

Spatial analysis of immune cells from multiplex immunofluorescence using R package spatialITIME and iTIME Shiny application

Jordan Creed

Alex Soupir

May 25, 2022

Statistical Methods in Imaging Conference

Nashville, TN





What will we cover in Part 1?

- Short background about why spatial analyses of immune cells from multiplex immunofluorescence (mIF)
- How are the data mIF data collected and processed
- Our implemented approach to combat holes in tissues
- Output from Vectra processed with HALO image analysis software
- Thing to prepare for Part 2:
 - Download github.com/FridleyLab/ASA_SMI_2022 repository zip
 - Go to link > Code > Download ZIP
 - Install R and R Studio if not already – will be using RMD document
 - Packages: **spatialTIME** and **tidyverse**

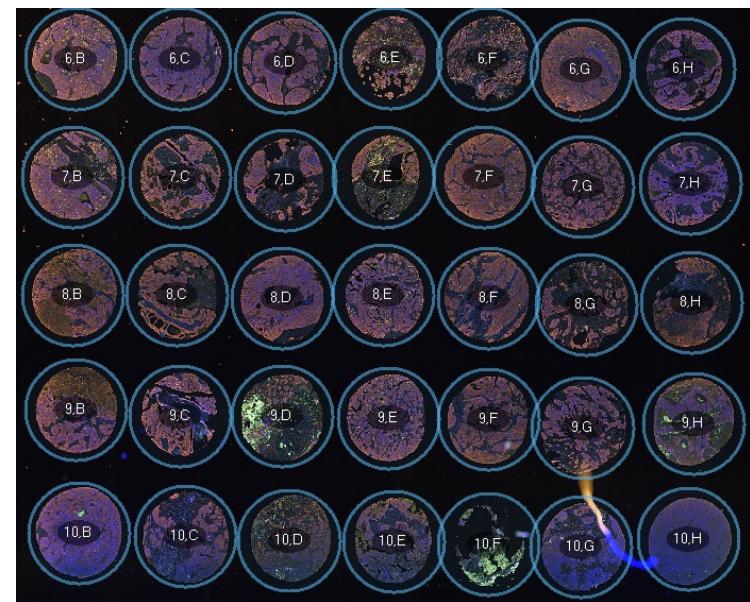


Background and Data Acquisition

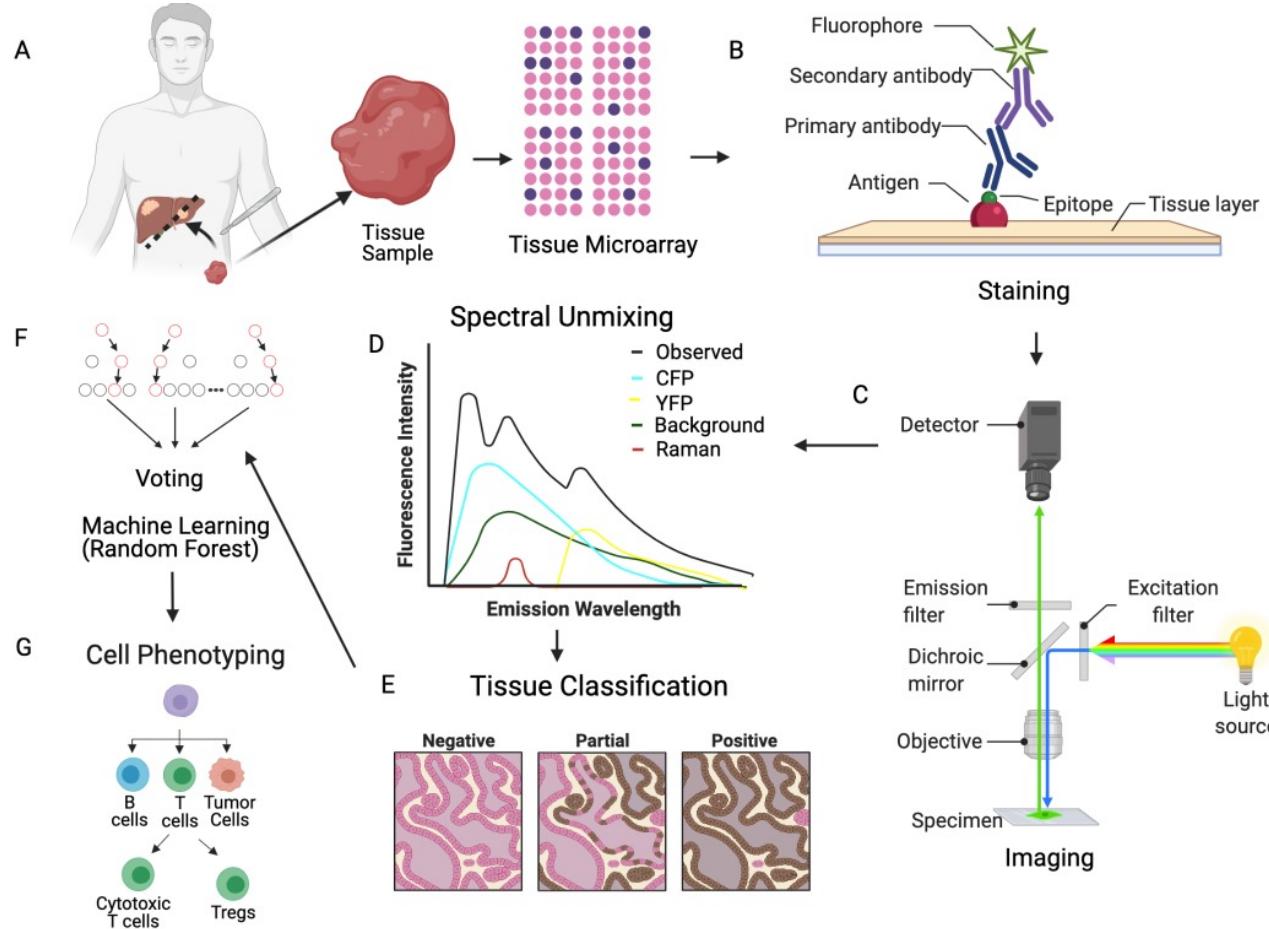
Cancer Immunology



- Immune modulation is considered a hallmark of cancer initiation and progression, and decades of research has offered promising opportunities for therapeutics
- Immunotherapies, agents that activate or act as a substitute for host antitumor immunity, has ushered in a new era of cancer treatment
- Novel approaches are that can be used to identify which cancer patients will benefit from immunotherapy
- Multiplex immunofluorescence (mIF) microscopy (and other technologies) combined with automated image analysis is a novel and increasingly used technique that allows for the assessment and visualization of the tumor immune microenvironment (TIME).



