

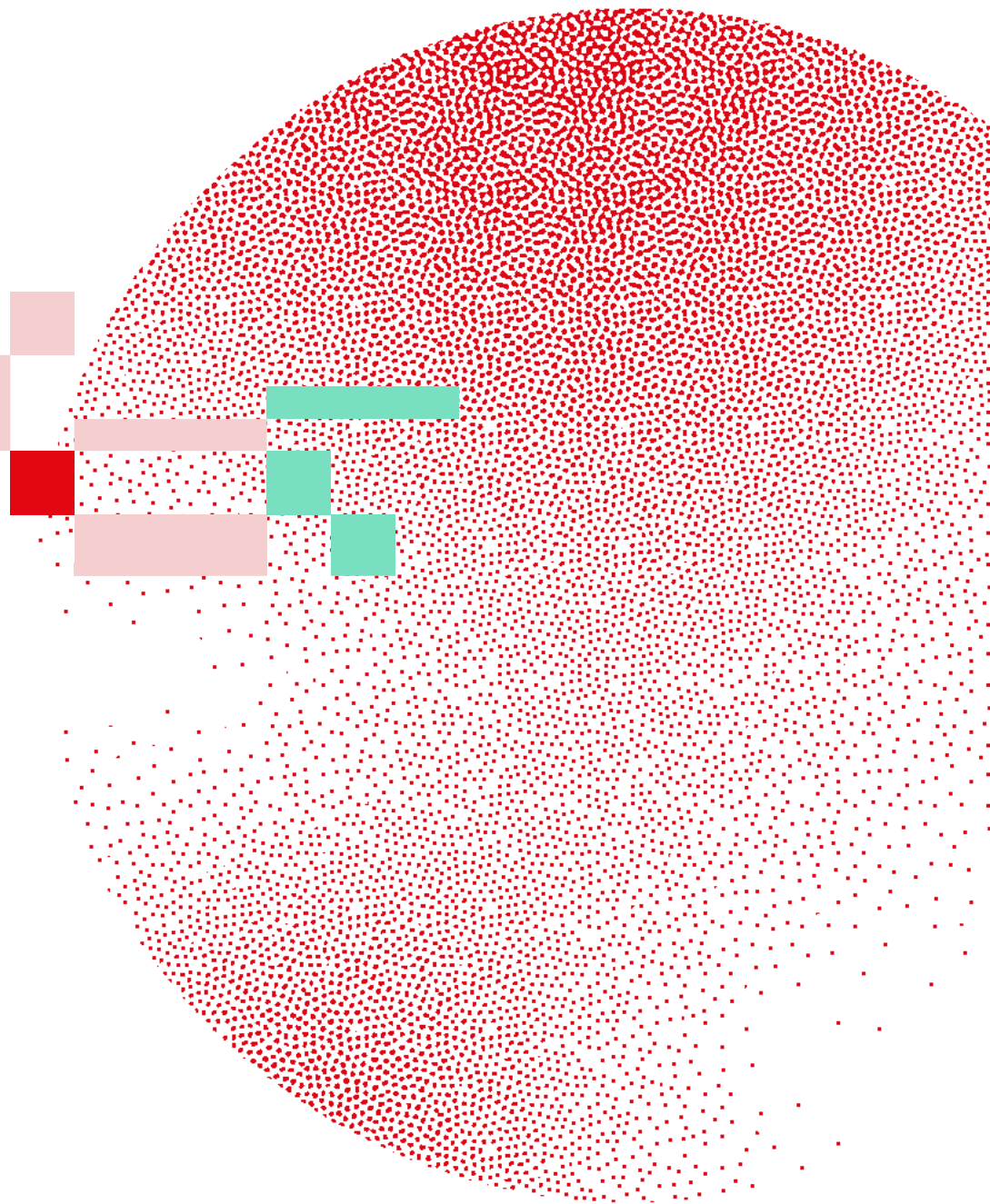


Swiss Institute of
Bioinformatics

INTRODUCTION TO SPATIAL TRANSCRIPTOMICS DATA ANALYSIS

QC (sequencing-based)

Deepak Tanwar
June 03-05, 2026



Which QC do you use for single-cell data?

QC for Visium HD is conceptually the same as single-cell QC – but computed *per High-Definition spot* rather than per cell.

Cell → spots/bins or cell segmentation

Why QC of spatial data?

High-density spots can capture technical artifacts (folds, bubbles, staining issues).

Why QC of spatial data?

High-density spots can capture technical artifacts (folds, bubbles, staining issues).

Artifacts are often **spatially coherent** – affecting contiguous regions, not just individual spots.

