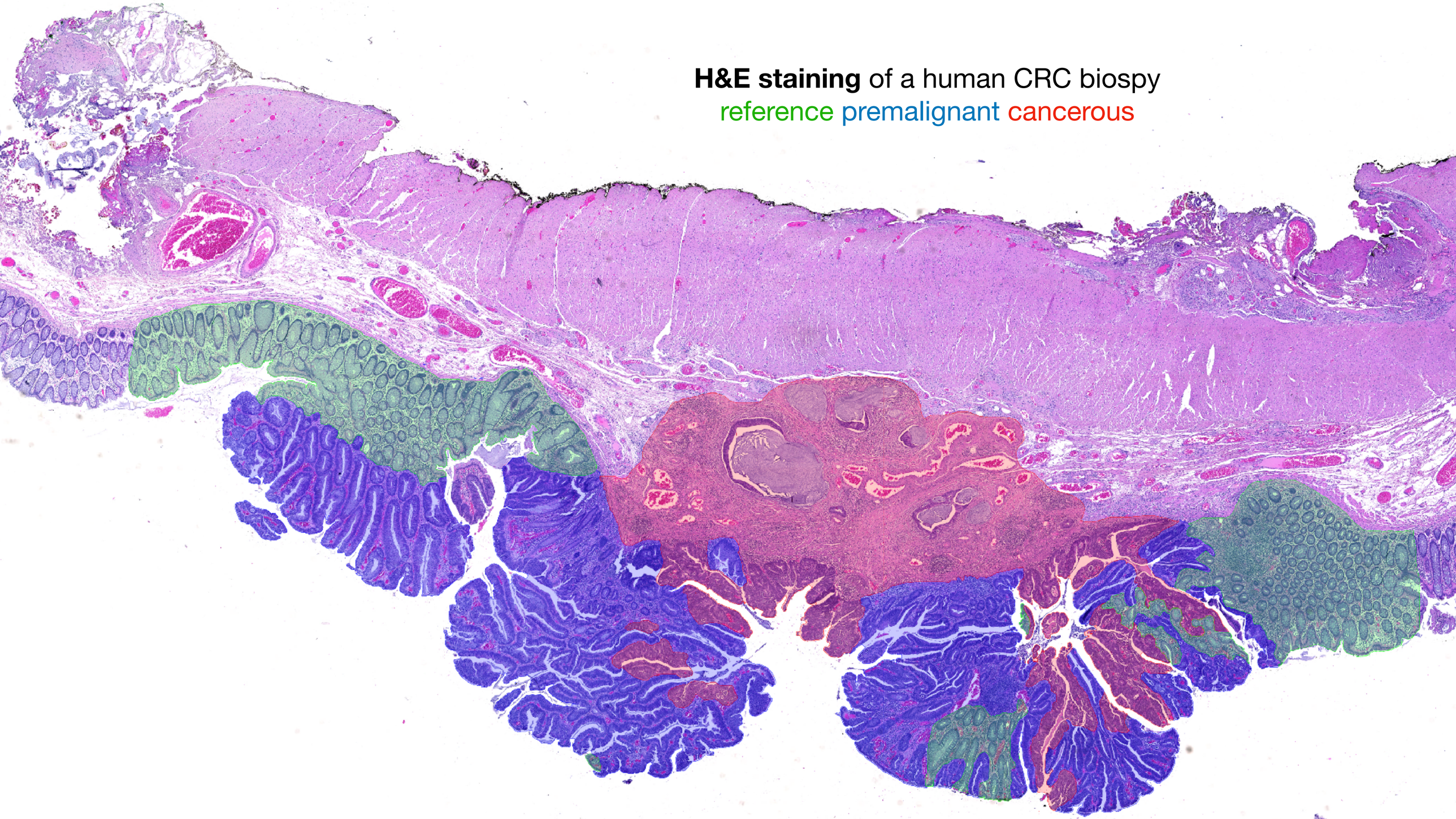
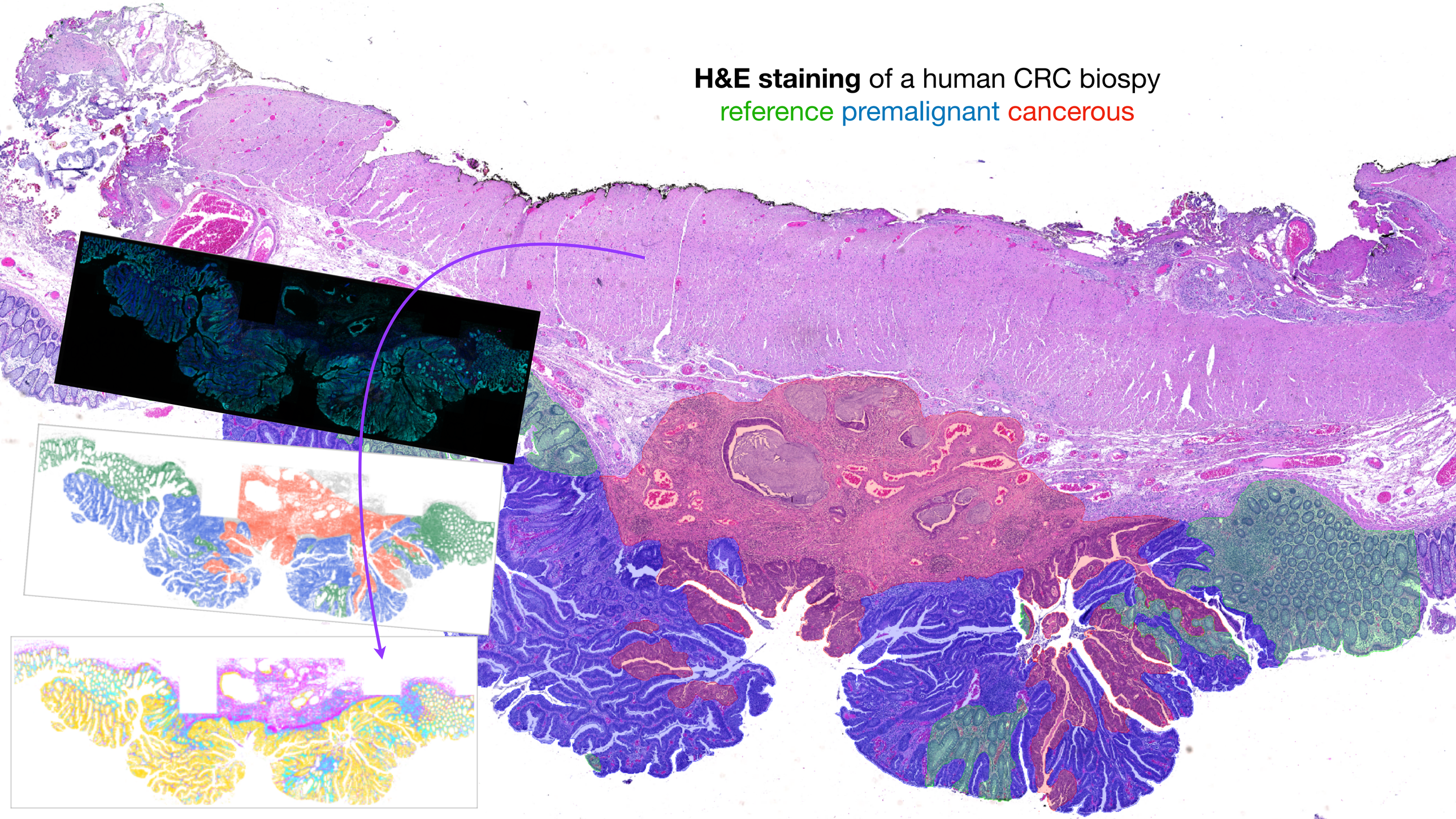
(pre)preprocessing of imaging-based spatial transcriptomics data

Swiss Institute of Bioinformatics
Spatial Omics Data Analysis

Helena L. Crowell — January 21st, 2025
in Lausanne, Switzerland

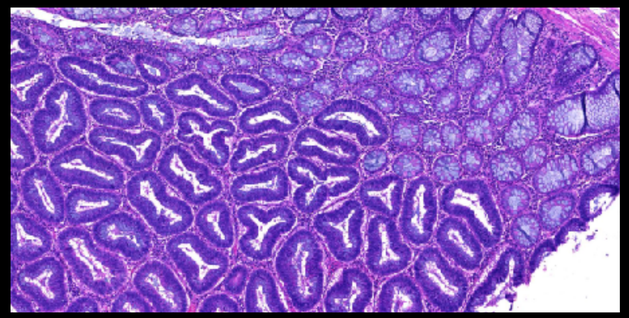

H&E staining of a human CRC biospy
reference premalignant cancerous


H&E staining of a human CRC biospy
reference premalignant cancerous

the path **from tissue to pathology** ain’t easy

(sometimes)
stats.-/
AI-based

pathology ← H&E staining ← tissue

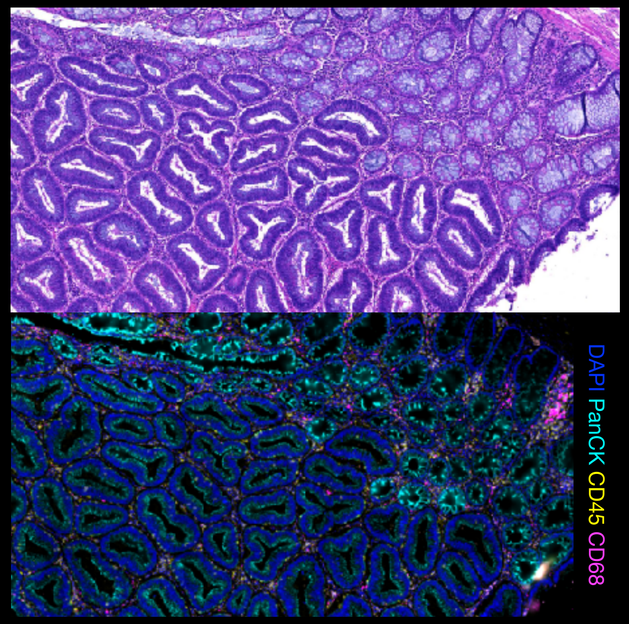


the path **from tissue to pathology** ain’t easy

(sometimes)
stats.-/
AI-based

pathology ← H&E staining ← tissue → IF staining

imaging

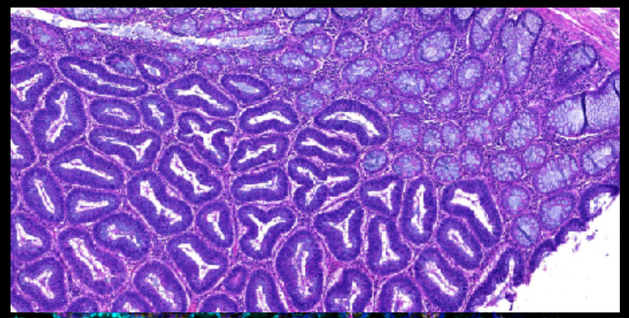


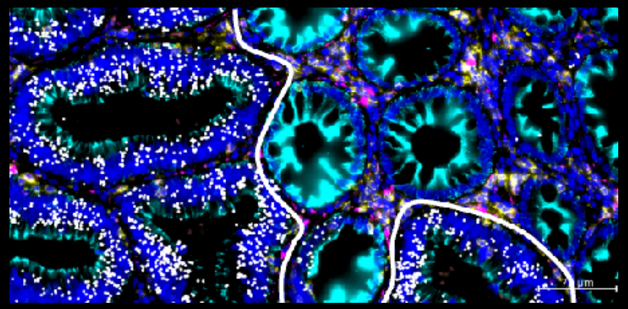
the path **from tissue to pathology** ain’t easy

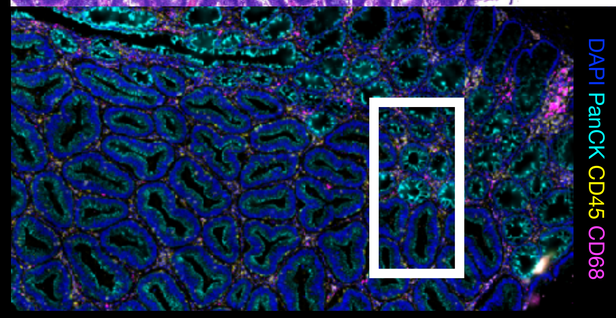
(sometimes)
stats.-/
AI-based

pathology ← H&E staining ← tissue → high-plex FISH → spot-calling → transcript molecules

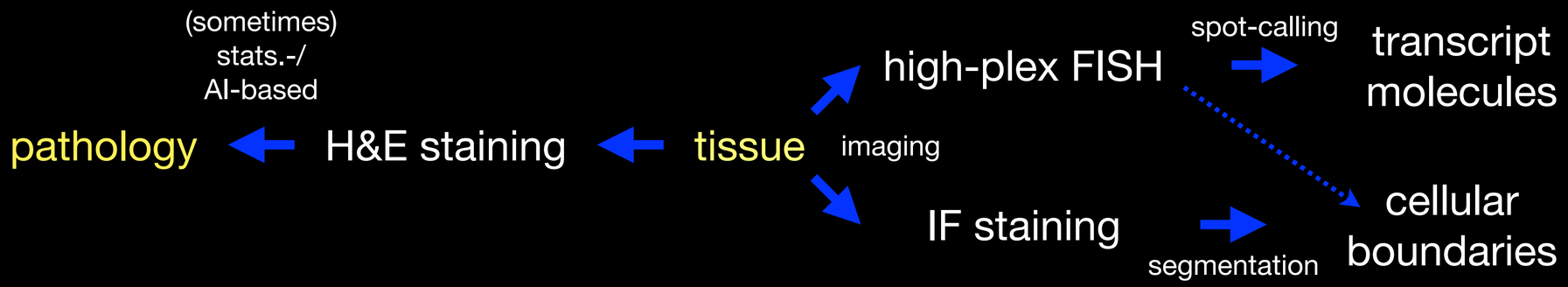
tissue → imaging → IF staining

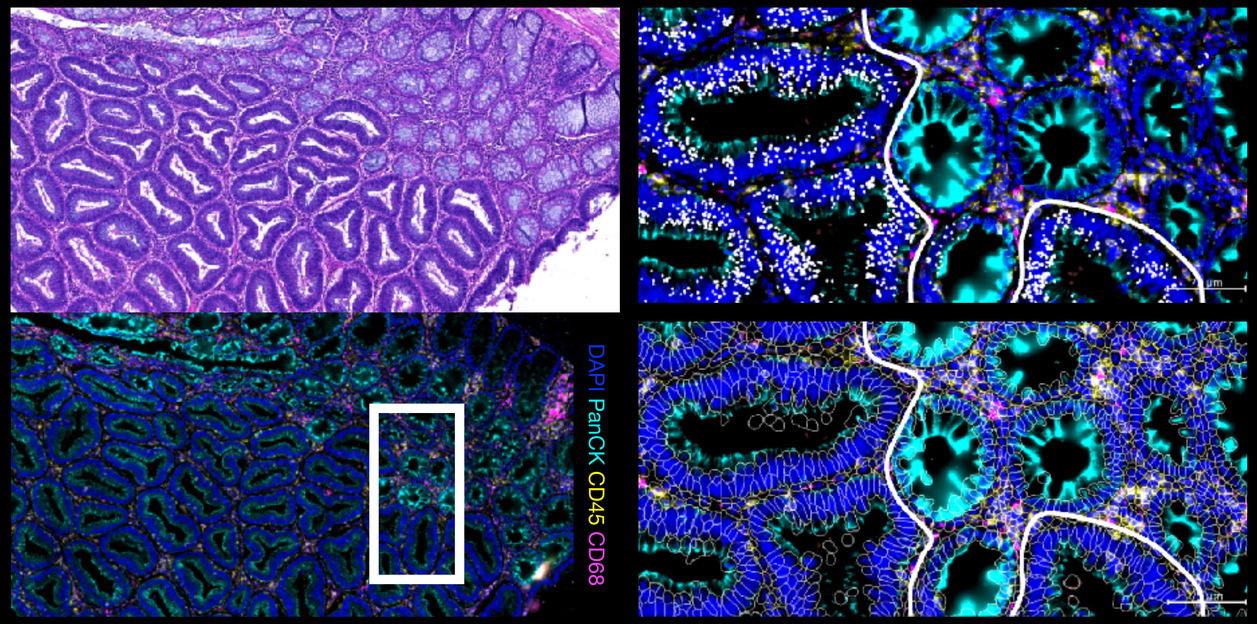




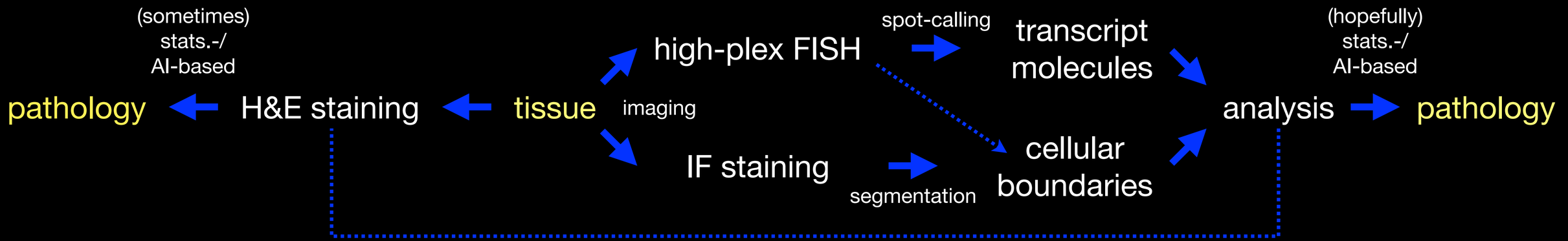


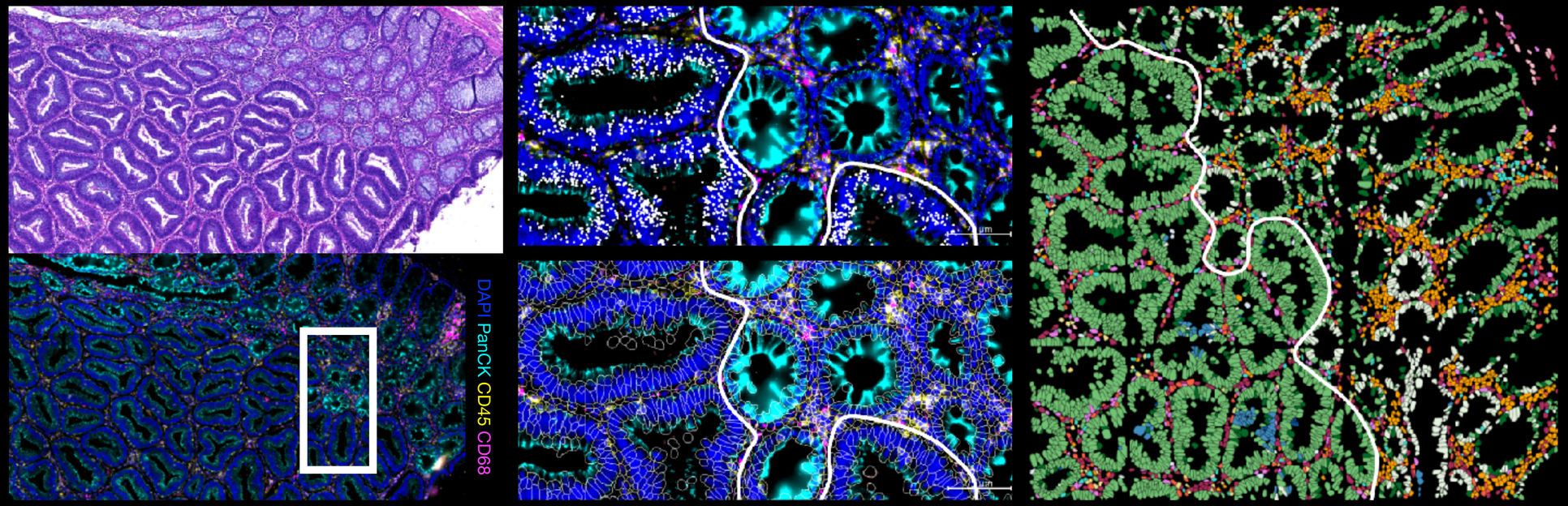
the path **from tissue to pathology** ain't easy



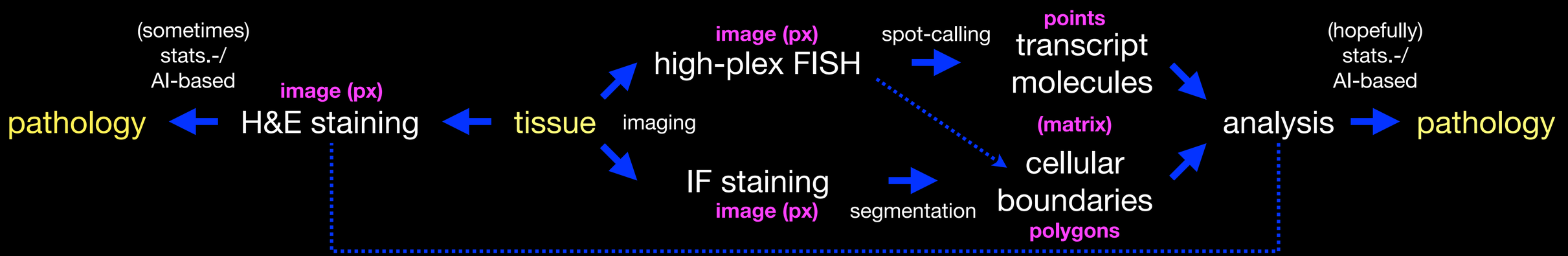


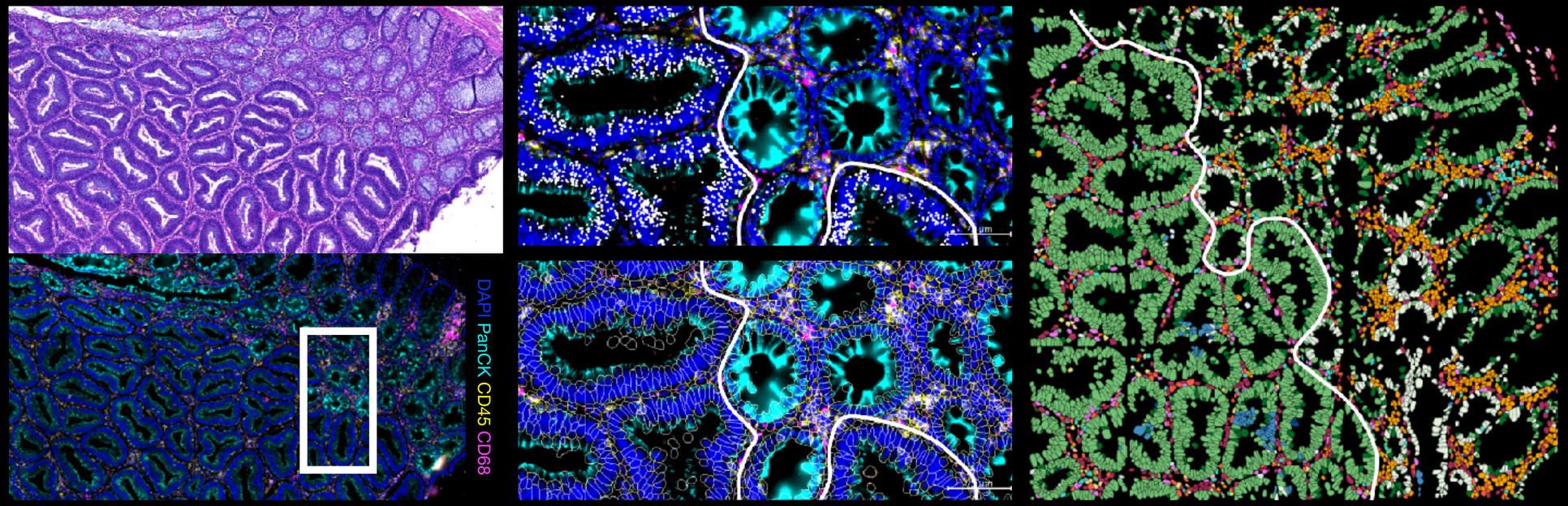
the path **from tissue to pathology** ain't easy





the path **from tissue to pathology** ain't easy





quality control for img-ST is ways from (sc)RNA-seq

high-throughput RNA sequencing

fragmentation, reverse transcription, mapping

↓

# of raw counts per transcript varies with transcript length, GC content, sequencing depth

↓

normalization strategies aim to minimize these effects & *“there’s awareness that misinterpretation of results where biological & technical effects are correlated”*

Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

quality control for img-ST is ways from (sc)RNA-seq

high-throughput RNA sequencing

fragmentation, reverse transcription, mapping

↓

# of raw counts per transcript varies with transcript length, GC content, sequencing depth

↓

normalization strategies aim to minimize these effects & *“there’s awareness that misinterpretation of results where biological & technical effects are correlated”*

imaging-based spatial transcriptomics

tissue preparation, chemistry, imaging

↓

tissue damage/detachment, image/transcript loss, varying detection across space & experiments

Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

quality control for img-ST is ways from (sc)RNA-seq

high-throughput RNA sequencing

fragmentation, reverse transcription, mapping

↓

# of raw counts per transcript varies with transcript length, GC content, sequencing depth

↓

normalization strategies aim to minimize these effects & *“there’s awareness that misinterpretation of results where biological & technical effects are correlated”*

imaging-based spatial transcriptomics

tissue preparation, chemistry, imaging

↓

tissue damage/detachment, image/transcript loss, varying detection across space & experiments

↓

“sources of these […] are known [but] it’s often unlcear how often errors occur, how to best detect & describe [them] & how [they] impact downstream analyses […]”

Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

quality control for img-ST is ways from (sc)RNA-seq

DISCLAIMER: We don't really know what's best (yet) —
I'll try to summarize some recent ideas, and personal pains.

high-throughput RNA sequencing

fragmentation, reverse transcription, mapping

↓

# of raw counts per transcript varies with
transcript length, GC content, sequencing depth

↓

normalization strategies aim to minimize
these effects & *"there's awareness that
misinterpretation of results where biological
& technical effects are correlated"*

imaging-based spatial transcriptomics

tissue preparation, chemistry, imaging

↓

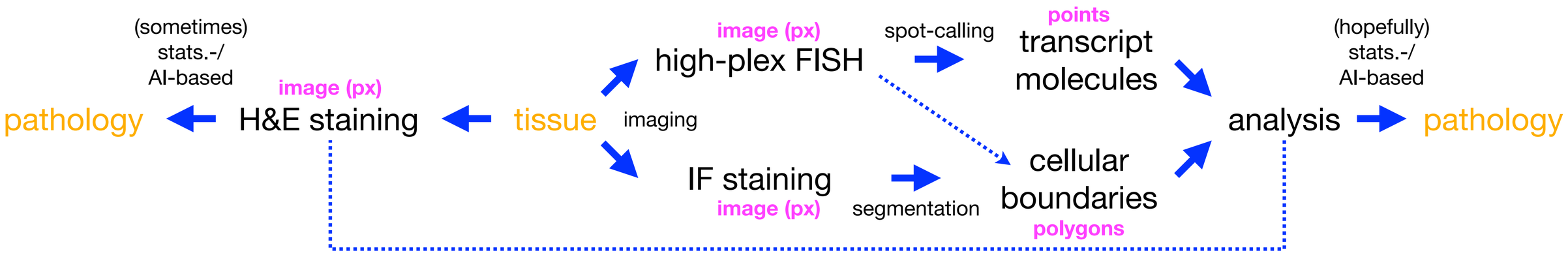
tissue damage/detachment, image/transcript loss,
varying detection across space & experiments

↓

*"sources of these […] are known [but]
it's often unlcear how often errors occur,
how to best detect & describe [them] &
how [they] impact downstream analyses […]"*

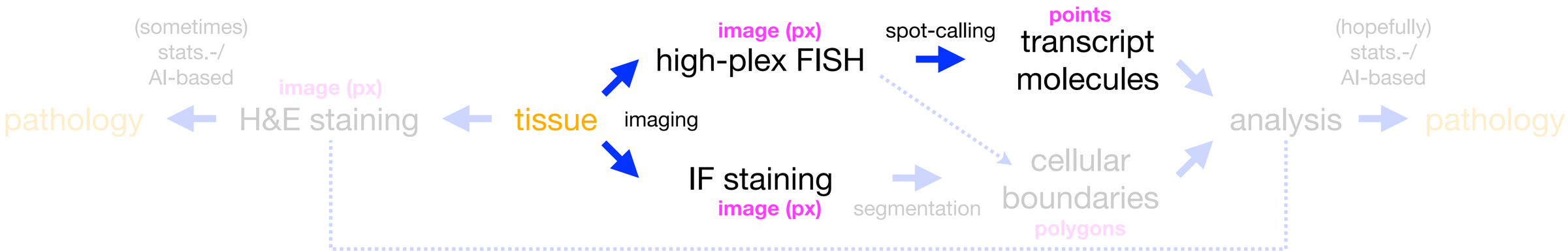
Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a **pixel classifier** & quality metrics



Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

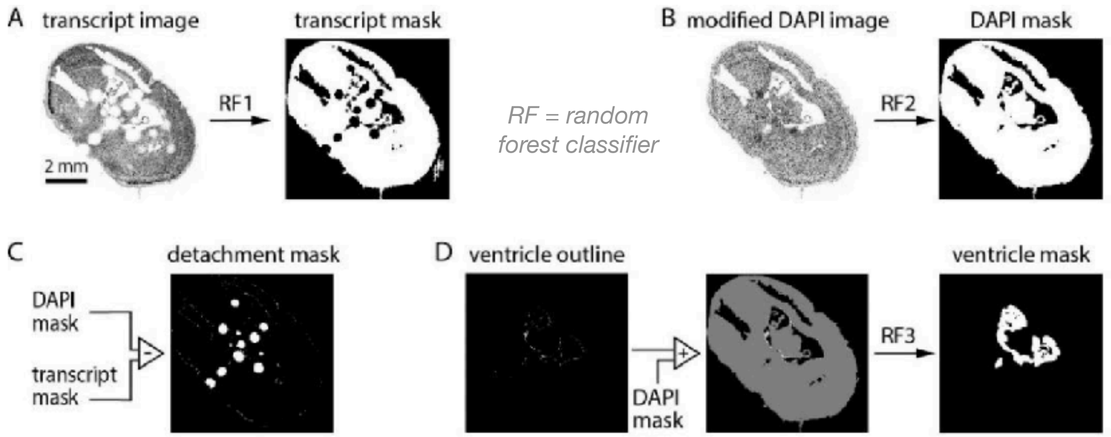
MerQuaCo proposes a **pixel classifier** & quality metrics



Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a **pixel classifier** & quality metrics

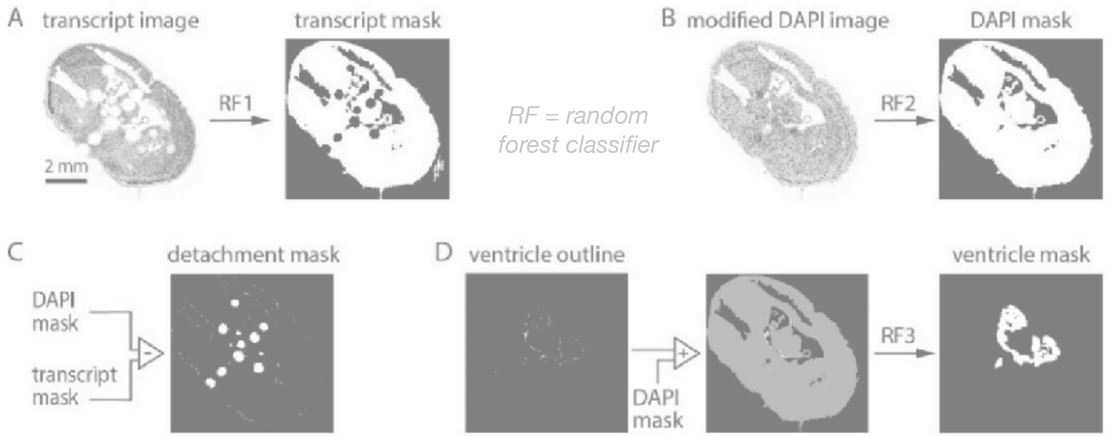
- input: transcript locations + DAPI staining
- series of **binary masks** (& combinations thereof), **trained on** few manually **annotated sections**