Multiplex-Immunoflorescence (mIF) Studies



- IF studies completed using regions of interest (ROIs) selected from a tissue slide or TMAs.
- Many cancer epidemiology studies have TMAs collected
- High-throughput way to assess the immune landscape in a tumor tissue

Cell types by immune markers



Type of cell	CD markers
stem cells	CD34+, CD31-, CD117
all leukocyte groups	CD45+
Granulocyte	CD45+, CD11b, CD15+, CD24+, CD114+, CD182+ ^[17]
Monocyte	CD4, CD45+, CD14+, CD114+, CD11a, CD11b, CD91+, ^[17] CD16+ ^[18]
T lymphocyte	CD45+, CD3+
T helper cell	CD45+, CD3+, CD4+
T regulatory cell	CD4, CD25, FOXP3 (a transcription factor)
Cytotoxic T cell	CD45+, CD3+, CD8+
B lymphocyte	CD45+, CD19+, CD20+, CD24+, CD38, CD22
Thrombocyte	CD45+, CD61+
Natural killer cell	CD16+, CD56+, CD3-, CD31, CD30, CD38

Data Acquisition

- Vectra (or other system) is used to assess multiple overlapping biomarkers (immune markers) on a single tissue section using multispectral imaging.
- Can do up to 9 markers (colors) per assay (aka “panel”).
 - DAPI marker = cell segmentation
 - Pancytokeratin/PCK = tissue segmentation (tumor / stroma)
- Often HALO or inForm software is then used to process the imaging data and generate locations of “events” and summaries at the tissue level.

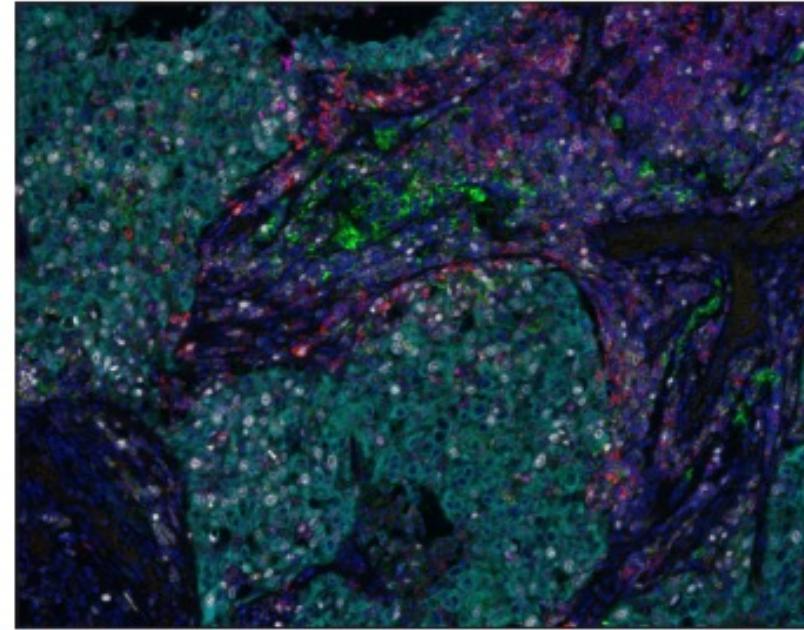


Figure 1. Unmixed 9-color multispectral image of human lung cancer tissue stained against CD20 (Opal 480, red), PD-L1 (Opal 520, green), CD8 (Opal 540, yellow), FoxP3 (Opal 570, orange), CD68 (Opal 620, magenta), PD-1 (Opal 650, pink), Ki67 (Opal 690, white), and PanCK (Opal 780, cyan).