Why QC of spatial data?

High-density spots can capture technical artifacts (folds, bubbles, staining issues).

Artifacts are often **spatially coherent** – affecting contiguous regions, not just individual spots.

Poor-quality regions can bias downstream analyses (clustering, spatial DE, trajectory analysis).

QC in Visium HD

Same metrics as single-cell RNA-seq:

- Total UMIs per spot

- Number of genes per spot

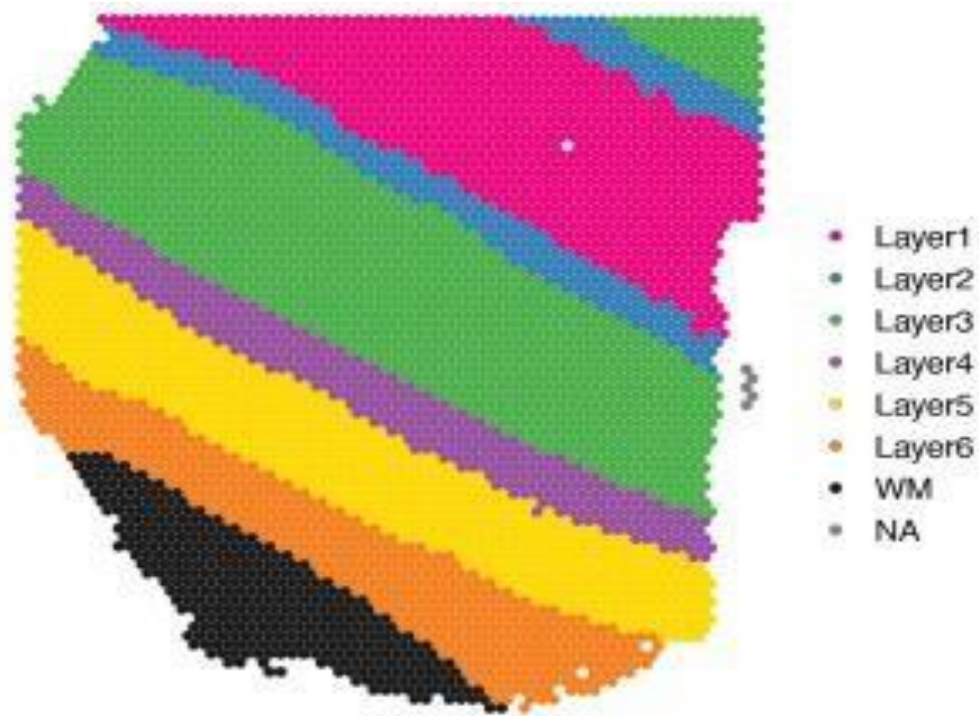
- Mitochondrial fraction

Key difference: unit is HD spot, not single cell.

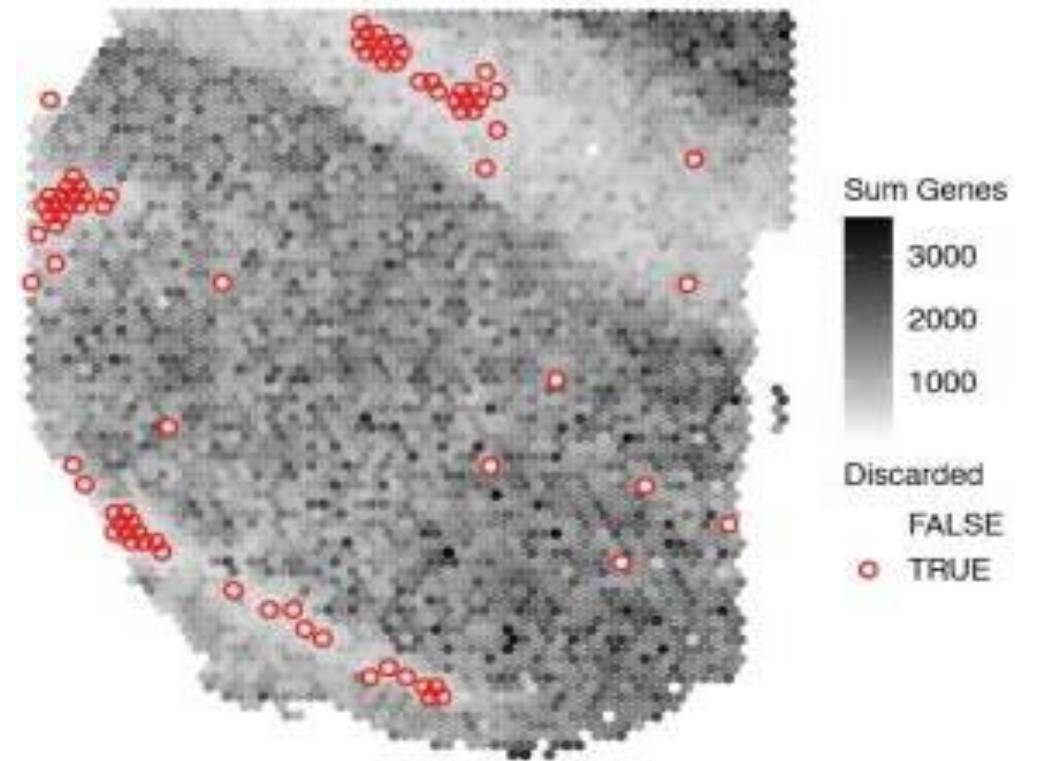
Neighboring spots share microenvironment → spatial artifacts can emerge

