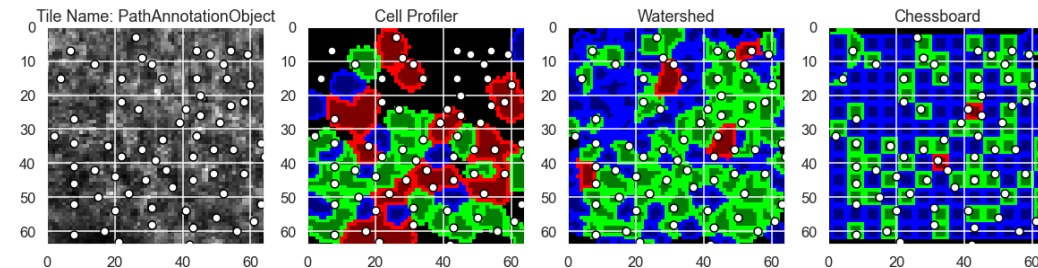


Manual Annotations (white dots)

1:1 “Singlets” (green)

Undersegmented cells (red)

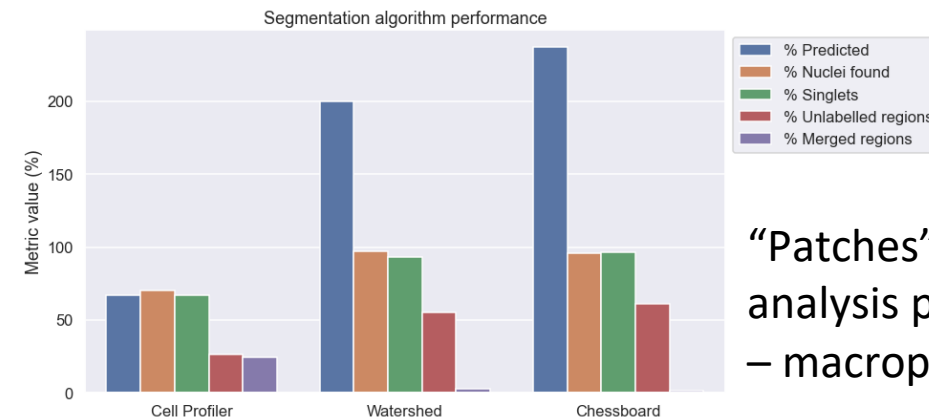
Oversegmented cells (blue)



pathomics.ai

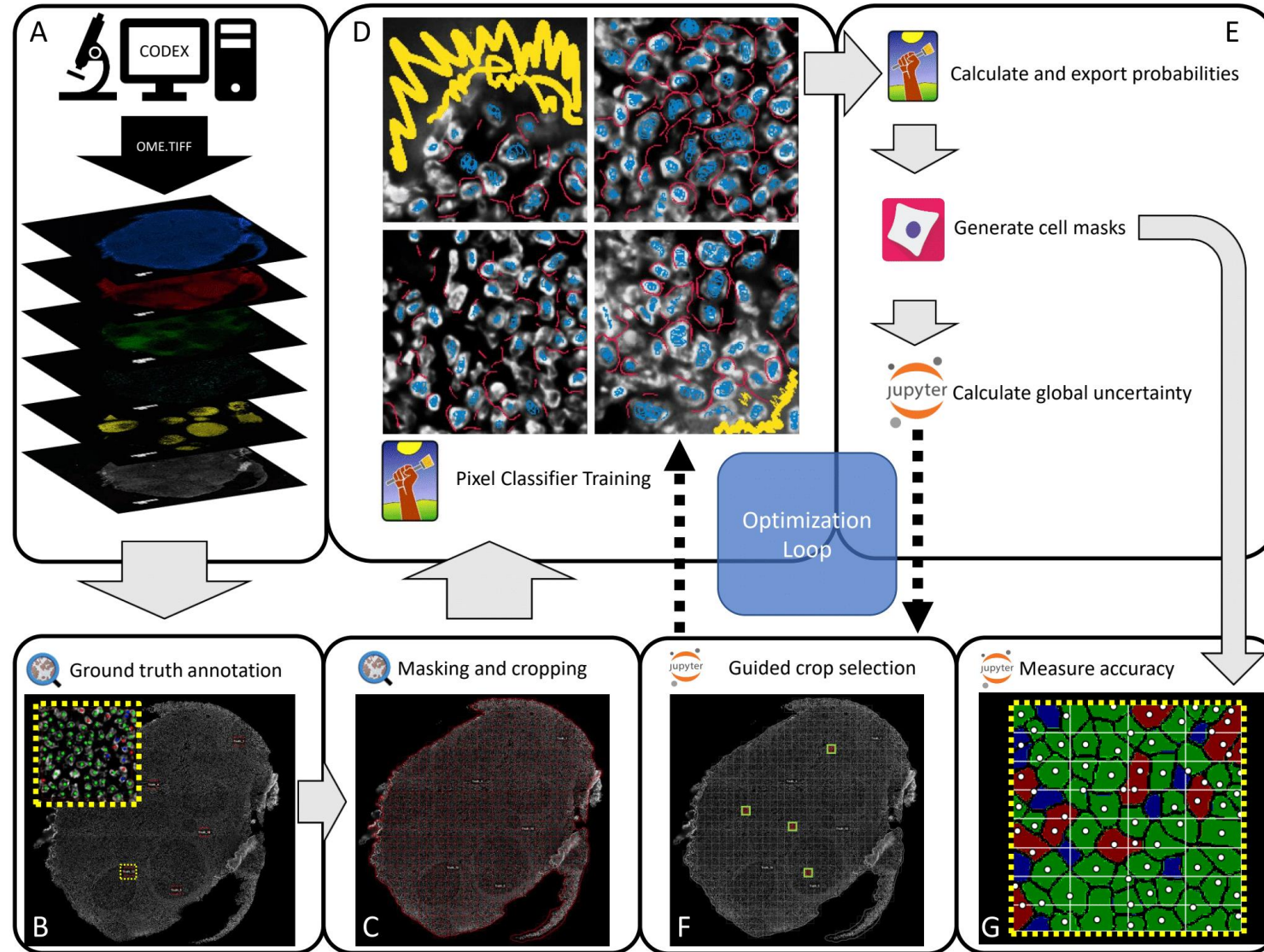


Quality Control



“Patches” or pixel-based analysis potentially useful – macrophage / astrocytes

LOGHAM: Looped Optimization Guided by Human Annotation and Machine Learning

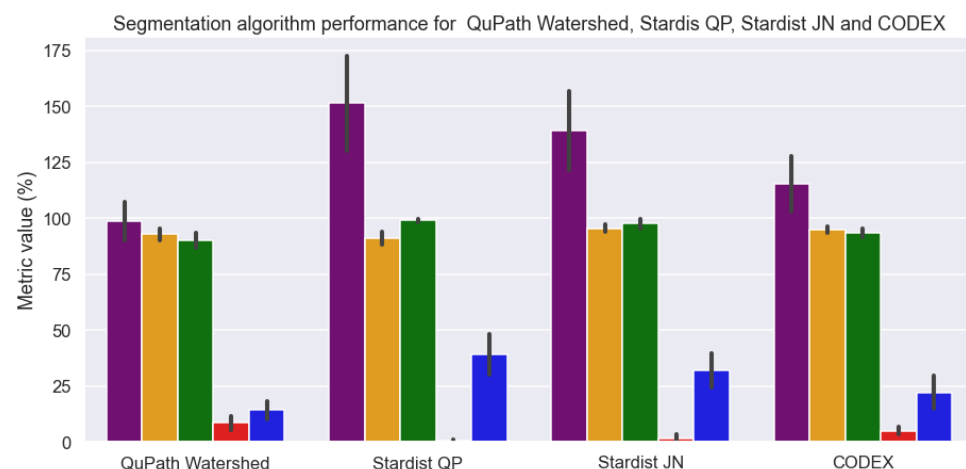
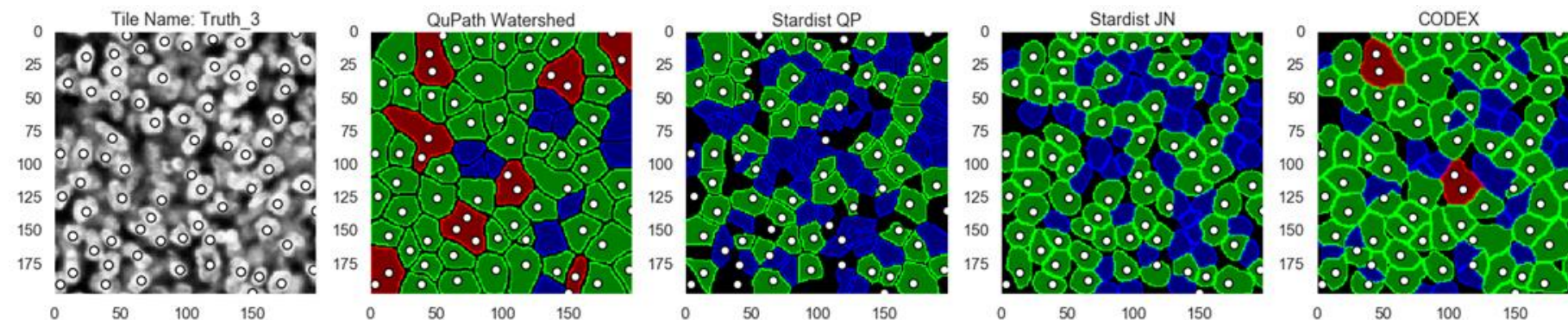