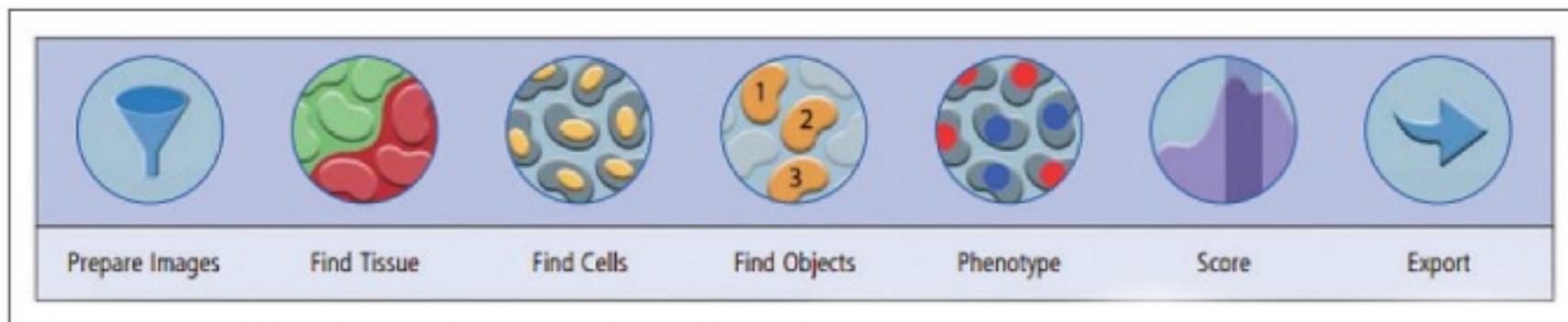
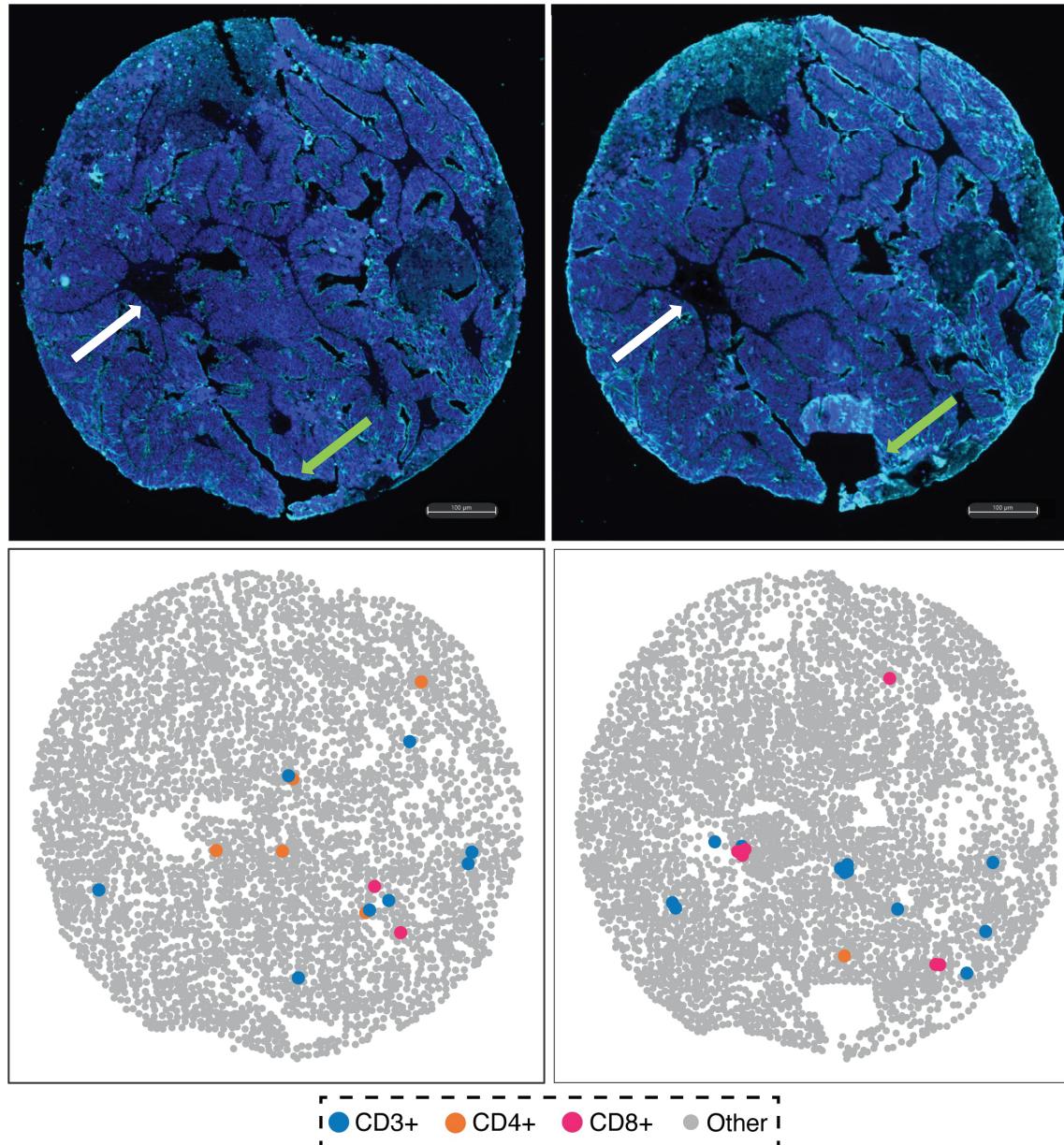


