# SEMINAR

**Dr Jiadong Mao**, School of Math & Stats

## Φ-Space ST: a platform-agnostic method to identify cell states in spatial transcriptomics studies

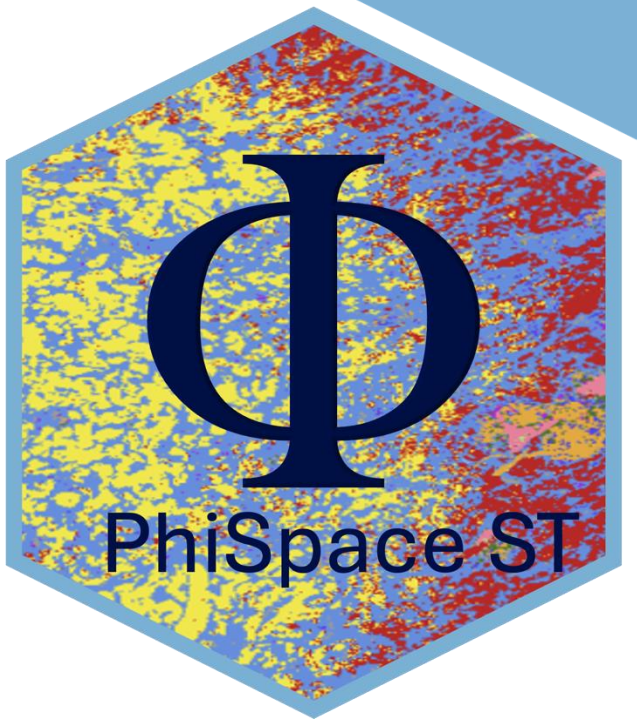
@jiadongm.bsky.social

**Visit Melbourne Integrative Genomics website for our events**

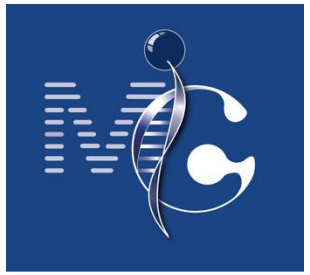
# A Tale of Two Spaces

*Discovery of spatial biology in phenotype space by Φ-Space ST*

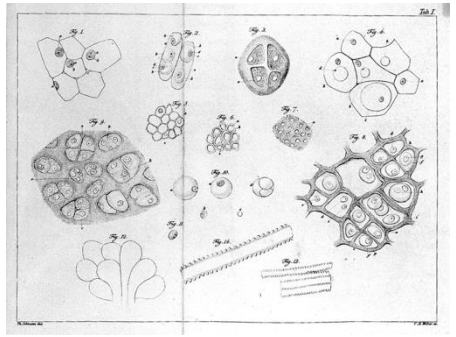
Dr Jiadong Mao
Melbourne Integrative Genomics





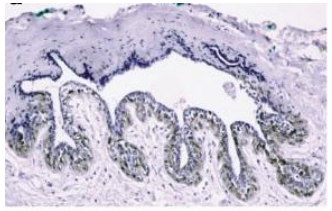
# ‘Cell type’



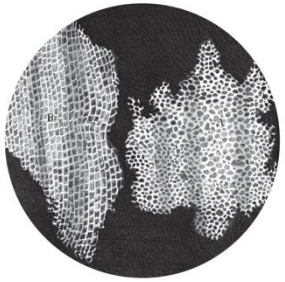
Schwann's *Mikroskopie*, 1839, Wellcome Collection

Describe cell type by

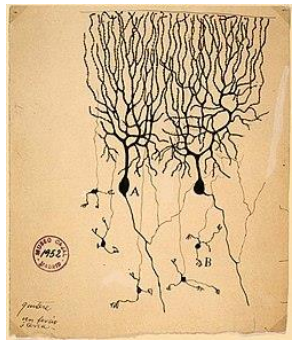
- Morphology
- function
- location



H&E staining, first introduced by Wissozky 1877, fig from Perou et al. (2000)