pathomics.ai

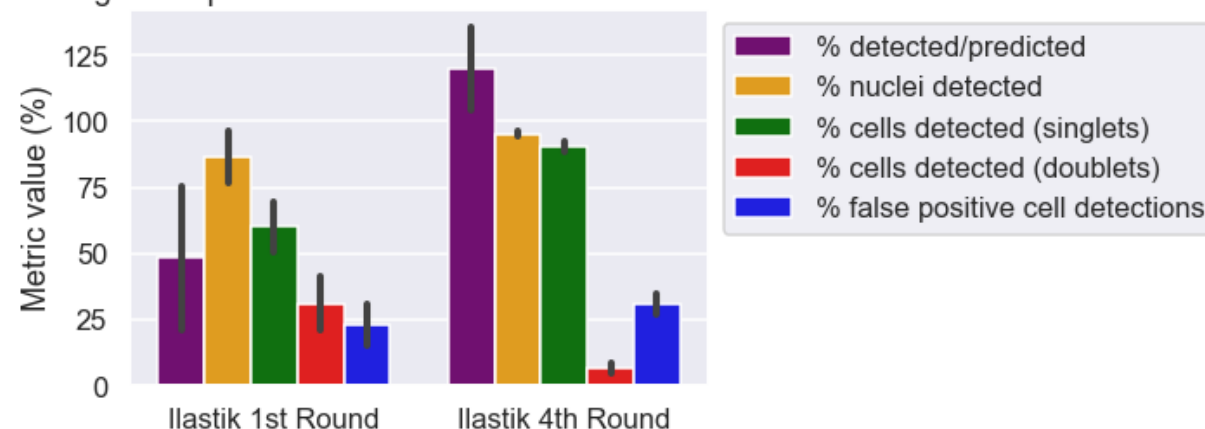


JC Villasboas

Head-to-head validation multiple AI/ML segmentation methods



Segmentation algorithm performance for Ilastik 1st Round and Ilastik 4th Round



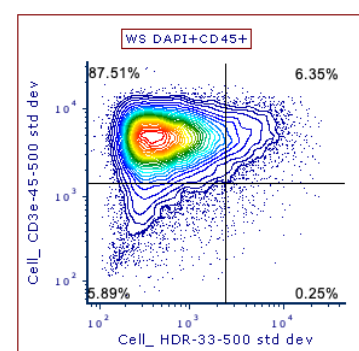
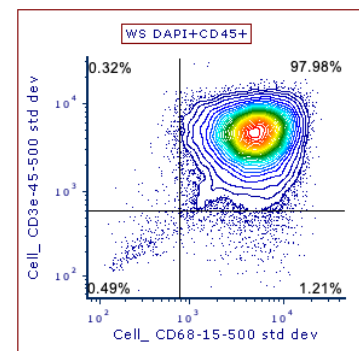
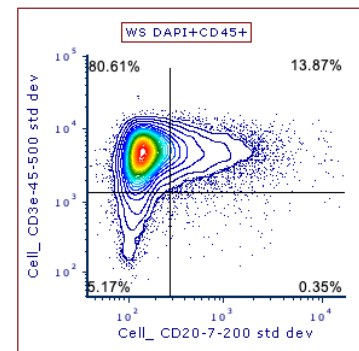
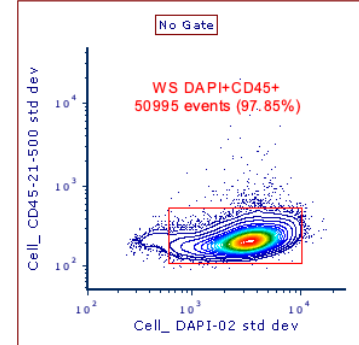
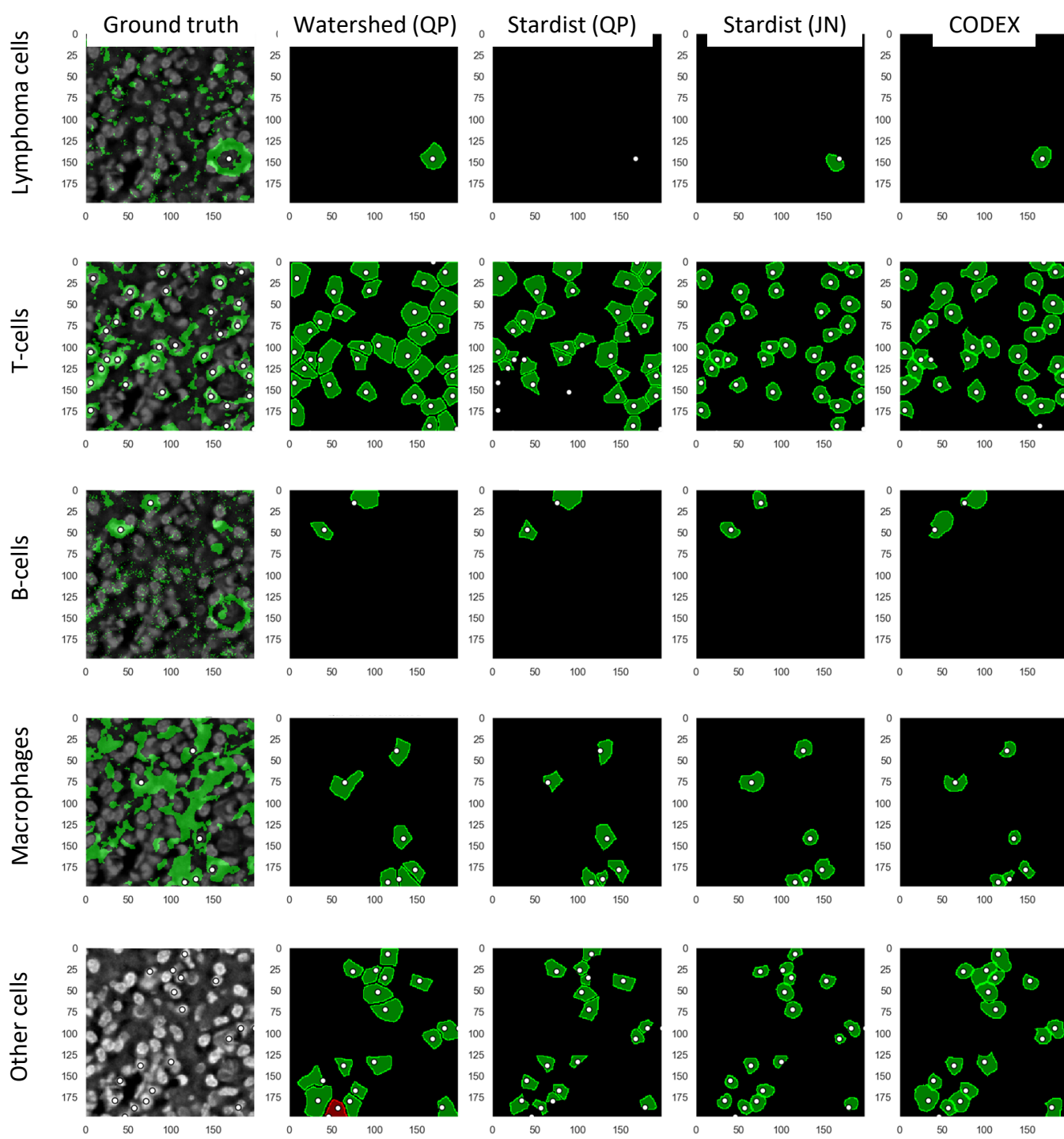
Quality Control

(paper in preparation)

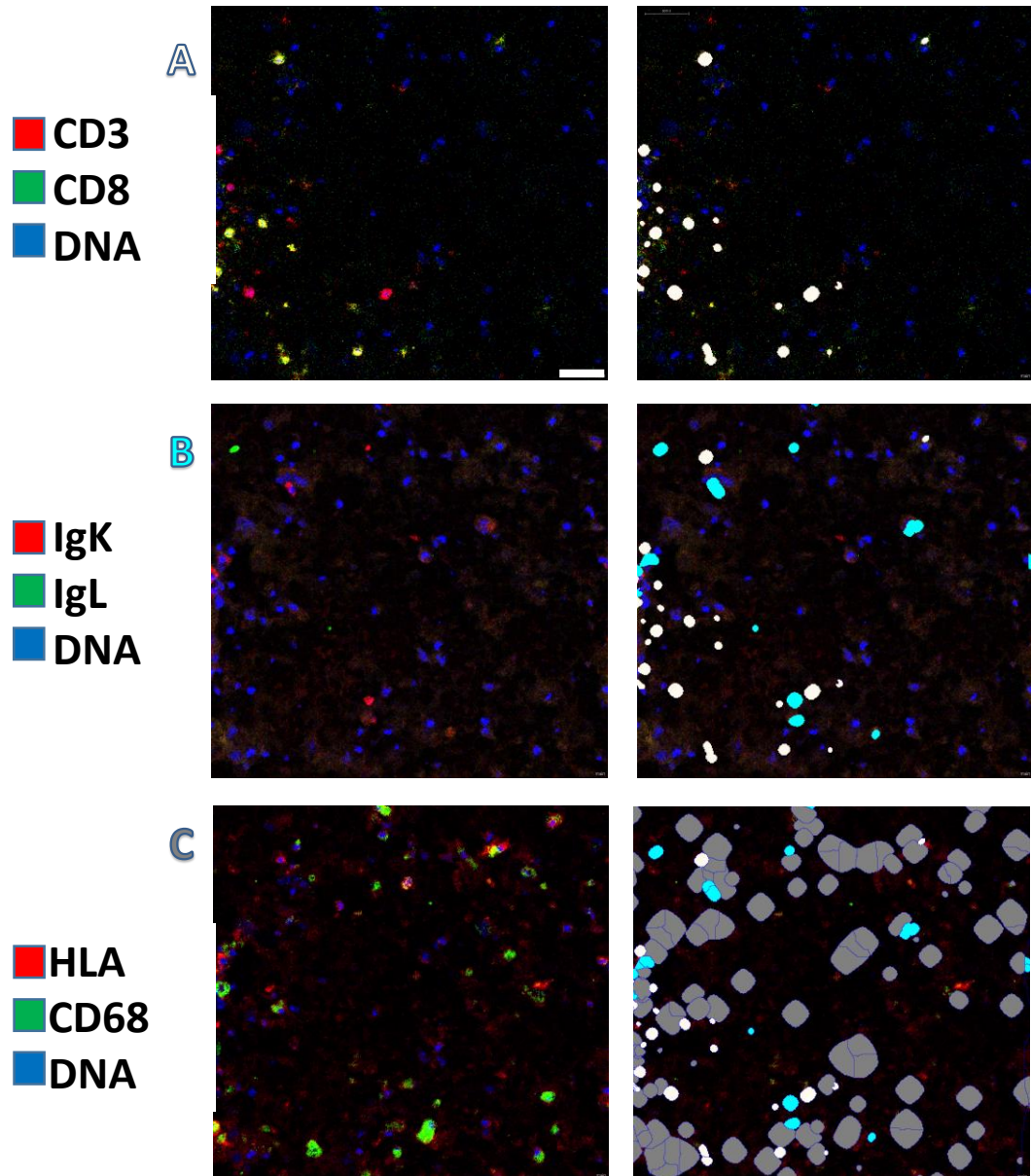
Quality Control

Cell Segmentation

Classification



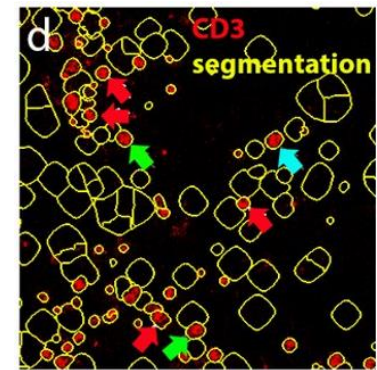
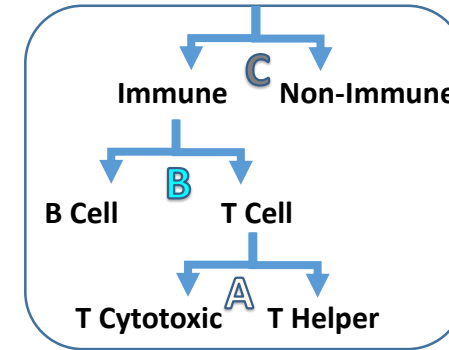
Using Biology To Guide Segmentation Strategy (w/help of Computer Vision)