Example of spots removal

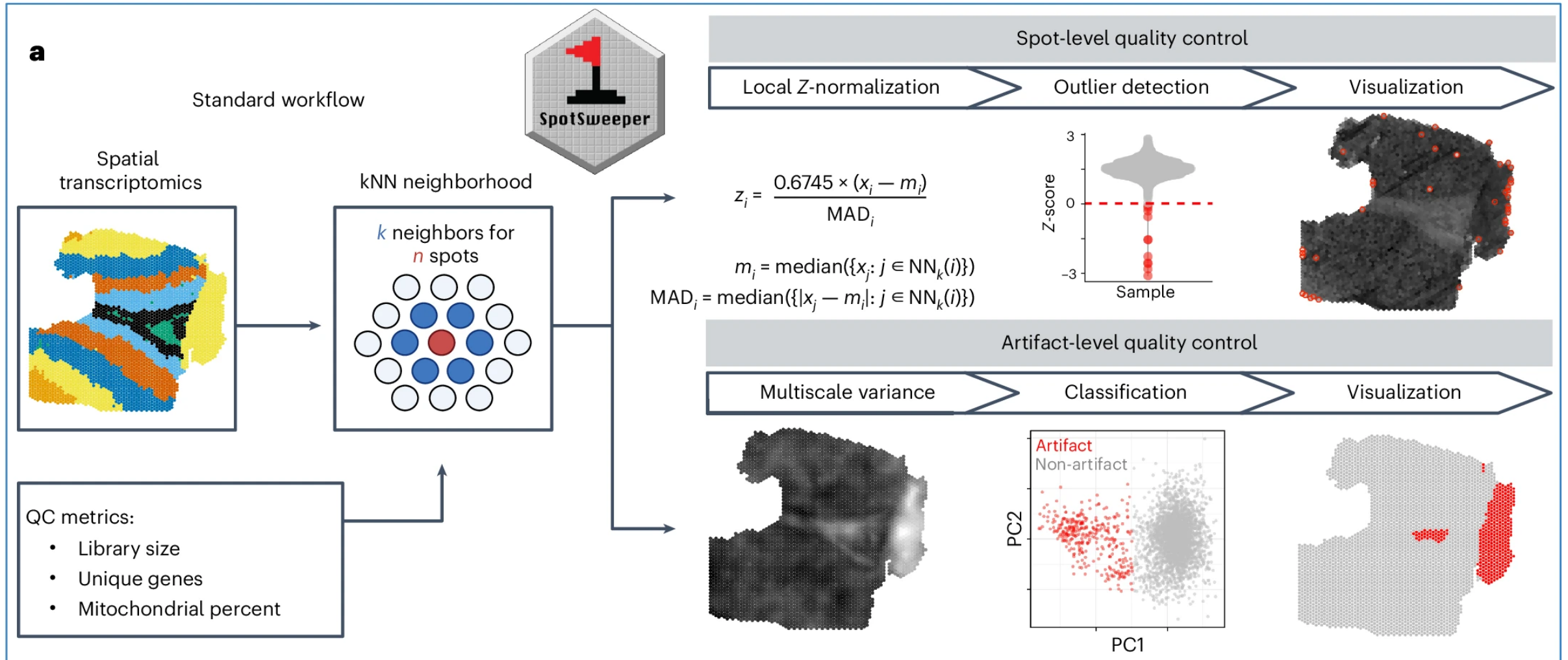
Annotation



Global

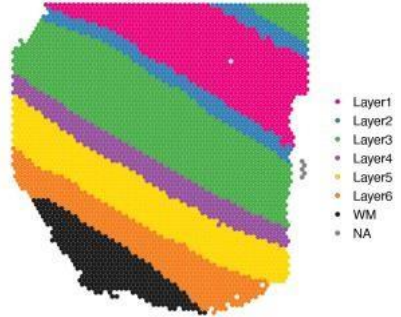


SpotSweeper: spatially aware QC

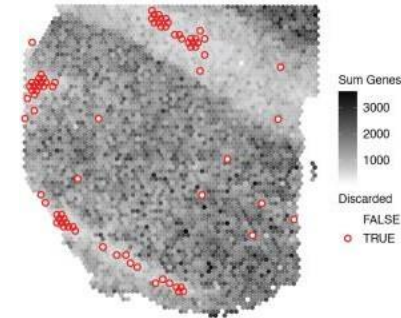


SpotSweeper (spot-level artifacts)

Annotation

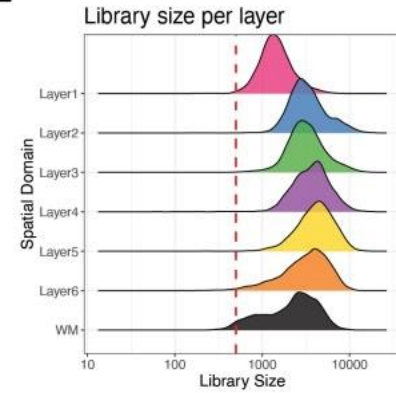