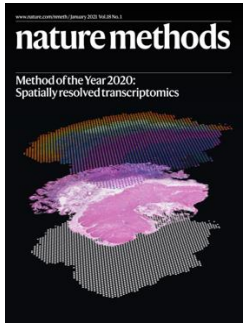


Dead cells in cork, from Hooke’s *Micrographia*, 1665, Science Museum UK



Drawing of Purkinje (1787–1869) cells by Cajal (1852–1934), Wikipedia



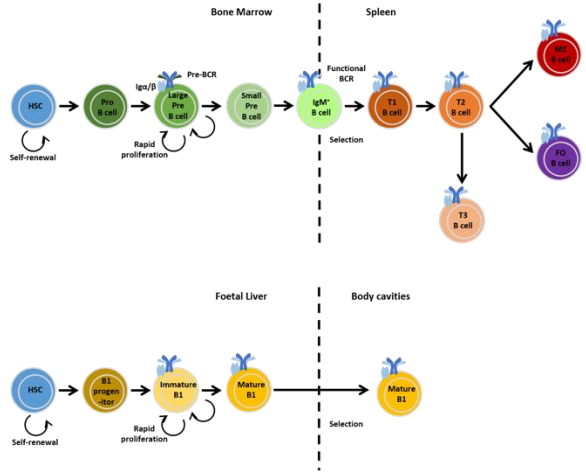




Describe cell type by

- Morphology
- function
- Location

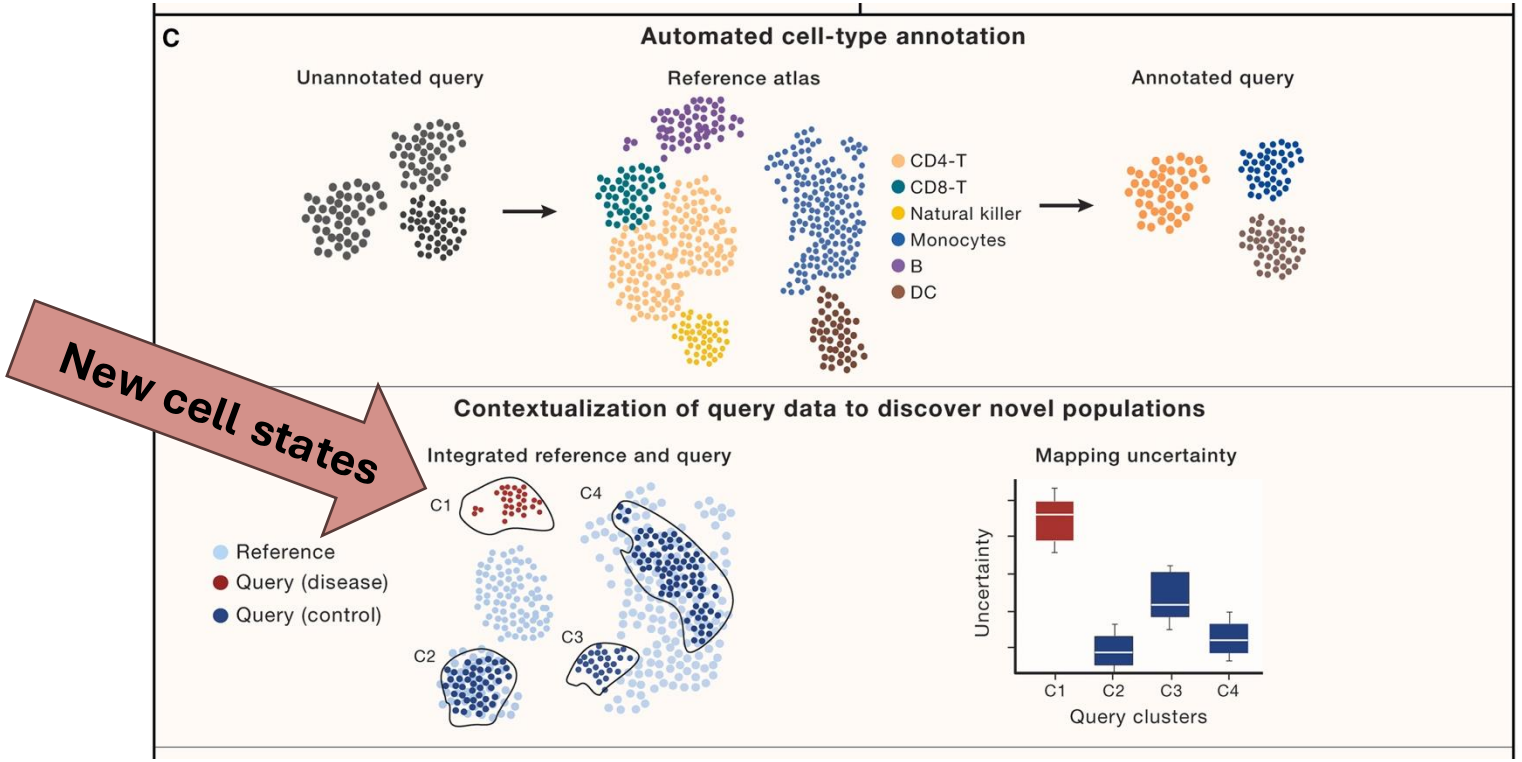
- Gene expression
- Protein expression
- Chromatin accessibility
- Histone modification
- …