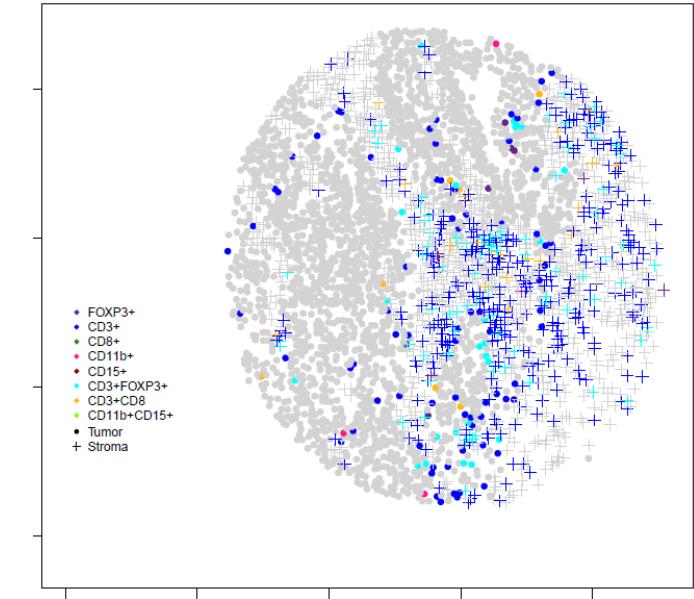
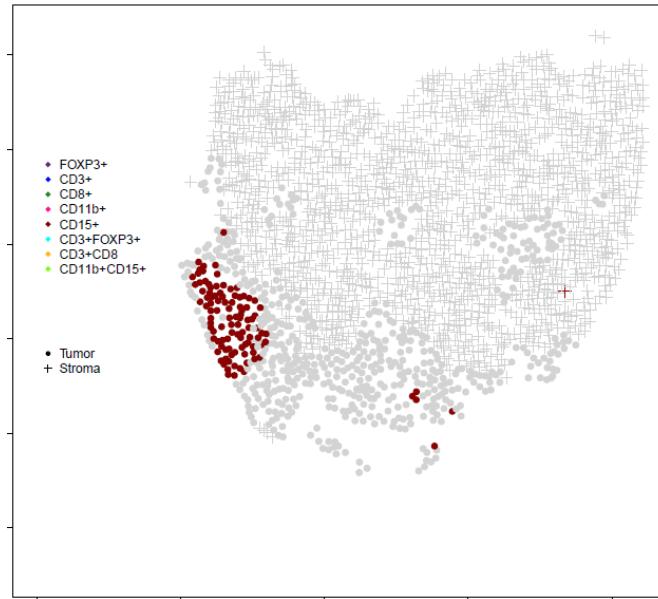
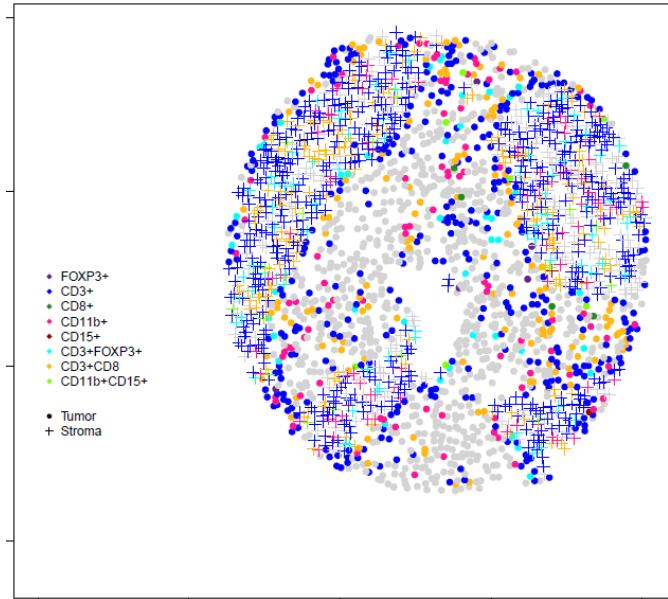
Figure 7. Some of the options available using the Configure Project tool within inForm.

How does the acquired data compare to image?



Tumor vs Stroma

- mIF is summarized for all cells, and then by tumor and stroma
- Different immune marker levels by tumor/stroma
 - Stratified analysis by cell type



HALO output from Vectra Image



	image.tag	XMin	XMax	YMin	YMax	CD3..FOXP3.	CD3..CD8.	CD11b..CD15.	Nuclei..DAPI..Positive	
1	Peres_P1_110188_A10_[40989,9689].tif	877	898	793	816	0	0	0	0	0
2	Peres_P1_110188_A10_[40989,9689].tif	793	815	794	811	0	0	0	0	0
						Nuclei..DAPI..Positive.Nucleus	Nuclei..Nucleus.Intensity	Nuclei..DAPI..Positive.Cytoplasm		
1						0	9.950563		0	
2						0	7.224355		0	
						Nuclei..DAPI..Cytoplasm.Intensity	FOXP3..Opal.540..Positive	FOXP3..opal.540..Positive.Nucleus		
1						7.426177	0	0	0	
2						5.719327	0	0	0	
						FOXP3..Opal.540..Nucleus.Intensity	FOXP3..Opal.540..Positive.Cytoplasm	FOXP3..opal.540..Cytoplasm.Intensity		
1						0.161433	0	0	0.120814	
2						0.381655	0	0	0.162026	
						CD3..opal.650..Positive	CD3..Opal.650..Positive.Nucleus	CD3..opal.650..Nucleus.Intensity	CD3..opal.650..Positive.Cytoplasm	
1						0	0	0.227083	0	
2						0	0	0.517002	0	
						CD3..opal.650..Cytoplasm.Intensity	CD8..opal.570..Positive	CD8..opal.570..Positive.Nucleus		
1						0.273951	0	0	0	
2						0.694598	0	0	0	
						CD8..opal.570..Nucleus.Intensity	CD8..Opal.570..Positive.Cytoplasm	CD8..opal.570..Cytoplasm.Intensity		
1						0.002307	0	0	0.007542	
2						0.092853	0	0	0.333973	
						CD11b..opal.620..Positive	CD11b..Opal.620..Positive.Nucleus	CD11b..opal.620..Nucleus.Intensity		
1						0	0	0.000000		
2						0	0	0.010526		
						CD11b..opal.620..Positive.Cytoplasm	CD11b..opal.620..Cytoplasm.Intensity	CD15..opal.520..Positive		
1						0	0	0	0	
2						0	0	0	0	
						CD15..opal.520..Positive.Nucleus	CD15..Opal.520..Nucleus.Intensity	CD15..opal.520..Positive.Cytoplasm		
1						0	0.089711	0	0	
2						0	0.088885		0	
						CD15..opal.520..Cytoplasm.Intensity	Cell.Area.. μm^2	Cytoplasm.Area.. μm^2	Nucleus.Area.. μm^2	Nucleus.Perimeter.. μm
1						0.123157	97.82485	49.90055	47.92429	32.80356
2						0.083942	59.78185	31.12609	28.65576	25.84523
						Nucleus.Roundness	Classifier.Label			
1						0.721109	Tumor			
2						0.602914	Tumor			