Global

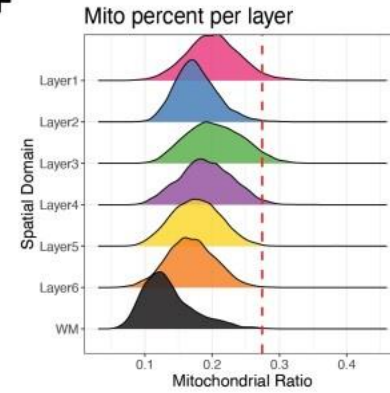


Global outliers

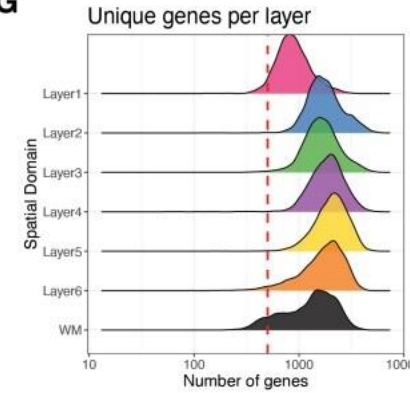
E



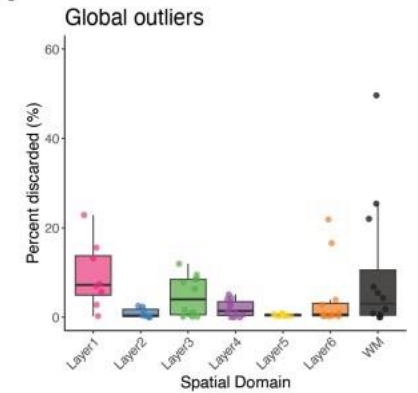
F



G

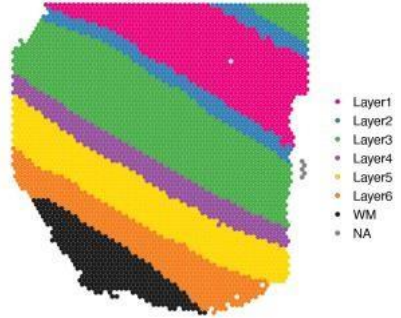


H

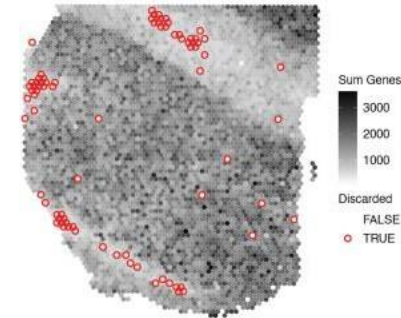


SpotSweeper (spot-level artifacts)

