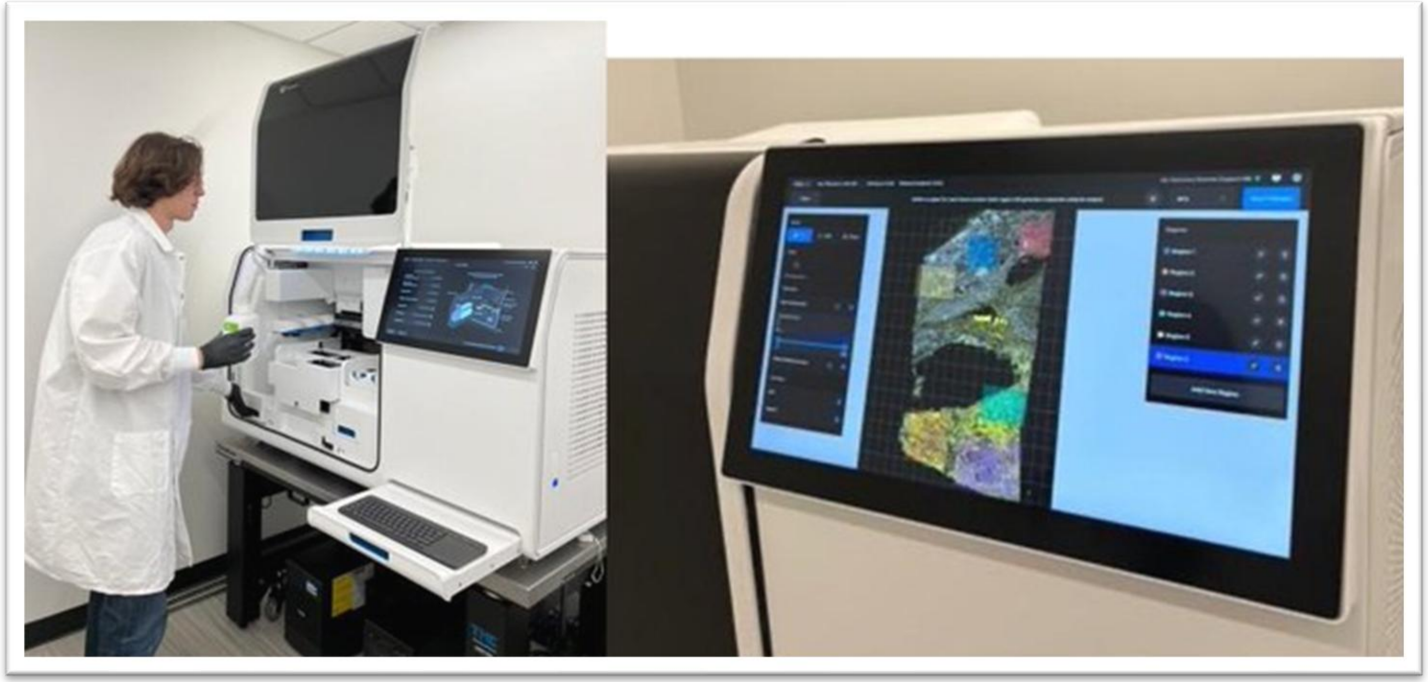
Designing A Spatial Transcriptomics Experiment