spatialTIME Approaches

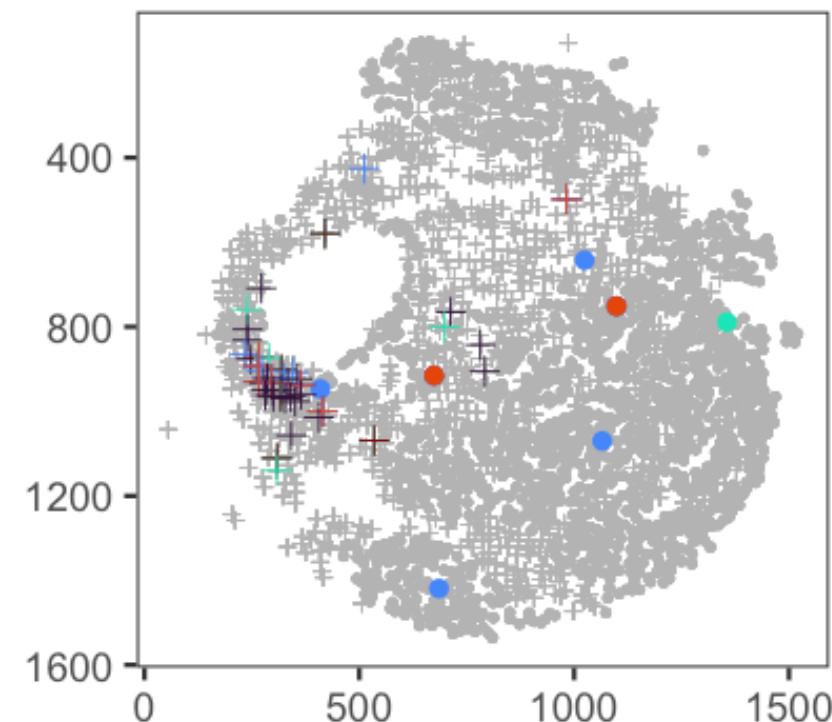
What can we do with spatialTIME?



Things like:

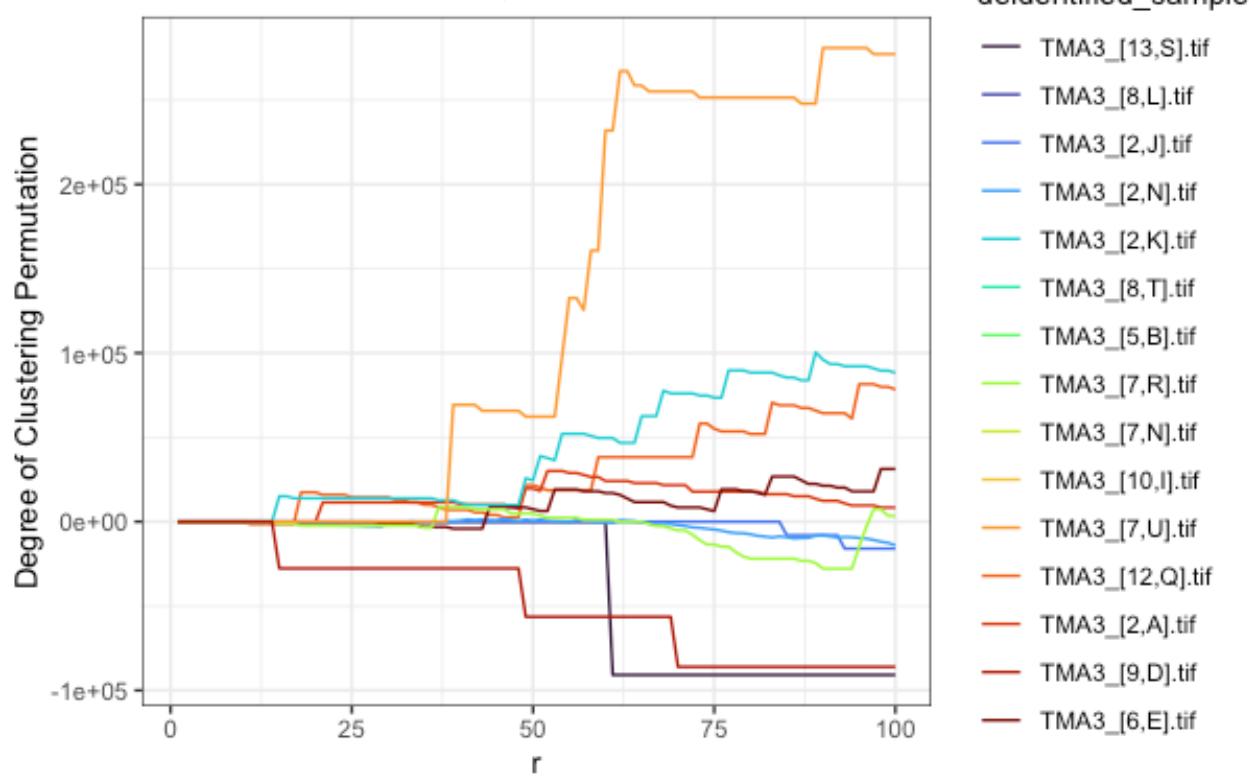
Make representative plots of samples

ID: TMA3_[7,U].tif



Check for colocalization of cell types

Anchor = CD3+ CD8+; Counted = CD3+ FOXP3+





What is Ripley's K count statistic?