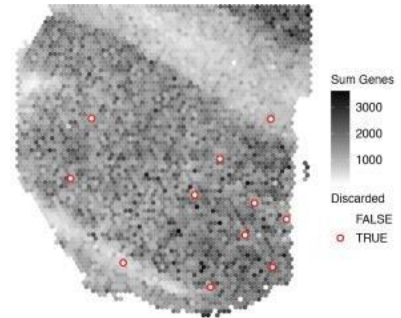
Annotation



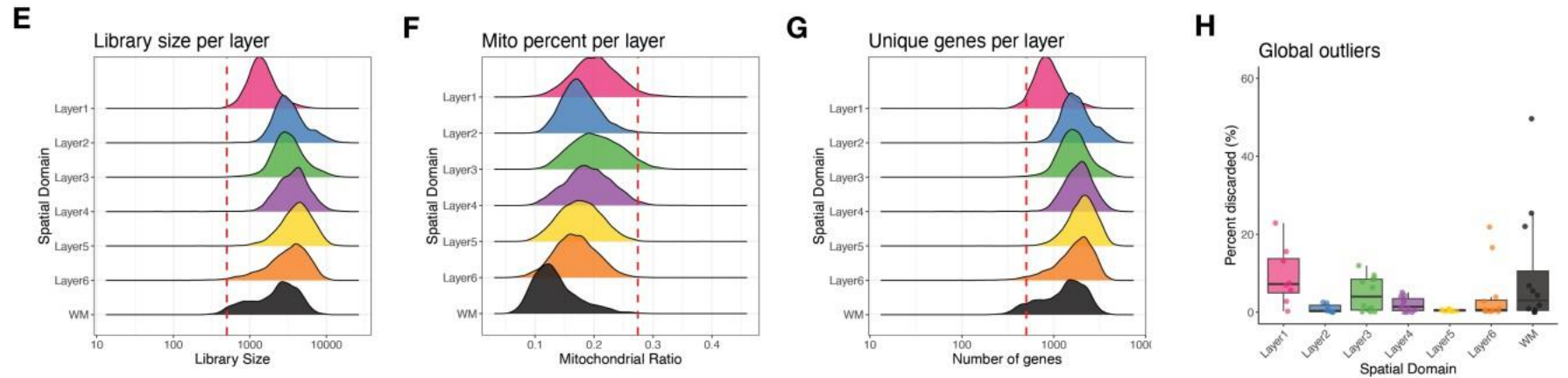
Global



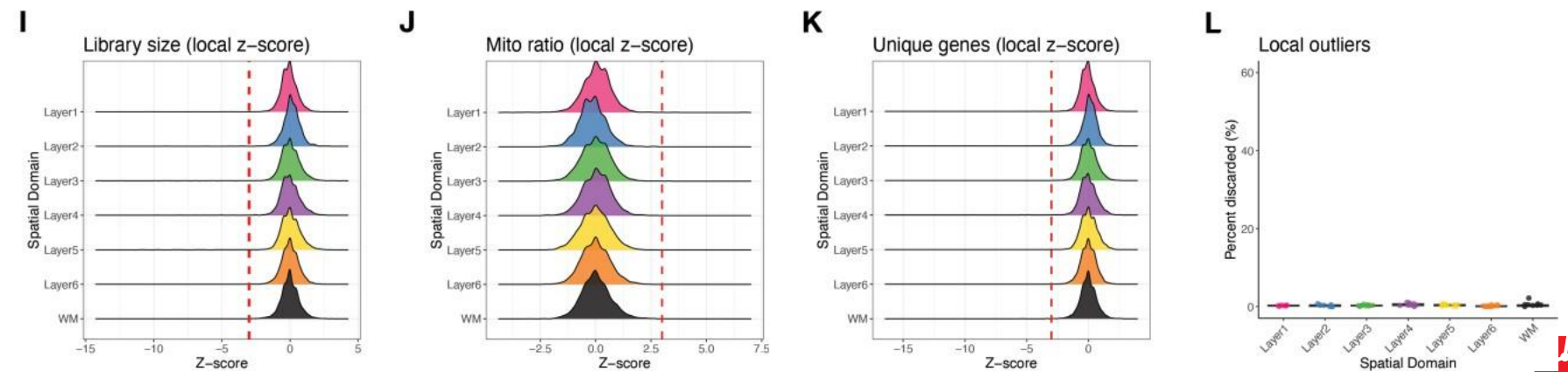
Local



Global outliers



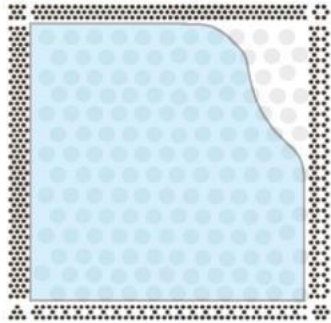
Local outliers



SpotSweeper (region-level artifacts)