B cells, Rebecca Newman, British Society for Immunology

# Cell type annotation using scRNA-seq

- Cell type: cells with similar gene expression profiles
- Reference based annotation
  - **Supervised**: train cell type classifier on **reference** data

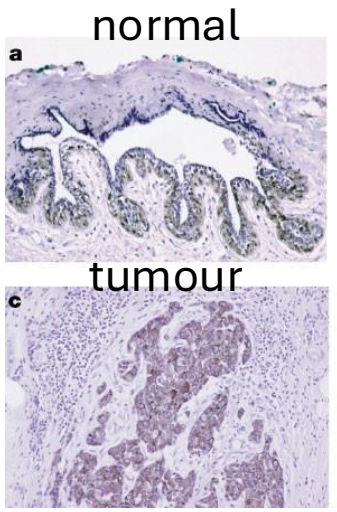


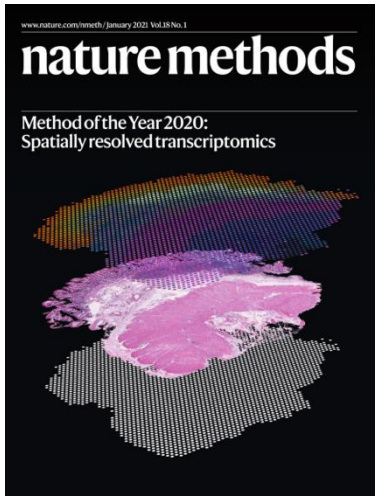

Lotfollahi et al. (2024). *Cell*.

# Adding spatial information

Biology has always been spatial

- 20-century tools: microscope, antibody staining
- But limited throughput (dimensionality)

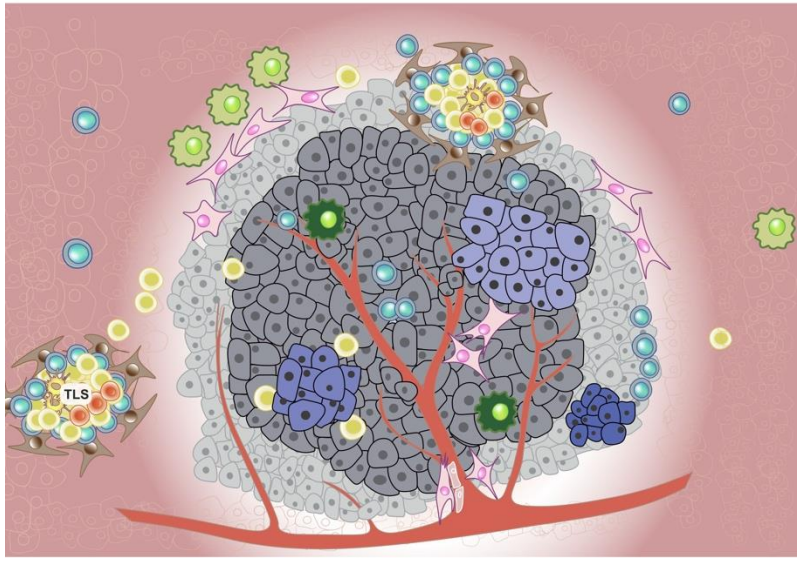



Perou et al. (2000). *Nature*.