- $\hat{K}(r) = \frac{|W|}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n 1\{d_{ij} \leq r\} e_{ij}$

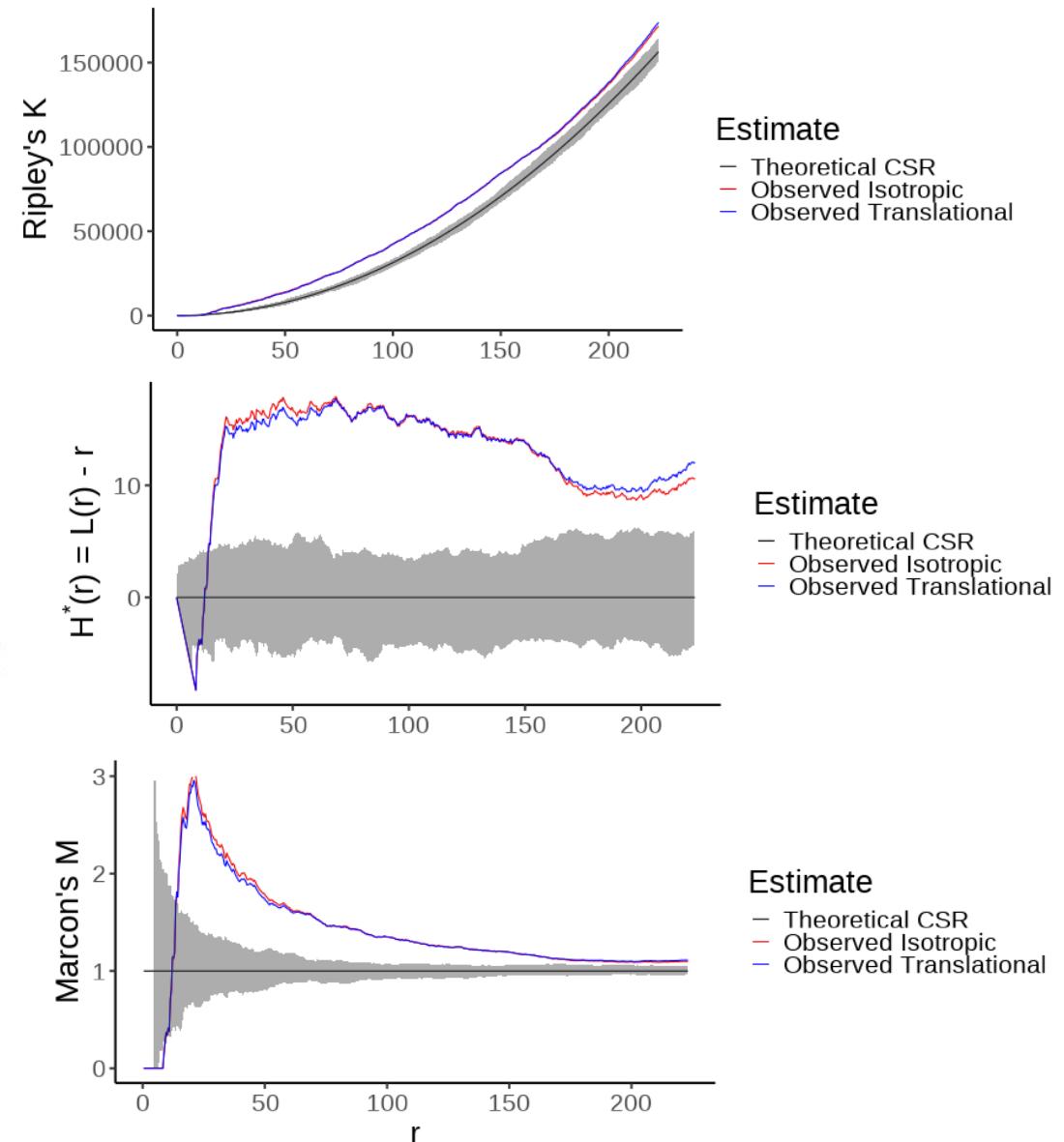
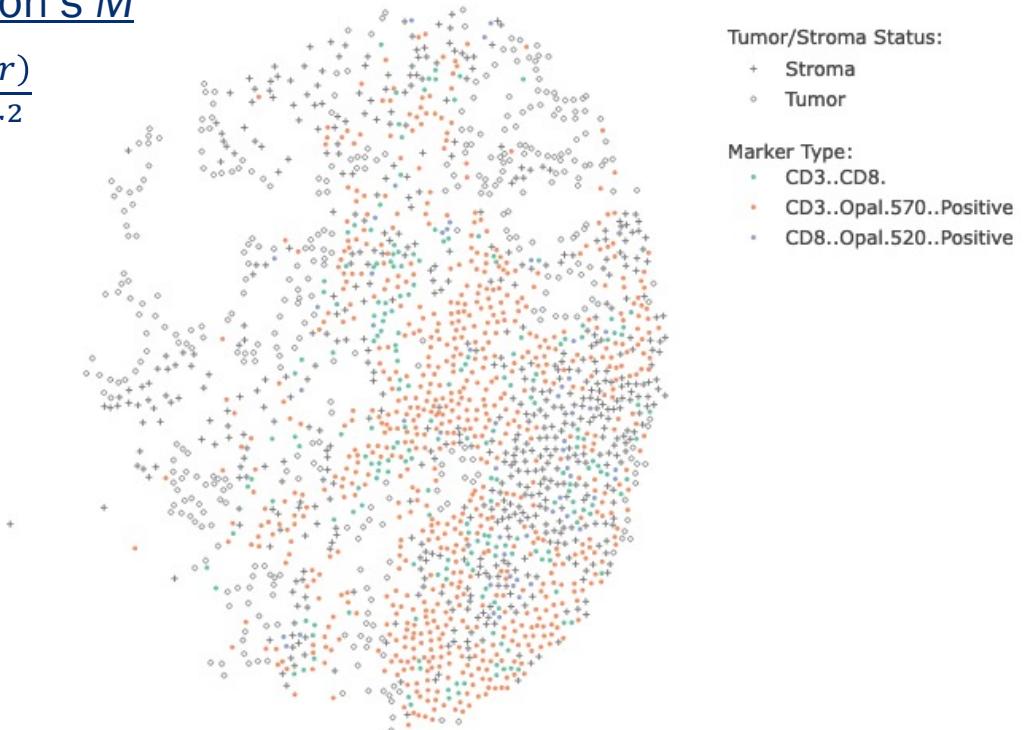
- Transformations