Incomplete coverage of Visium array

Liquid reagent

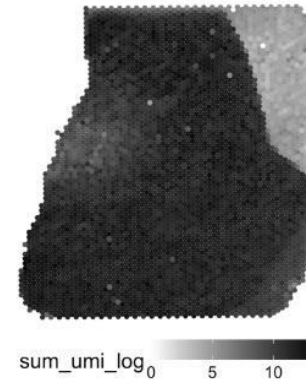


Dry spots result in smaller library size and fewer genes detected

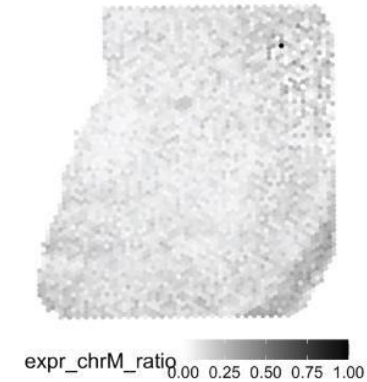
Br3942_mid



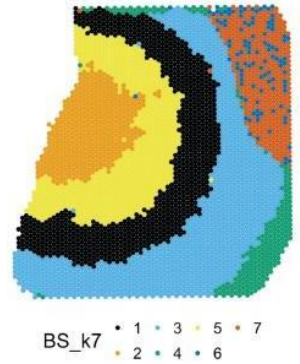
Library size



Mito Ratio



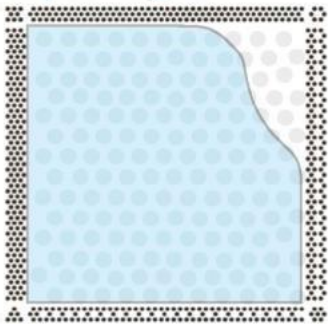
BayesSpace k=7



SpotSweeper (region-level artifacts)