Putative T Cell

Putative T Cell
Putative B Cell

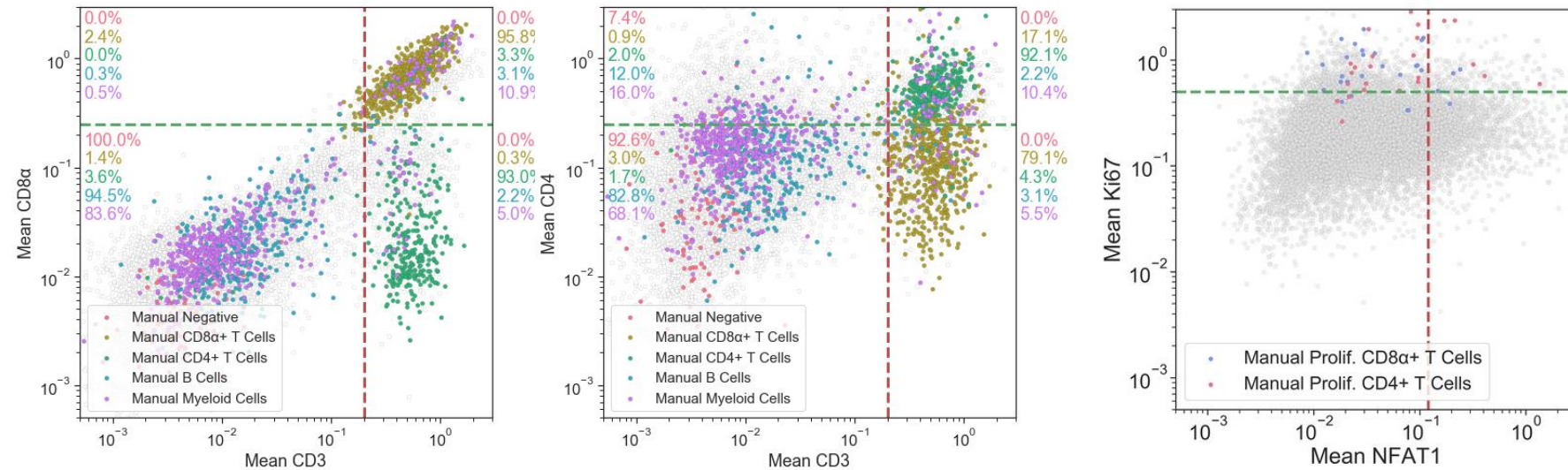
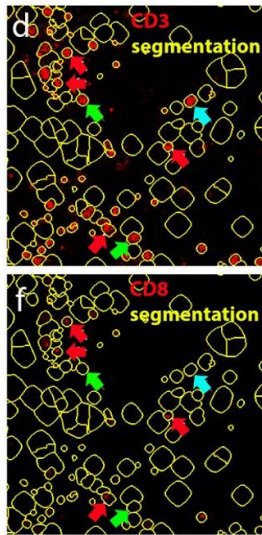
Putative T Cell
Putative B Cell
Putative Other Cell



Validation:
Manual inspection
to label particular
cell types

Ramaglia et al, eLife 2019

To Supervise: Multiplexed Cellular Classification using Manual Cytometry Gating



```
df.loc[:, "class_name"] = "Nucleus" # Reset all to nucleus

negative_cond = L0["IgKappa"] & L0["IgLambda"] & L0["IgM"] #& L0["HLA"]# & L0["CD68"]
df.loc[negative_cond, "class_name"] = "IgM- λ- κ-"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"], "class_name"] = "T Cells"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & HI["CD8a"] & L0["CD4"], "class_name"] = "Cytotoxic T Cells (Tc)"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & L0["CD8a"] & HI["CD4"], "class_name"] = "Helper T Cells (Th)"

df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & HI["CD8a"] & L0["CD4"] & HI["KI67"], "class_name"] = "Proliferating Tc"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & L0["CD8a"] & HI["CD4"] & HI["KI67"], "class_name"] = "Proliferating Th"

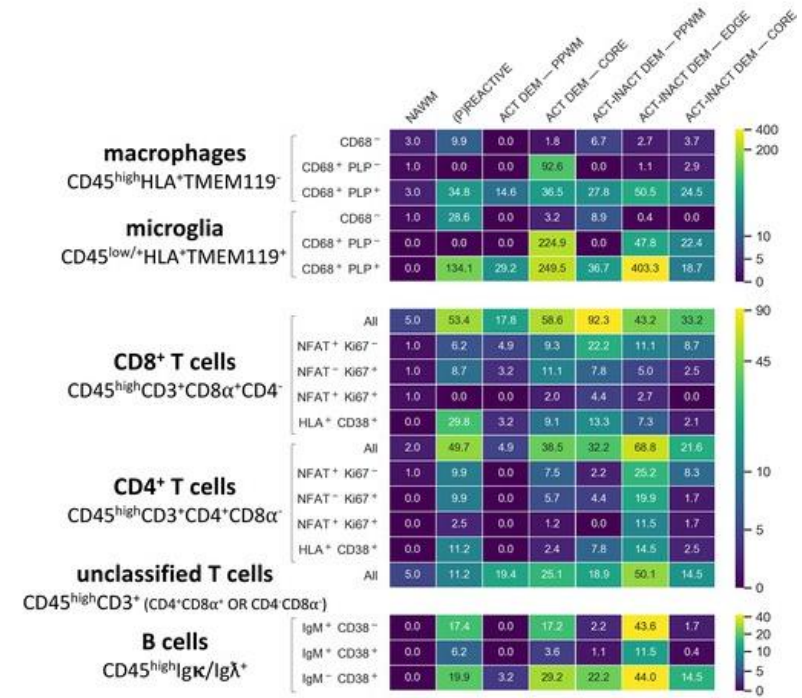
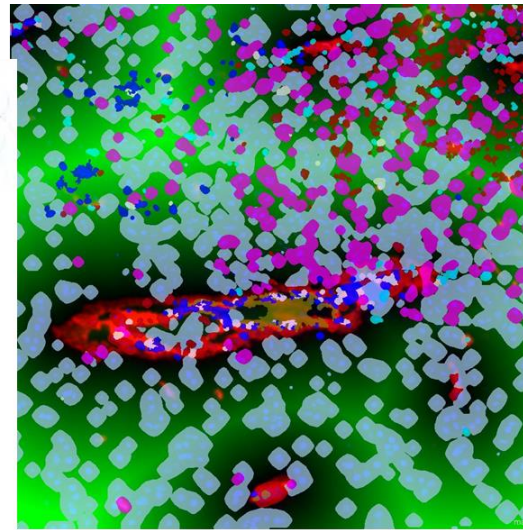
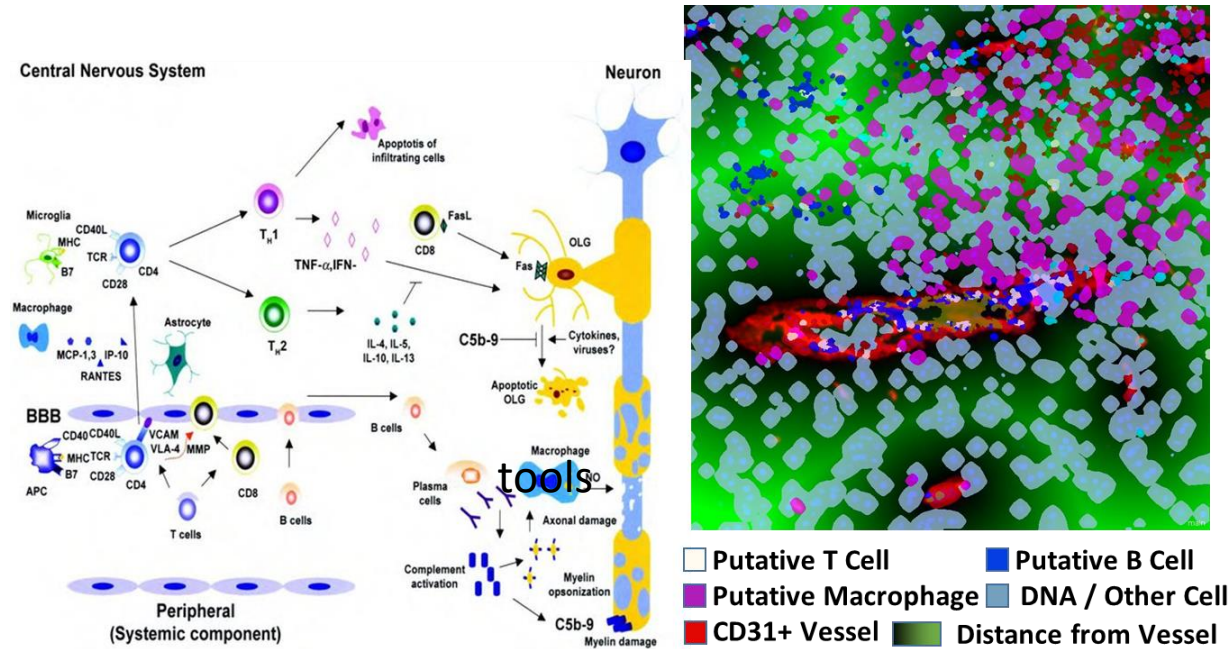
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & HI["CD8a"] & L0["CD4"] & HI["NFAT"], "class_name"] = "Activated Tc"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & L0["CD8a"] & HI["CD4"] & HI["NFAT"], "class_name"] = "Activated Th"

df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & HI["CD8a"] & L0["CD4"] & HI["KI67"] & HI["NFAT"], "class_name"] = "Prolif. & Activ. Tc"
df.loc[negative_cond & HI["CD3"] & HI["CD45_T"] & L0["CD8a"] & HI["CD4"] & HI["KI67"] & HI["NFAT"], "class_name"] = "Prolif. & Activ. Th"
```

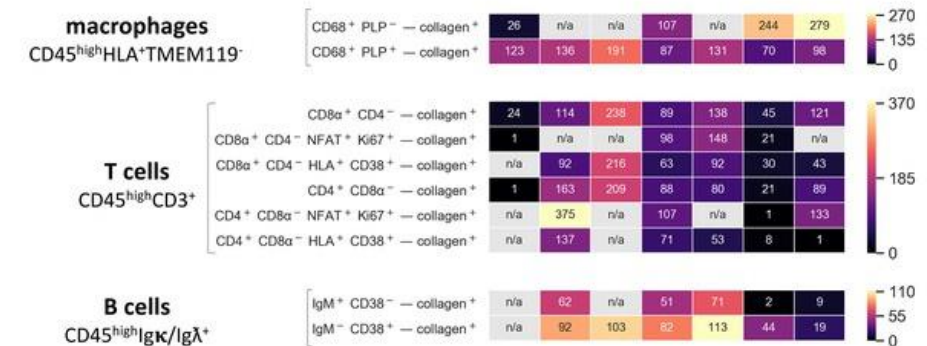
Ramaglia et al, eLife 2019

To Supervise: Quantitative Spatial Immune Invasion in Multiple Sclerosis Brain Lesions