Why ST

- High-throughput measurement of RNA
- Retaining spatial locations of RNA
- scRNA-seq: which cells are doing what
- ST: which cells are doing what, and **where they are doing it**

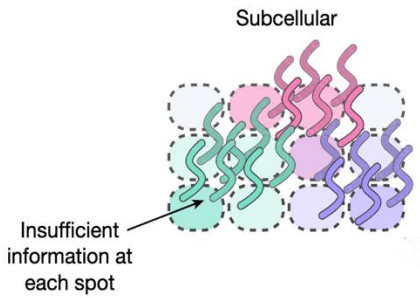



Cell type annotation is still the key

# Wait … but what is a cell in ST?

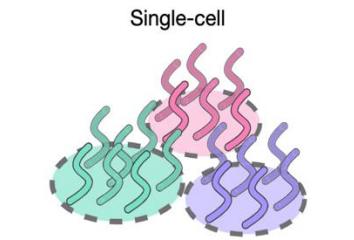
- Fast evolving tech, different spatial resolutions
  - Super-cellular: Visium
  - Sub-cellular: Xenium, CosMx, Visium HD, Stereo-seq, …
  - Nearly-cellular: Slide-seqV2
  - But lack of exact single cells
- Common problem: **mixture of cell identities**

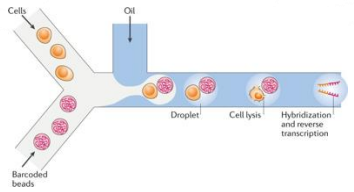
https://www.inveniagroup.com/blog/spatial-omics-technology/



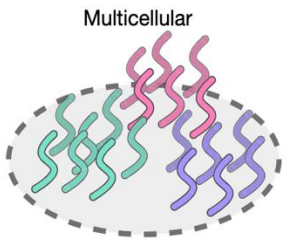

Mini-bulks by cell segmentation or binning

lacking




Cell type deconvolution




Images adapted from Benjamin et al. (2024). *Nature*.

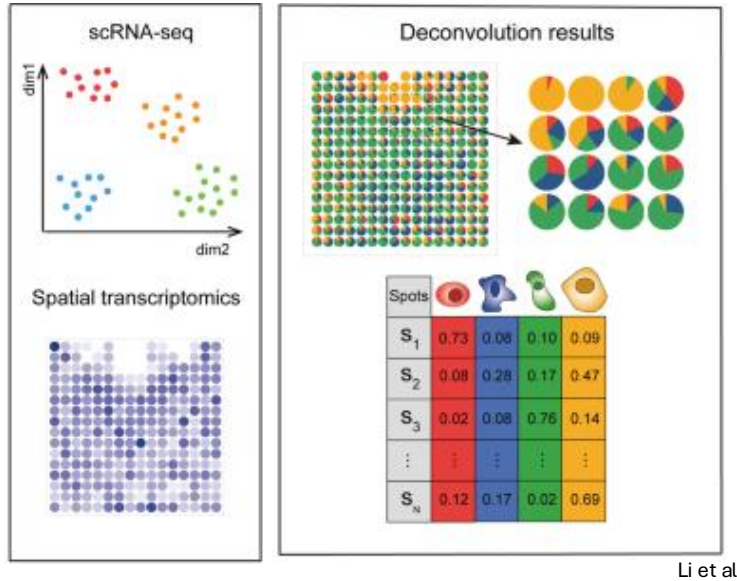
# Cell type deconvolution

Conventional deconvolution

- Using scRNA-seq to decompose spot into cell type proportions, eg

  1 = 0.7 CD4T + 0.2 CD8T + 0.1 NK

- Goal: understanding spatial distribution of (known) cell types




Li et al. (2023). Nat Comms

Φ-Space ST

- Use scRNA-seq to deconvolute every '**Cell-like objects**':

  Segmented cells, multi-cellular spots, spatial bins, ...