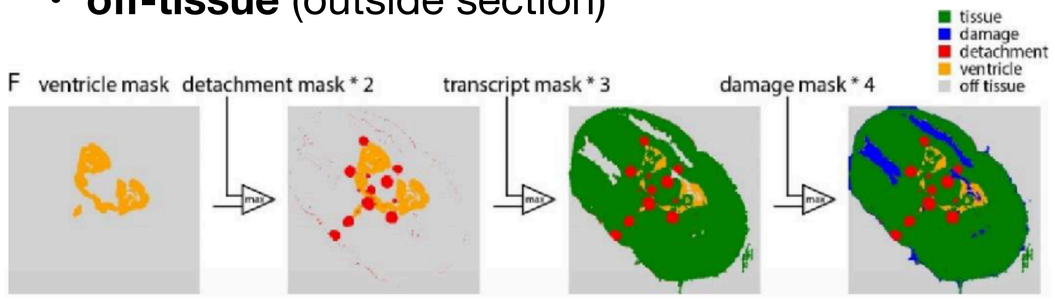
Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a **pixel classifier** & quality metrics

- input: transcript locations + DAPI staining
- series of **binary masks** (& combinations thereof), **trained on** few manually **annotated sections**



- each location is classified into one of **five categories**
  - **tissue** within image volumen
  - **detachment** (tissue not imaged)
  - **ventricle** (no tissue but no loss)
  - **damage** (no tissue due to loss)
  - **off-tissue** (outside section)



Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a pixel classifier & **quality metrics**

perfusion rate

- log files can reveal inconsitencies (e.g., blockage) of volume per time during solution exchange

data loss

- iterative comparison of transcript counts between neighboring fields of view (FOVs)

detection efficiency

- across section: periodicity,
  through section: p6/p0 ratio

transcript density

- should vary across section, but little between sections of comparable quality

Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a pixel classifier & **quality metrics**

perfusion rate

- log files can reveal inconsitencies (e.g., blockage) of volume per time during solution exchange

data loss

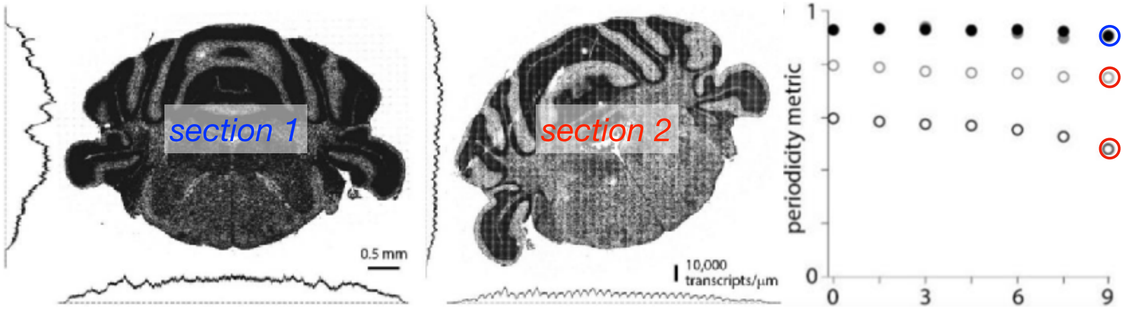
- iterative comparison of transcript counts between neighboring fields of view (FOVs)

detection efficiency

- across section: periodicity,
  through section: p6/p0 ratio

transcript density

- should vary across section, but little between sections of comparable quality



Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

MerQuaCo proposes a pixel classifier & **quality metrics**

perfusion rate

- log files can reveal inconsitencies (e.g., blockage) of volume per time during solution exchange

data loss

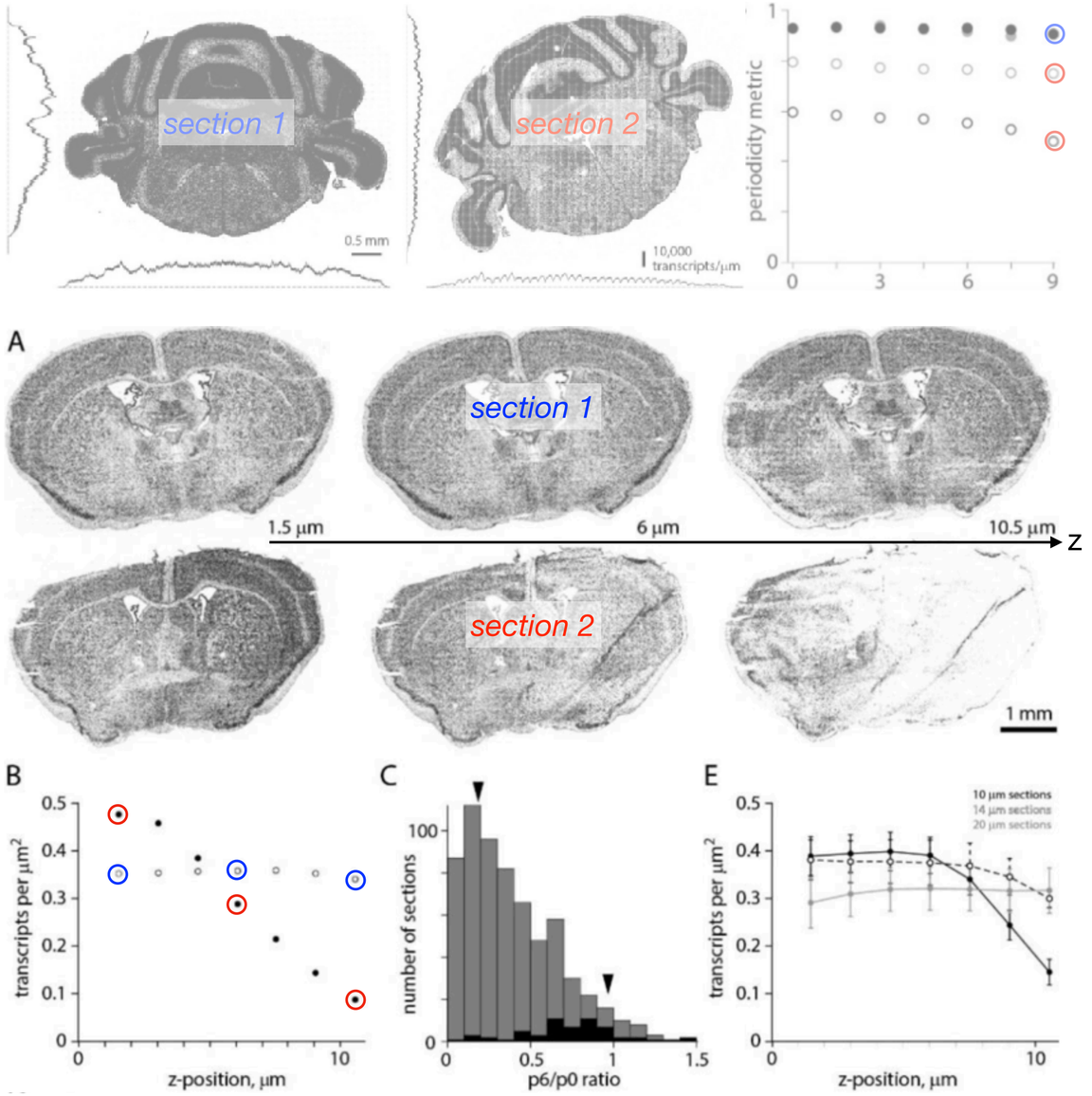
- iterative comparison of transcript counts between neighboring fields of view (FOVs)

detection efficiency

- across section: periodicity,
  through section: p6/p0 ratio

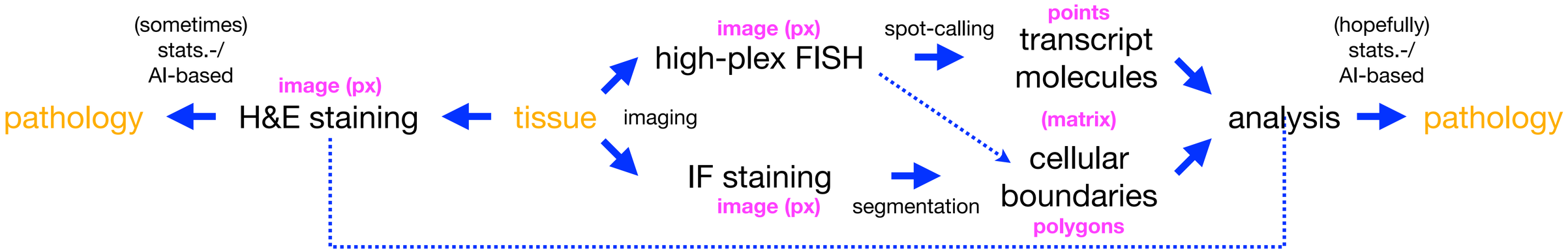
transcript density

- should vary across section, but little between sections of comparable quality

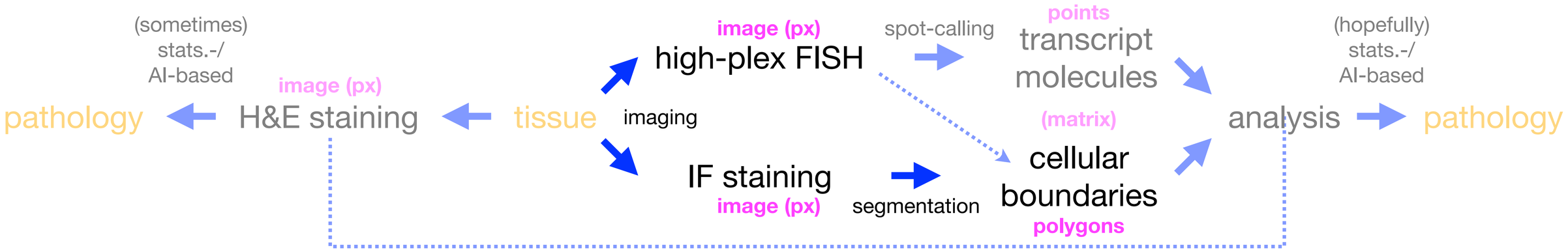


Martin *et al.* (2024), *bioRxiv 2024.12.04.626766*

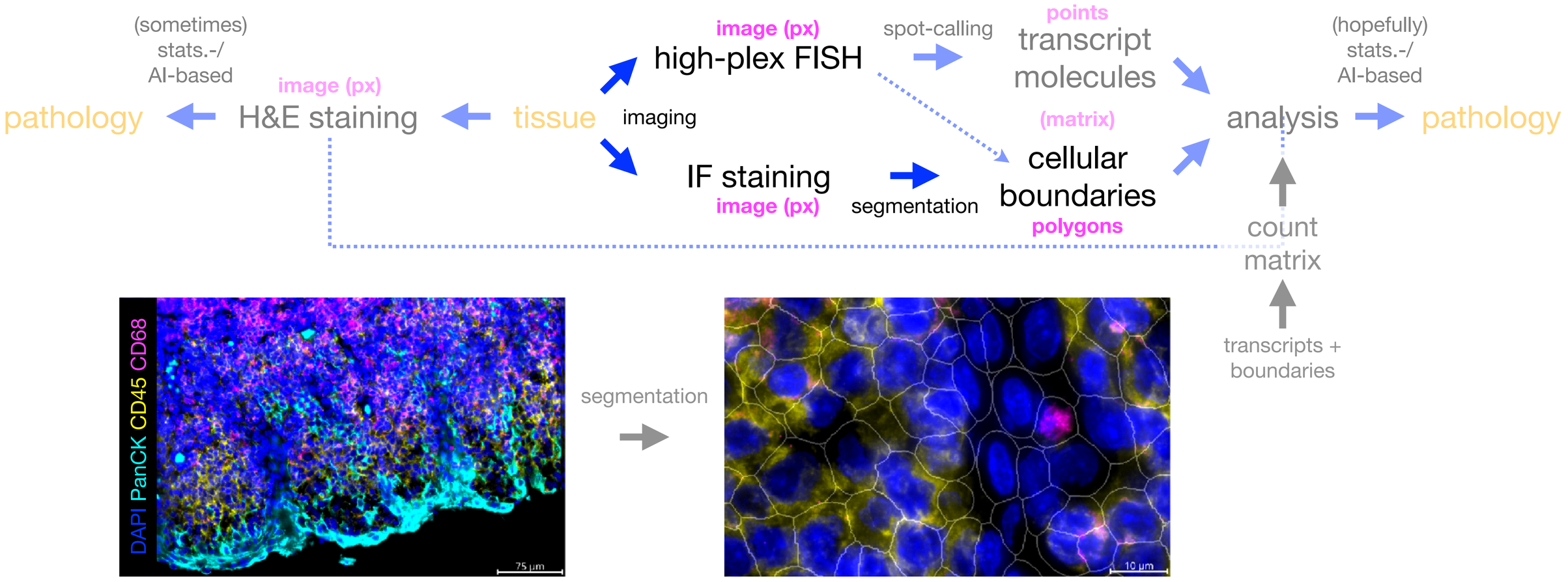
for (most) analyses, we need to go **from images to counts**



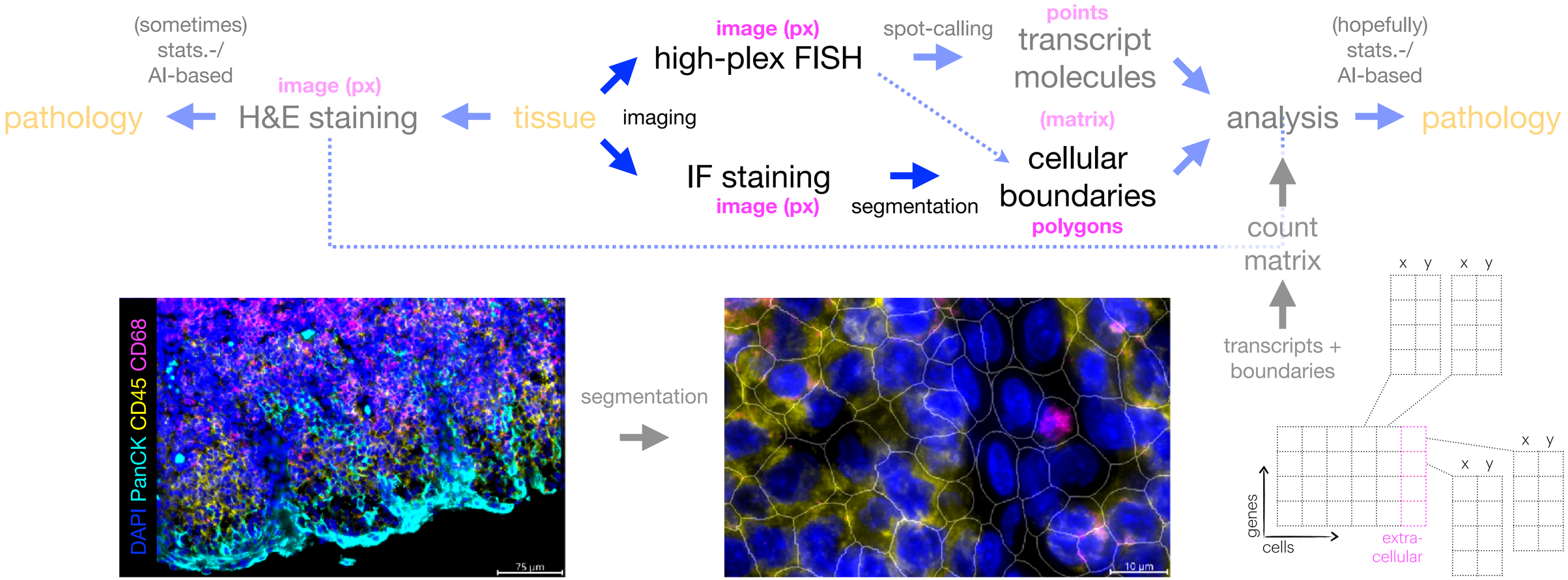
for (most) analyses, we need to go **from images to counts**



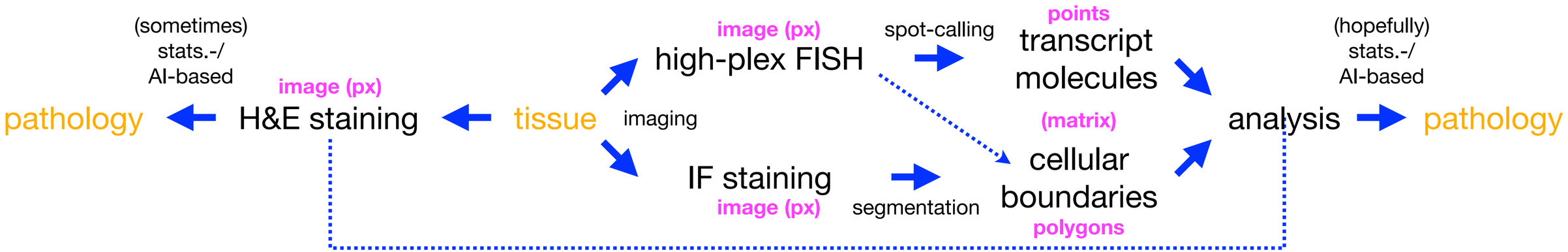
for (most) analyses, we need to go **from images to counts**



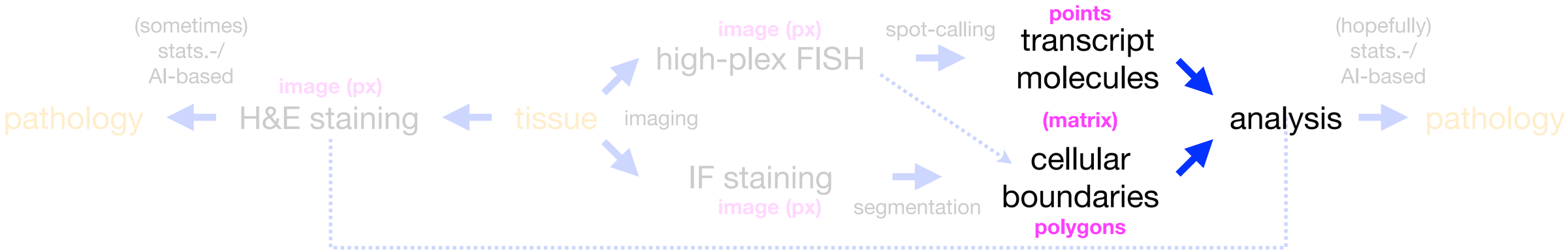
for (most) analyses, we need to go **from images to counts**



beware standard QC metrics — counts relate area AND biology

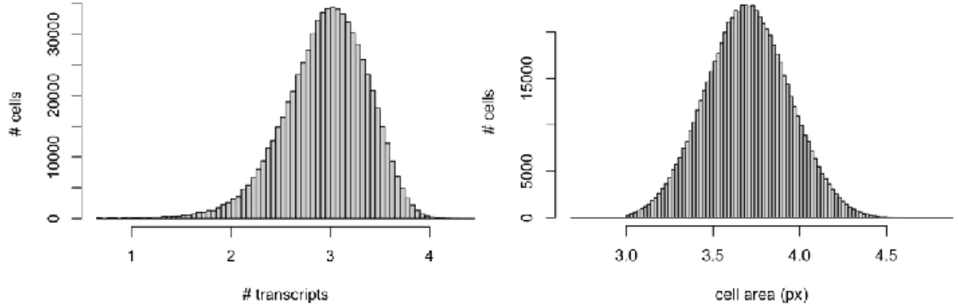


beware standard QC metrics — counts relate area AND biology



beware standard QC metrics — counts relate area AND biology

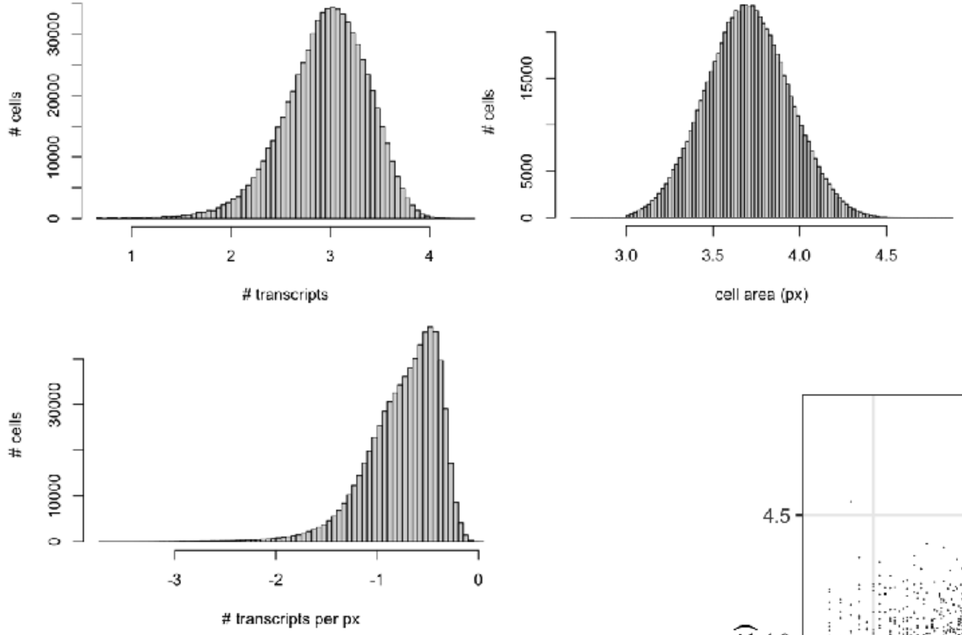
all axes $\log_{10}$-transformed (…what we'd usually filter on)



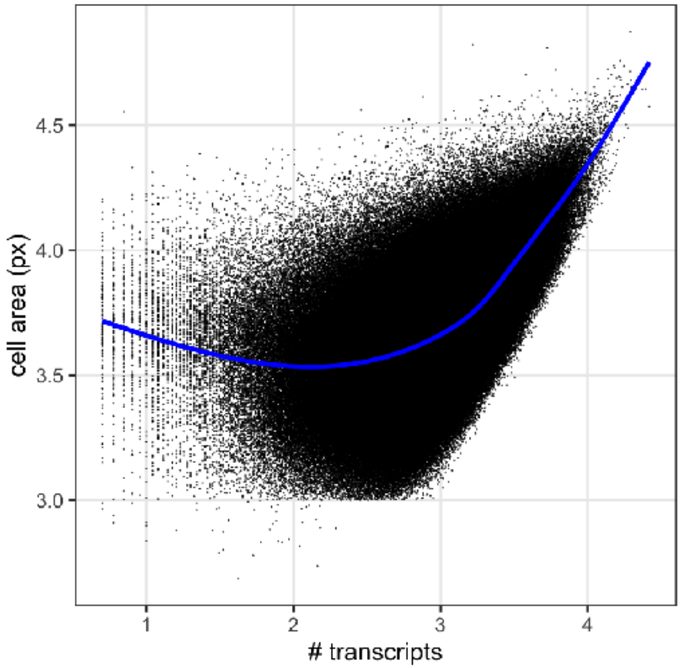
- filtering based on standard scRNA-seq QC metrics may bias again smaller & transcriptionally less complex cells

beware standard QC metrics — counts relate area AND biology

all axes $\log_{10}$-transformed (…what we'd usually filter on)

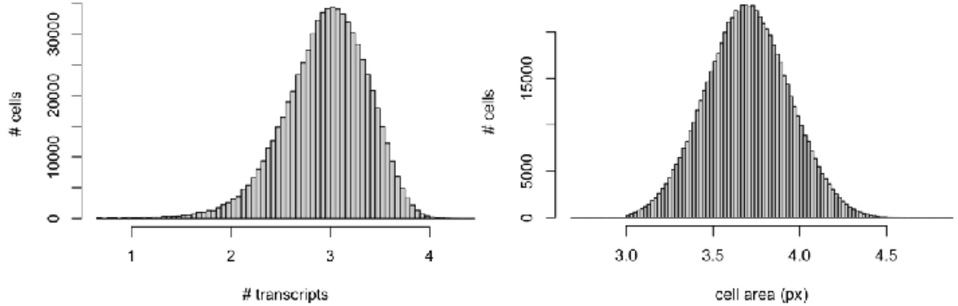


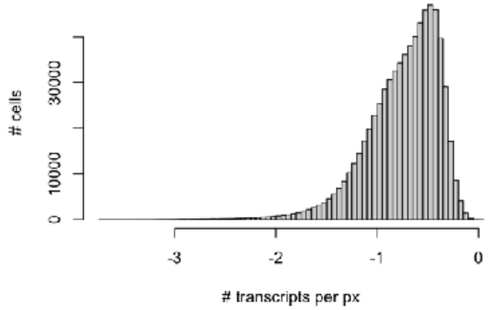
- filtering based on standard scRNA-seq QC metrics may bias again smaller & transcriptionally less complex cells



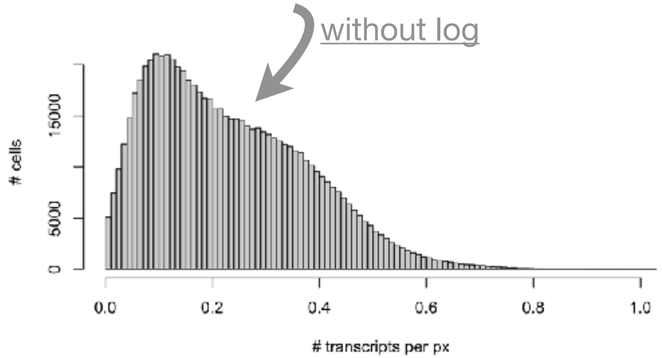
beware standard QC metrics — counts relate area AND biology

all axes $\log_{10}$-transformed (…what we'd usually filter on)

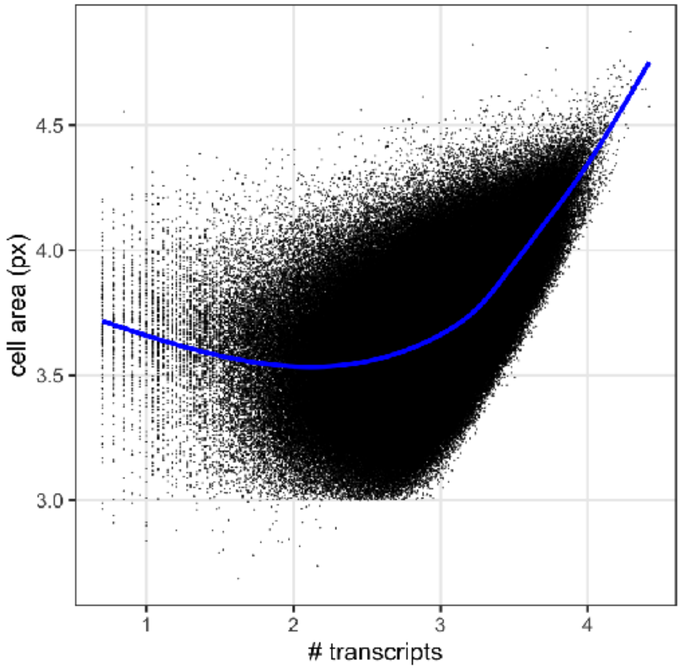




without log

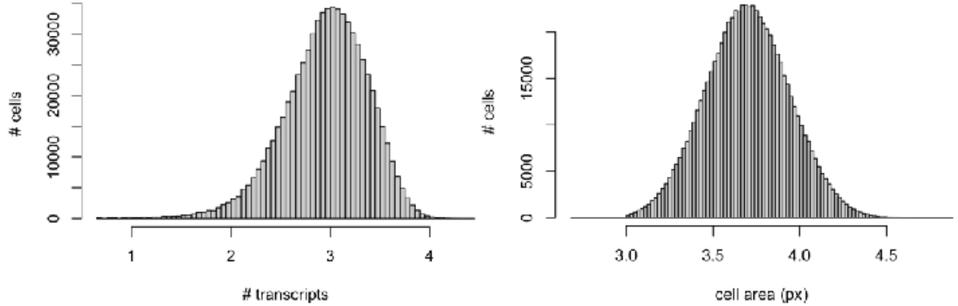


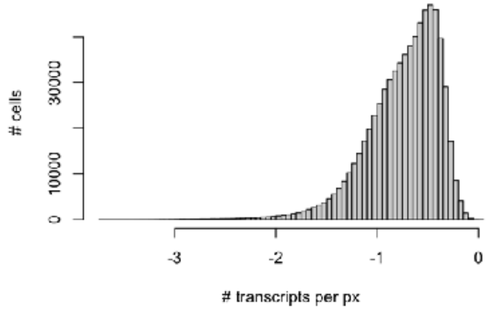
- filtering based on standard scRNA-seq QC metrics may bias again smaller & transcriptionally less complex cells



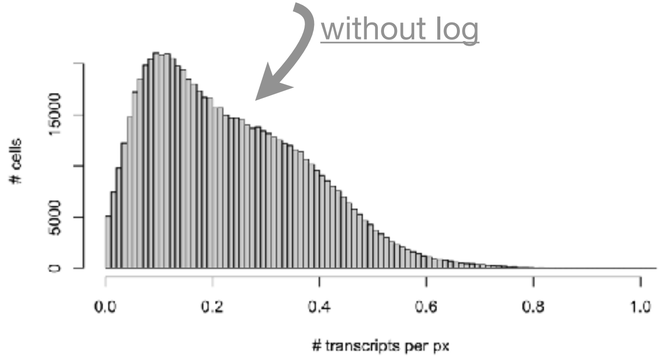
beware standard QC metrics — counts relate area AND biology

all axes $\log_{10}$-transformed (…what we'd usually filter on)

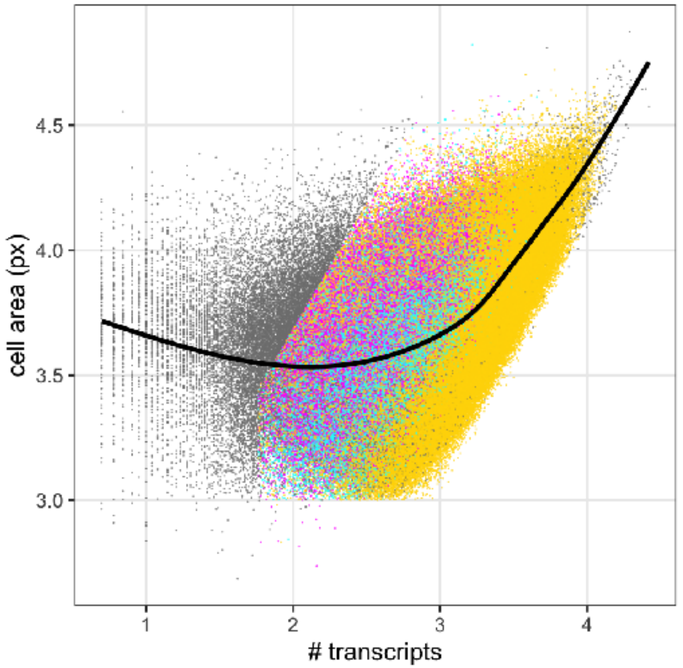


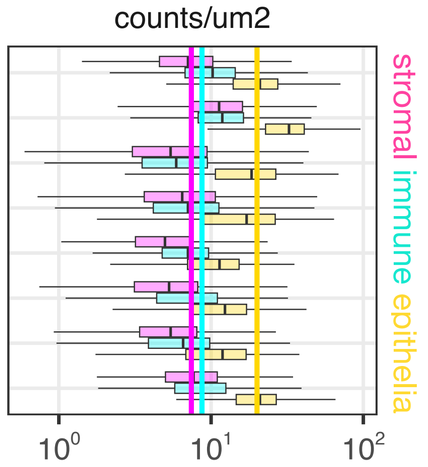


without log



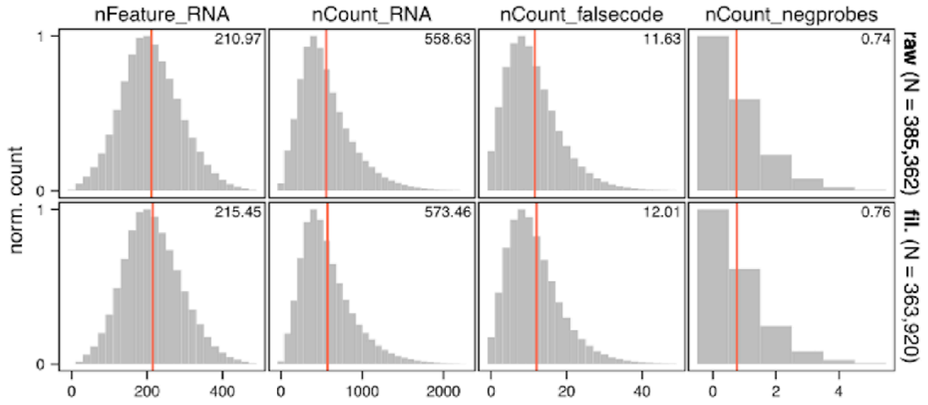
- filtering based on standard scRNA-seq QC metrics may bias again smaller & transcriptionally less complex cells