Spatial Heatmap

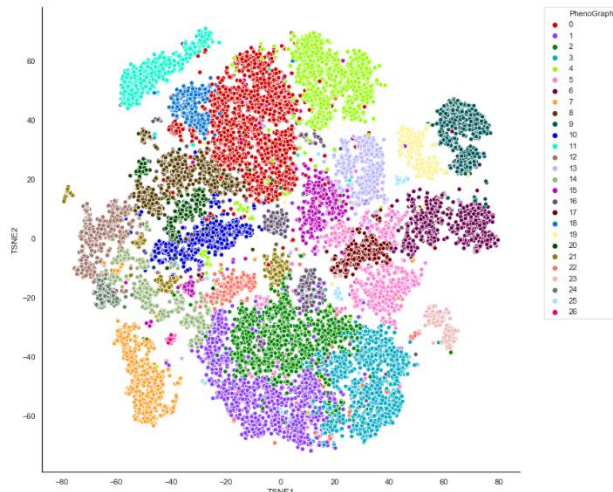


- Segment individual immune cell subtypes, determined by manual gating strategies relative to “gold standard” immunologist annotations
- Report immune cell subtype density (cells / mm²) as heatmap within MS lesions with discrete biological activity (inactive, active, slowly expanding lesions)
- In addition, can report per-cell distance to nearest (in-plane) blood vessel, to interrogate immune cell influx into lesion microenvironment

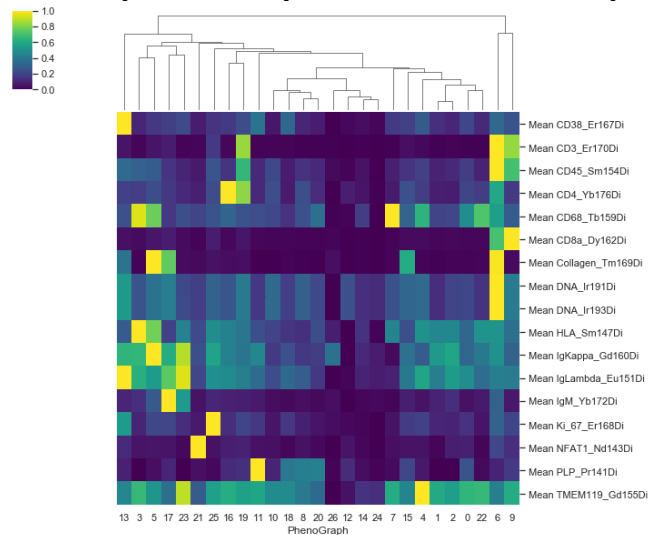


Unsupervised classification: Dimensionality reduction & PhenoGraph clustering

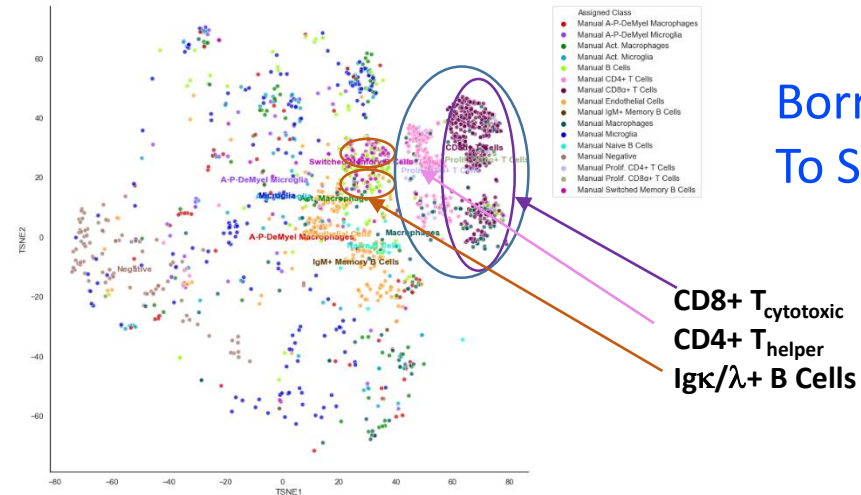
t-SNE plot of cells, colored by PhenoGraph cluster



PhenoGraph scatter plot: Distinct marker profiles

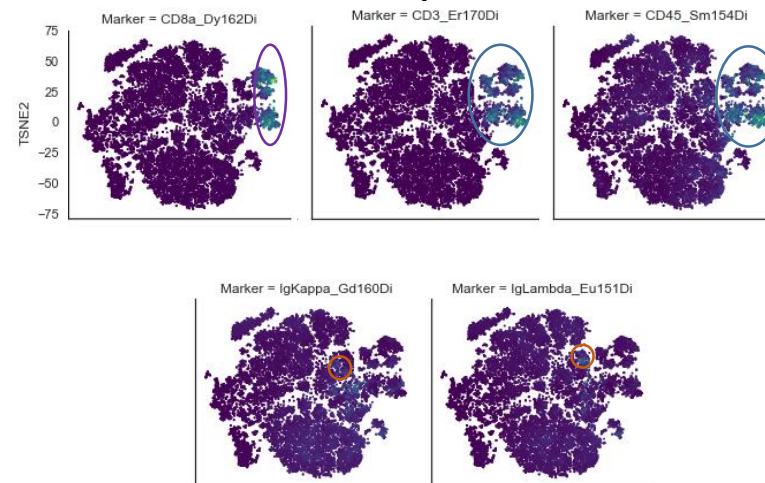


t-SNE plot: Just manually-labeled cells



Borrow from:
To Supervise

Individual Marker Intensities plotted on t-SNE



Combine Supervised with
Unsupervised?

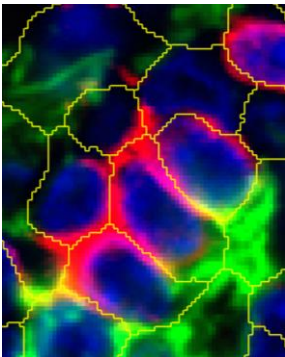
- Identify new immune subsets within major immune cell clusters
- Distances from T Cells to Tumor / Myelinated Cells?

Sometimes Segmentation isn't Enough: Bleed-through from neighbors

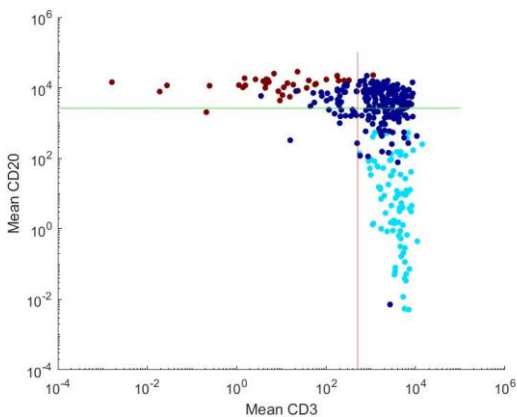
Touching
Membranes?

Oblique Cell
Cut?

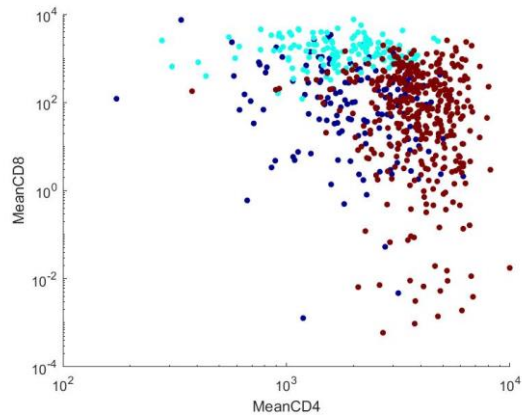
Segmented Cells



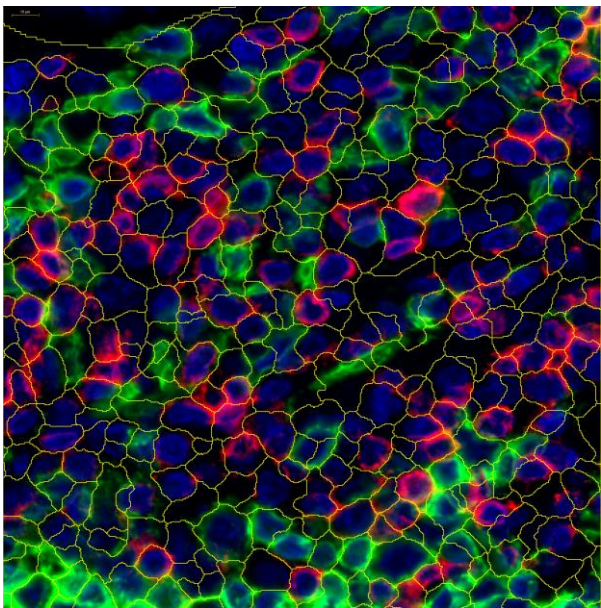
B Cell / T Cell Scatterplot



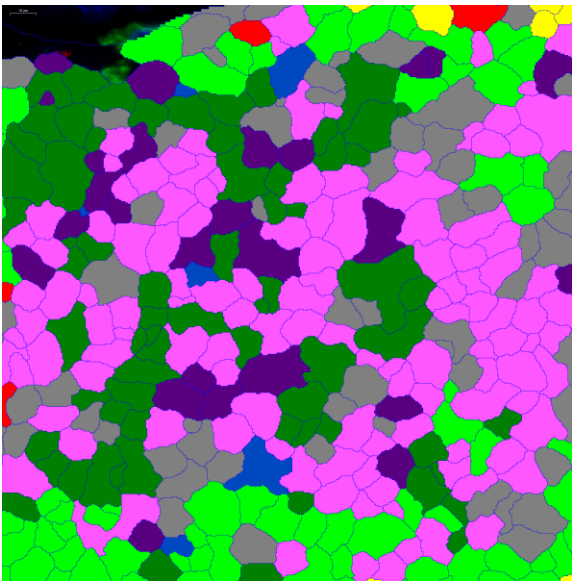
Thelper vs TCytotoxic Scatterplot



Original Image



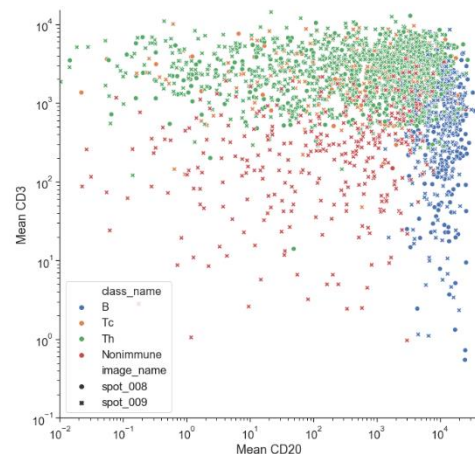
Manually Labeled Cells



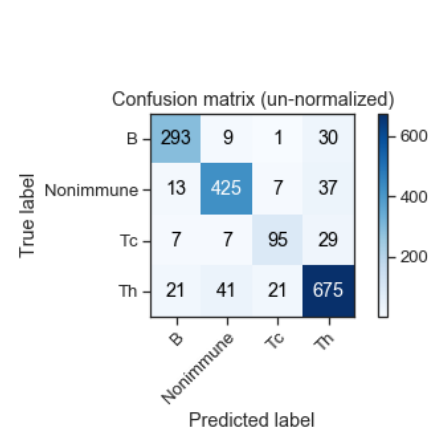
- Red – CD3 Intensity (T Cells)
- Green – CD20 intensity (B Cells)
- Yellow – Segmented Cell Border
- B cells based on threshold
- Actual B cells (based on manual investigation)
- Actual TC cells (based on manual CD8 investigation)
- Actual TH cells (based on manual CD4 investigation)
- and ■ Non-immune cells