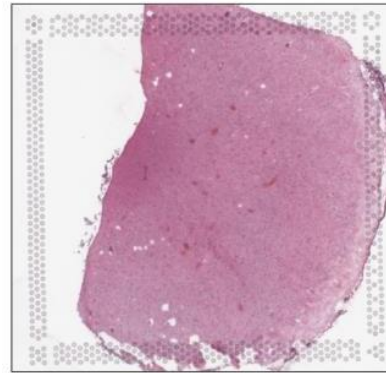
Incomplete coverage of Visium array

Liquid reagent

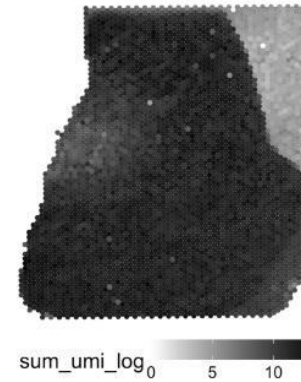


Dry spots result in smaller library size and fewer genes detected

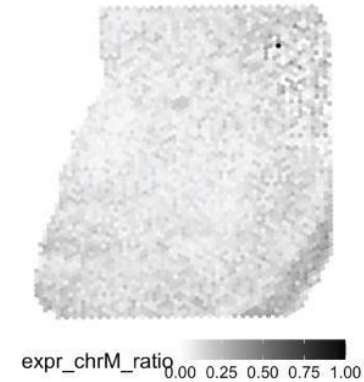
Br3942_mid



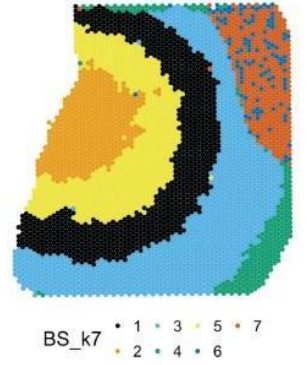
Library size



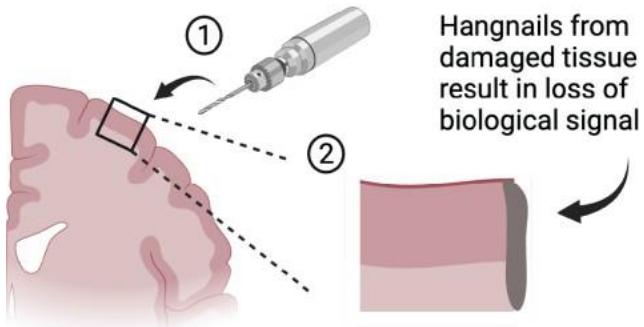
Mito Ratio



BayesSpace k=7



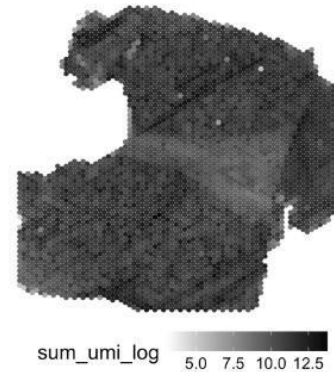
Tissue damage during dissection



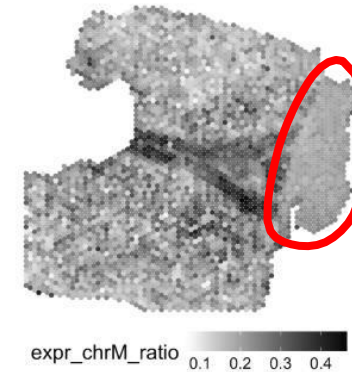
Br8325_ant



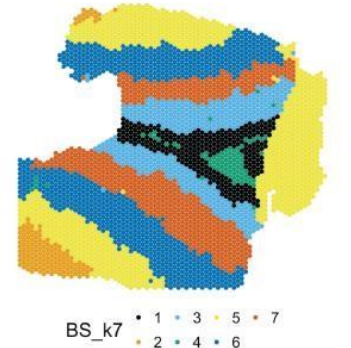
Library size



Mito Ratio

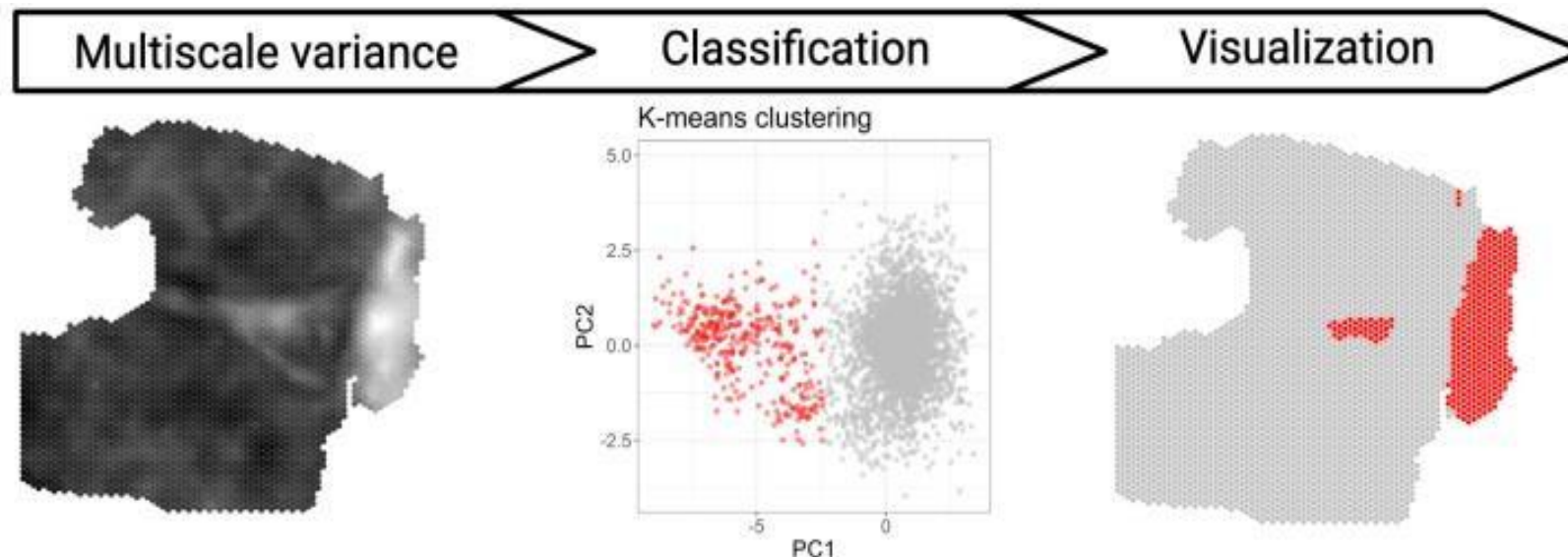


BayesSpace k=7



SpotSweeper (region-level artifacts)

1. The k-NN for each spot are identified based on the spatial coordinates
2. For each neighborhood size (i.e., scale), local variance of the mitochondrial ratio is calculated and adjusted for a mean-variance relationship using linear regression
3. Perform PCA on the mean-corrected local variances of all neighborhood sizes
4. Apply k-means clustering (k=2) in the first two PCs to identify regional artifacts compared to high-quality tissue