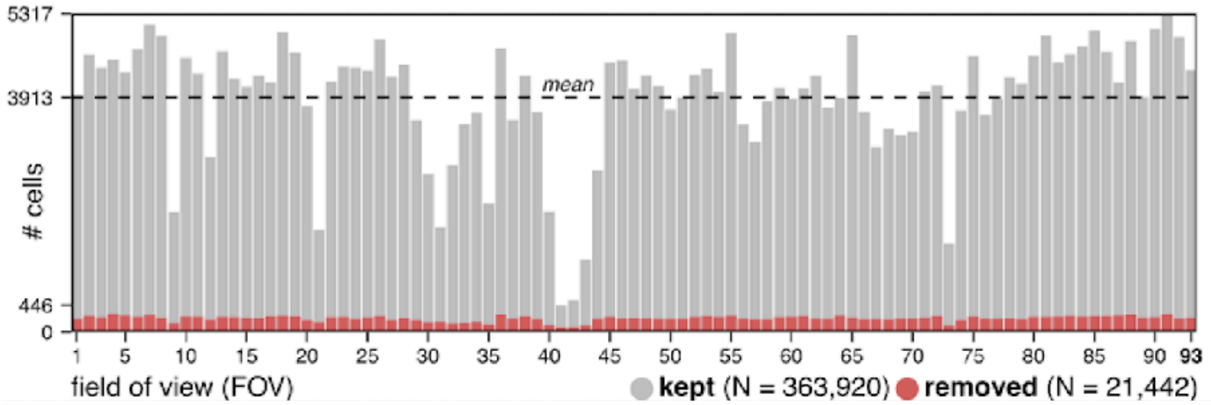


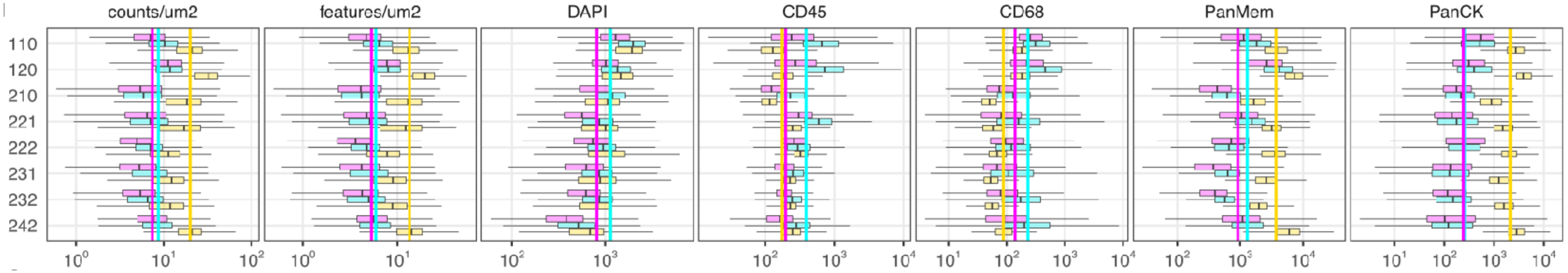
but there's more we could be looking at… some **QC examples**

standard QC metrics



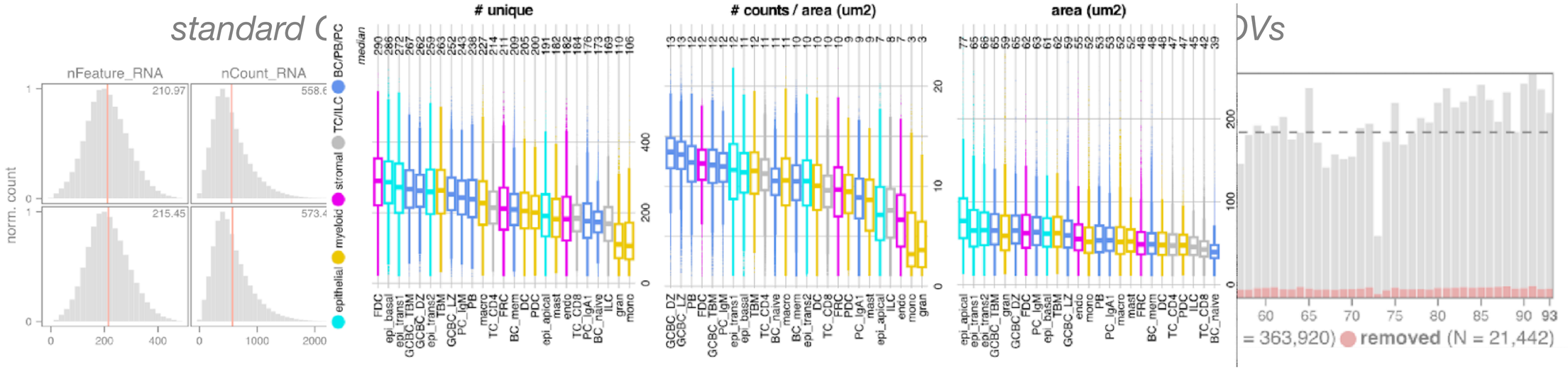
 # cells across FOVs

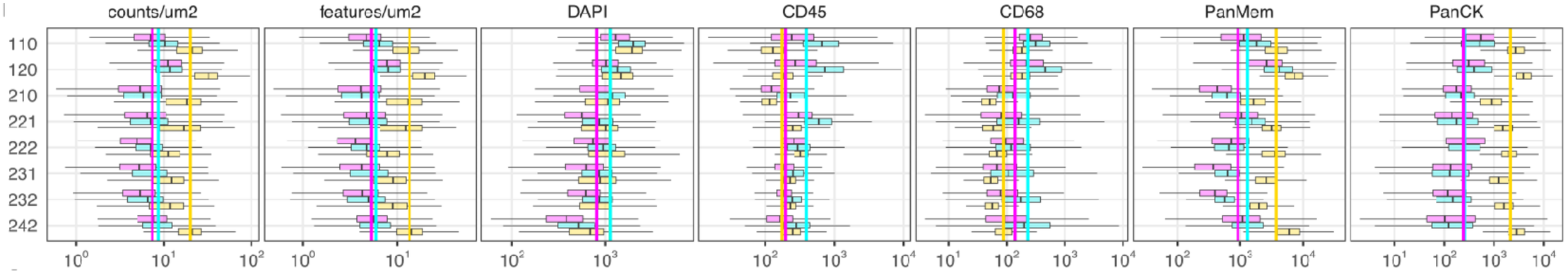




RNA per area & IF markers

but there's more we could be looking at… some **QC examples**

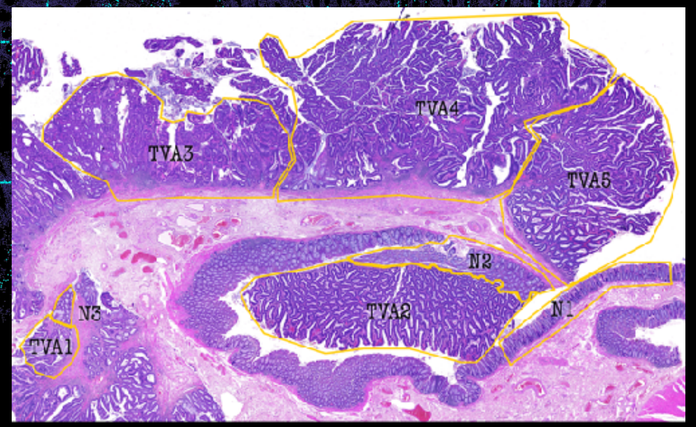




RNA per area & IF markers

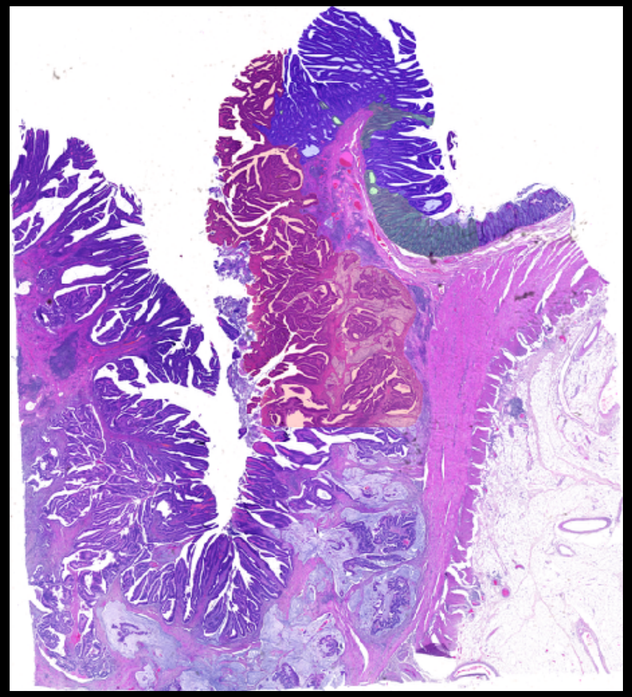
spatial organization of excluded cells might indicate a **bias against specific types** (but it depends!)

retained
excluded



img-ST data stem from serial imaging of **fields of view (FOVs)**

H&E staining

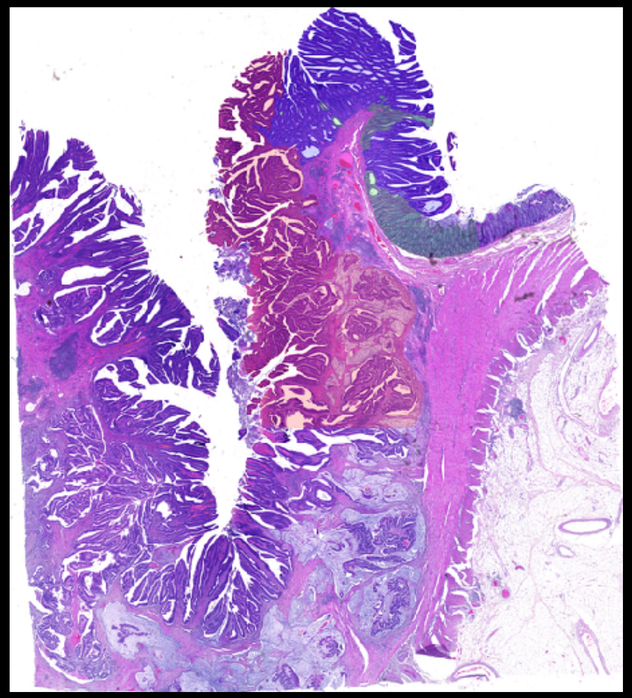


REF TVA CRC

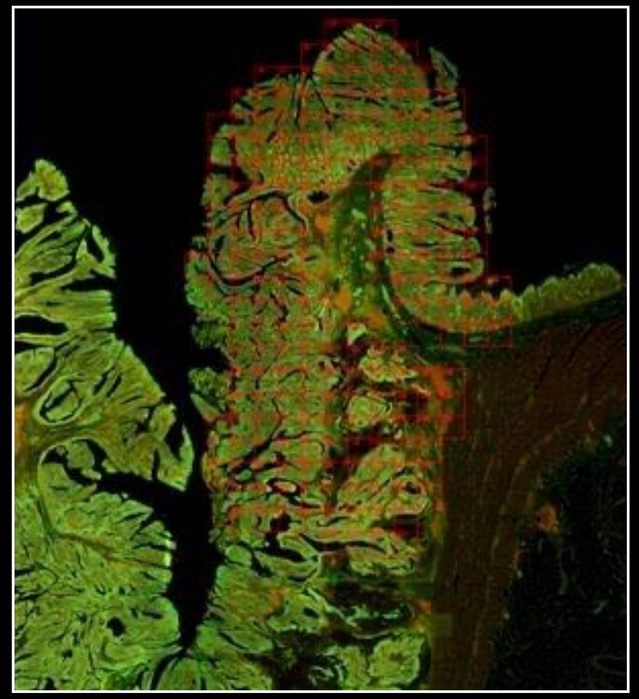
Crowell *et al.* Tejpar & Heyne (in prep.)

img-ST data stem from serial imaging of **fields of view (FOVs)**

H&E staining → FOV placement



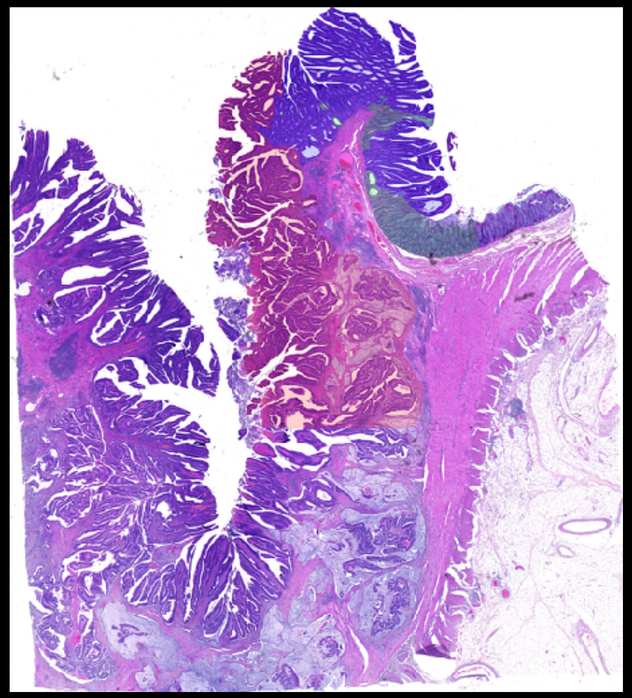
REF TVA CRC



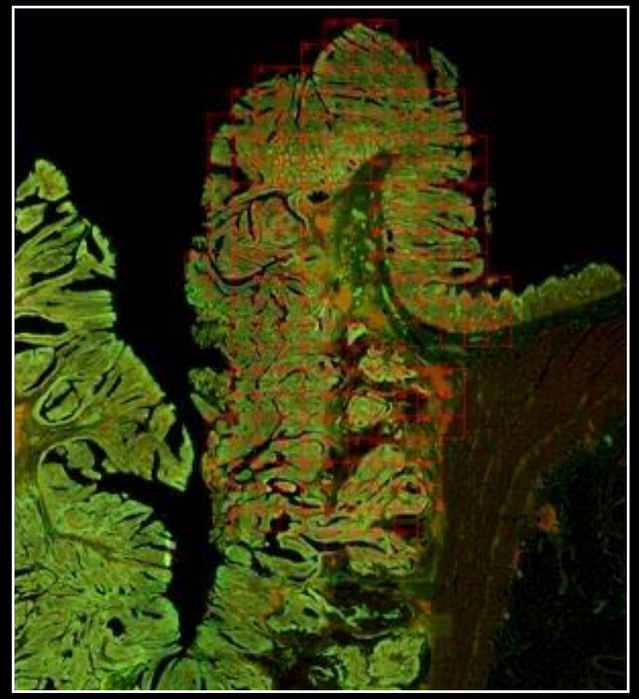
Crowell *et al.* Tejpar & Heyne (in prep.)

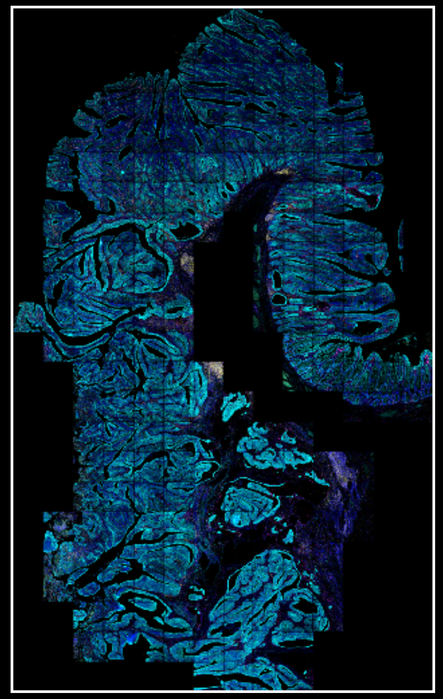
img-ST data stem from serial imaging of **fields of view (FOVs)**

H&E staining → FOV placement → imaging



REF TVA CRC



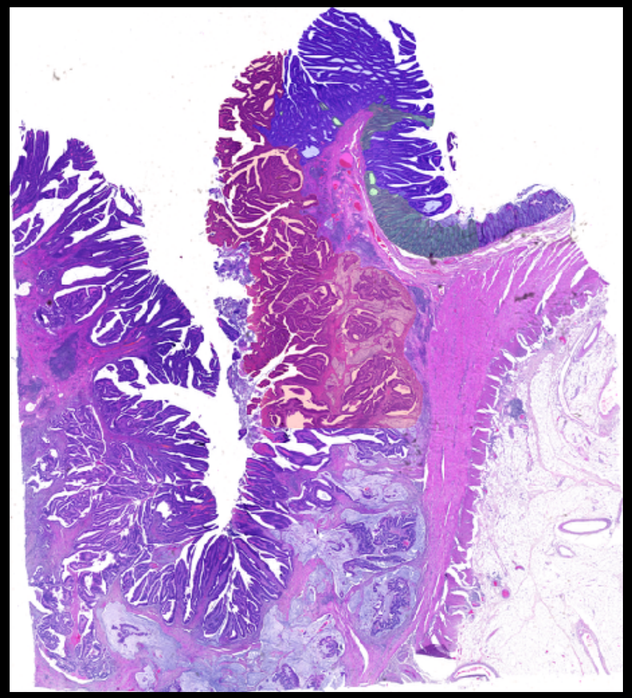


DAPI PanCK CD45 CD68

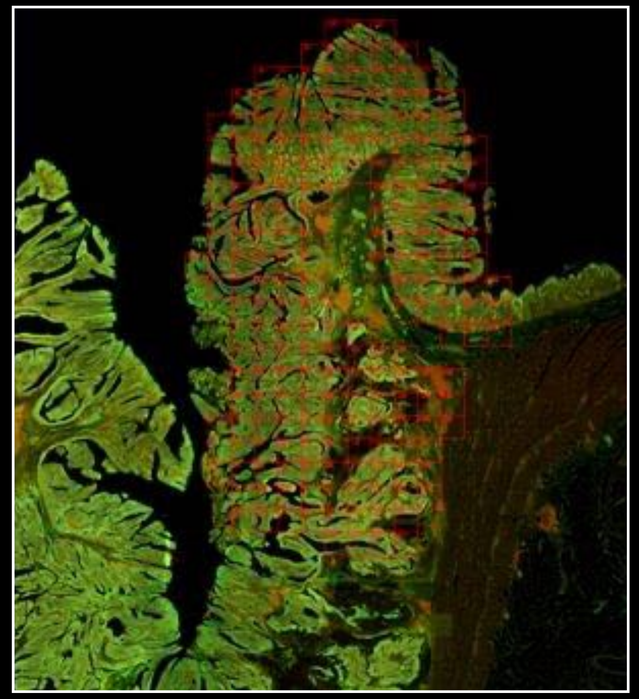
Crowell *et al.* Tejpar & Heyne (in prep.)

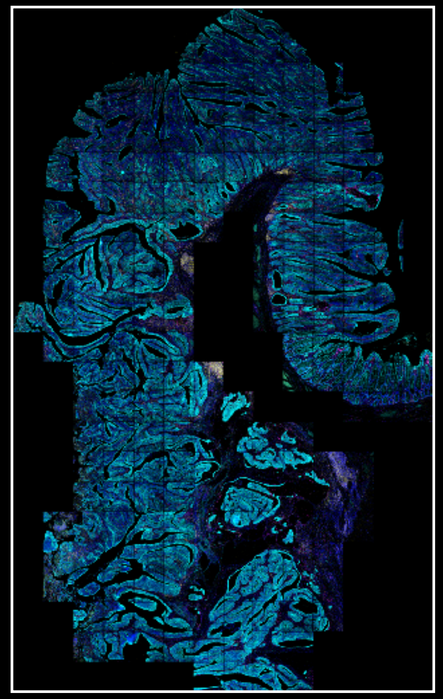
img-ST data stem from serial imaging of **fields of view (FOVs)**

H&E staining → FOV placement → imaging → analysis

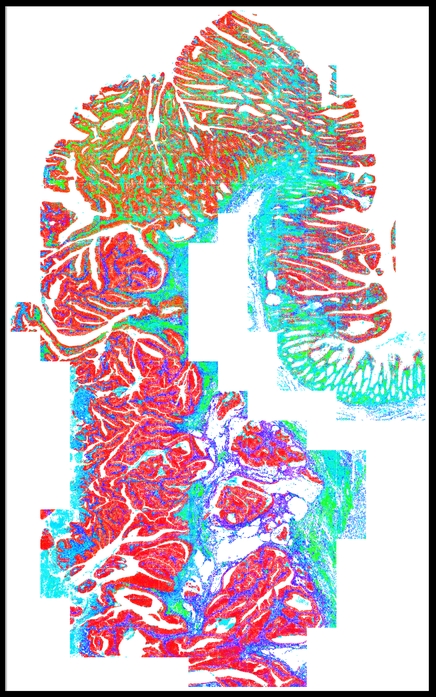


REF TVA CRC





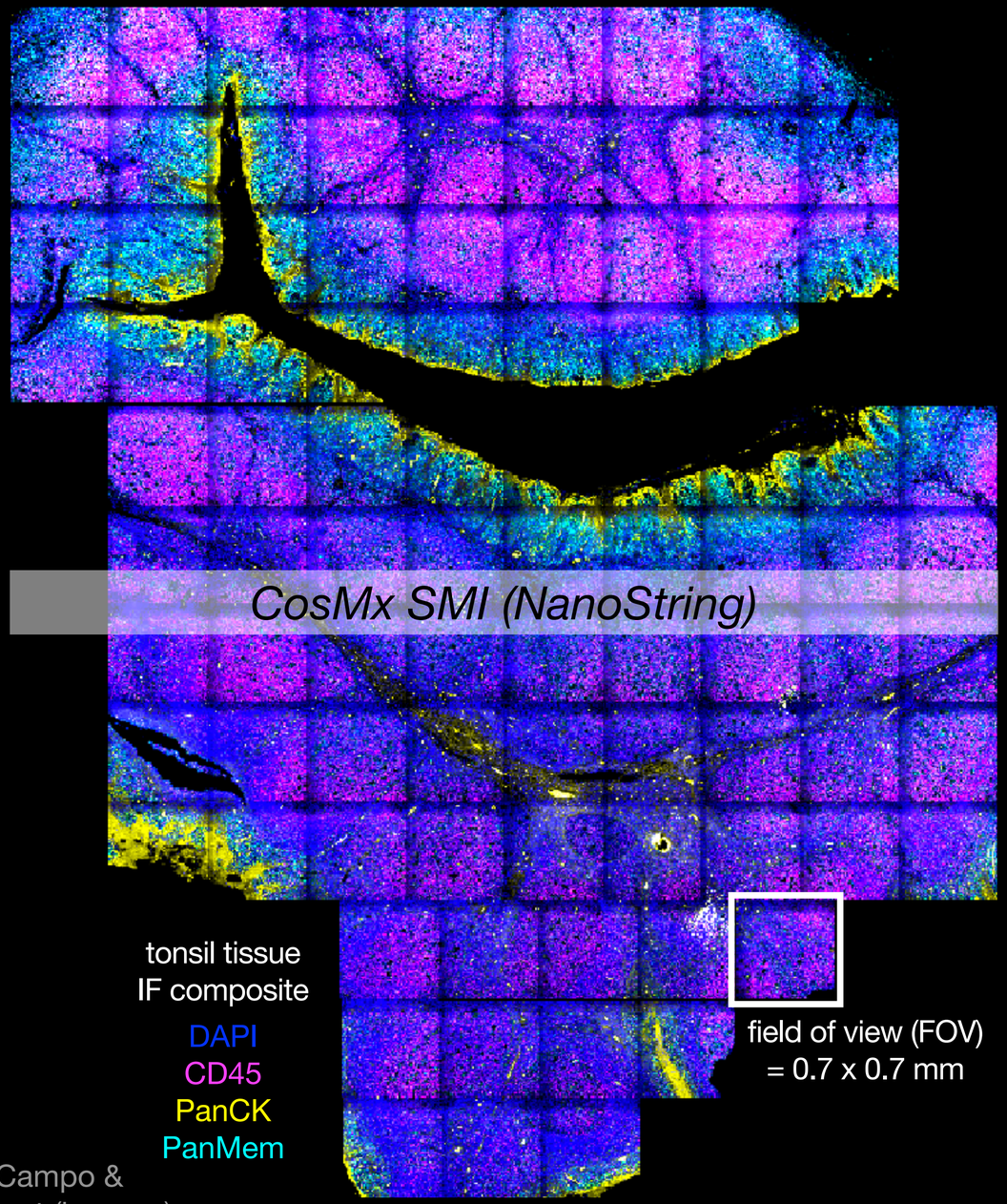
DAPI PanCK CD45 CD68



PC1 PC2 PC3

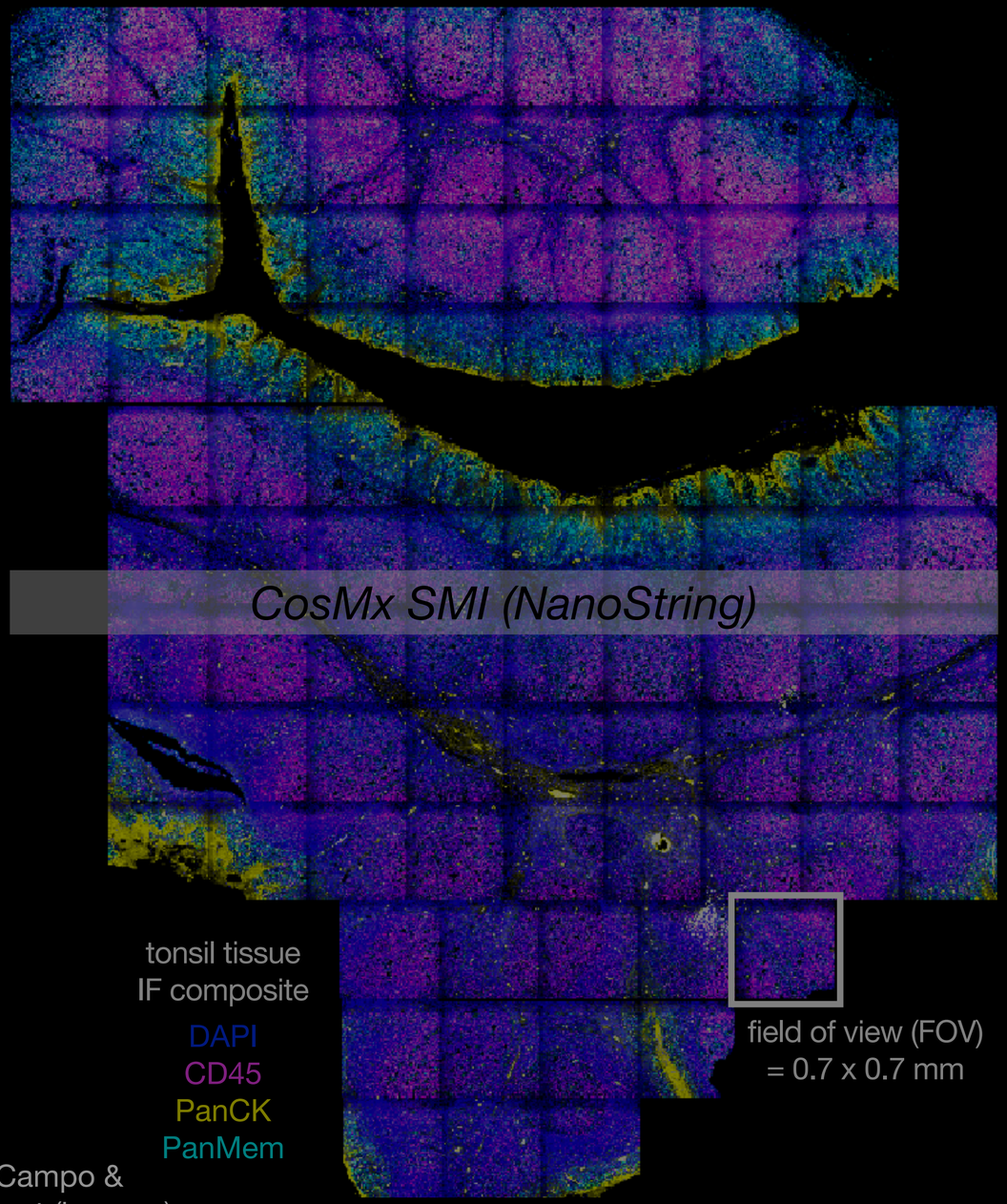
Crowell *et al.* Tejpar & Heyne (in prep.)