To Supervise: Pathologist Annotation-guided Random Forest Cell Classifier

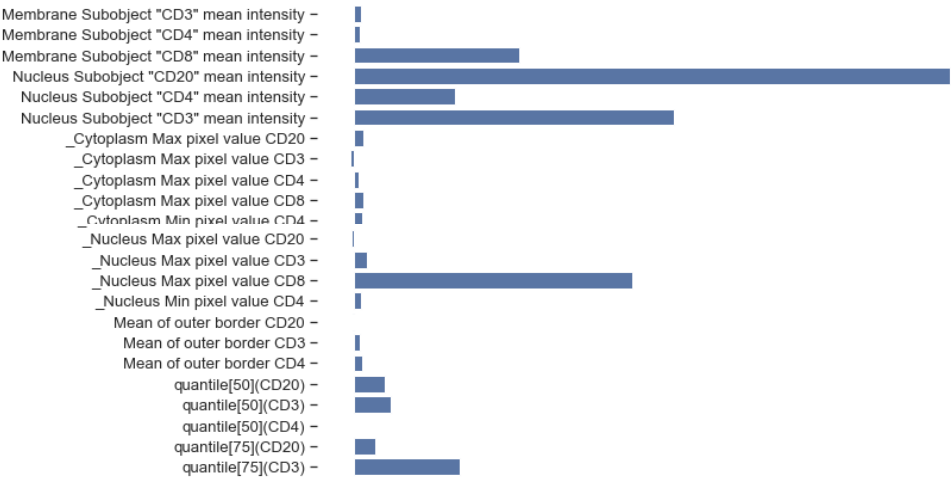
Scatterplot of Classified Cells



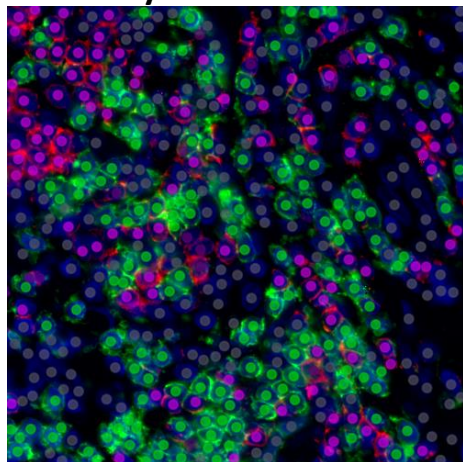
Validation: Confusion Matrix



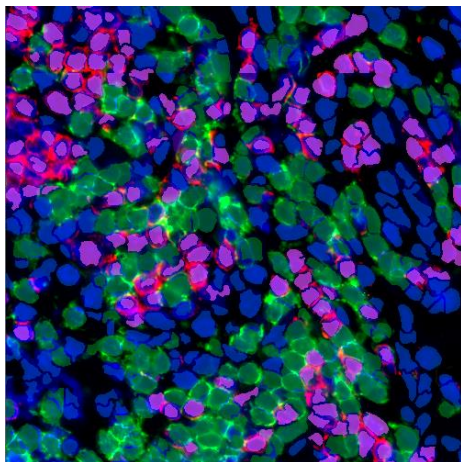
Relative Feature Importance for Classification



Manually Labeled Cells



Trained Random Forest Cell Classifier

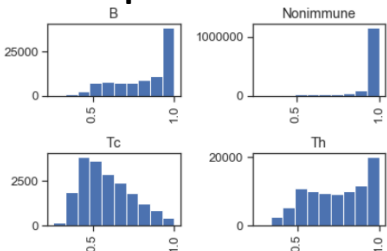


Red – CD3 Intensity (T Cells) Green – CD20 intensity (B Cells)

Total Counts in this ROI

	Manual	Classifier
B Cell	145	165
T Helper	102	126
T Cytotoxic	11	2
Non-Immune	175	141

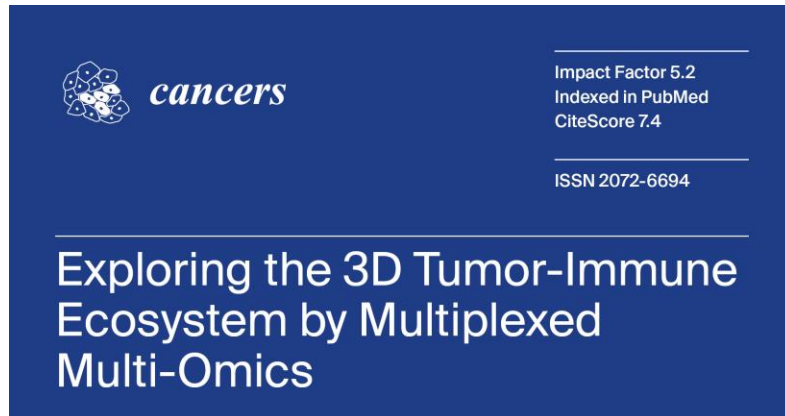
Classifier per-class confidence



- Trained classifier achieves an accuracy of 0.87, f1-score (macro) of 0.84; Precision of: [0.86, 0.88, 0.74, 0.89], and recall of [0.88 0.88 0.74 0.88] when trained on labeled cells from representative tumor microarray cores

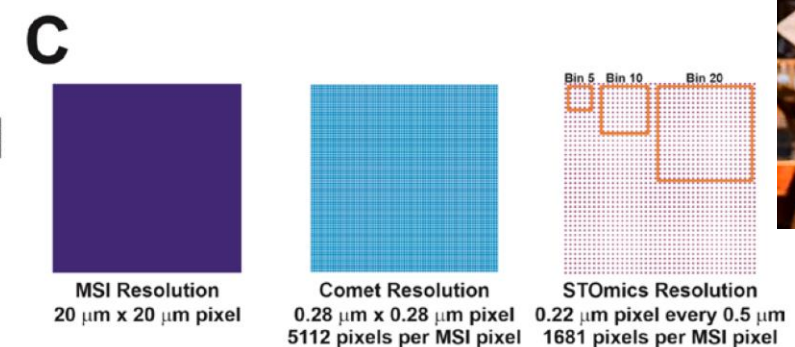
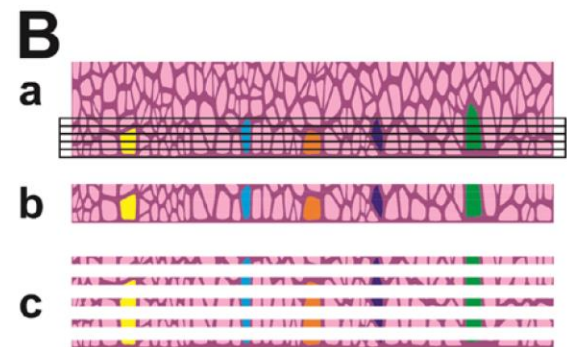
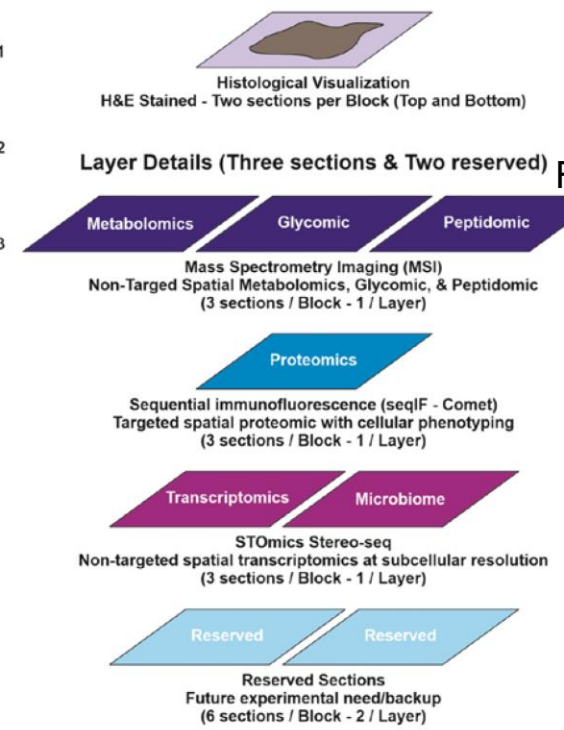
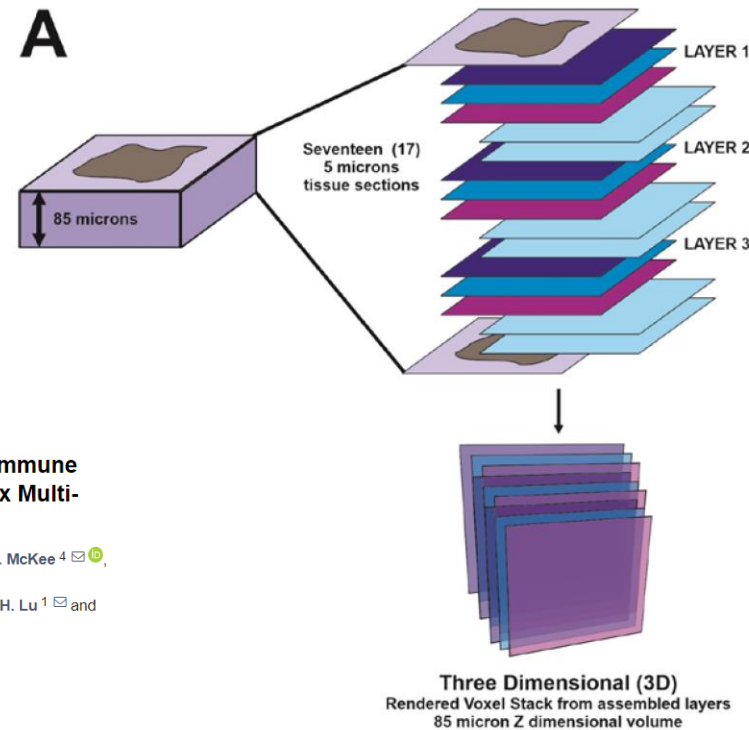
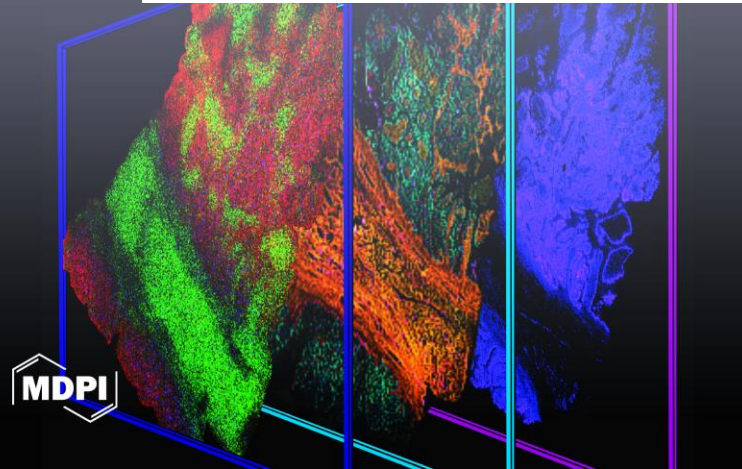
Kanwar, Baldy et al., Canc Res 84:6196 (2021)