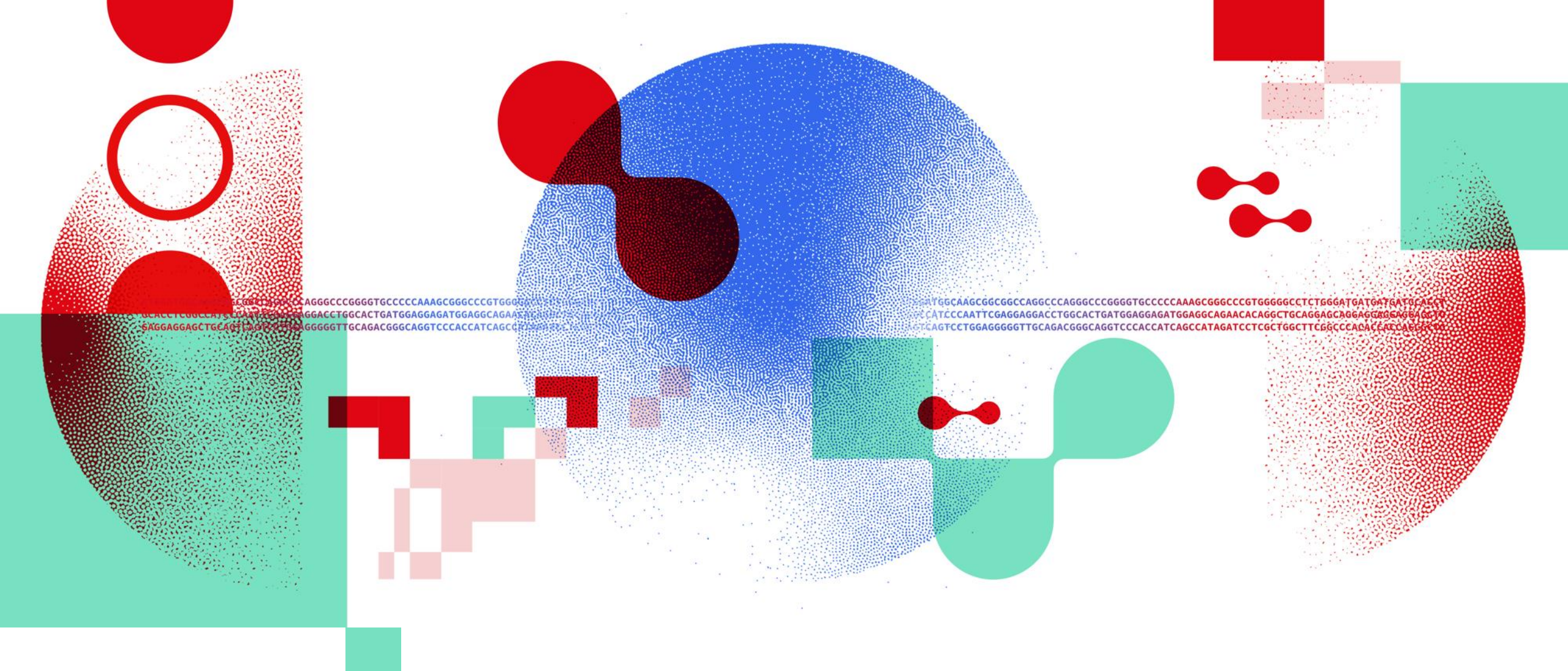
Preserve biological variability while removing technical artifacts: **SpotSweeper**

Detects **spatially coherent low-quality regions** in Visium HD slides.

Uses **per-spot QC metrics** (mainly mitochondrial fraction) and **spatial neighborhood info**.

Flags artifacts for downstream masking/removal.

Works across **multiple neighborhood scales** to detect both small and large defects.



...AGGGCCCGGGTGCCCCAAAGCGGGCCGTGGG...
...GACCTCGCCATGCTAATTCAGGACCTGGCACTGATGGAGGAGATGGAGGCAGAA...
...SAGGAGGAGCTGCAGTAACTTCAGGGGGTTGCAGACGGGCAGGTCCACCATCAGCC...
...TGGCAAGCGGGCCAGGCCAGGGCCCGGGTGCCCCAAAGCGGGCCGTGGGGCCTCTGGGATGATGATGATGCACT...
...CATCCCAATTCGAGGAGGACCTGGCACTGATGGAGGAGATGGAGGCAGAACACAGGCTGCAGGAGCAGGAGGAGGAGG...
...TCAGTCCTGGAGGGGGTTGCAGACGGGCAGGTCCACCATCAGCCATAGATCCTCGCTGGCTTCGGCCCAACACAGGCT...

Thank you

DATA SCIENTISTS FOR LIFE

sib.swiss