Irvine Tatenda Mutiti

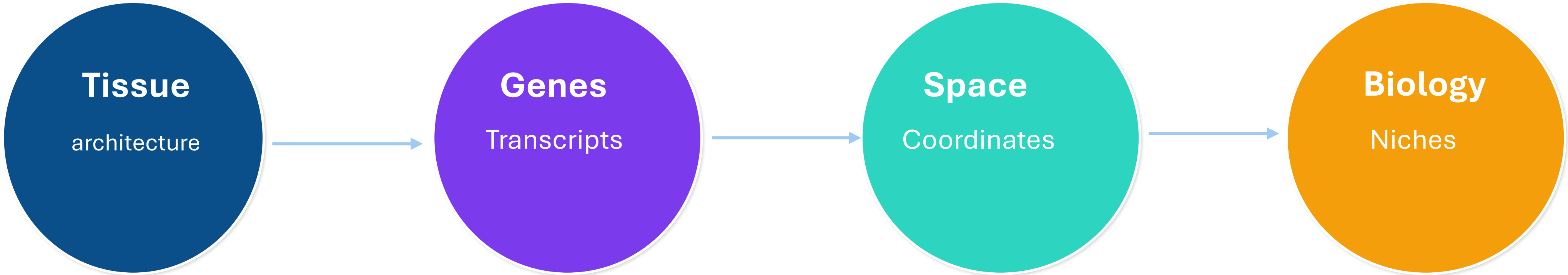
Learning Objectives

- By the end, students should be able to explain and defend a spatial transcriptomics experimental design:

1) Define when spatial transcriptomics is biologically appropriate
    - ➢ intact tissue + viable RNA + spatially meaningful question
2) Distinguish hypothesis testing from hypothesis generation
    - ➢ targeted high-resolution validation vs unbiased discovery
3) Explain the major technical trade-offs
    - ➢ resolution, gene number, capture efficiency, sample throughput
4) Interpret tissue suitability and sample quality issues
    - ➢ FFPE/fresh-frozen, DV200, autofluorescence, tearing
5) Justify companion datasets
    - ➢ scRNA-seq references, H&E, DAPI, boundary stains
6) Build a simple decision checklist
    - ➢ question → tissue → quality → platform → controls → analysis

The Central Idea

- Spatial transcriptomics asks: where are transcripts located within tissue architecture?
- Classical RNA-seq asks “what genes are expressed?” Spatial transcriptomics adds “where?” and “next to what?”.



Tissue

architecture

Genes

Transcripts

Space

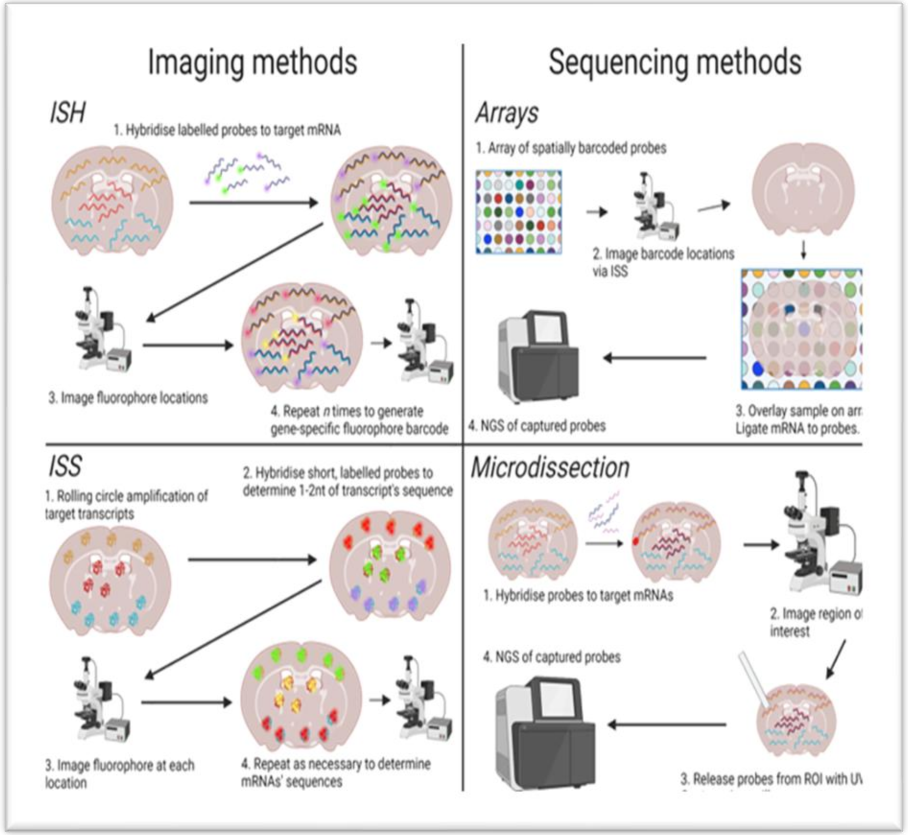
Coordinates

Biology

Niches

A good experiment preserves all four: tissue integrity, molecular signal, spatial localisation and interpretable biological context.