Not To Supervise: Multi-Spatial Omics Data



Molecular, Metabolic, and Subcellular Mapping of the Tumor Immune Microenvironment via 3D Targeted and Non-Targeted Multiplex Multi-Omics Analyses

by Sammy Ferri-Borgogno ^{1,*}, Jared K. Burks ^{2,*}, Erin H. Seeley ³, Trevor D. McKee ⁴, Danielle L. Stolley ², Akshay V. Basi ², Javier A. Gomez ², Basant T. Gamal ¹, Shamini Ayyadthury ⁴, Barrett C. Lawson ⁵, Melinda S. Yates ⁶, Michael J. Birrer ⁷, Karen H. Lu ¹ and Samuel C. Mok ^{1,*}



Sammy Ferri-Borgogno



Jared Burks



MD Anderson

Not To Supervise: Complete Genomics STOmics Spatial Transcriptomics: Dimensionality Reduction & Clustering

- 225nm Spatial Transcriptomics Stereo-Seq Data
- Bin, Normalize Data, perform Leiden Clustering
- Images show remarkable pathological features – stromal clusters, stroma / tumor interface clusters, distinct tumor clusters

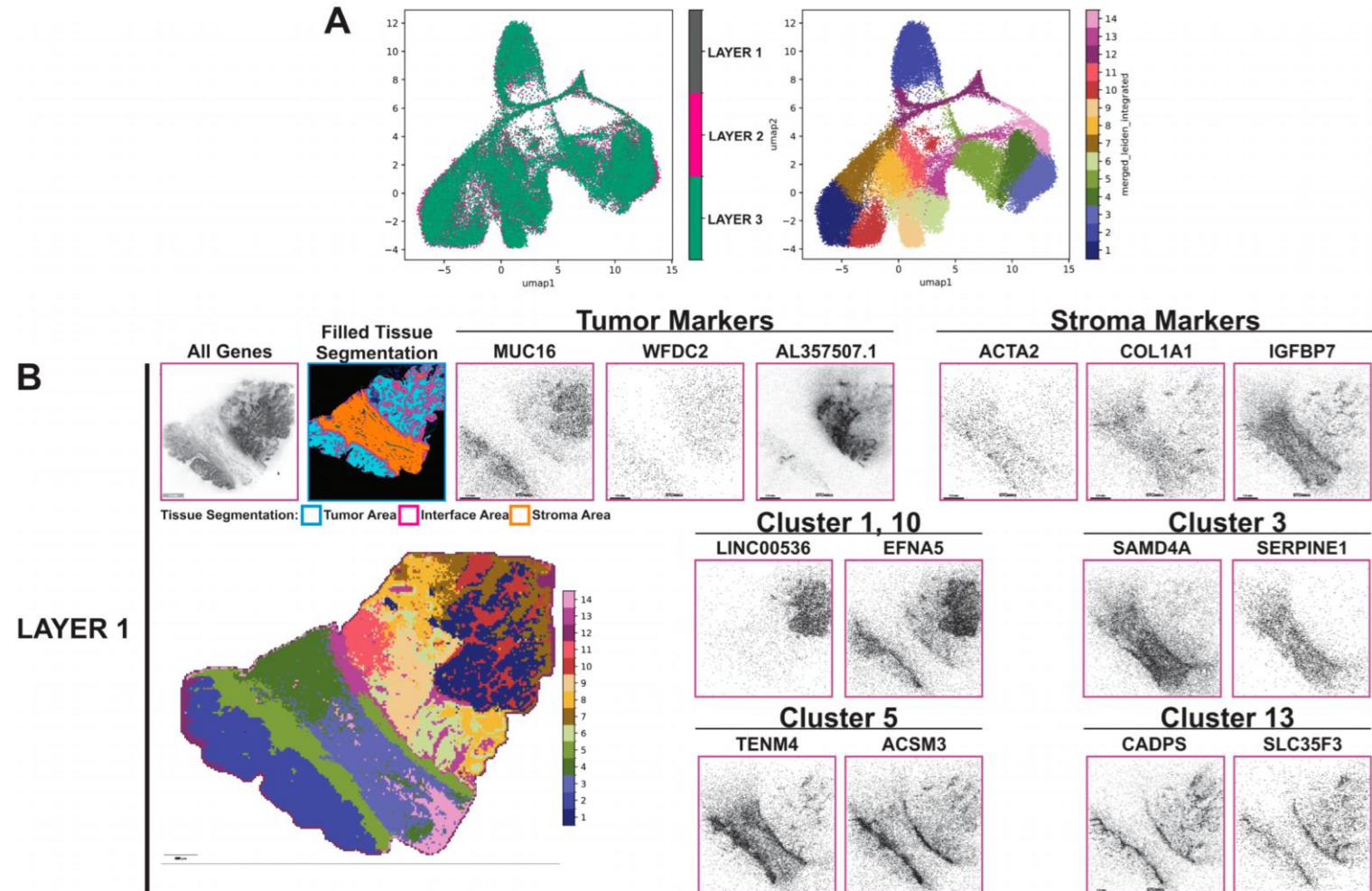
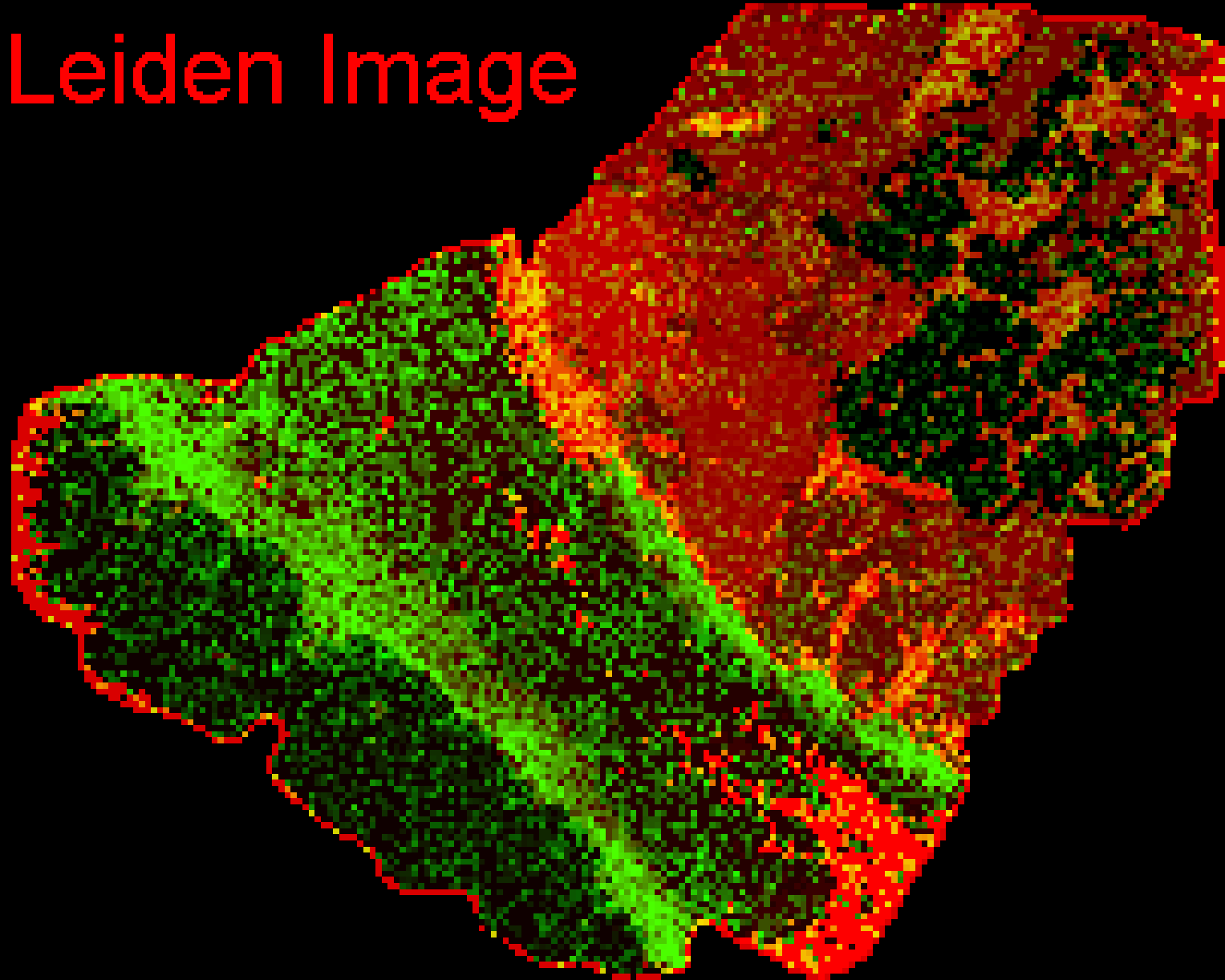