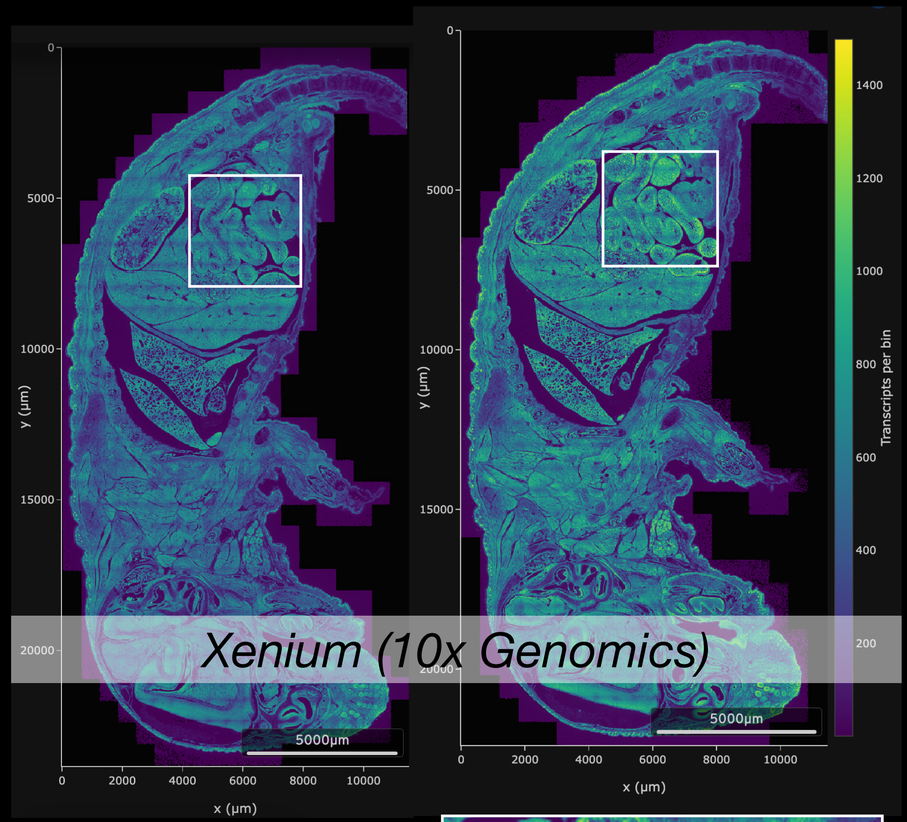
changes in optical performance lead to **lower optical resolution at FOV edges**

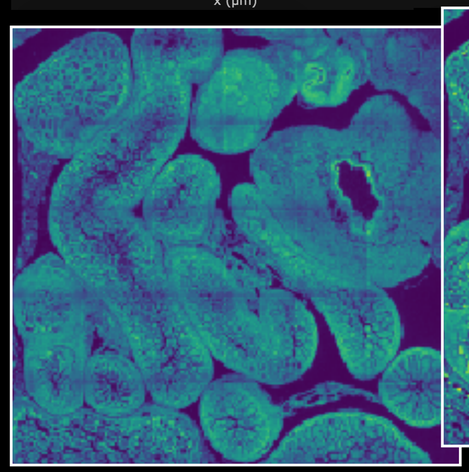


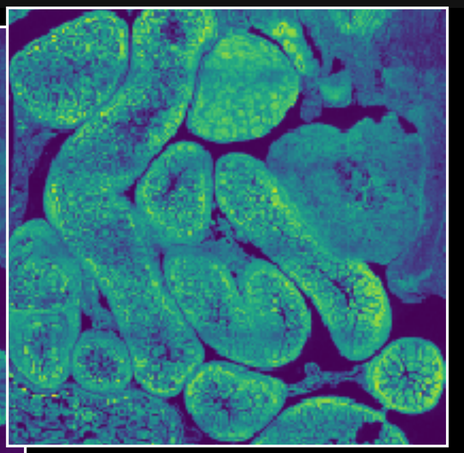
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

changes in optical performance lead to **lower optical resolution at FOV edges**



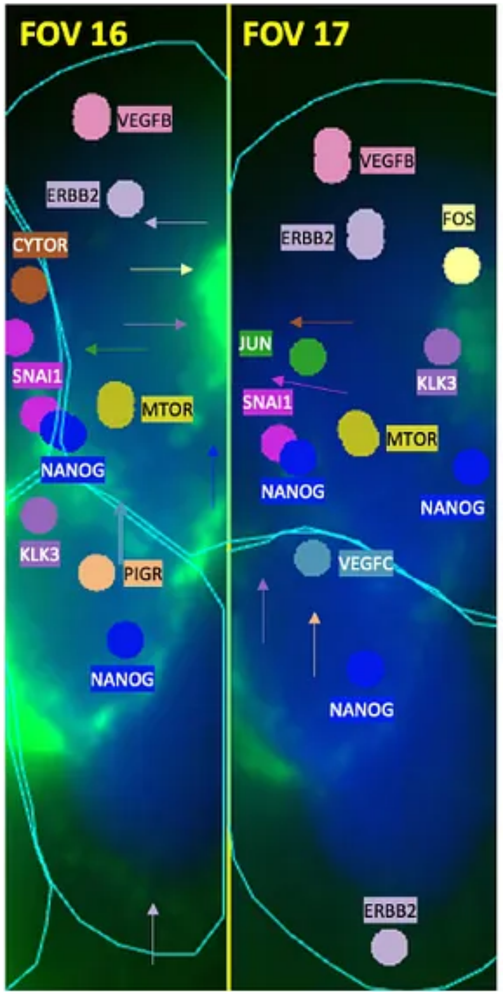


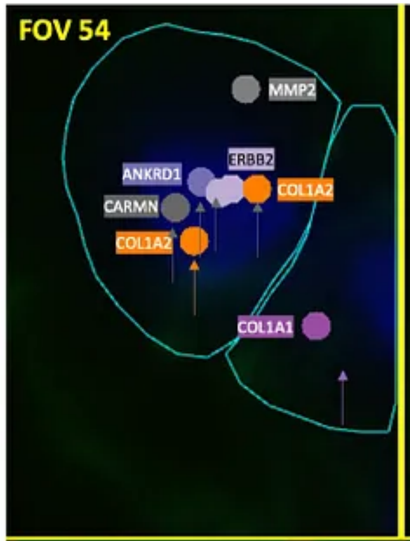


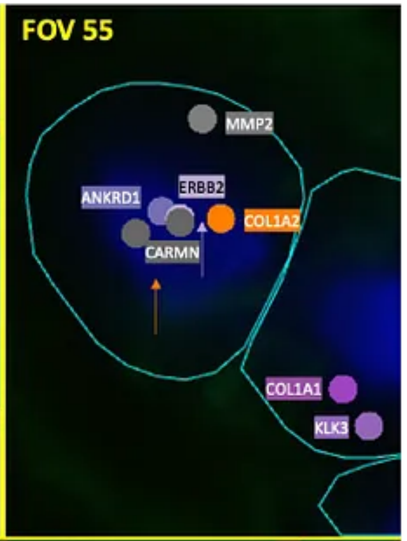


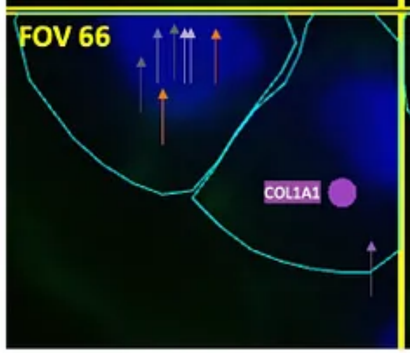
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

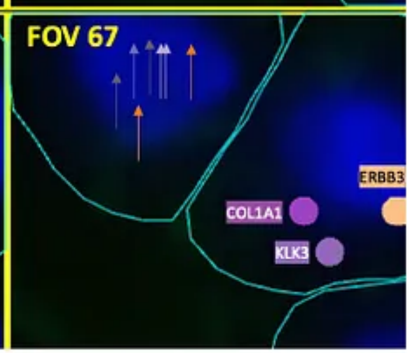
lack of stiching leads to cell **fragmentation, duplication & inconsistencies**

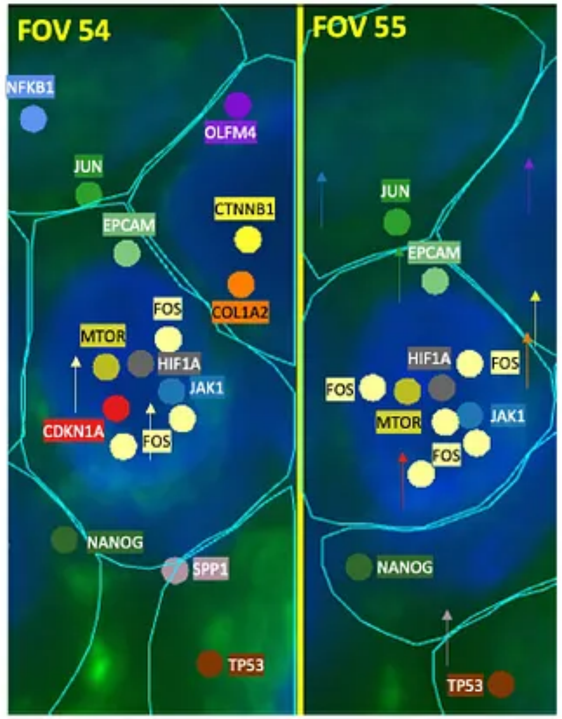


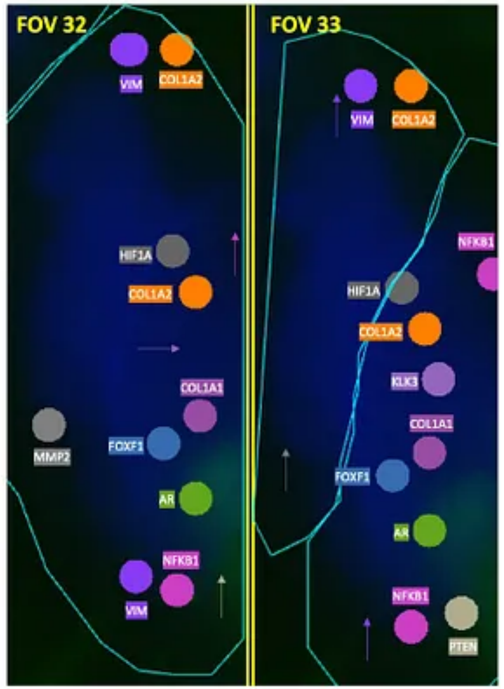








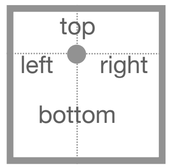




- arrows = missing transcripts
- partial cells are the same between FOVs, but come with unique IDs

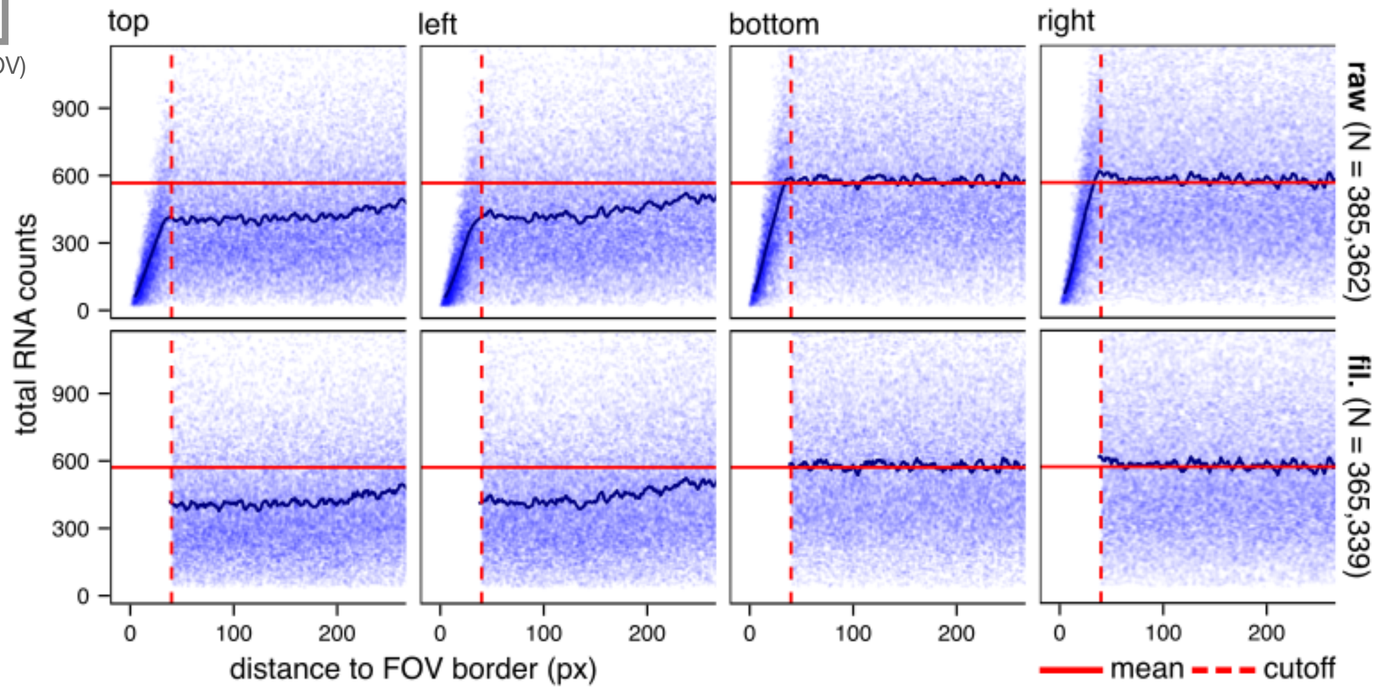
https://albertvilella.substack.com/p/comparison-between-10x-genomics-xenium

border cells have fewer counts & can highlight other artefacts



field of view (FOV)

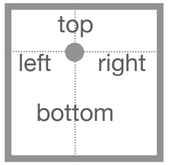
- compute each cell's distance to each FOV border
- plot counts vs. distance, stratified by direction



points = cells (across all FOVs), lines = running median

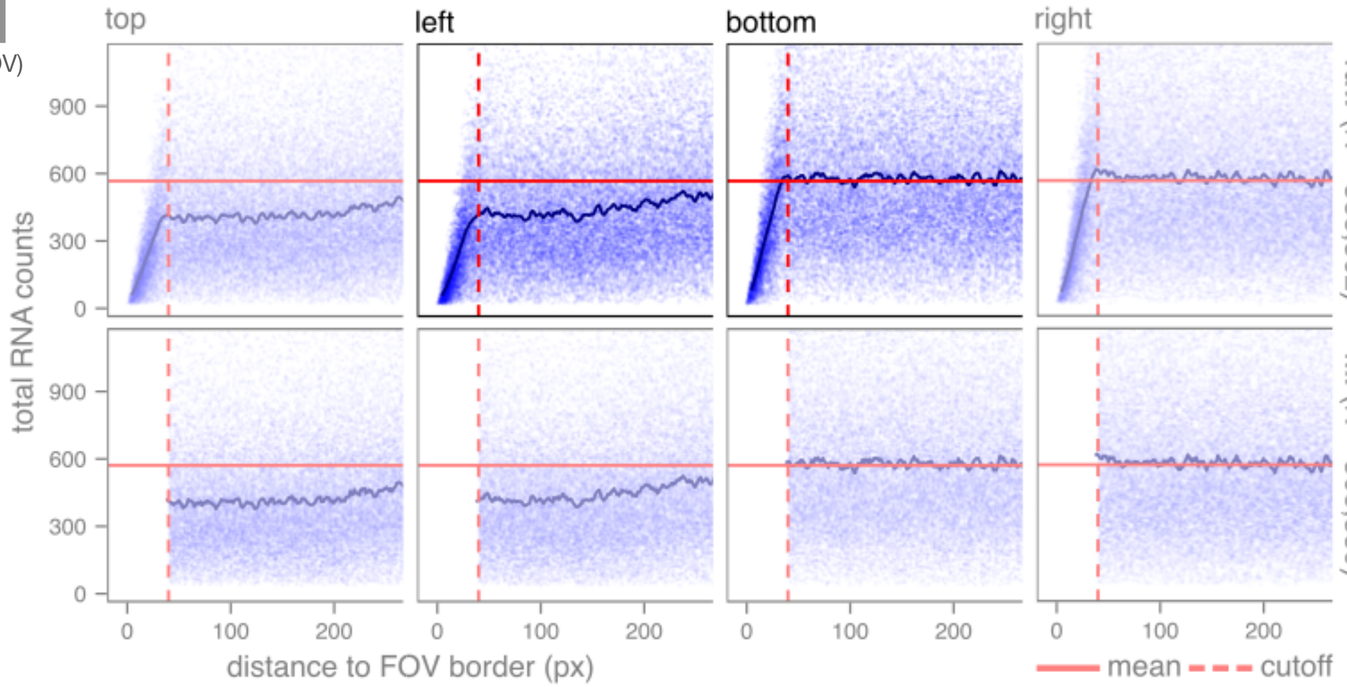
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

border cells have fewer counts & can highlight other artefacts

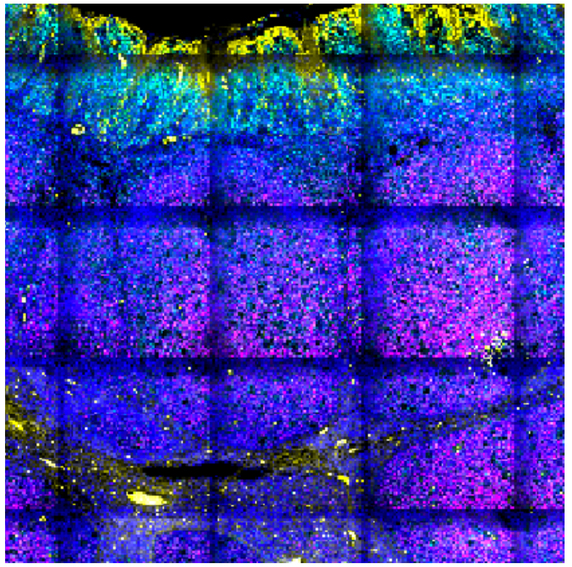


field of view (FOV)

- compute each cell's distance to each FOV border
- plot counts vs. distance, stratified by direction



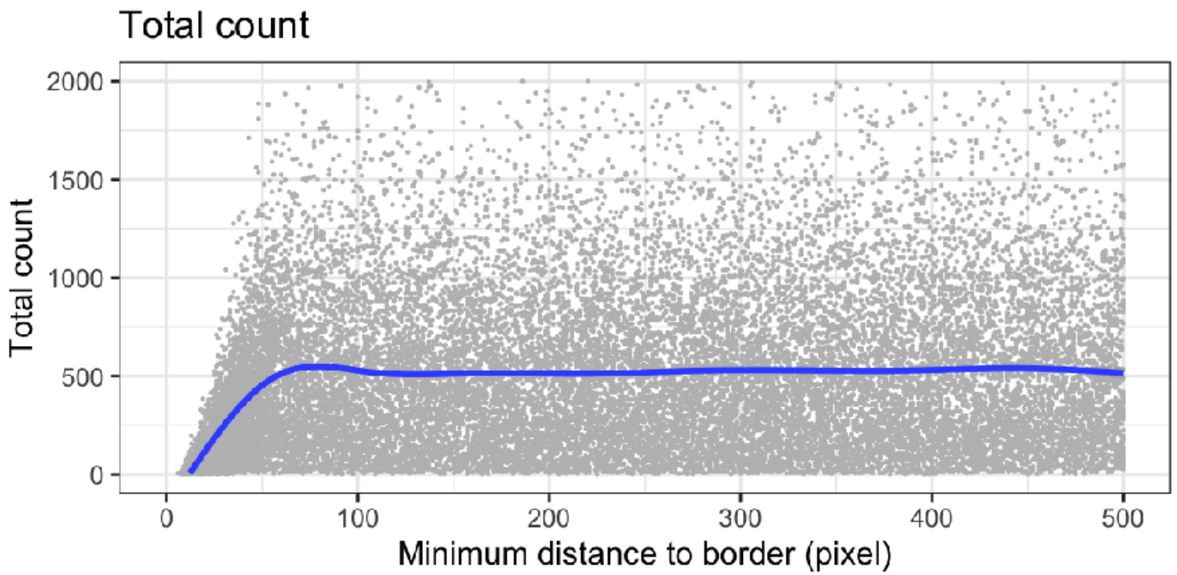
points = cells (across all FOVs), lines = running median

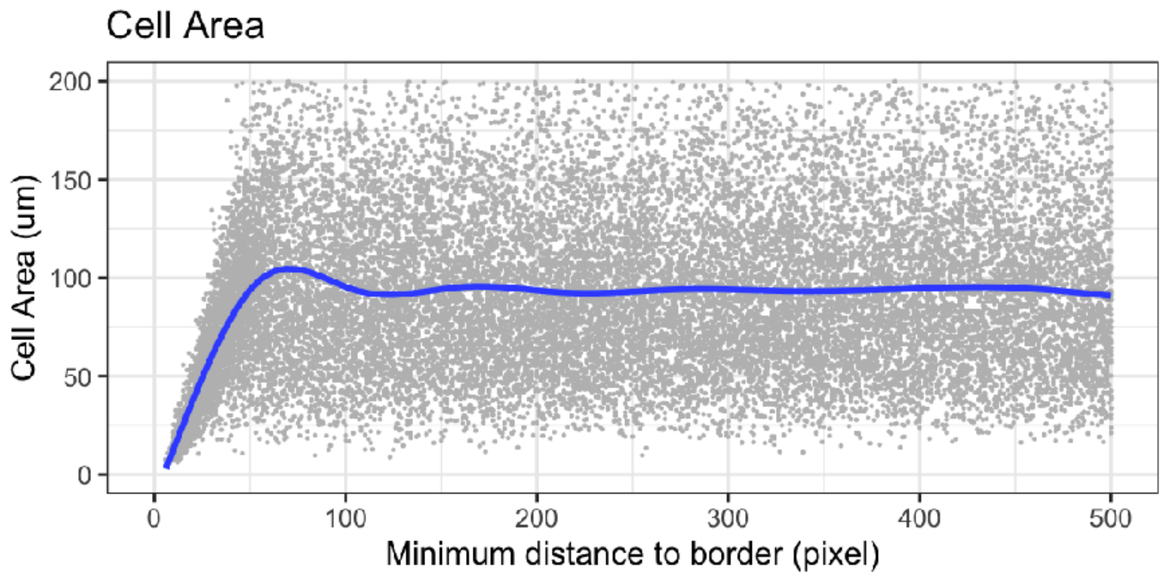


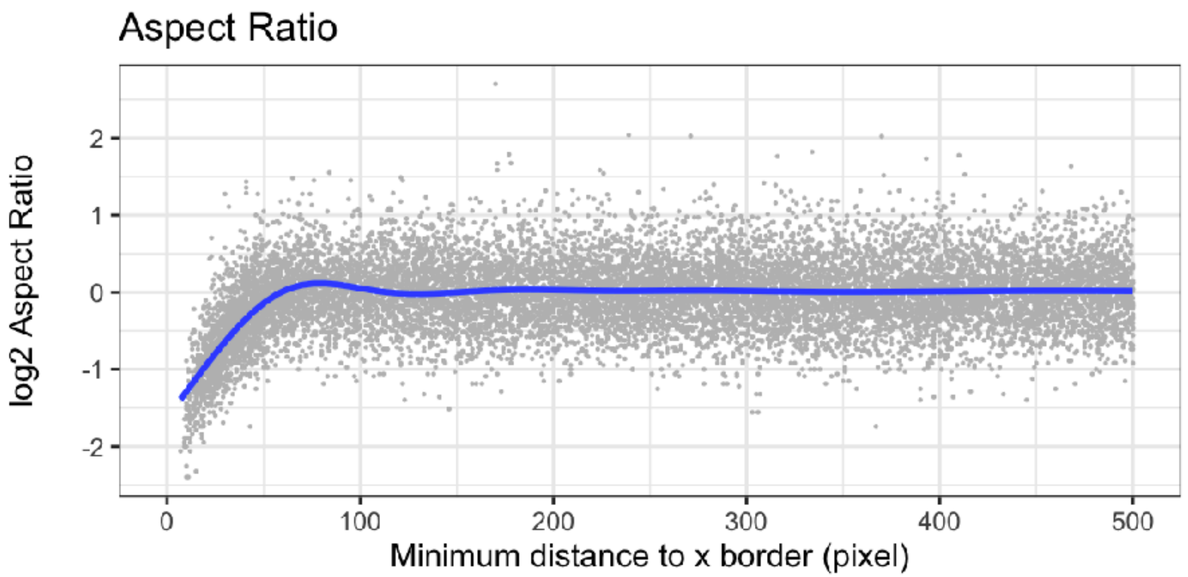
top-left corners exhibit dimmer IF signals

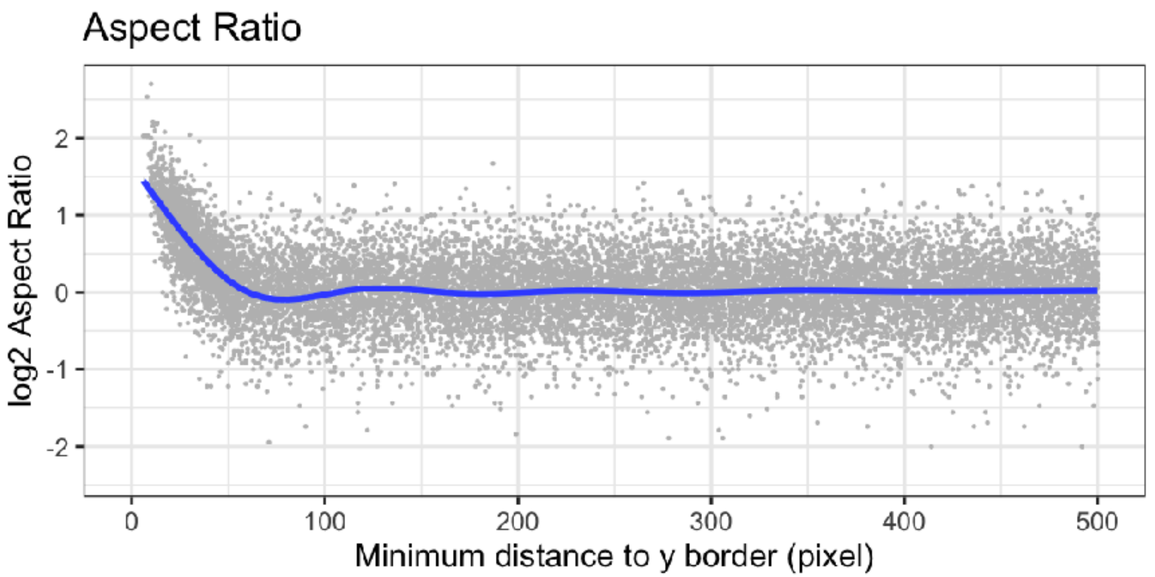
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

besides fewer counts, **border cells are smaller & slimmer**





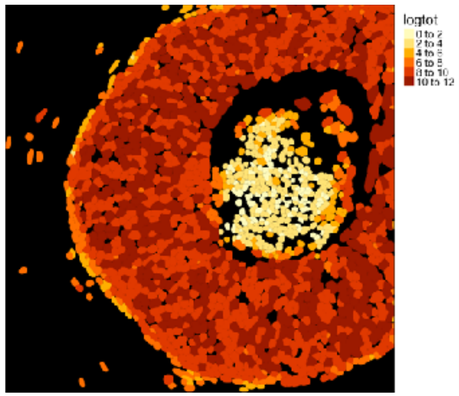




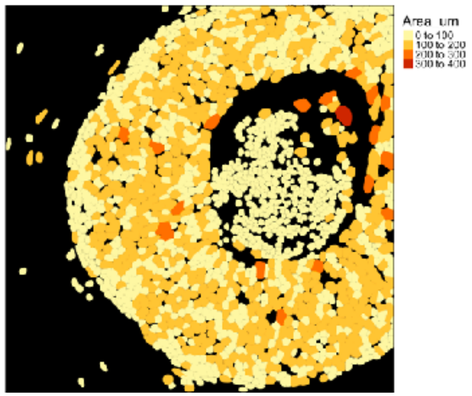
plots courtesy of Davide Risso

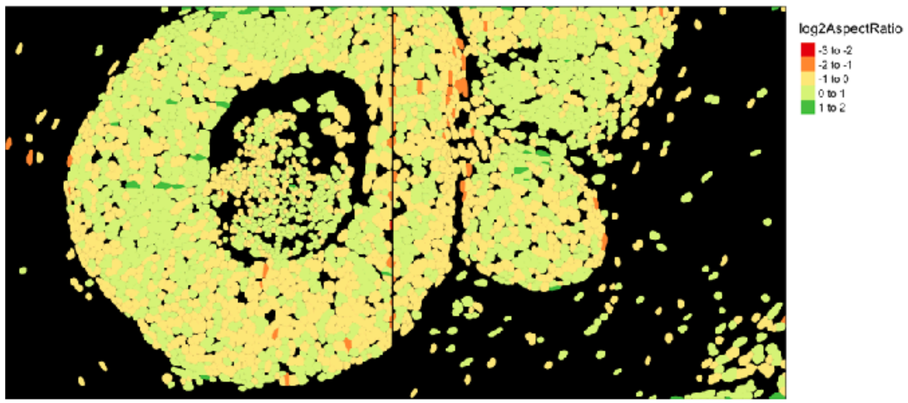
SpaceTrooper proposes a *flag score* combining several metrics

counts



area



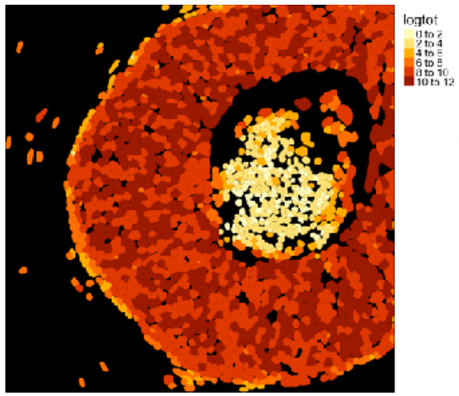


aspect ratio

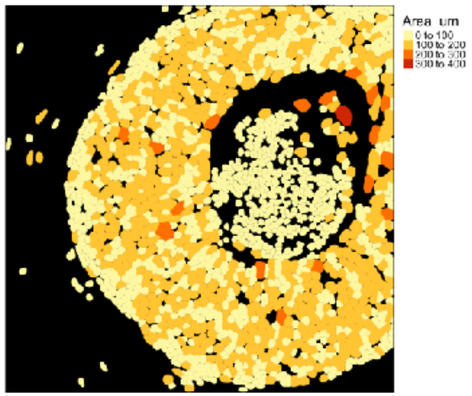
https://github.com/drighelli/SpaceTrooper

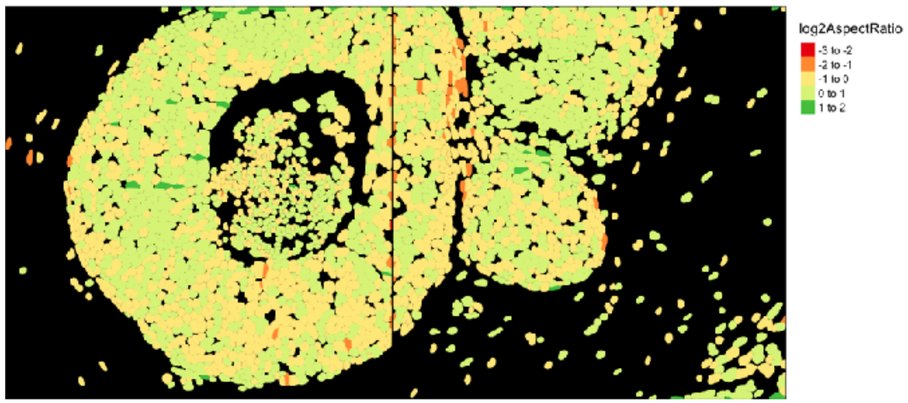
SpaceTrooper proposes a *flag score* combining several metrics

counts



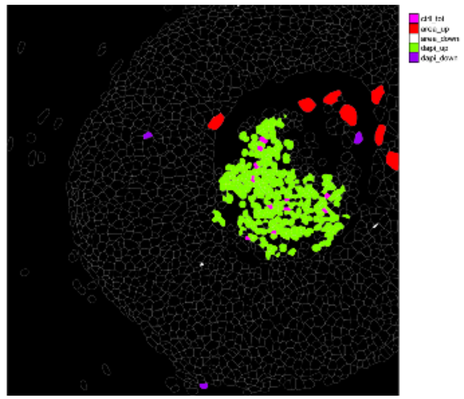
area





aspect ratio

outliers



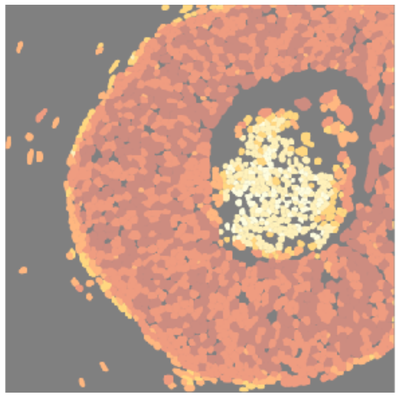
(e.g., thresholding on MADs of univariate distributions)

https://github.com/drighelli/SpaceTrooper

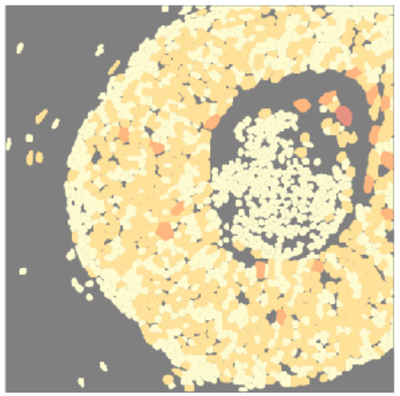
SpaceTrooper proposes a *flag score* combining several metrics

$$\text{logit}(F) = \beta_0 + \beta_1 \log \frac{\text{count}}{\text{area}} + \beta_2 \left| \log(\text{aspect ratio}) \right| \cdot I_{\{d<\text{threshold}\}}$$