Spatial Transcriptomic Technologies



ISH-based imaging

- MERSCOPE, Esper, CosMx
- Gene-specific probes bind RNA in tissue; repeated imaging localises transcripts. Usually targeted but very high resolution.

ISS-based imaging

- Xenium lineage technologies
- Transcripts are amplified or sequenced in place; imaging detects sequence information within the tissue section.

Array-based capture

- Visium, STomics
- mRNAs are released from tissue and captured at barcoded positions on a slide or array, then sequenced.

Microdissection/ROI

- GeoMx
- Selected regions of tissue are profiled; stronger spatial context than bulk RNA-seq but usually ROI-level resolution

Decision Tree Overview

Is there intact tissue with usable mRNA?	• No → spatial transcriptomics is unsuitable or risky. Yes → continue.
Is the main aim hypothesis testing or generation?	• Testing → targeted high-resolution methods. Generation → unbiased array/ROI methods
Does the tissue require optimisation?	• Check permeabilisation, autofluorescence, fragile handling, FFPE RNA quality.
Are rare cells, small structures or lowly expressed genes central?	• If yes, prioritise resolution and capture efficiency; validate with staining.
Can you add companion data?	• scRNA-seq, H&E, DAPI, boundary staining and controls increase interpretability.

Can I Use Spatial Transcriptomics?

- Spatial transcriptomics is appropriate when the tissue itself carries biological information that would be lost if cells were dissociated
- Suitable starting point
  - Intact tissue section + measurable RNA + spatially meaningful structure.
- Unsuitable starting point
  - Old, degraded, distorted or very fragile sections where signal or geometry may fail.
- Unsuitable question
  - If location does not matter, bulk or sRNA-seq may answer the question more directly.
- Viable mRNA does not mean living tissue; preserved tissue can work if RNA fragments are still adequate.
- The key is whether spatial organisation is part of the biological hypothesis.
- A pilot or adjacent histology section can confirm that the target region or feature is present.

When Spatial Transcriptomics may be Impractical

- Too many samples
  - Imaging-based methods may require long instrument time because samples are repeatedly imaged.
- Too costly at scale
  - Sequencing methods accrue spatial reagents, NGS costs and often companion scRNA-seq costs.
- Poor representativeness
  - A single human tissue section may not represent the disease state unless carefully pre-screened

Tissue-Specific Optimisation

- Some platforms must be tailored to the tissue before the main experiment
- Why optimise?
  - Different tissues release RNA differently. Under-permeabilisation gives weak signal; over-permeabilisation can cause lateral diffusion.
  - A range of permeabilisation times can be tested by labelling and imaging cDNA to find strong signal with minimal spread, Eg Visium
  - It protects spatial accuracy and RNA recovery.
  - Maximise cDNA signal while minimising lateral diffusion, because diffusion weakens spatial accuracy.
  - The best condition is not simply the brightest signal, but the best signal with the least movement away from the original tissue location.

Tissue section → Permeabilise multiple times → Capture cDNA → Fluorescent image → Choose best condition

Rare Cells and Small Structures

- Rare cells
  - ➢ If rare cells are mixed into a larger capture area, their RNA may be diluted below detection or difficult to assign.
- Small structures
  - ➢ A structure made of only a few cells can be blurred in methods that aggregate RNA over a spot or ROI.
- Low-expression genes
  - ➢ Transcription factors may be biologically powerful but detected weakly if capture efficiency is low.
- If the key biology is rare, small or lowly expressed, favour high spatial resolution, high capture efficiency and validation by staining adjacent sections.

Is the main aim Hypothesis Testing or Generation?