Besag's L

- $L(r) = \sqrt{\frac{K(r)}{\pi}}$

Marcon's M

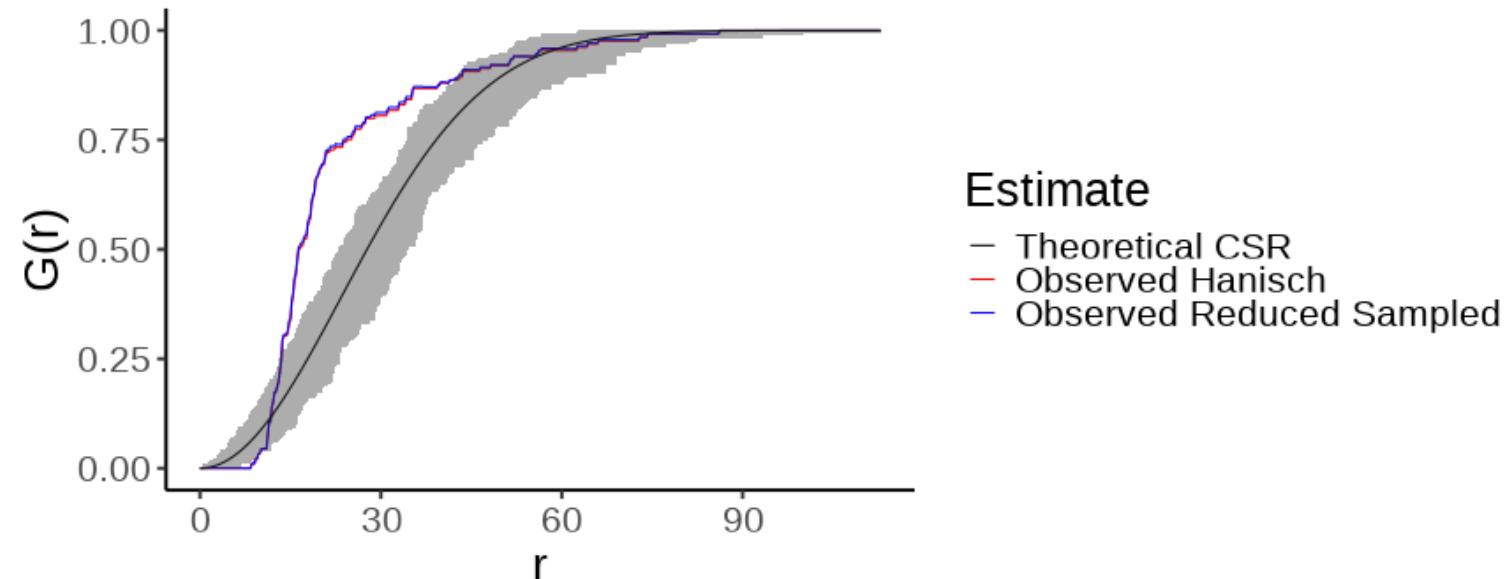
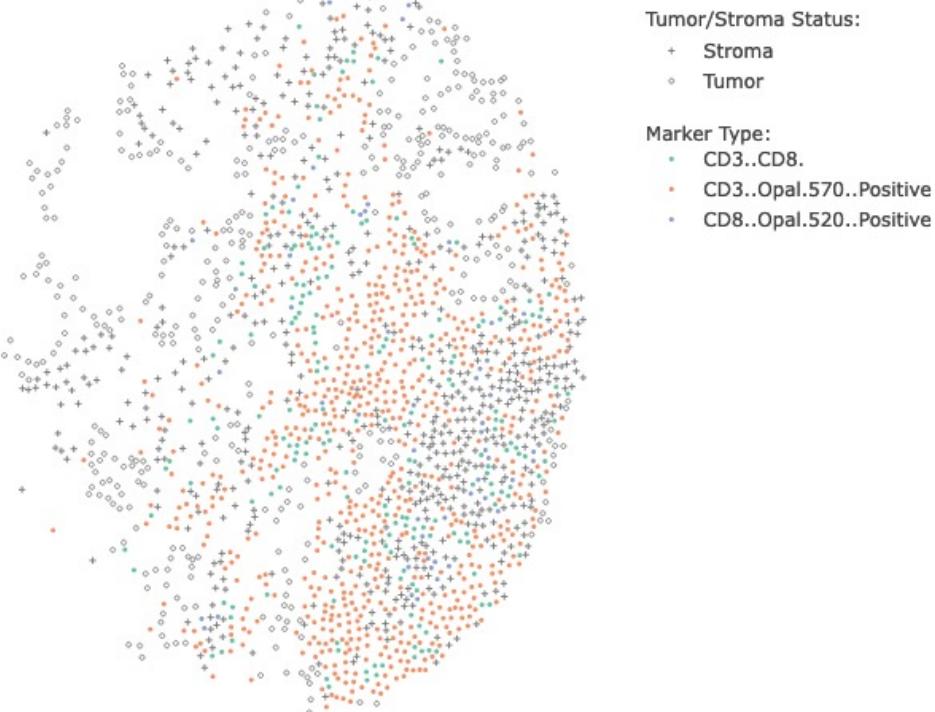
- $\frac{K(r)}{\pi r^2}$





What is Nearest Neighbor G function?