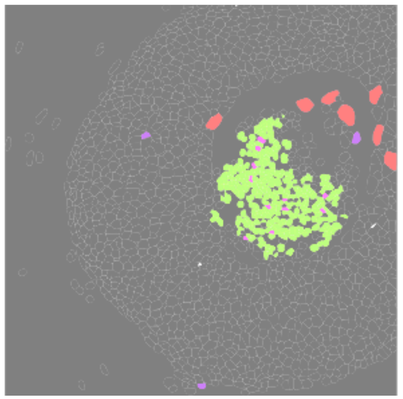
counts



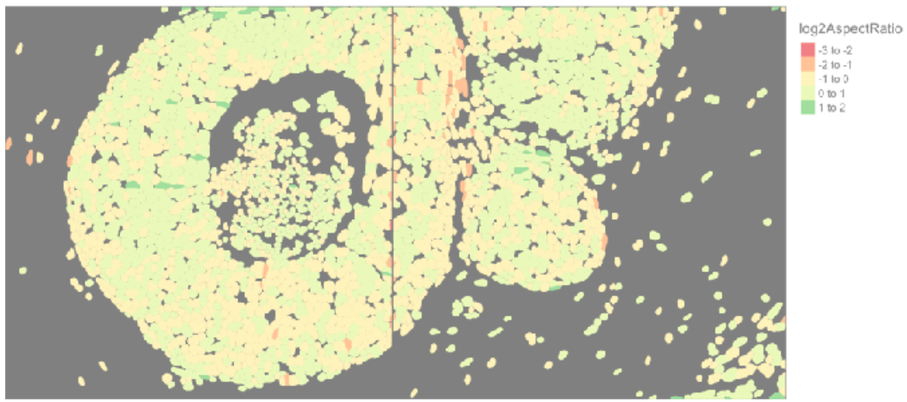
area



outliers



(e.g., thresholding on MADs of univariate distributions)



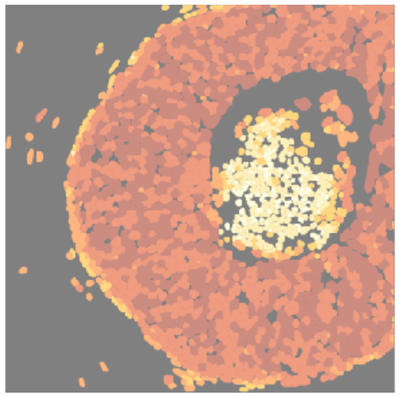
aspect ratio

https://github.com/drighelli/SpaceTrooper

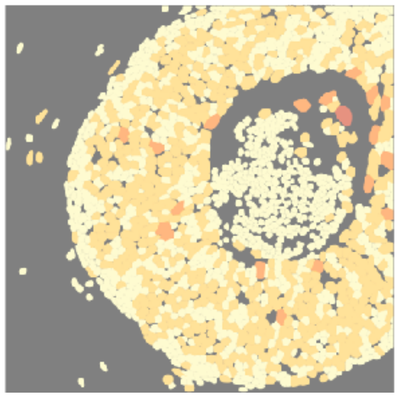
SpaceTrooper proposes a *flag score* combining several metrics

$$\text{logit}(F) = \beta_0 + \beta_1 \log \frac{\text{count}}{\text{area}} + \beta_2 \left| \log(\text{aspect ratio}) \right| \cdot I_{\{d<\text{threshold}\}}$$

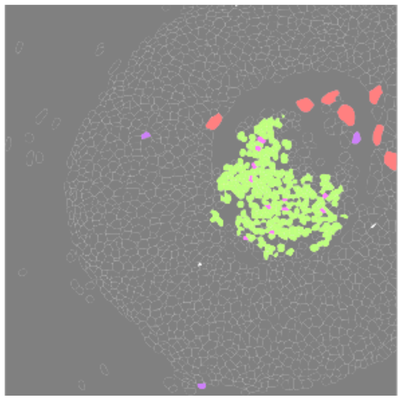
counts

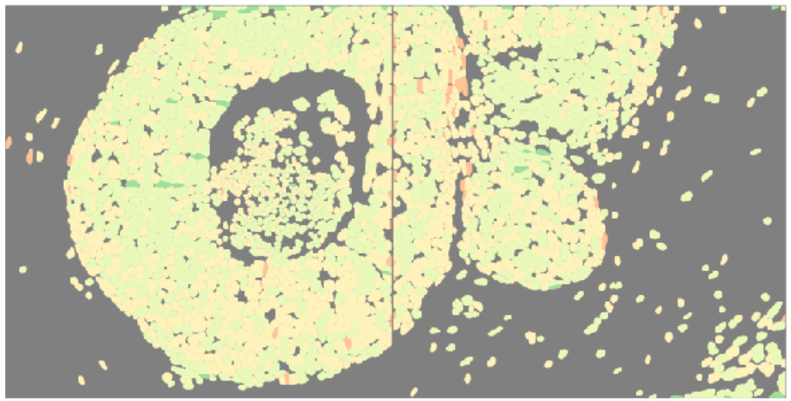


area

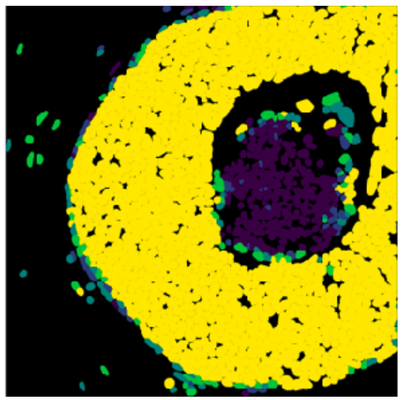


outliers





aspect ratio



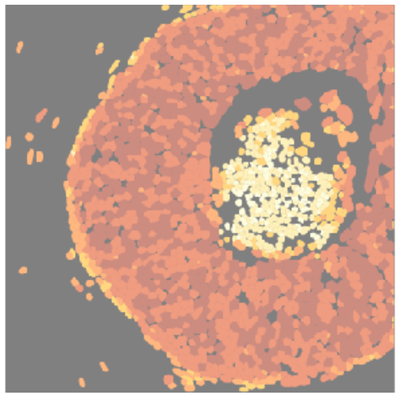
flag score

https://github.com/drighelli/SpaceTrooper

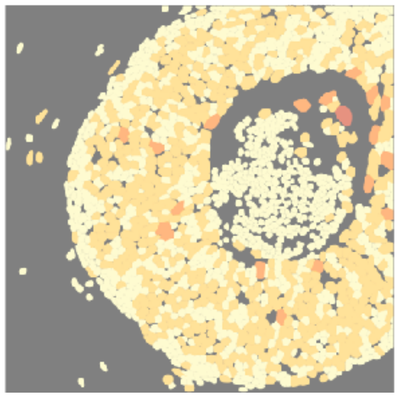
SpaceTrooper proposes a *flag score* combining several metrics

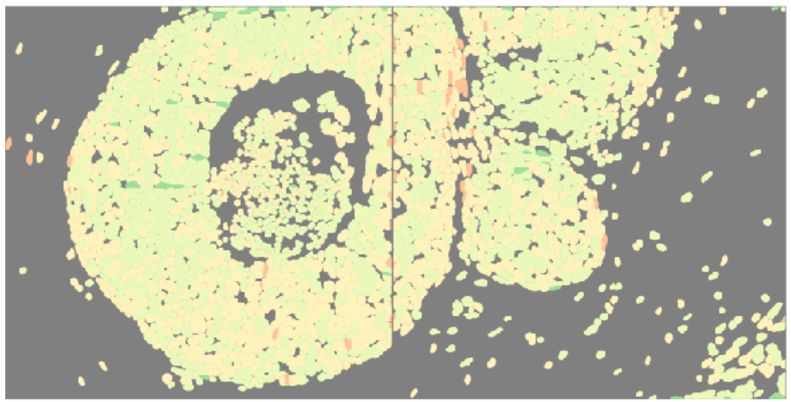
$$\text{logit}(F) = \beta_0 + \beta_1 \log \frac{\text{count}}{\text{area}} + \beta_2 \left| \log(\text{aspect ratio}) \right| \cdot I_{\{d<\text{threshold}\}}$$

counts



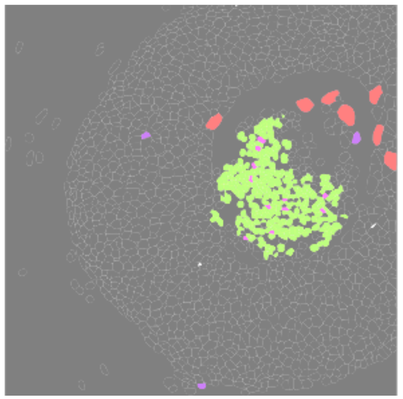
area

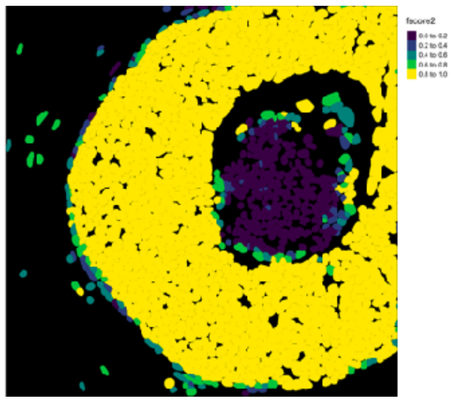




aspect ratio

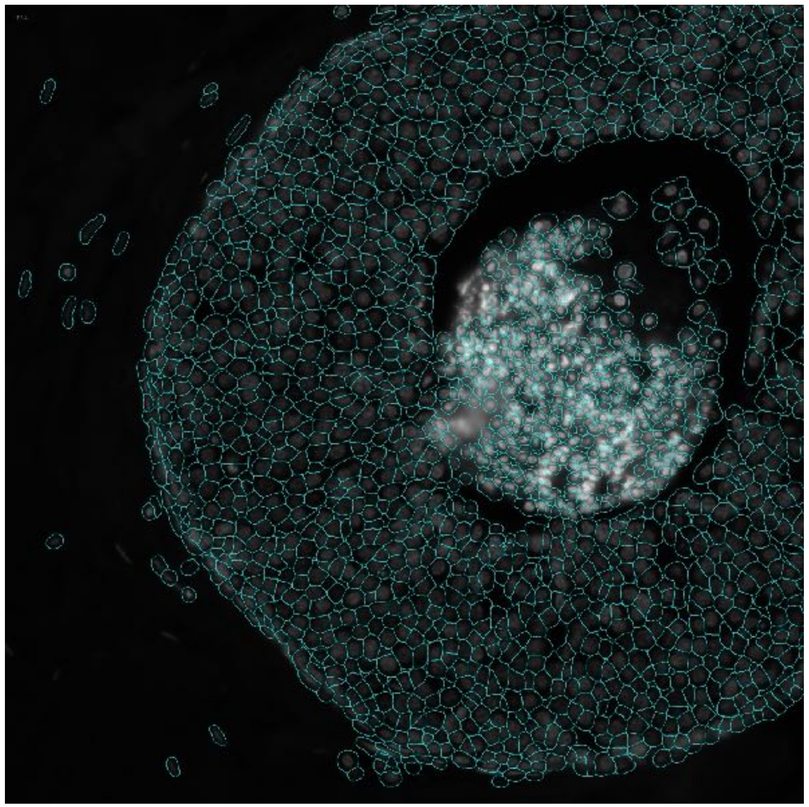
outliers





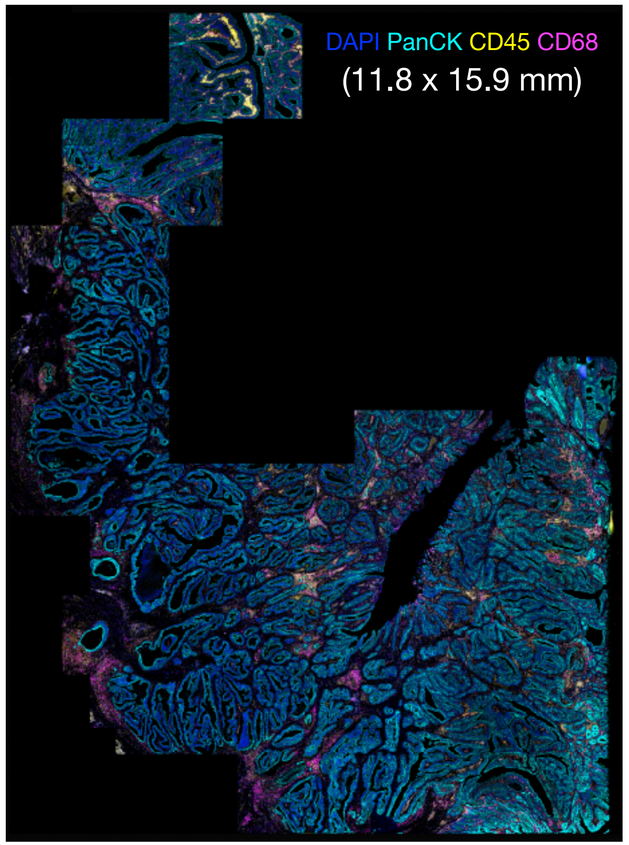
flag score

DAPI + segmentation



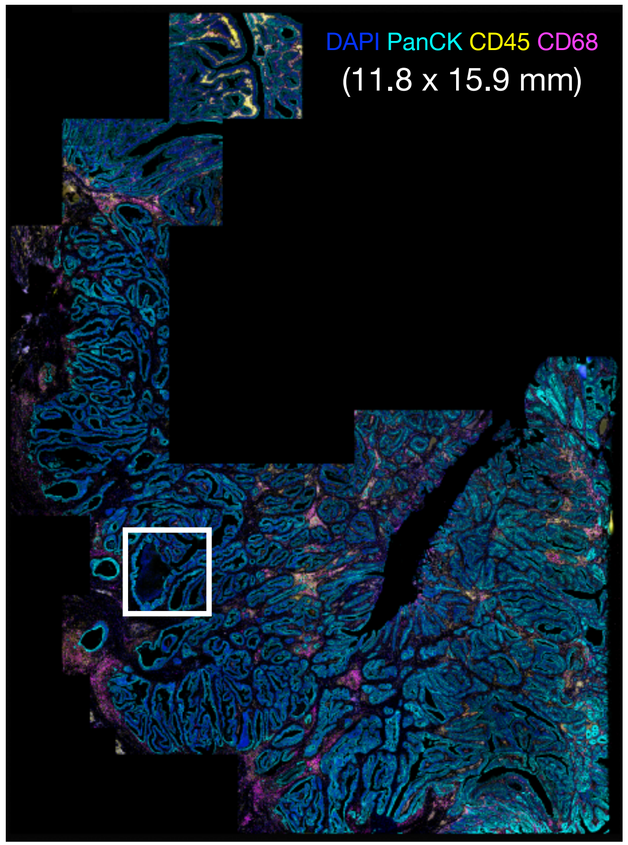
https://github.com/drighelli/SpaceTrooper

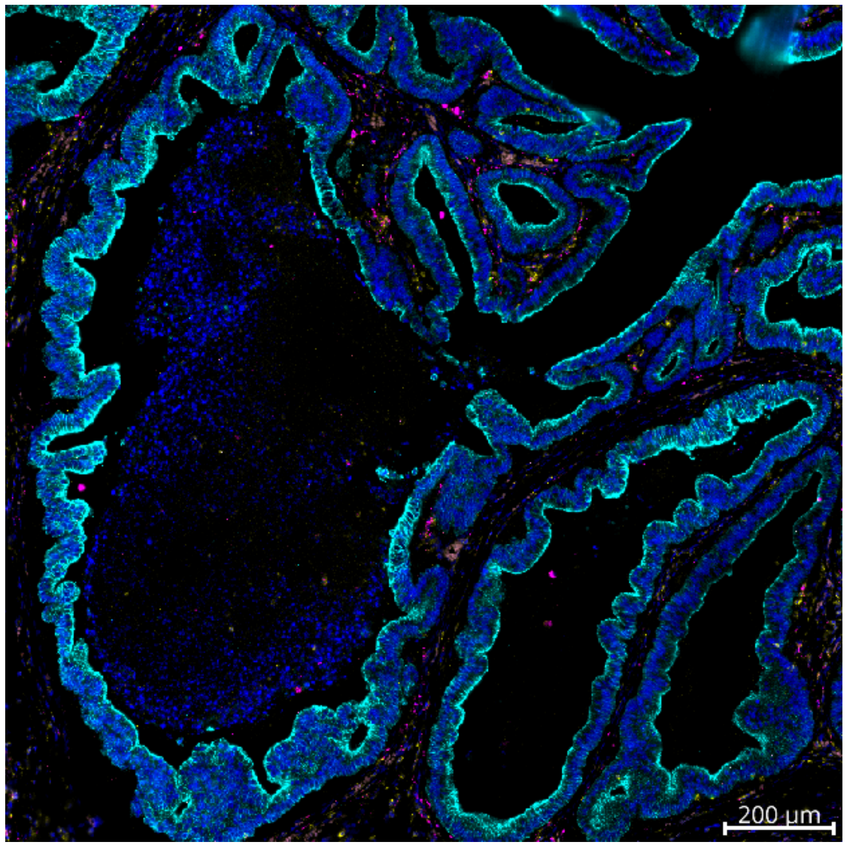
debris is usually segmented, yet hard to distinguish



Crowell *et al.* Tejpar & Heyne (in prep.)

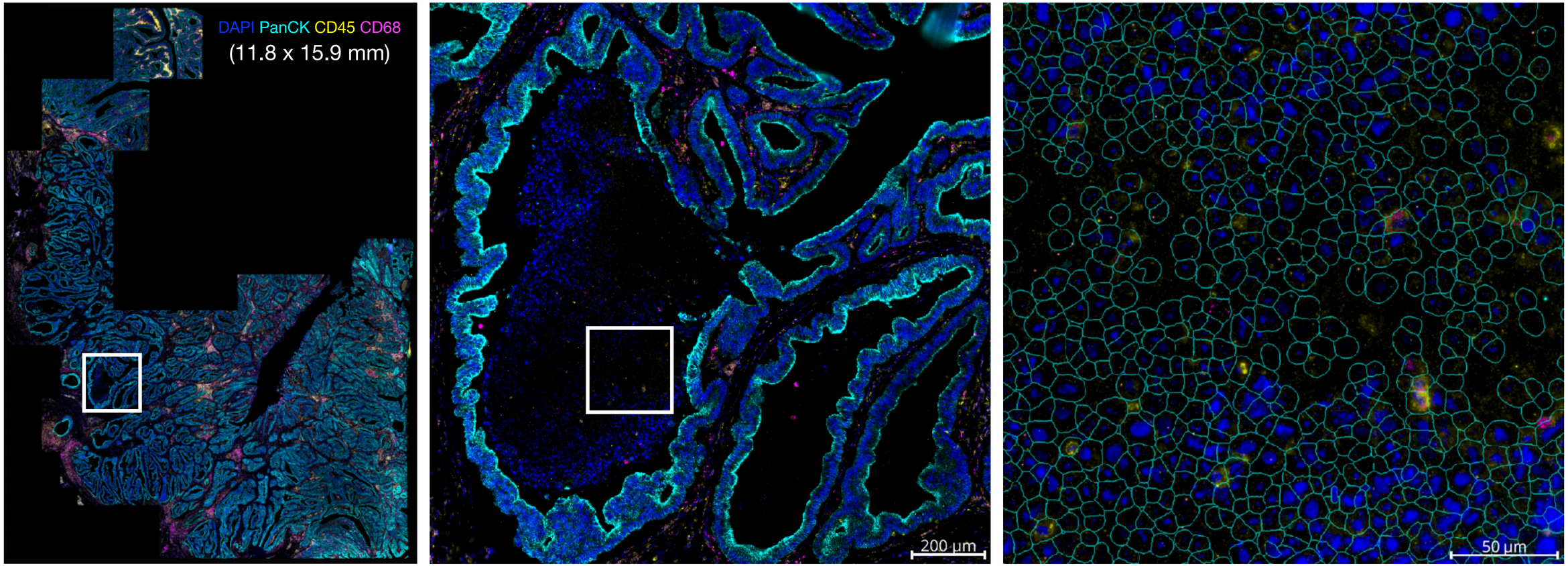
debris is usually segmented, yet hard to distinguish





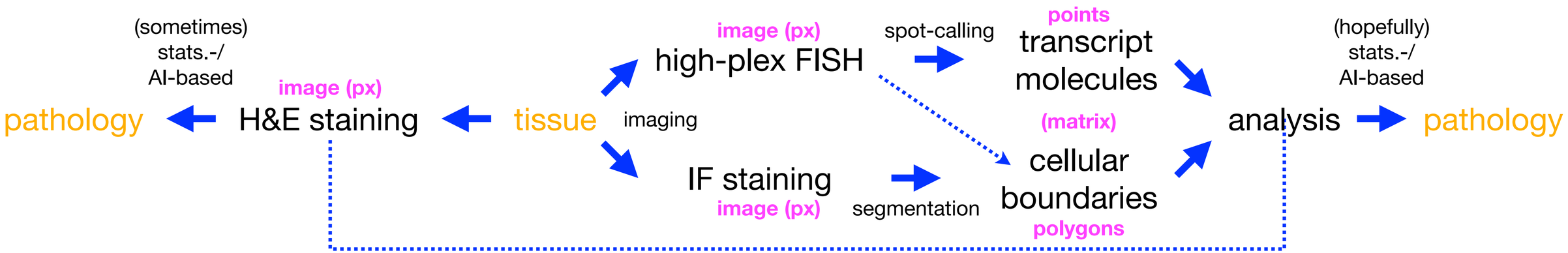
Crowell *et al.* Tejpar & Heyne (in prep.)

debris is usually segmented, yet hard to distinguish



Crowell *et al.* Tejpar & Heyne (in prep.)

library size normalization on non-representative panels introduces biases



Atta *et al.* (2024), *Genome Biology* 25:153

library size normalization on non-representative panels introduces biases

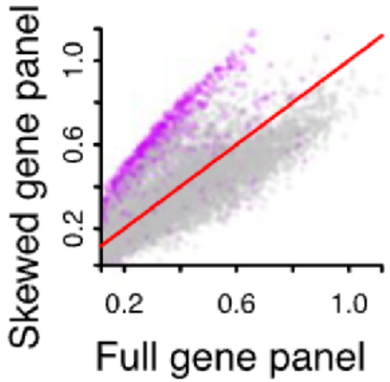
Atta *et al.* (2024), *Genome Biology* 25:153

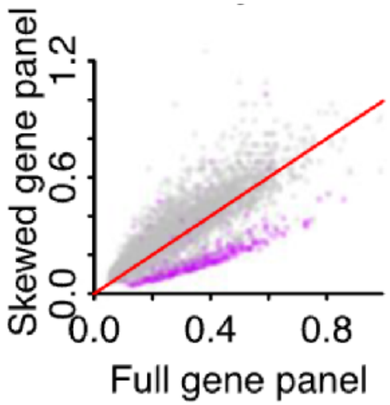
library size normalization on non-representative panels introduces biases

- **scaling factors** for cells in R are systematically **larger**

100-gene panel skewed towards some region R

- **gene expression** for cells in R are systematically **smaller**





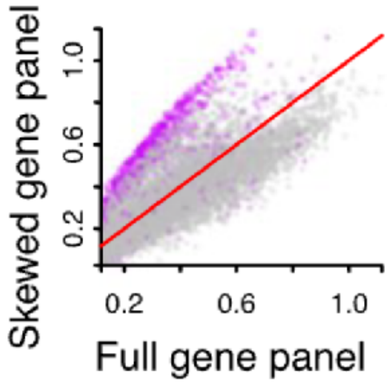
Atta *et al.* (2024), *Genome Biology* 25:153

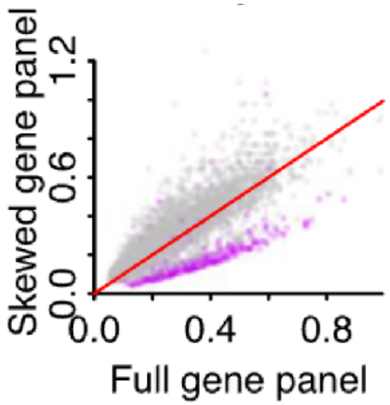
library size normalization on non-representative panels introduces biases

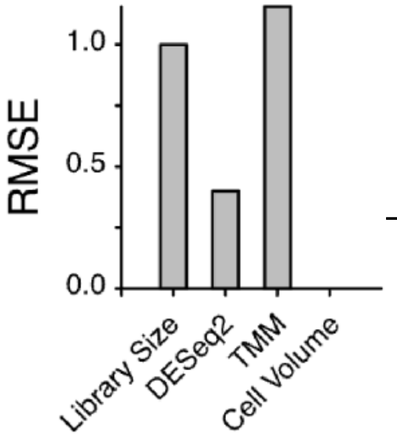
- **scaling factors** for cells in R are systematically **larger**

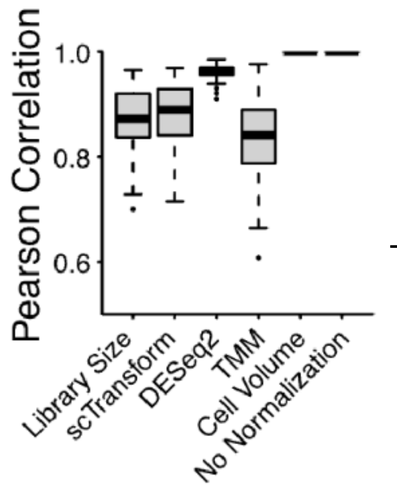
100-gene panel skewed towards some region R

- **gene expression** for cells in R are systematically **smaller**





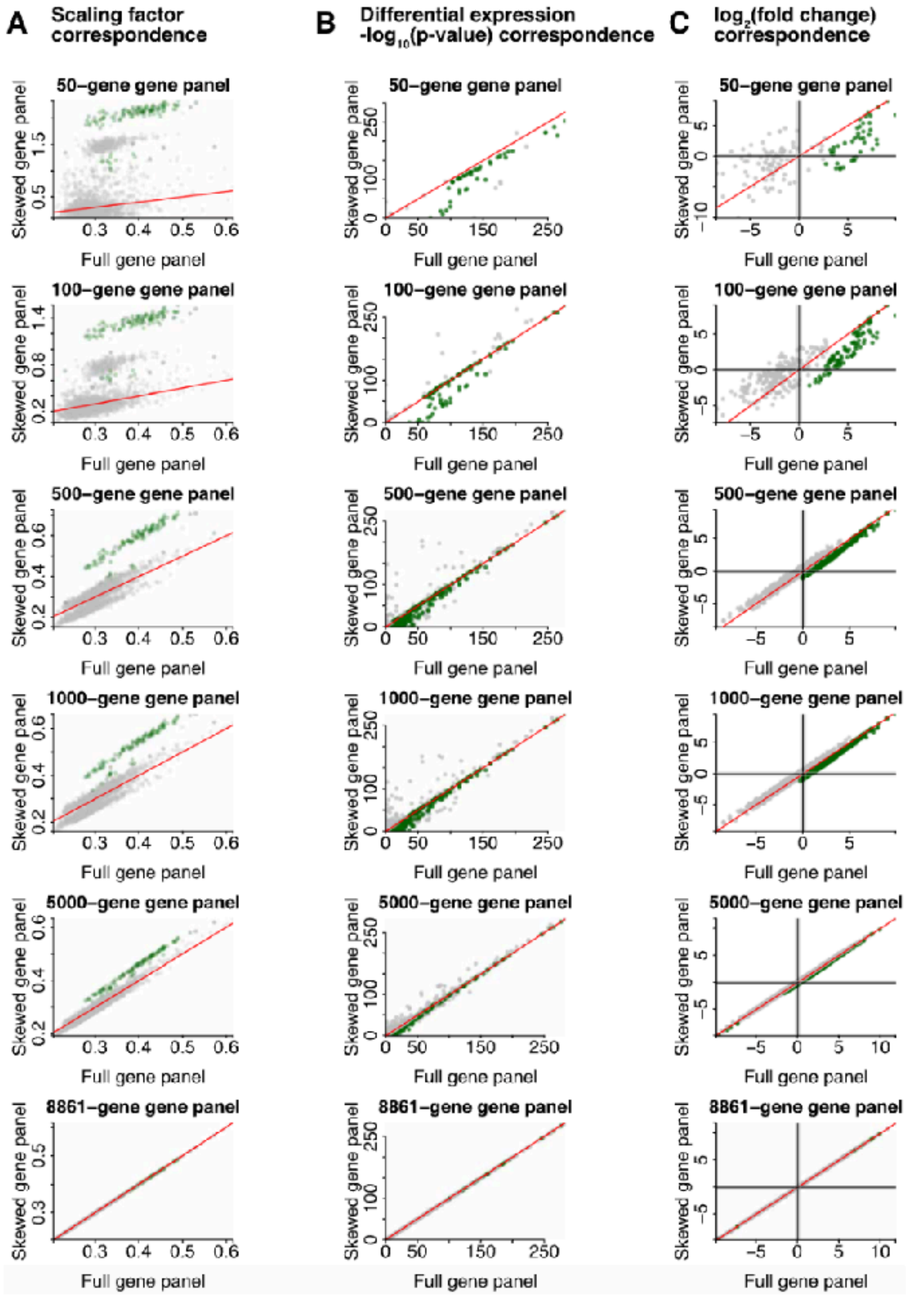




- systematic biases affect analyses to evaluate **differential gene expression & spatially variable genes**

Atta *et al.* (2024), *Genome Biology* 25:153

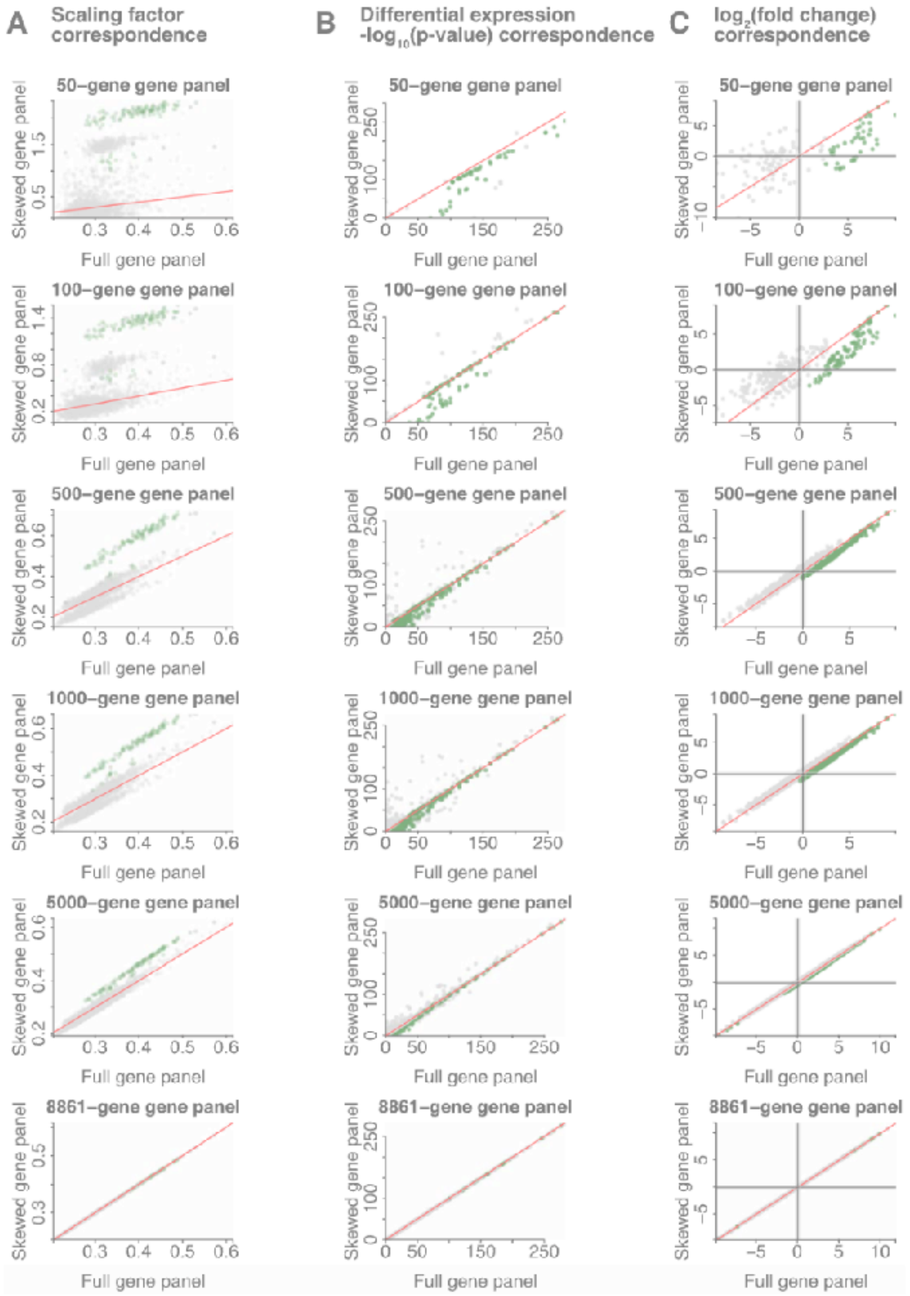
larger/more representative panels help **mitigate region-specific effects**