Hypothesis testing

- Defined genes or pathways
- Need high spatial resolution
- Often targeted panels are enough
- Example: validate spatial expression near a tumour margin

➢ Best fit: ISH/ISS imaging or targeted high-resolution spatial approaches

Hypothesis generation

- Unbiased gene discovery
- Whole-transcriptome profiling
- Large tissue areas or atlases
- Example: discover new pathways in an uncharacterised tissue

➢ Best fit: array-based or microdissection-based methods with broad coverage

Tissue Preservation: Fresh-Frozen vs FFPE

Fresh-frozen tissue

- Stored cold to slow RNA degradation. Often strong for transcriptome profiling, but delicate tissues may tear or distort during sectioning and placement.

FFPE tissue

- Formalin-fixed, paraffin-embedded tissue is common in archives and clinical pathology. Many newer methods can profile FFPE, but RNA fragmentation must be assessed.

DV200: A Practical RNA-Quality Metric

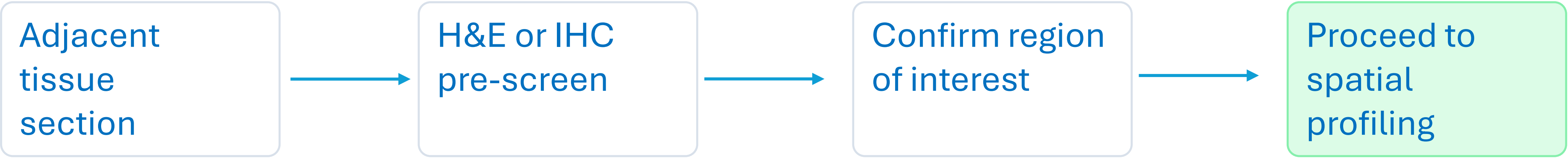
- DV200 = percentage of mRNA fragments longer than 200 nucleotides
- DV200 ≥ 50% usually recommended threshold for Visium FFPE sample suitability.
- DV200 < 50%: Higher risk of poor sequencing output; one replicate may be insufficient especially for array-based approaches.
- NB: Age of the tissue block often correlates with degradation, but storage conditions can accelerate degradation - so testing is preferred over assumptions.

DV200: A Practical RNA-Quality Metric

- Sections tear, fold and distort: Replicates are not only about statistics; they also protect against sample failure
- Tearing
  - Thin tissue sections can physically tear while being mounted onto a slide or array
- Distortion
  - Stretching, folding or uneven placement can compromise spatial coordinates and morphology
- Low output
  - Poor RNA quality or failed chemistry can produce inadequate sequencing or imaging signal.
- Build in replicates whenever possible. On some platforms, multiple sections can be profiled on the same slide

Pre-Screening and Controls

- Before running the experiment, prove that the section contains the feature you need



Adjacent tissue section

H&E or IHC pre-screen

Confirm region of interest

Proceed to spatial profiling

- Controls help separate true spatial biology from tissue damage, poor capture, background signal or missing target features.
- Use immunohistochemistry or histology on parallel sections to confirm the rare structure or disease region before profiling.

Variables to Consider

- Sample number
- Spatial resolution
- Number of genes
- Tissue area

Sample Number

- Throughput differs strongly by technology class
- ISH / ISS imaging methods
  - ➢ Samples are typically imaged individually and repeatedly on specialised instruments. Throughput can be limited by imaging time.
- Array-based methods
  - ➢ Multiple tissue sections may be assessed on one array.
  - ➢ Eg Visium: can have four sections per slide.
- Planning question: How many biological samples, tissue sections and technical replicates are needed to make the result relevant?

Spatial Resolution

- Resolution determines whether you see regions, cells or subcellular RNA locations
- ROI / region
  - larger than cells
  - GeoMx-like ROI profiling
- Spot / capture area
  - may contain mixtures
  - Visium-like array capture
- Single-cell / subcellular
  - individual mRNA positions
  - ISH/ISS imaging; STomics in excerpt reaches <1 μm.
- Ask yourself: What is the smallest biological structure you must distinguish: a tissue domain, a neighbourhood, a cell, or subcellular localisation?