- Nearest Neighbor Distance Distribution function
- $d_i = \min_{j \neq i} d_{ij}$ or $d_i = d(x_i, X \setminus x_i)$ finds the nearest neighbor
- $G(r) = \mathbb{P}\{d(u, X \setminus u) \leq r | X \text{ has a point } u\}$



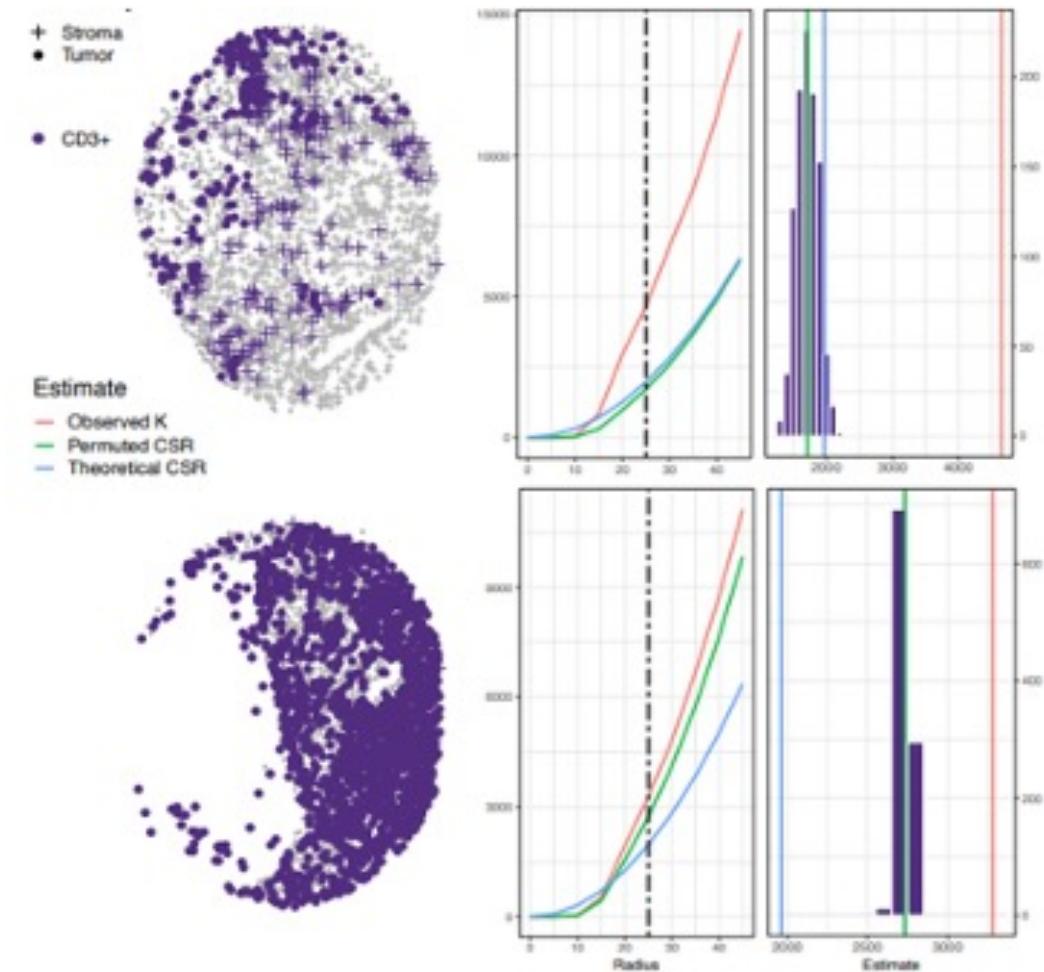
Estimate

- Theoretical CSR
- Observed Hanisch
- Observed Reduced Sampled



Why the Need for Permutations to Assess CSR?

- Assumption of continuous spatial point process
- Properly compare the observed measure between different cores
- Ability to determine whether the degree of clustering, that is, subtracting the permuted CSR from the observed metric, is significantly different from CSR
- Accounts for large holes that may appear, and allows for compartment specific analysis (tumor/stroma)





spatialTIME package vs iTIME interface

spatialTIME

Benefits

- CRAN R package
- Scripting interface
- Compute Ripley's K or nearest neighbor G at any *radii*
- Process many HALO output files at once
- Empirical Ripley's K , Degree of Clustering (Theoretical and Permuted)

iTIME

Benefits

- Easy interface to visualize abundance data
- Ability to see how different cut-offs split data
- Beta-binomial modeling of **abundances** with covariates
- Single spatial visualization with Plotly
- Empirical Ripley's K , not degree of clustering



R package and iTIME shiny application supplement each other



Part 2 Overview

- Remember to have R and R Studio installed
- Download the data and code RMD:
 - github.com/FridleyLab/ASA_SMI_2022
- Samples from a real mIF (Vectra) study for patients with prostate cancer
 - Samples have been subset to include 15 samples from 14 individuals
 - Patient ID and status are masked from actual



Thank you!