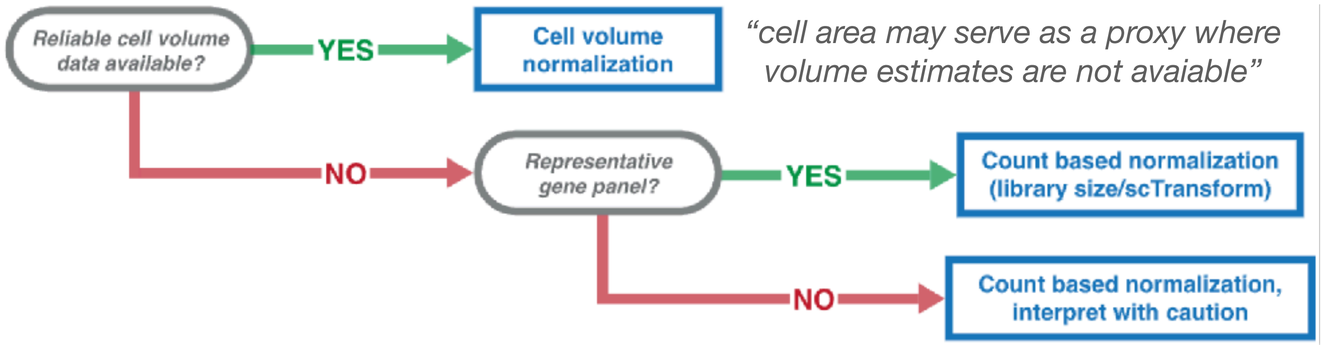
- skewed panels of 50…5,000 vs. all genes (simulated based on scRNA-seq data)
- differences are observed for skewed panels of all sizes, but their extent decreases as panel size increases

Atta *et al.* (2024), *Genome Biology* 25:153

larger/more representative panels help **mitigate region-specific effects**

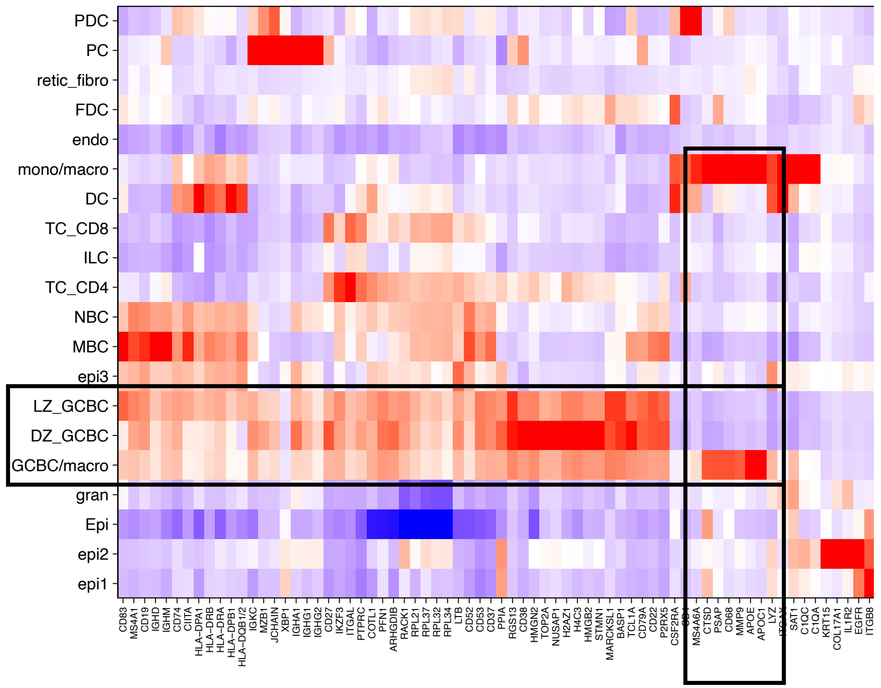


- skewed panels of 50…5,000 vs. all genes (simulated based on scRNA-seq data)
- differences are observed for skewed panels of all sizes, but their extent decreases as panel size increases



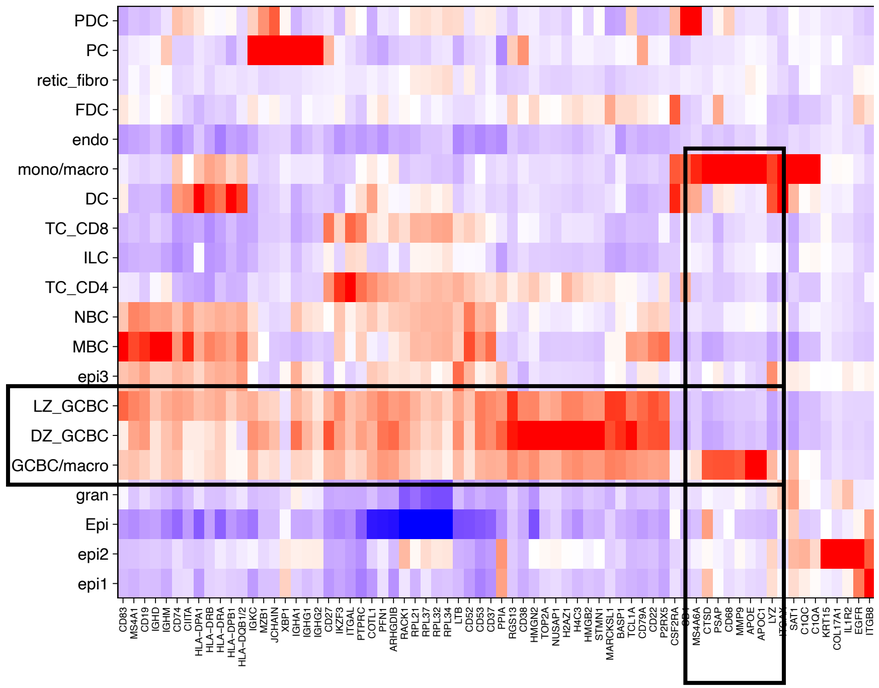
Atta *et al.* (2024), *Genome Biology* 25:153

CD68+ B cells — what's going on here?

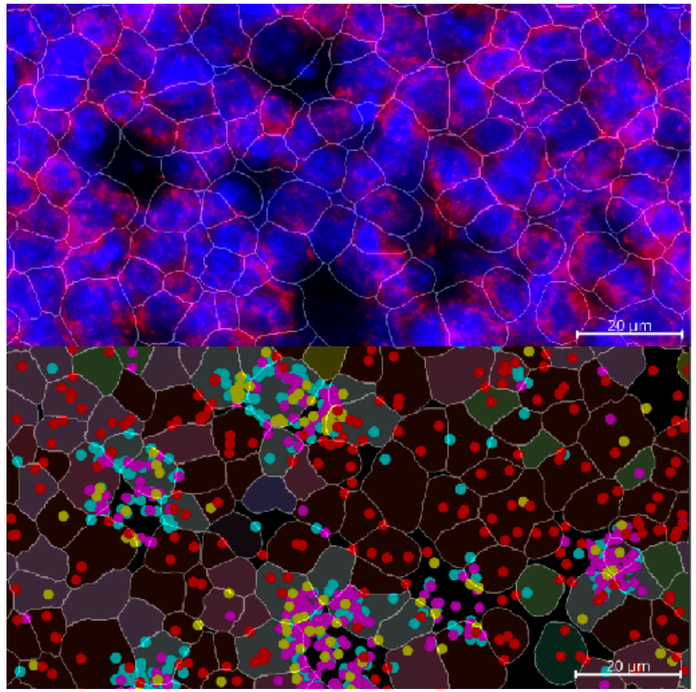


Crowell *et al.* Campo & Pascual-Reguant (in prep.)

CD68+ B cells — what's going on here?



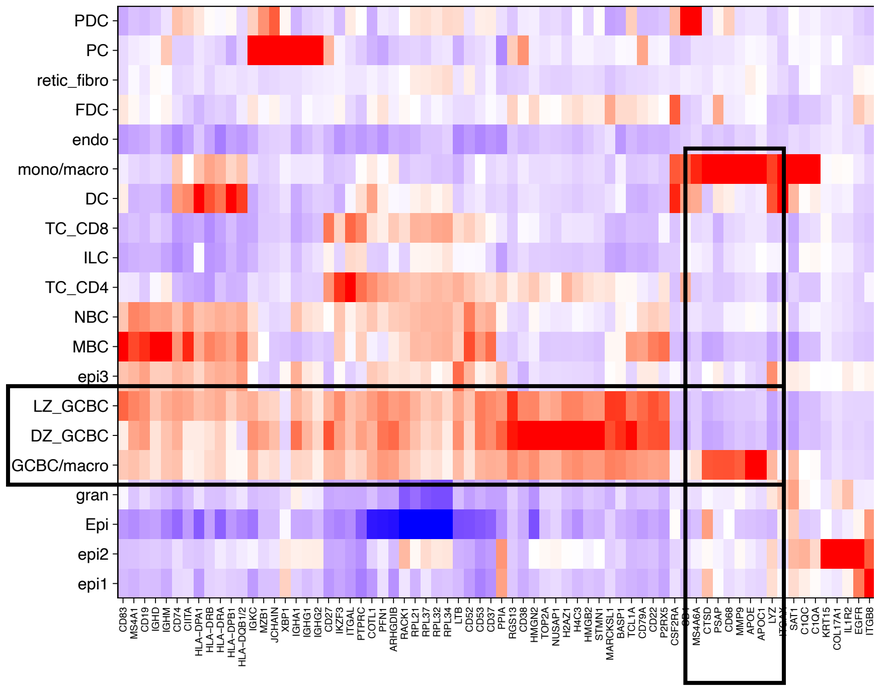
DAPI CD45



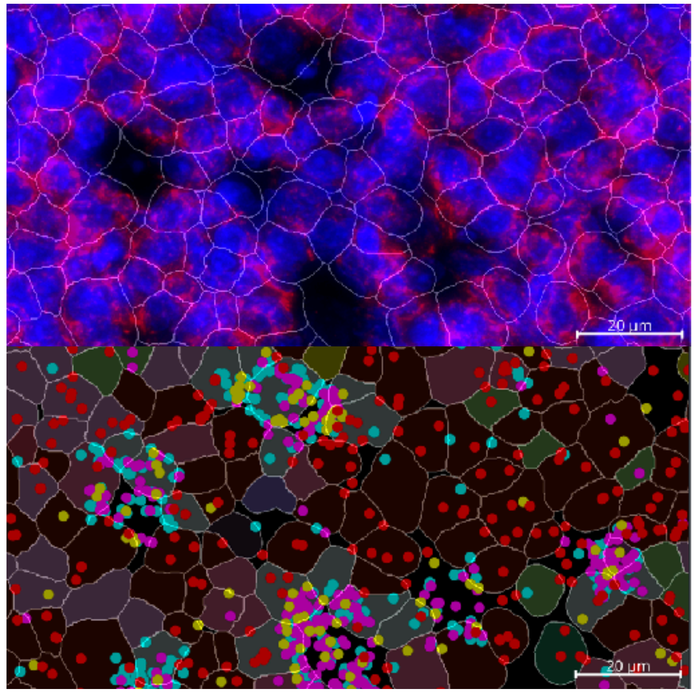
CD19 MMP9 CD68 LYZ

Crowell *et al.* Campo & Pascual-Reguant (in prep.)

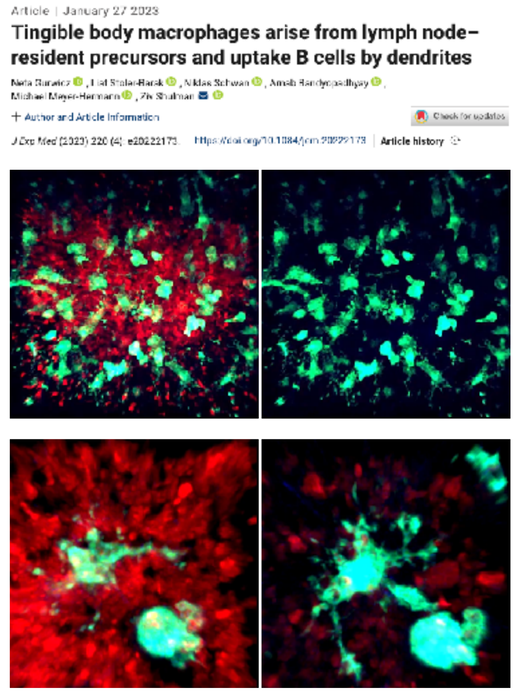
CD68+ B cells — what's going on here?



DAPI CD45



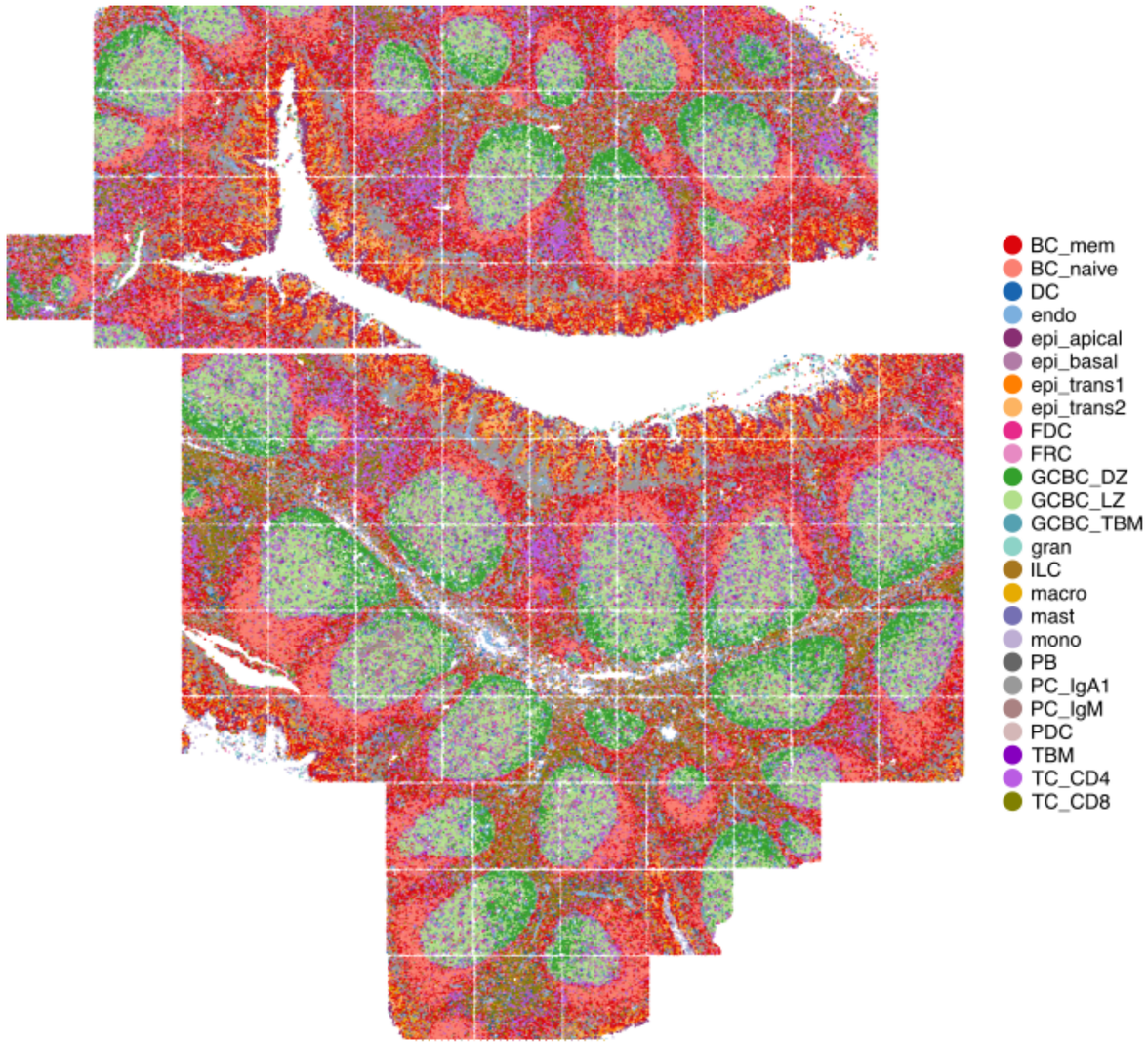
CD19 MMP9 CD68 LYZ



B cells macrophages

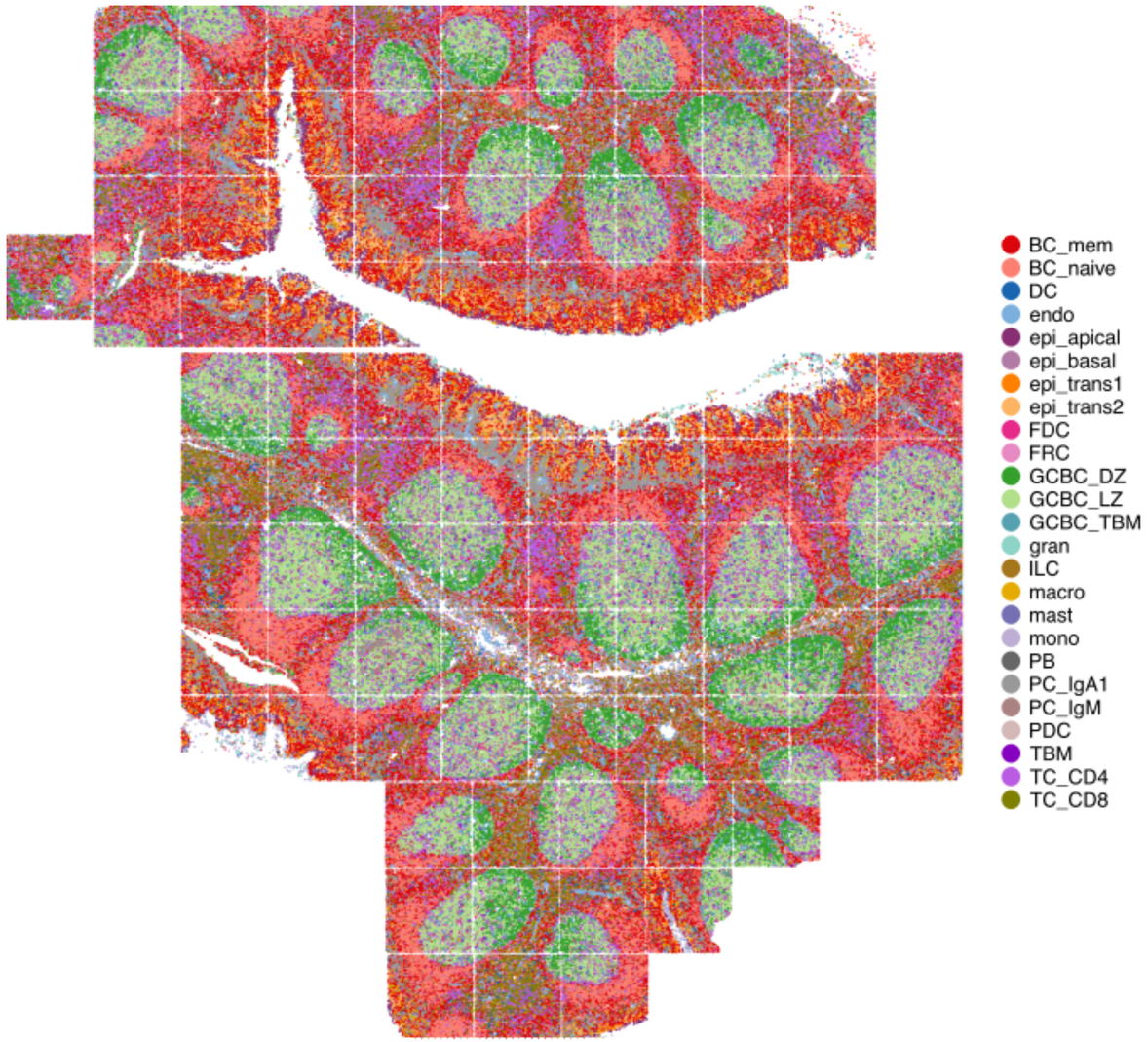
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

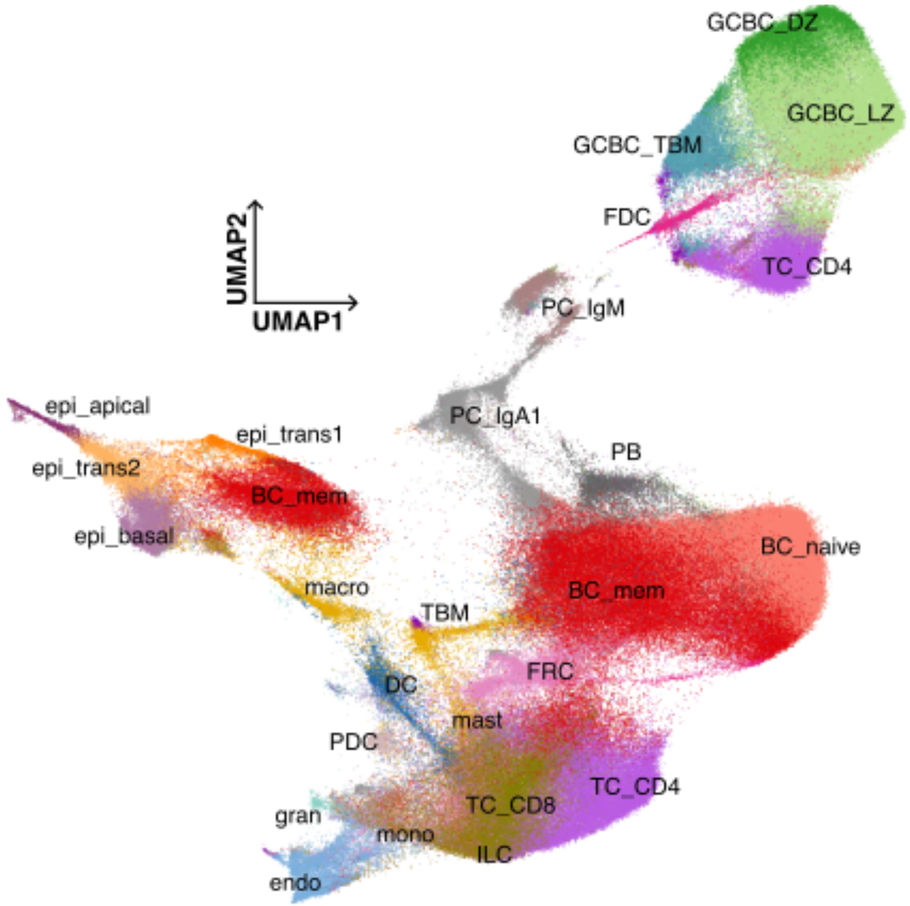
spatial bleeding manifests in RNA counts, hence PCs, and UMAPs



Crowell *et al.* Campo & Pascual-Reguant (in prep.)

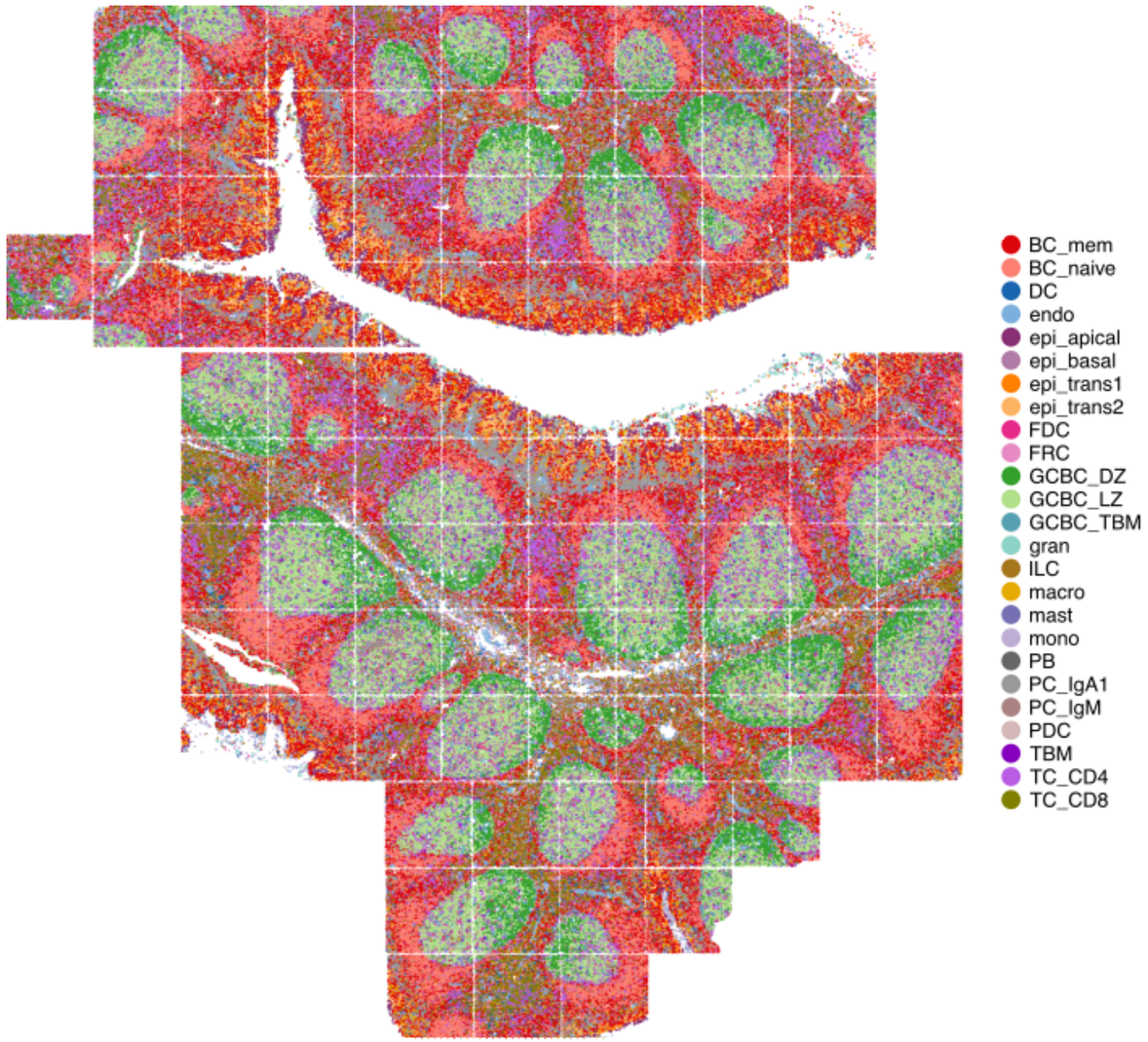
spatial bleeding manifests in RNA counts, hence PCs, and UMAPs

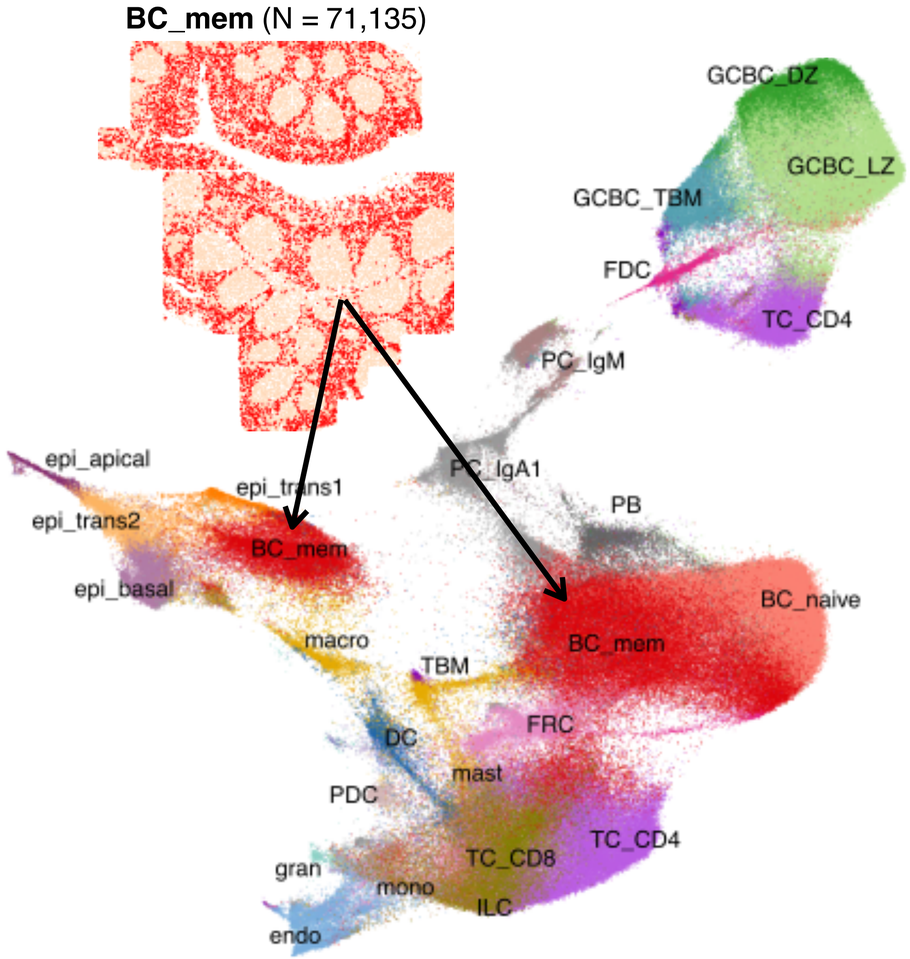




Crowell *et al.* Campo & Pascual-Reguant (in prep.)

spatial bleeding manifests in RNA counts, hence PCs, and UMAPs

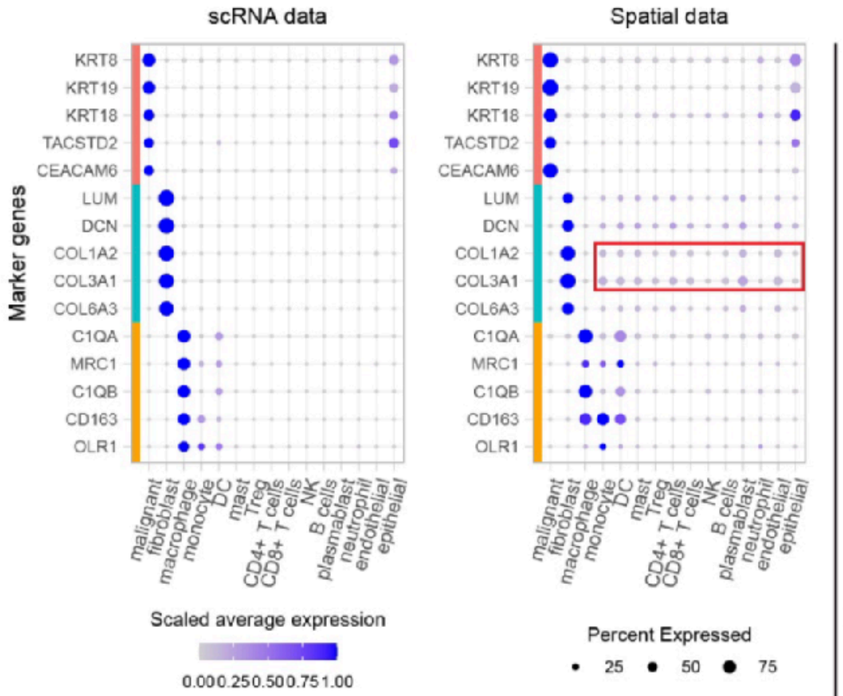




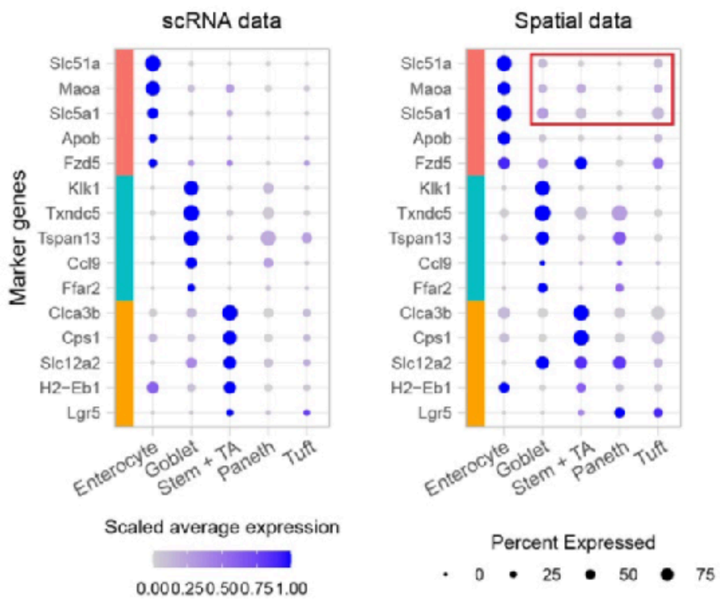
Crowell *et al.* Campo & Pascual-Reguant (in prep.)