Number of Genes Profiled

- Targeted panels and whole-transcriptome profiling answer different questions
- MERSCOPE
  - Targeted imaging: panels up to 1000 genes
- Esper
  - Targeted imaging: up to about 5000 genes
- Array methods
  - Untargeted sequencing gives genome-scale coverage but usually with lower resolution/capture trade-offs.
- NB:
  - A targeted panel is efficient when you already know what genes/pathways matter.
  - Whole-transcriptome profiling is better when you do not know what mechanisms will emerge.
  - Panel design is itself a scientific decision: missing genes cannot be spatially interpreted later unless inferred from references.

Tissue Area Profiled

- Some questions need high-resolution local detail; others need broad anatomical coverage
- Small field, high detail
  - ➢ Best for rare interactions, cell boundaries, spatial validation, and high-resolution molecular localisation.
- Large area, broad discovery
  - ➢ Best for atlas generation, regional heterogeneity, anatomical gradients, and unbiased pathway discovery.
- Key question before selecting a platform: Is the central biological signal expected to be local and microscopic, or tissue-wide and anatomical?

What Instruments are Needed?

- The required infrastructure depends on the technology class
- Imaging-based ISH/ISS
  - Requires specialist microscopes/imagers, repeated imaging cycles and image-processing software.
- Array-based sequencing
  - Requires molecular capture workflow and access to next-generation sequencing; Visium does not need bespoke imager
- ROI/microdissection
  - Requires instruments to select or capture regions of interest before molecular profiling.
- A practical design question: Can your laboratory realistically run the platform at the sample scale and timeline required?

scRNA-seq Reference

- Many spatial methods are not truly single-cell by default
- Targeted spatial panels
  - A single-cell reference can help infer unmeasured genes or support assignment of imaged RNAs to cells.
- Array-based spots
  - Each capture area may contain RNA from multiple cells; references help deconvolve cell types and proportions.
- Low-efficiency capture
  - References can help interpret weakly profiled genes and cell states, though they cannot fully rescue poor data.
- If the tissue can be dissociated without destroying the relevant biology, consider generating a matched single-cell RNA-seq reference dataset.

Deconvolution

Spatial spot

A spot contains RNA from several neighbouring cells.

scRNA-seq references

Cell types/states are measured without spatial coordinates.

Estimated proportions

Computational model estimates which cell types contributed to each spot.

NB:

- Deconvolution is inference, not direct observation. It depends on reference quality, gene overlap, cell-state annotation and model assumptions.

Auxiliary staining and segmentation

- Nuclear staining
  - DAPI-like staining identifies nuclei, which helps locate cells in imaging-based methods.
- Boundary staining
  - Cell membrane or boundary markers improve cell segmentation and transcript assignment.
- Histology staining
  - H&E and related images can support tissue-domain identification and morphology-aware analysis.
- Segmentation: Computationally estimating where one cell ends and another begins, so individual mRNA coordinates can be assigned to reconstructed cells.

Summary

- Spatial transcriptomics is justified when tissue location is part of the biology.
- Platform choice depends on experimental aim: targeted high-resolution testing versus broad hypothesis generation.
- Tissue type, RNA quality, section integrity and autofluorescence can determine success or failure.
- Resolution, capture efficiency, gene number, sample throughput and cost are interdependent trade-offs.
- Single-cell references and staining often turn spatial data into interpretable cell-level biology.

Test your understanding

1. You need to validate five genes at a tumour margin. Which class of method is most appropriate and why?
2. A 15-year-old FFPE block has DV200 < 50%. What design changes could reduce risk?
3. Why might a matched scRNA-seq reference be valuable for a Visium-like experiment?
4. Why can a method with lower gene coverage still be better for a focused hypothesis-testing experiment?

References

- Williams *et al*. (2022). An introduction to spatial transcriptomics for biomedical research. *Genome medicine*, 14(1), 68.
- Marx, V. (2021). Method of the Year: spatially resolved transcriptomics. *Nature methods*, 18(1), 9-14.