Image Conversion

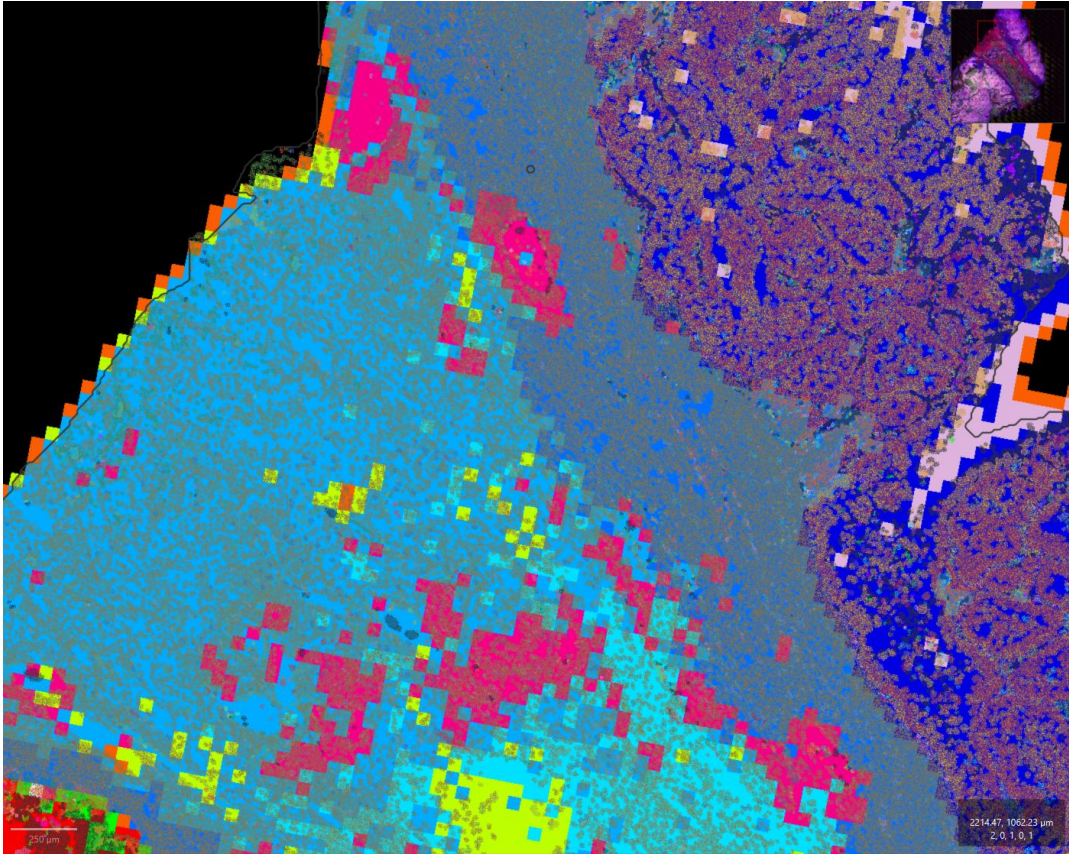
- AnnData Object –
convert to ome.tiff
- Generate image stack
that can be registered
with other modalities
- Ome.tiff stack –
imported back into
Visiopharm – Registered
with COMET and
MALDI-MSI data

ACSM3

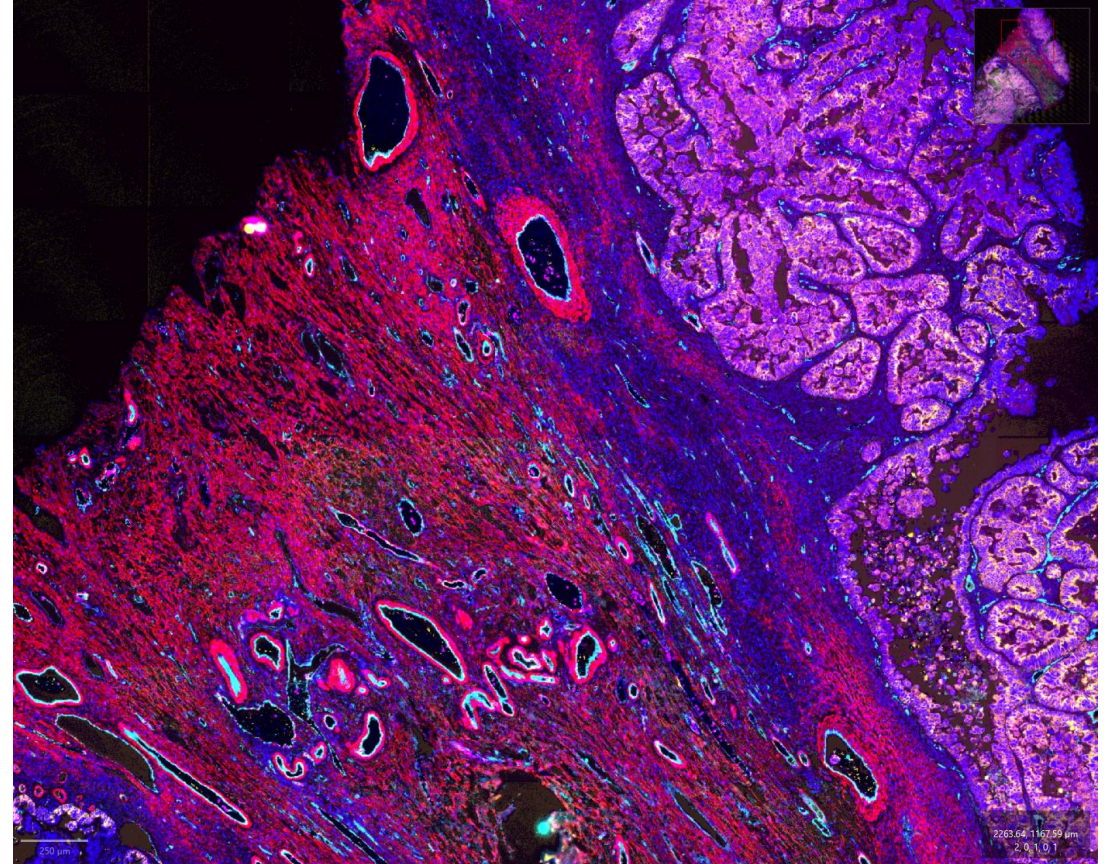
Leiden Image



Not To Supervise: Blood Vessel Clusters

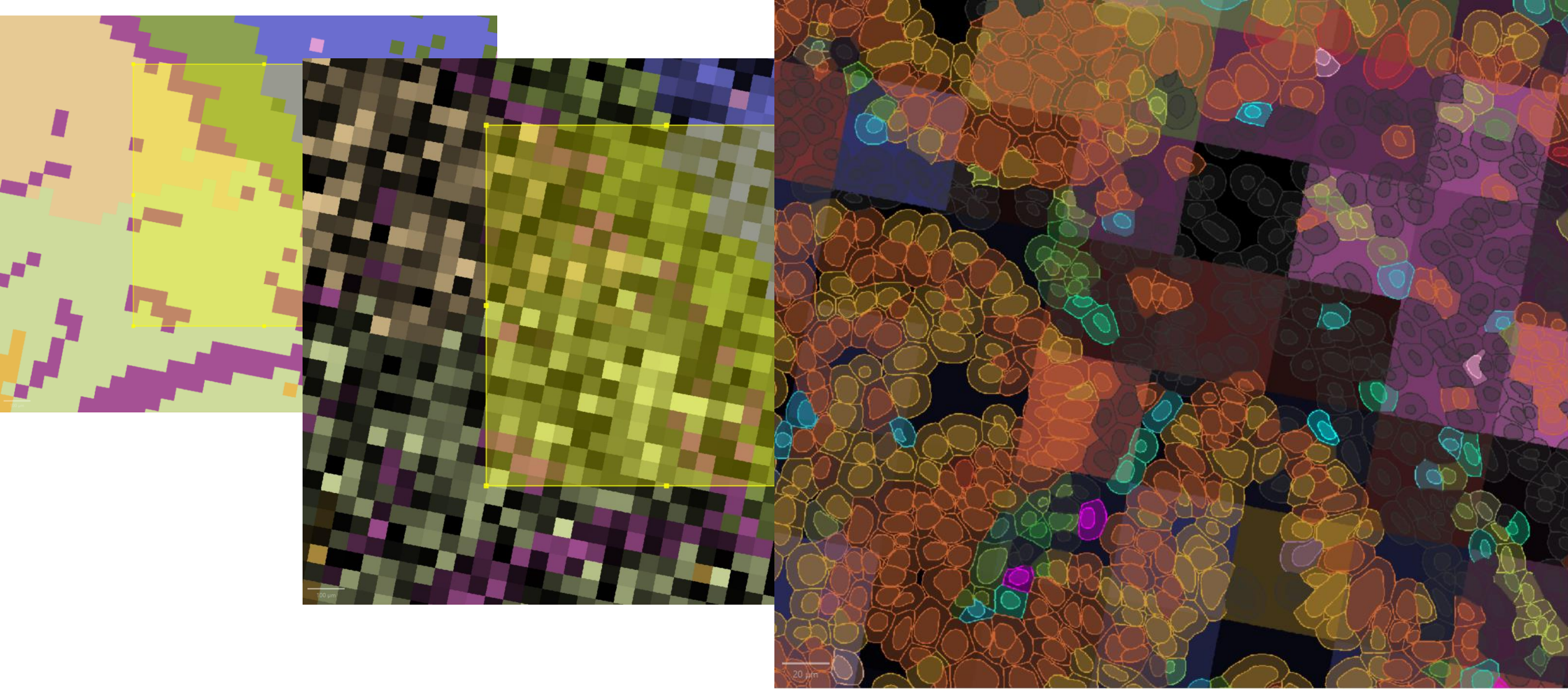


- STOmics Leiden Clusters BV?