bleeding occurs in 3D — around, above & below cells

malignant fibroblast macrophage



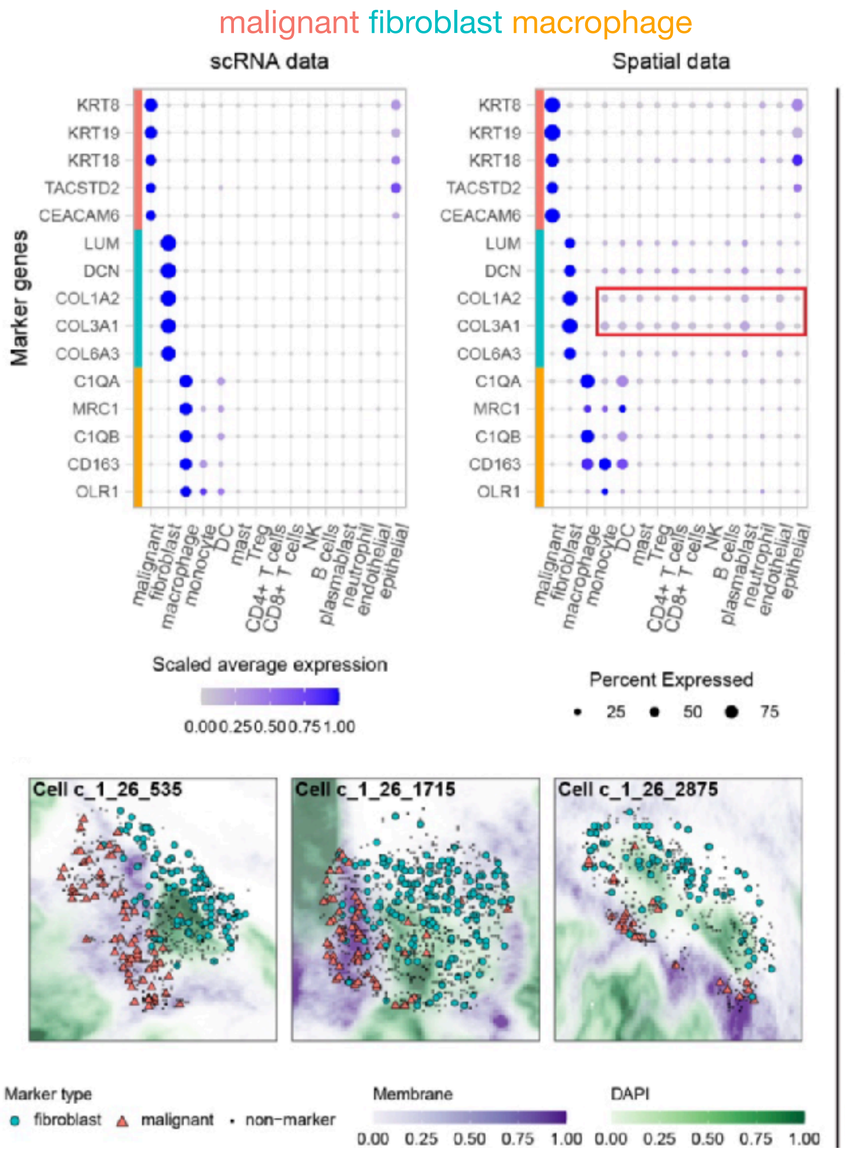
enterocyte goblet stem+TA

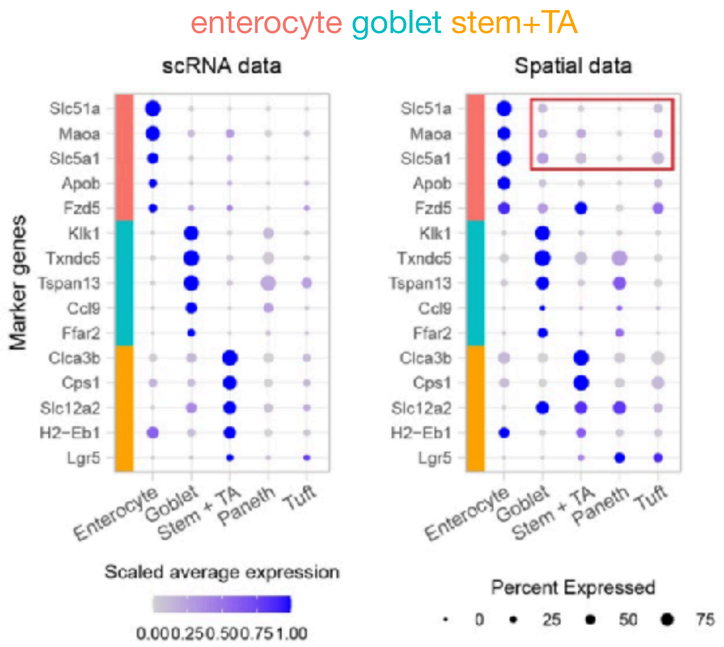


Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

bleeding occurs in 3D — around, above & below cells

admixtures occur at **cellular periphery**

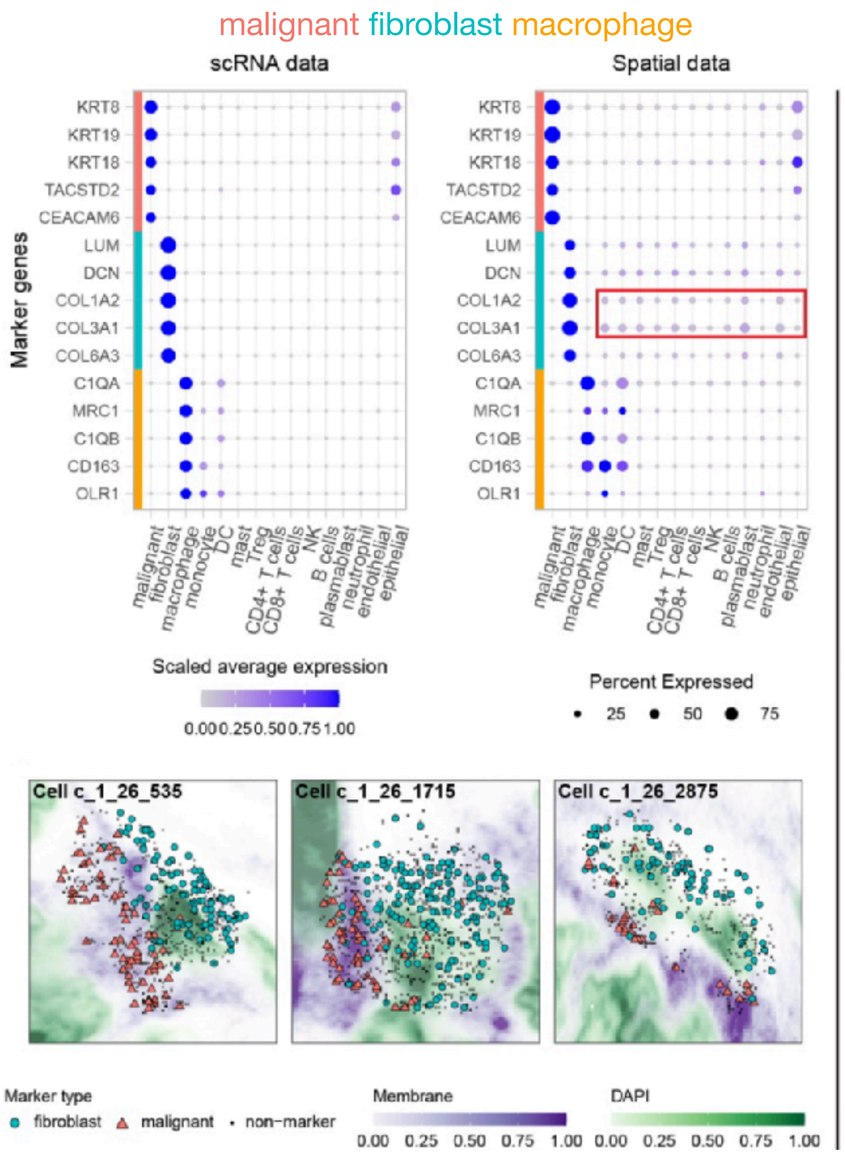


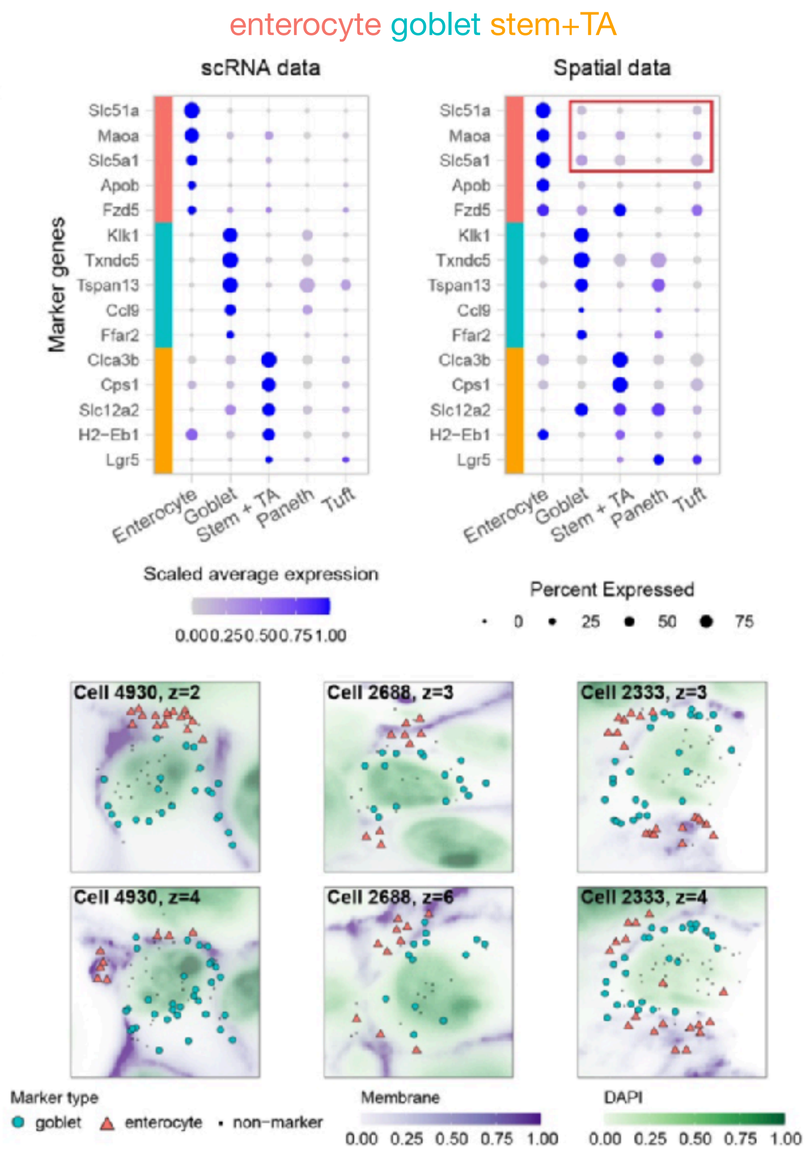


Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

bleeding occurs in 3D — around, above & below cells

admixtures occur at **cellular periphery**



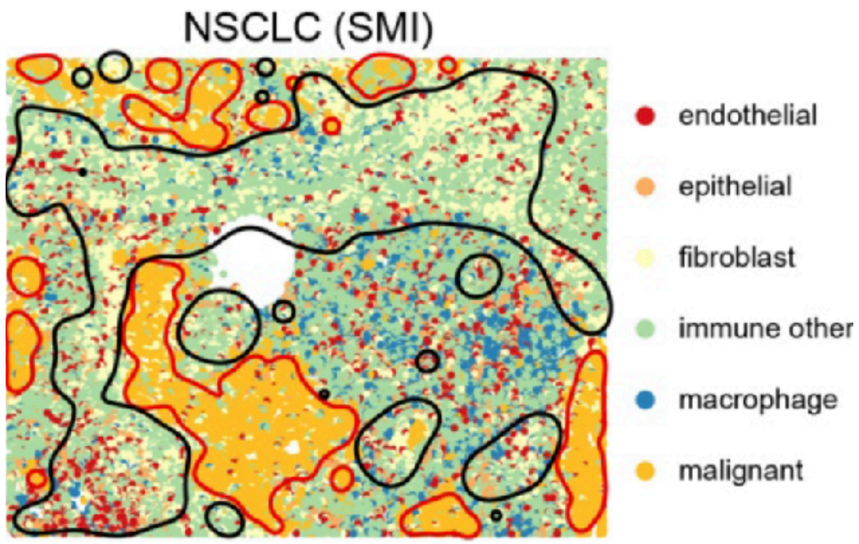


admixtures occur at **upper/lower z-plane**

Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

DE genes between regions reflect compositional differences (not differences in state)

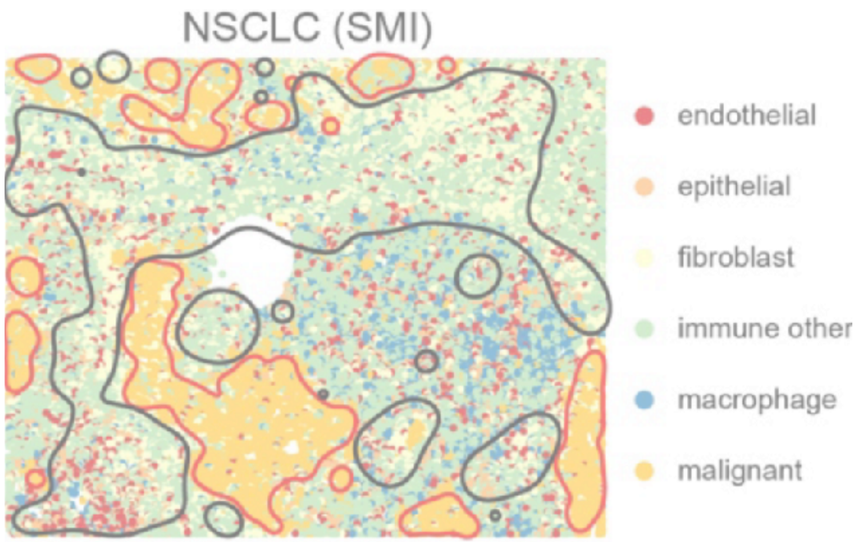
- manual annotation into **tumor** & **stromal** regions



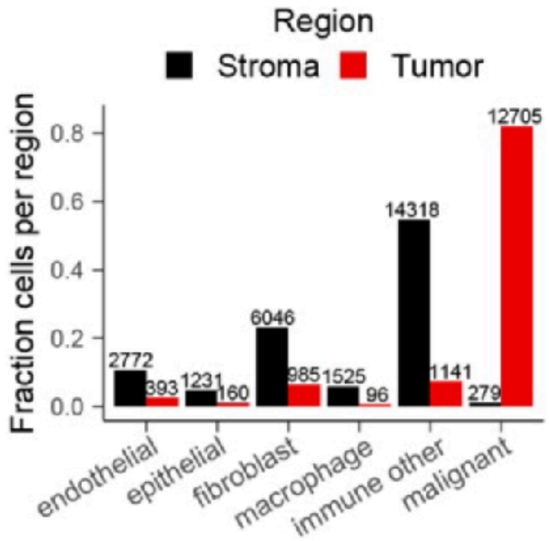
Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

DE genes between regions reflect compositional differences (not differences in state)

- manual annotation into **tumor** & **stromal** regions



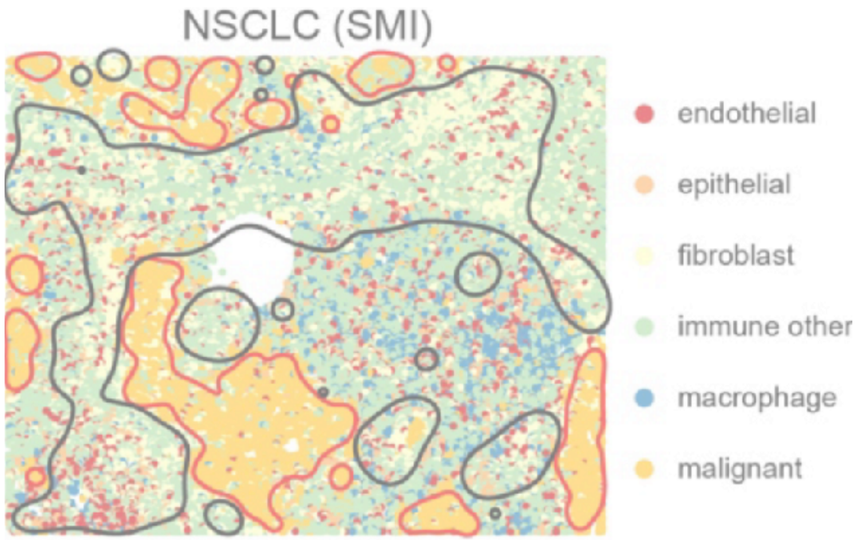
- tumor is dominated by malignant, stroma is dominated by other cells



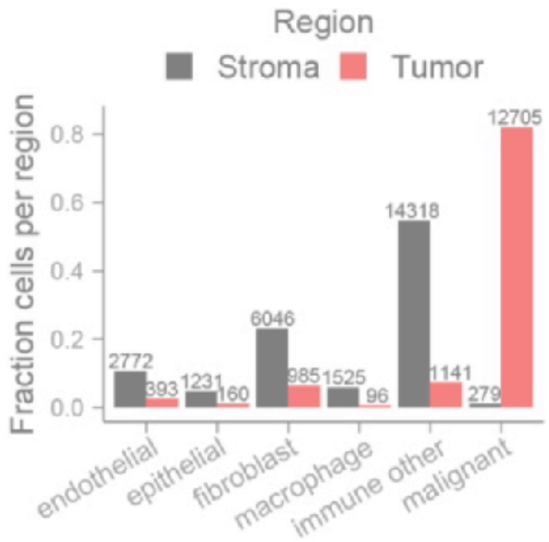
Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

DE genes between regions reflect compositional differences (not differences in state)

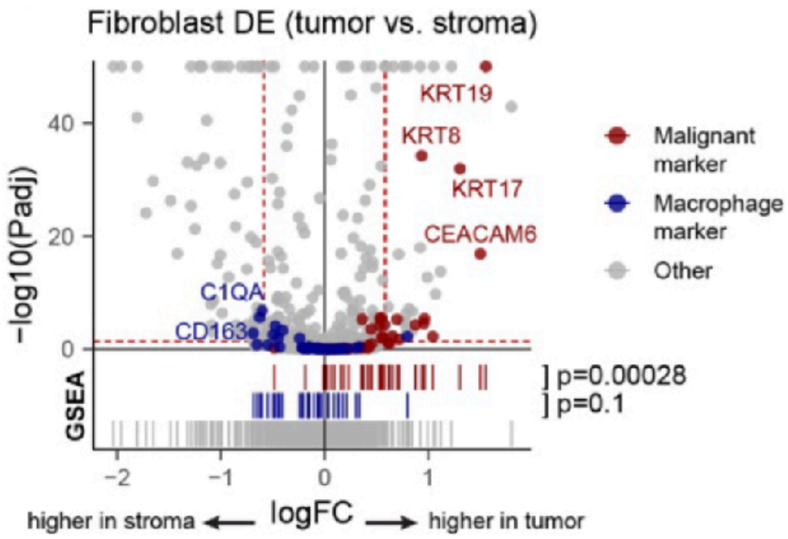
- manual annotation into **tumor** & **stromal** regions



- tumor is dominated by malignant, stroma is dominated by other cells



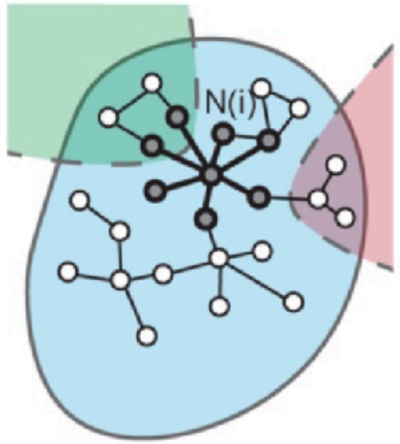
- comparing regions, genes upregulated in fibroblasts are **epithelial markers**



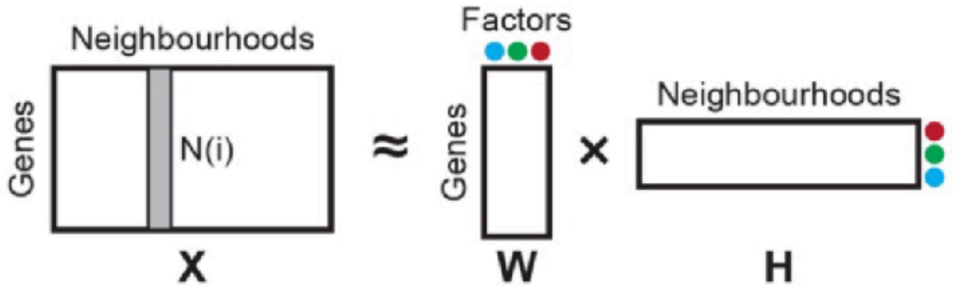
Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

Mitchel *et al.* propose **NMF + CRF clean-up** to mitigate spatial bleeding

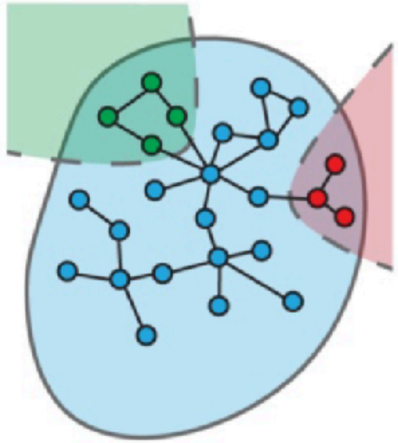
1. Construct KNN graph per cell



2. Recover pure expression profiles using weighted NMF



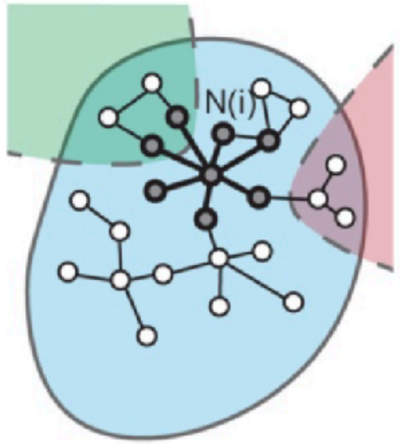
3. Label admixture molecules using CRF



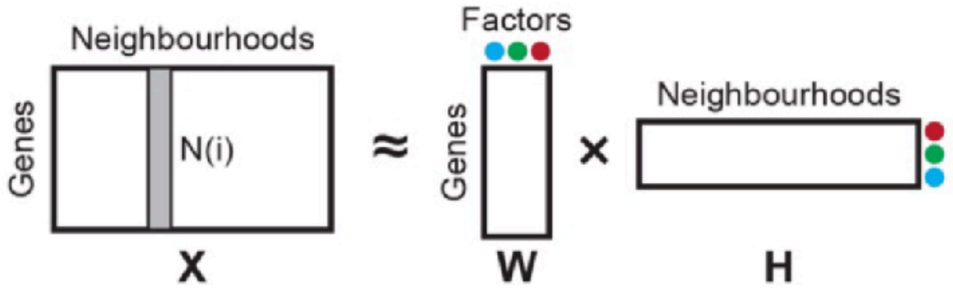
Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

Mitchel *et al.* propose **NMF + CRF clean-up** to mitigate spatial bleeding

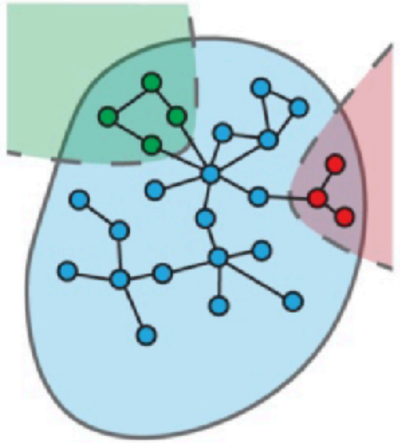
1. Construct KNN graph per cell

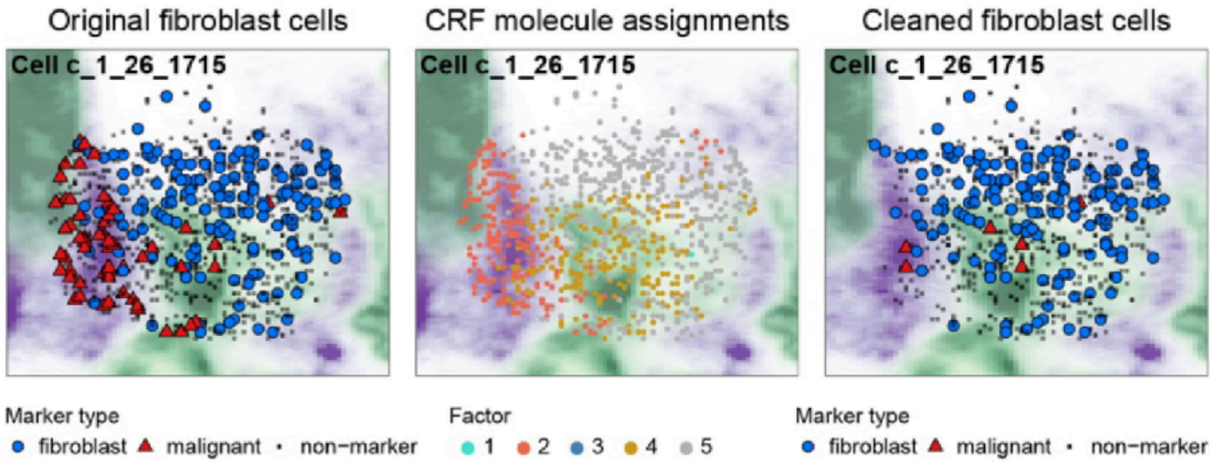


2. Recover pure expression profiles using weighted NMF



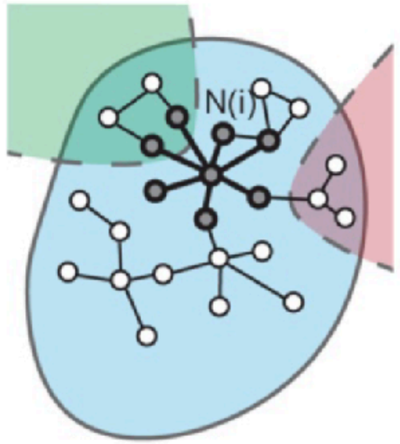
3. Label admixture molecules using CRF



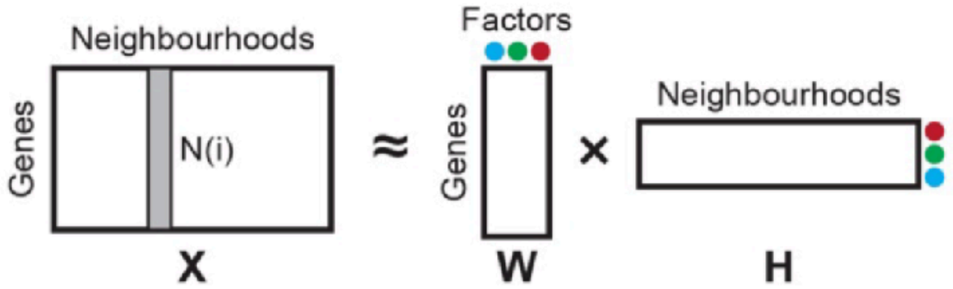


Mitchel *et al.* propose **NMF + CRF clean-up** to mitigate spatial bleeding

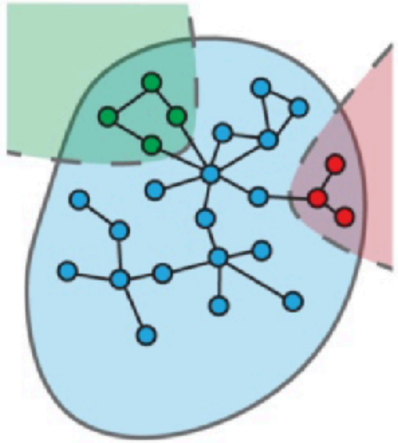
1. Construct KNN graph per cell



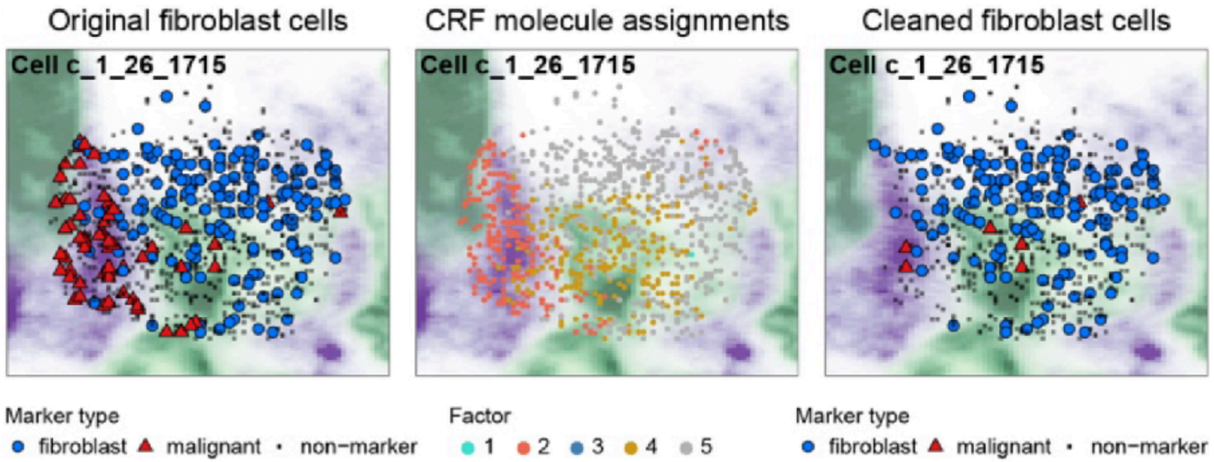
2. Recover pure expression profiles using weighted NMF



3. Label admixture molecules using CRF



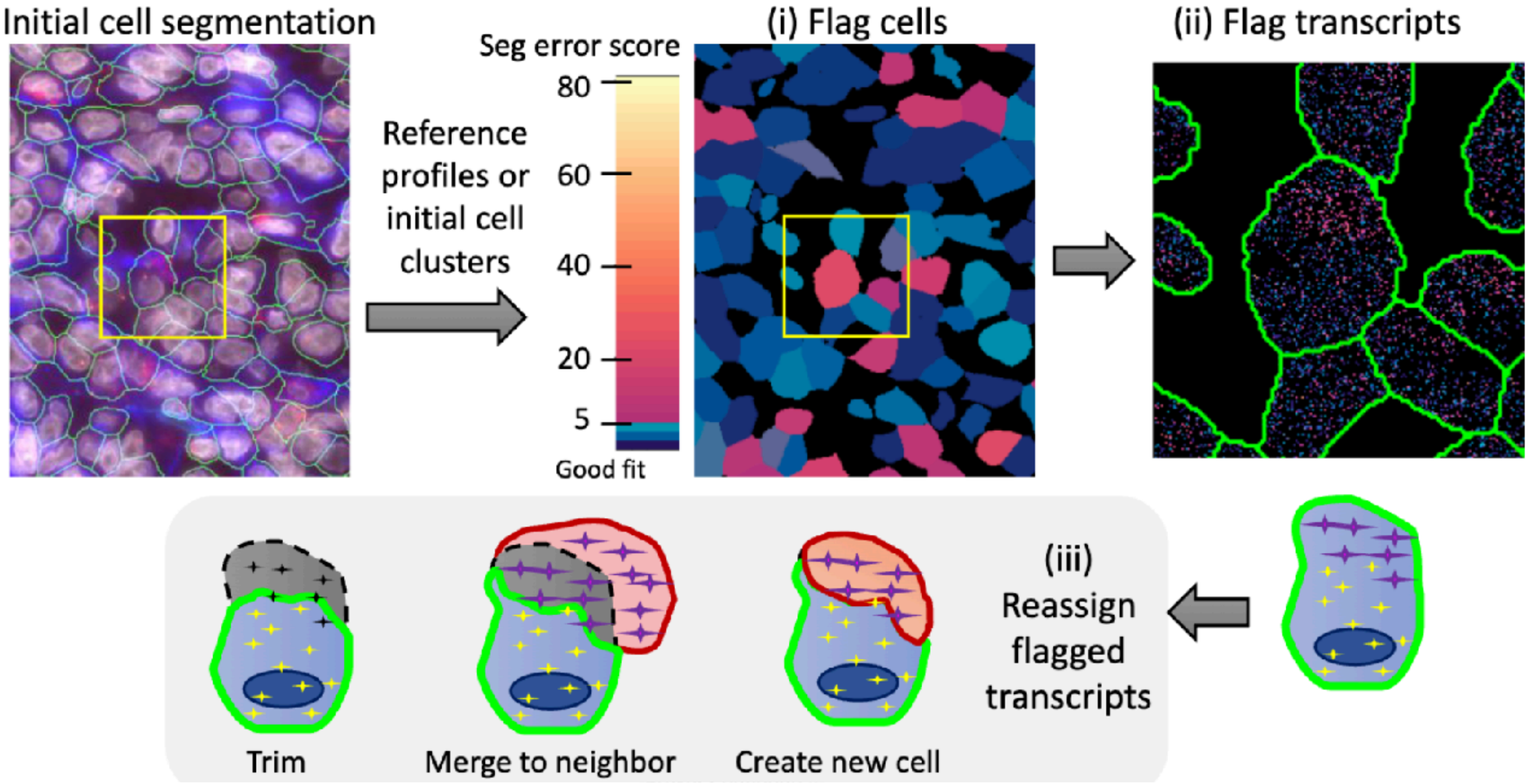
- **subcellular features** can include
  - recurrent **admixture patterns** (e.g., between frequently co-occuring cell types)
  - true **cellular structures** (e.g., ER, nuclei, polarization)



Mitchel *et al.* (2024), *BioRxiv* 2025.01.02.631135

FastReseg uses transcript locations to refine img-based segmentation

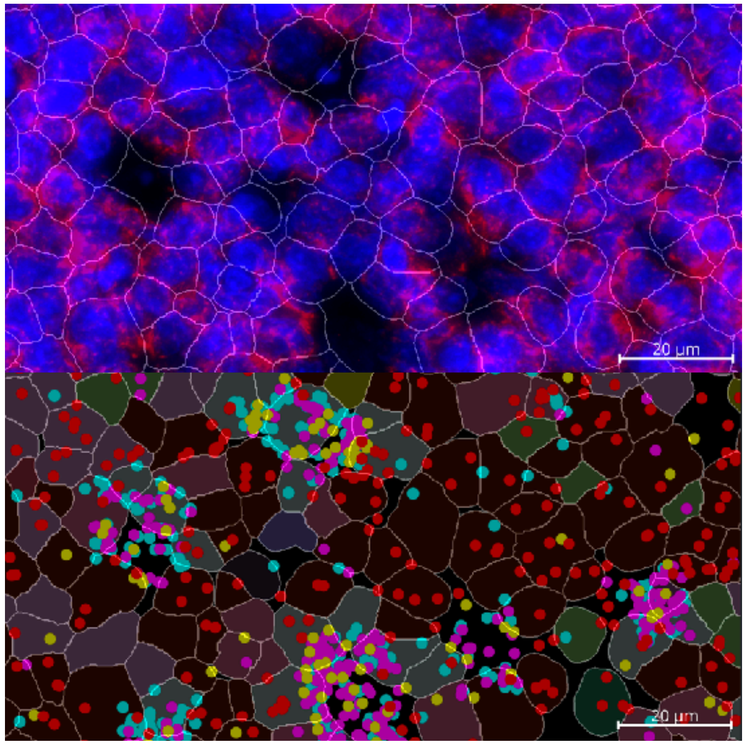
- **transcript scoring**
  based on initial host cell
- **flag spatial doublets**
  as putative segmentation errors
- **flag misassigned transcripts**
  within flagged cells only
- **correct counts** (but not segmentation boundaries)



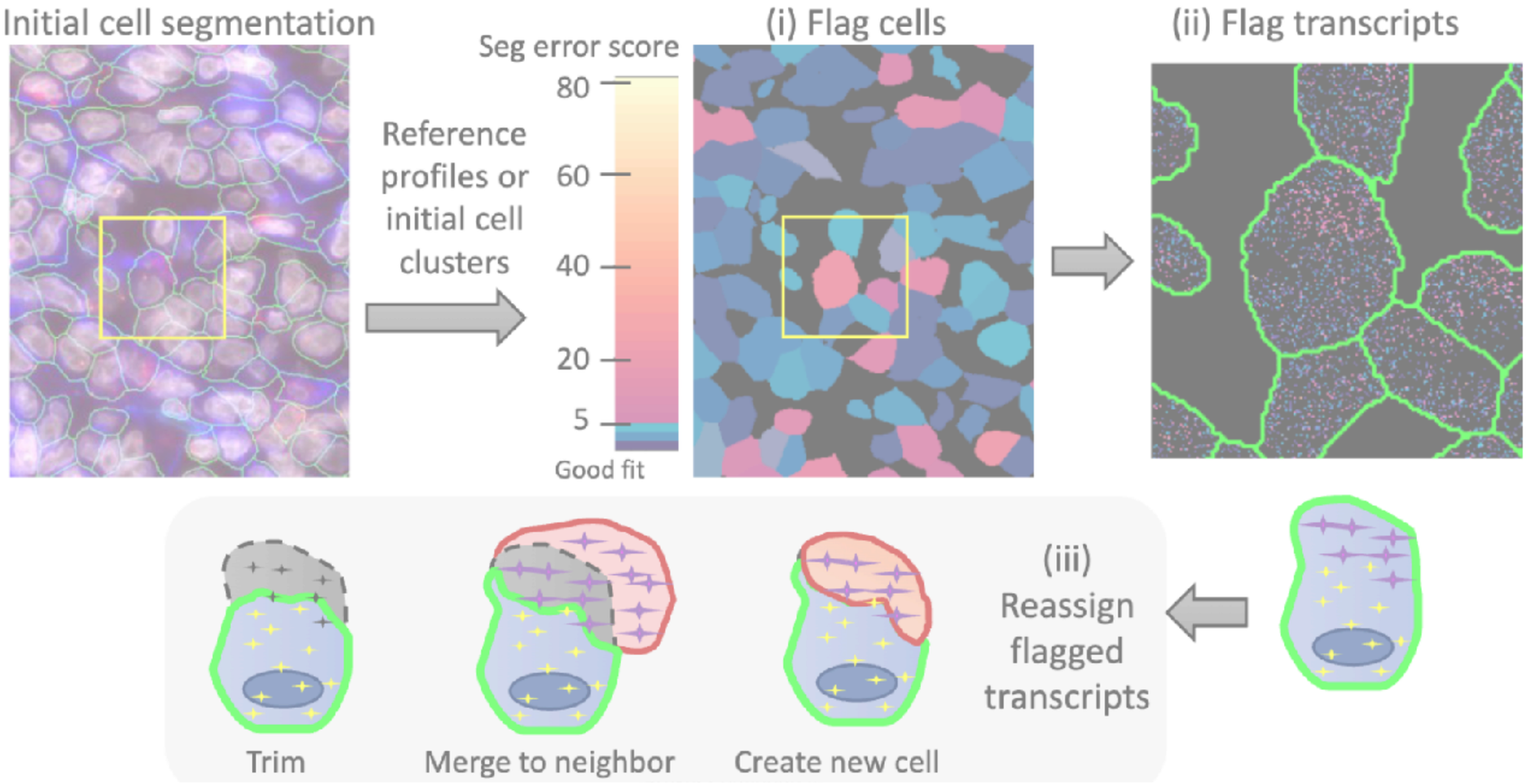
Wu *et al.* (2024), *BioRxiv* 2024.12.05.627051

FastReseg uses transcript locations to refine img-based segmentation

- flag B cells surrounding black holes
- flag macrophage-related genes
- correct B cell counts & create new macrophages

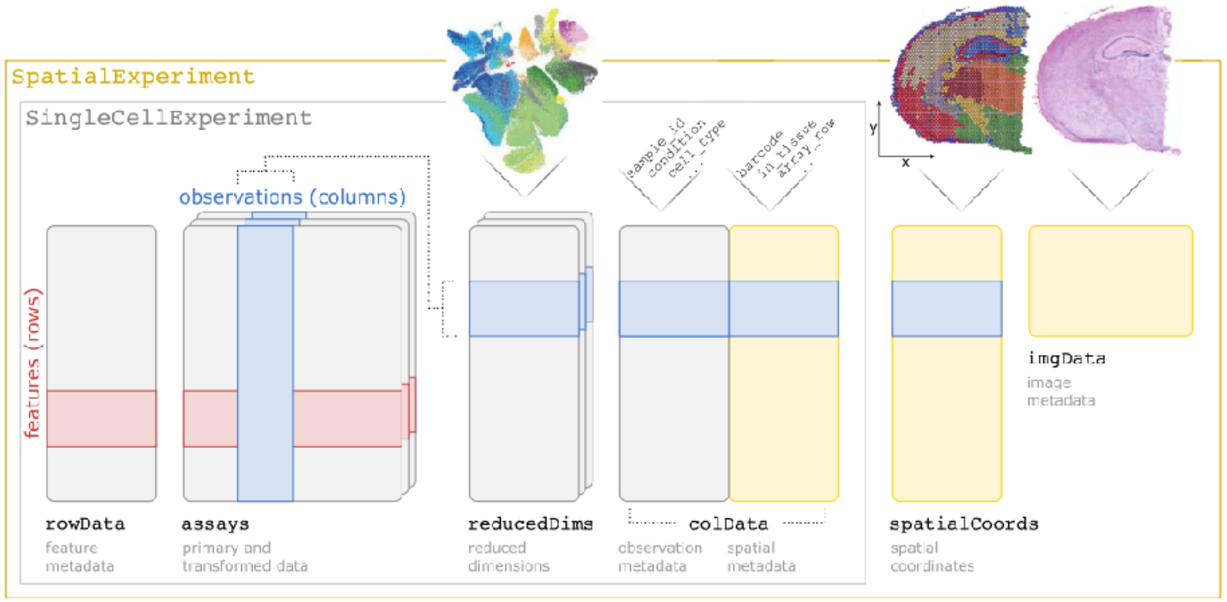


CD19 MMP9 CD68 LYZ



Wu *et al.* (2024), *BioRxiv* 2024.12.05.627051

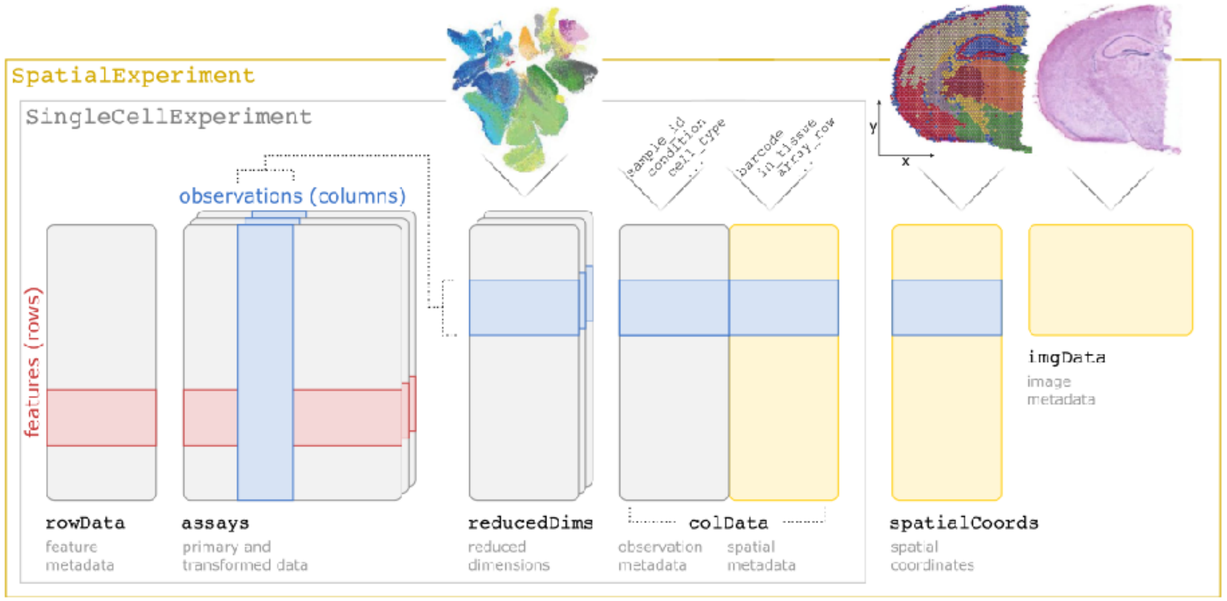
infrastructure for handling img-ST data **in R**



Righelli, Weber, Crowell *et al.* (2022)
Bioinformatics 38(11):3128-3131

SpatialExperiment

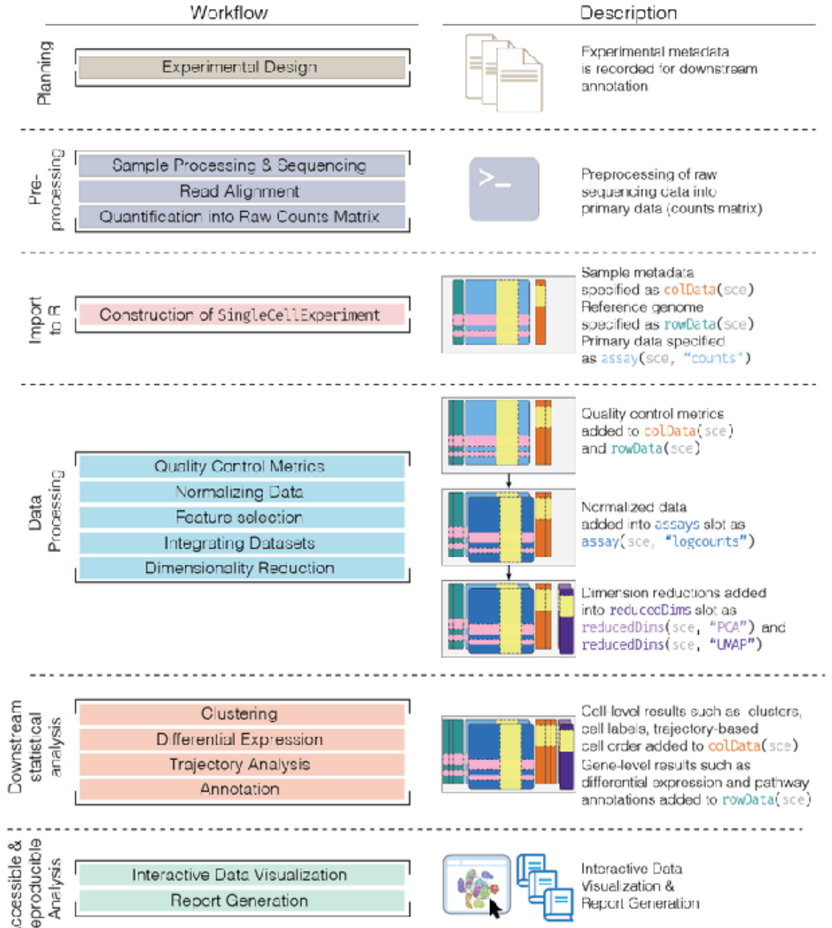
infrastructure for handling img-ST data **in R**



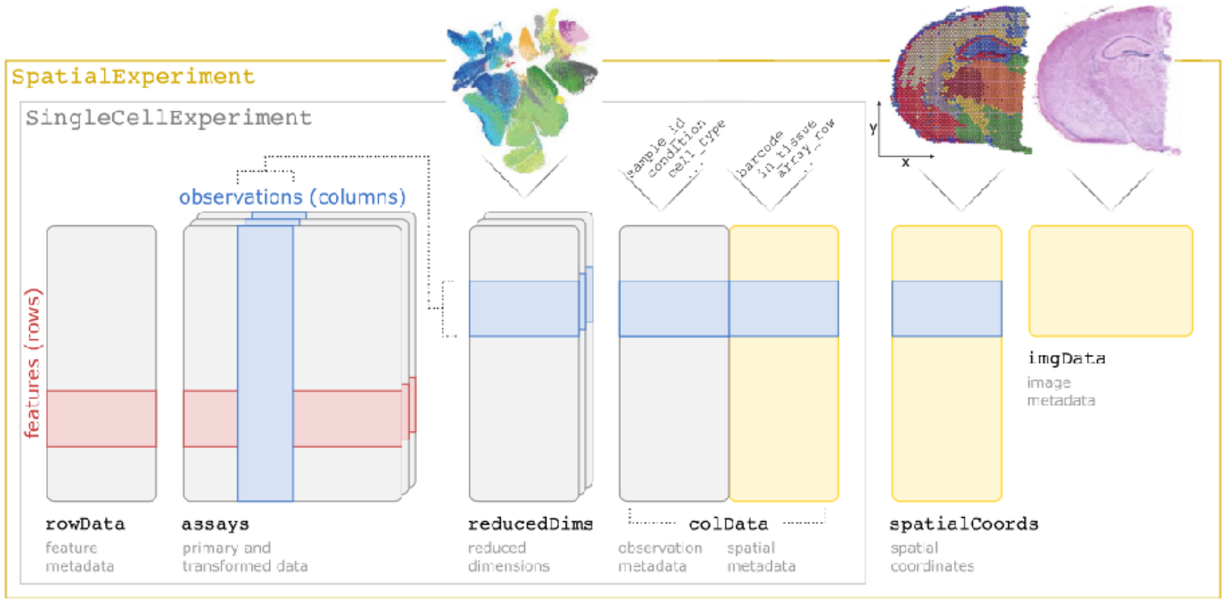
Righelli, Weber, Crowell *et al.* (2022)
Bioinformatics 38(11):3128-3131

SpatialExperiment

Why? Because infrastructure around single-cell data is large!



infrastructure for handling img-ST data **in R**

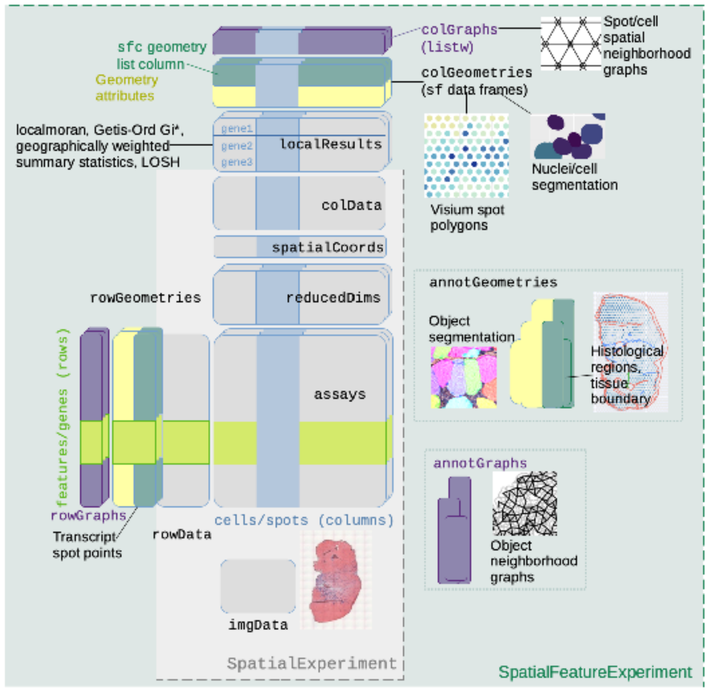


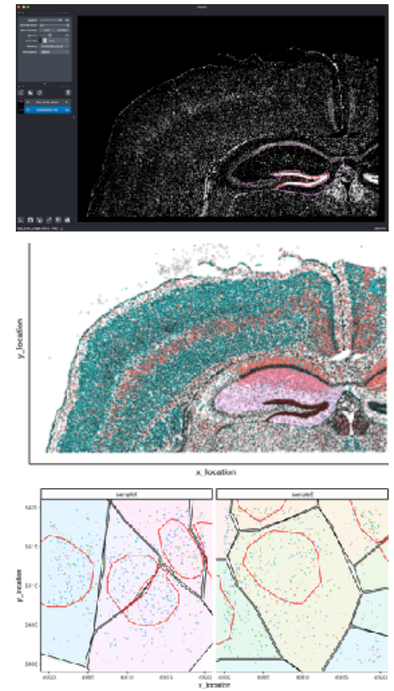
Righelli, Weber, Crowell *et al.* (2022)
Bioinformatics 38(11):3128-3131

SpatialExperiment

Moses *et al.* (2023) *bioRxiv* 2023.07.20.549945

SpatialFeatureExperiment

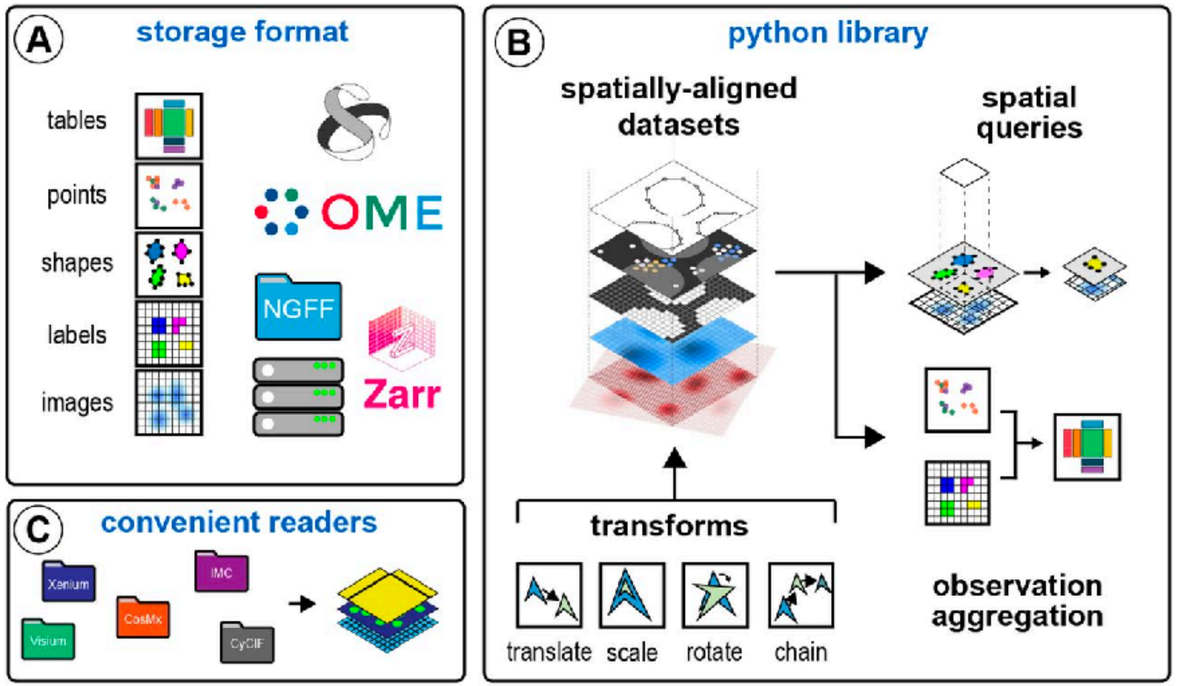




MoleculeExperiment

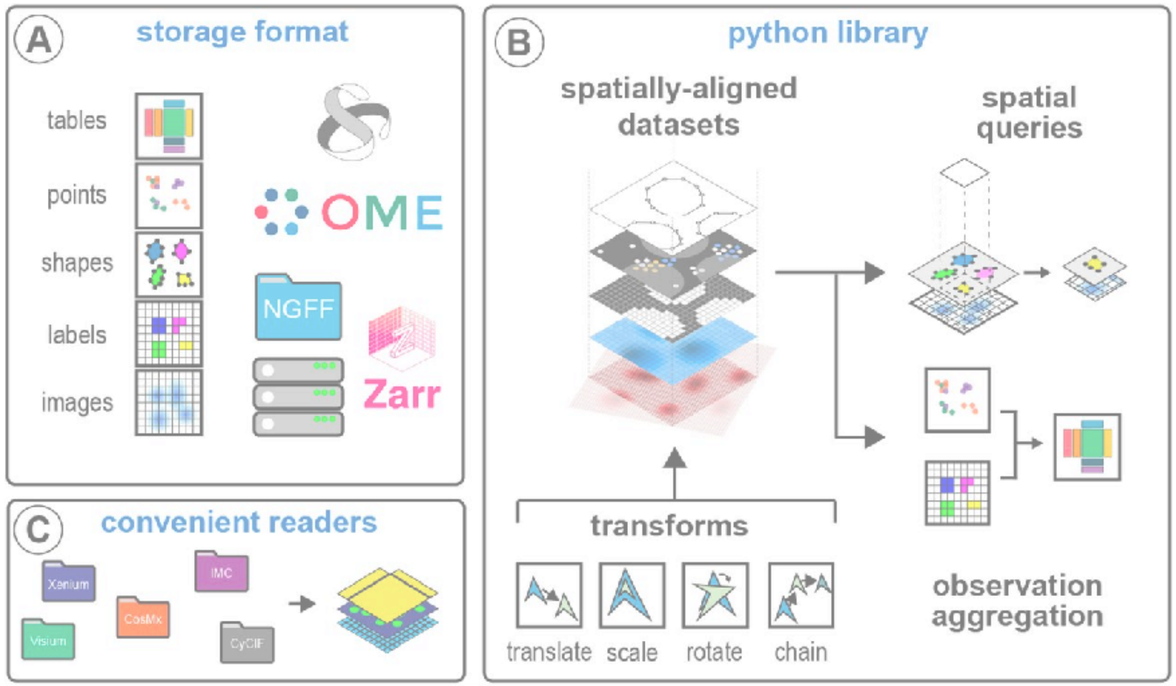
Couto *et al.* (2023) *bioRxiv* 2023.05.16.541040

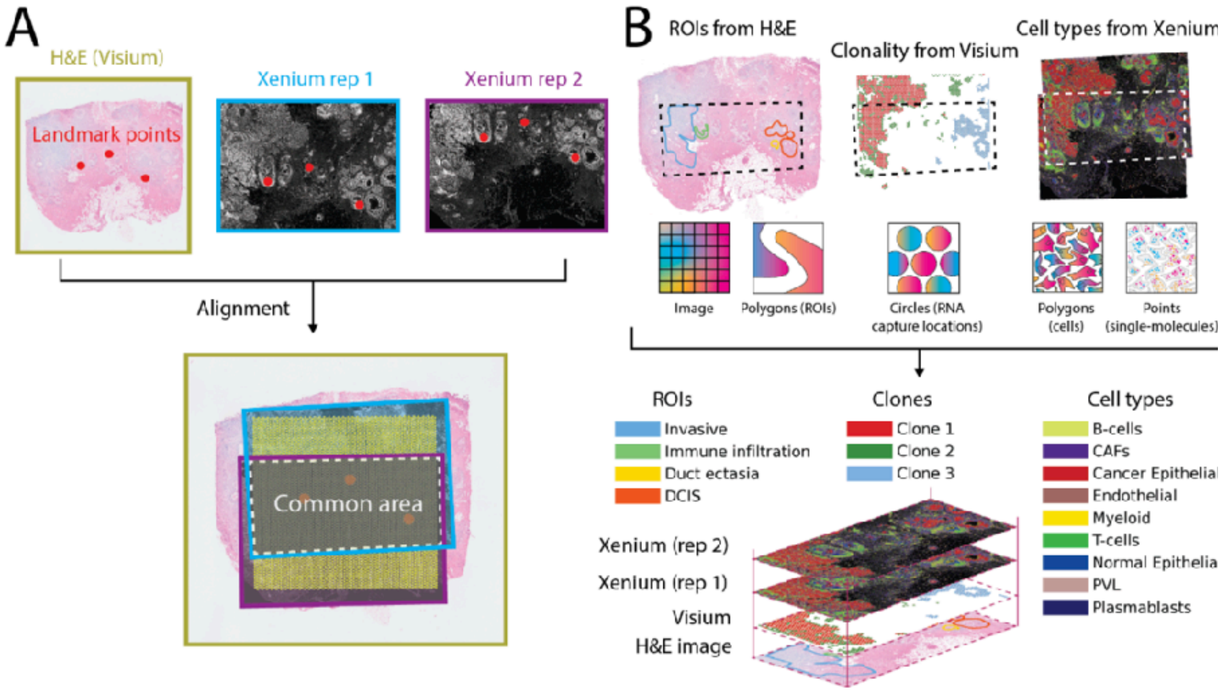
infrastructure for handling img-ST data **in Python**



Marconato *et al.* (2024) *Nature Methods* s41592-024-02212-x

infrastructure for handling img-ST data **in Python**





Marconato *et al.* (2024) *Nature Methods* s41592-024-02212-x

appendix — references & resources

technology
- He *et al.* (2022). High-plex Multiomic Analysis in FFPE at Subcellular Level by Spatial Molecular Imaging. *bioRxiv* 2021.11.03.467020
- Khafizov *et al.* (2024). Sub-cellular imaging of the entire protein-coding human transcriptome (18933-plex) on FFPE tissue using SMI. *bioRxiv* 2024.11.27.625536

segmentation
- Stringer *et al.* (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods* 18(1):100-106
- Petukhov *et al.* (2021). Cell segmentation in imaging-based spatial transcriptomics. *Nature Biotechnology* 40:345-354
- Wu *et al.* (2024). FastReseg: using transcript locations to refine image-based cell segmentation results in spatial transcriptomics. *bioRxiv* 2024.12.05.627051
- Mitchel *et al.* (2024). Impact of Segmentation Errors in Analysis of Spatial Transcriptomics Data. *bioRxiv* 2025.01.02.631135

miscellaneous
- Martin *et al.* (2024). MerQuaCo: a computational tool for quality control in image-based spatial transcriptomics. *bioRxiv* 2024.12.04.626766
- NanoString scratch space: *https://nanostring-biostats.github.io/CosMx-Analysis-Scratch-Space*

normalization
- Bhuva *et al.* (2024): Library size confounds biology in spatial transcriptomics data. *Genome Biology* 25:99
- Atta *et al.* (2024). Gene count normalization in single-cell imaging-based spatially resolved transcriptomics. *Genome Biology* 25:153

infrastructure
- Righelli, Weber, Crowell *et al.* (2022). SpatialExperiment: infrastructure for spatially resolved transcriptomics data in R using Bioconductor. *Bioinformatics* 38(11):3128-3131
- Couto *et al.* (2023). MoleculeExperiment enables consistent infrastructure for molecule-resolved spatial transcriptomics data in Bioconductor. *bioRxiv* 2023.05.16.541040
- Moses *et al.* (2023). Voyager: exploratory single-cell genomics data analysis with geospatial statistics. *bioRxiv* 2023.07.20.549945
- Marconato *et al.* (2024). SpatialData: an open and universal data framework for spatial omics. *Nature Methods* s41592-024-02212-x

benchmarks
- Wang *et al.* (2023). Systematic benchmarking of imaging spatial transcriptomics platforms in FFPE tissues. *bioRxiv* 2023.12.07.570603
- Cook *et al.* (2023): A comparative analysis of imaging-based spatial transcriptomics platforms. *bioRxiv* 2023.12.13.571385
- Rademacher *et al.* (2024): Comparison of spatial transcriptomics technologies using tumor cryosections. *bioRxiv* 2024.04.03.586404
- Ren *et al.* (2024). Systematic Benchmarking of High-Throughput Subcellular Spatial Transcriptomics Platforms. *bioRxiv* 2024.12.23.630033