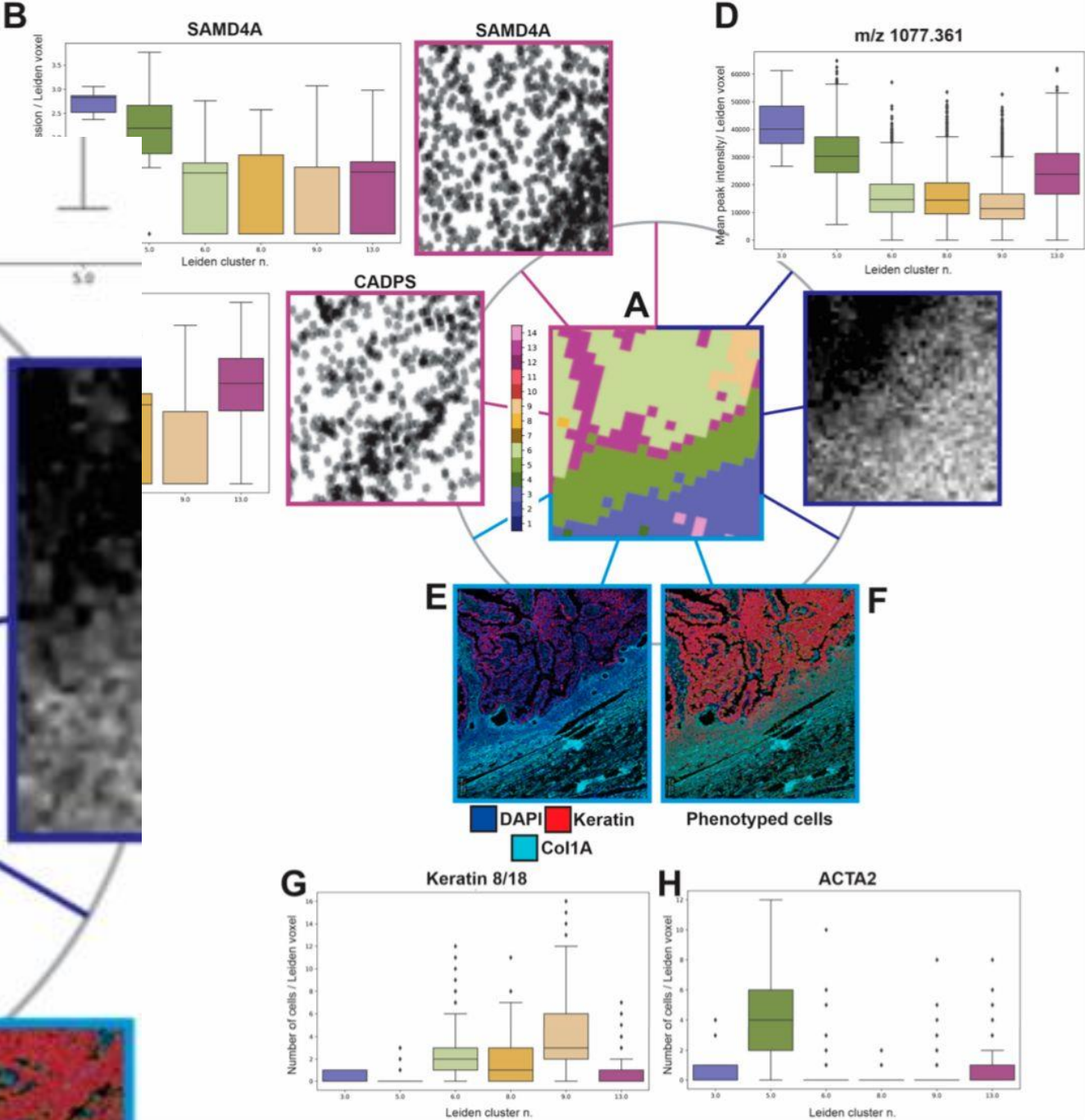
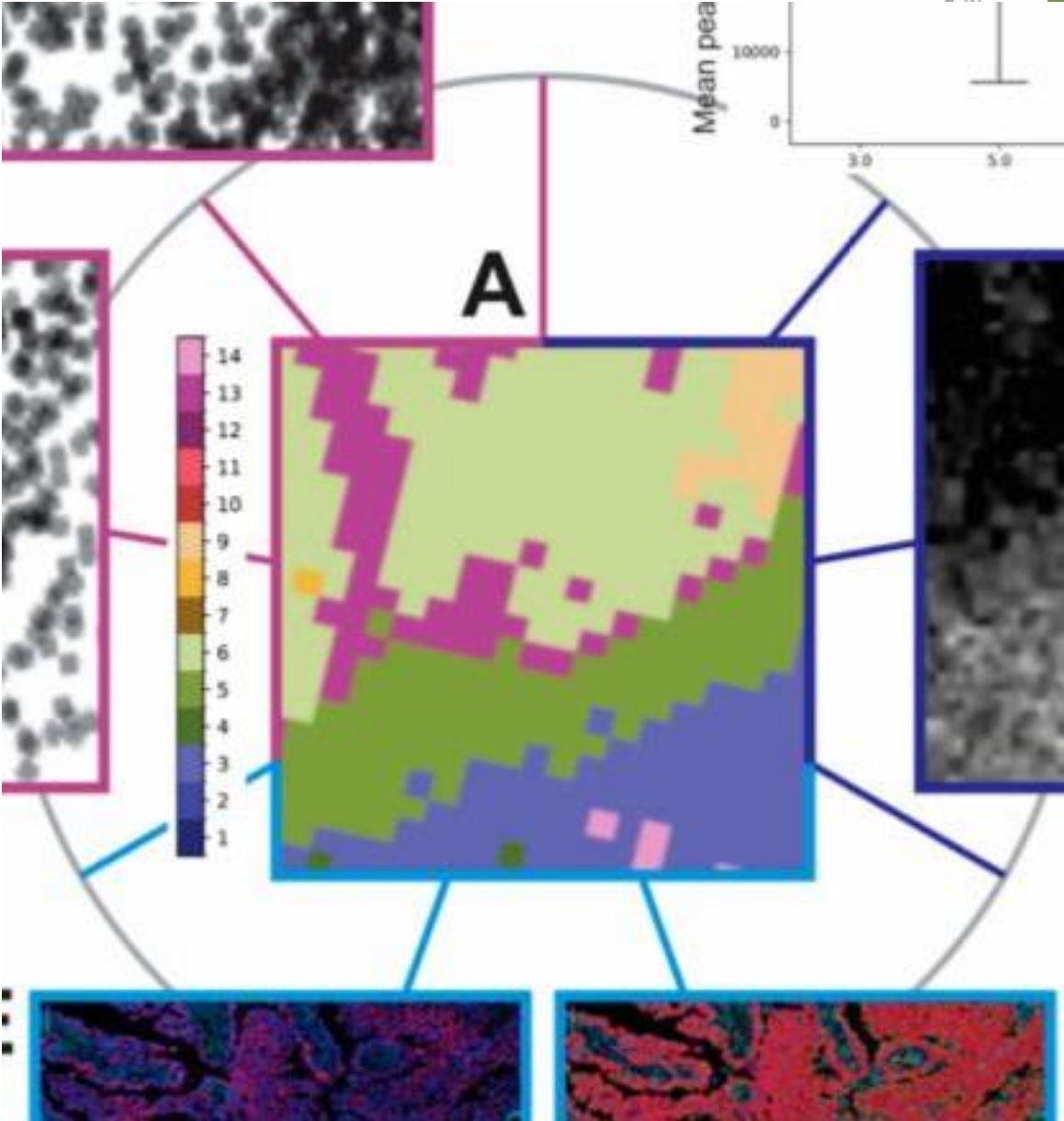


aSMA / CD31 / PanCK

Spatial Multi-Omics: Leiden “Voxels” to move between modalities



“One Ring to Rule them All”



To Supervise or Not To Supervise?

- Supervised: Adapt “one color at a time” Human-in-the-loop supervised cellular classification – just like IHC DAB protocol
- Sometimes its not possible – amend either segmentation or classification to adapt to the biology in question
- Co-registration with multiple modalities tells you more about the biology underlying your data
- Talk to us – we would love to help! info@pathomics.ai

Contact Us

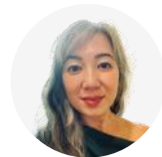


Contact Information



Trevor D. McKee, PhD
CTO

trevor.mckee@pathomics.ai



Trang Vu, PhD
COO

trang.vu@pathomics.ai



Alex Mardon
CCO

alex.mardon@pathomics.ai



<https://pathomics.ai>



info@pathomics.ai

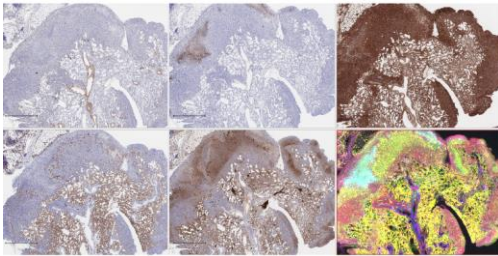
Problem: Research is Drowning in Image Data

- **Pharma & Biotech Research:** Extensive use of pathology data, by non-pathology stakeholders, with high demand for explainable, simple outputs from complex initial datasets
- **Service Providers:** CROs, Antibody Manufacturers, Slide Scanning manufacturers **seek products that solve image analysis problems**, both internally and for customers
- There has been an explosion in both volume and complexity of high quality images and a lack of easy to use tools that get to key answers quickly.

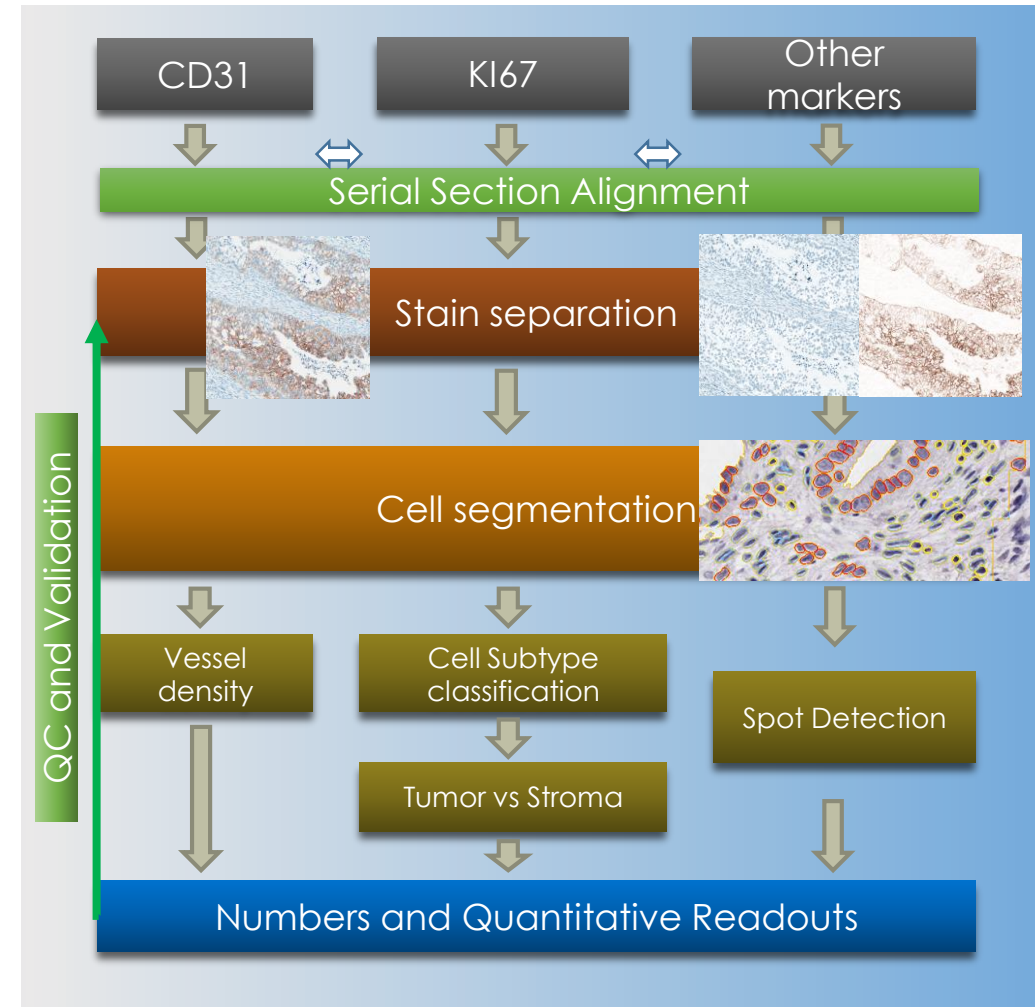
Our Solution: Designing Robust, Reliable Analytical Pipelines

- Quantitative analysis requires several steps, incorporated in an end-to-end pipeline for analysis
- Single color IHC (or other modalities) can be “virtually multiplexed” with serial section alignment

“Virtual”
multiplex
5 markers



- Individual Modules, like: Stain separation, cellular segmentation; Segmentation-free methods; and tissue / cell classification can be combined
- Quantitative readouts are identified based on desired outputs – a “standardized spatial report”
- Critical to success is a robust QC and validation strategy to ensure accuracy – results are only as good as the weakest link in this chain



Goal: Combine the best of open-source digital pathology tools with existing commercial products, to configure optimal pipeline for accurate, useful results