- Scoring without restriction, eg

  1.5 = 0.8 Epithelial + 0.5 Stem + 0.2 Immune (tumour region)

  0.9 = 0.8 Epithelial + 0.1 Immune (stromal region)

- Goal: understanding spatial distribution of (known and unknown) **cell states**

Cell type divergence

PhiSpace ST

# Φ-Space ST as unified approach

Limitations of existing ST deconvolution methods

- Specialised in specific types of cell-like objects
- Reliance on a single reference dataset
- Specialised in well-defined cell types
- Computational efficiency
- Restrictive parametric assumptions (eg variational autoencoders)

Φ-Space ST:

- Platform-agnostic
- Multi-reference
- Divergent cell states
- Very fast to compute
- Nonparametric

} Extension to emerging tech

Φ-Space: Continuous phenotyping of single-cell multi-omics data

Jiadong Mao[1], Yidi Deng[1], Kim-Anh Lê Cao[1,*]

Φ-Space ST: a platform-agnostic method to identify cell states in spatial transcriptomics studies

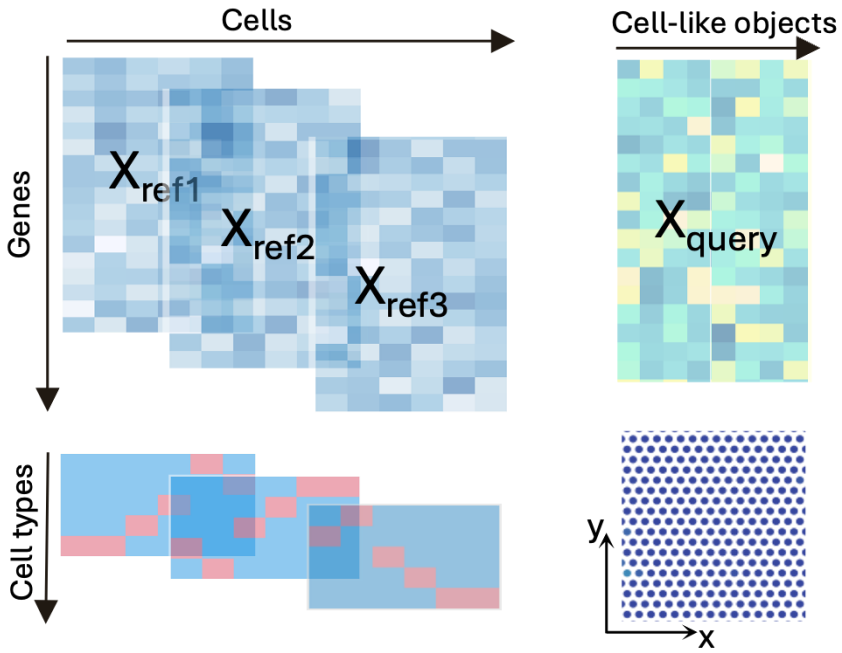
Jiadong Mao[1], Jarny Choi[2,&], Kim-Anh Lê Cao[1,*,&]

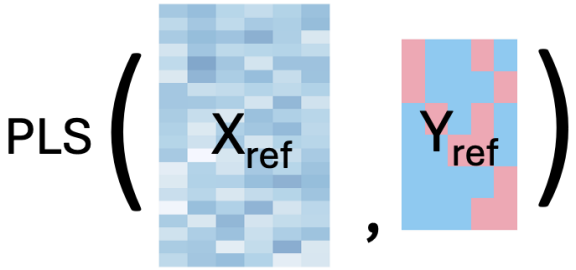
Capybara imitator Yidi

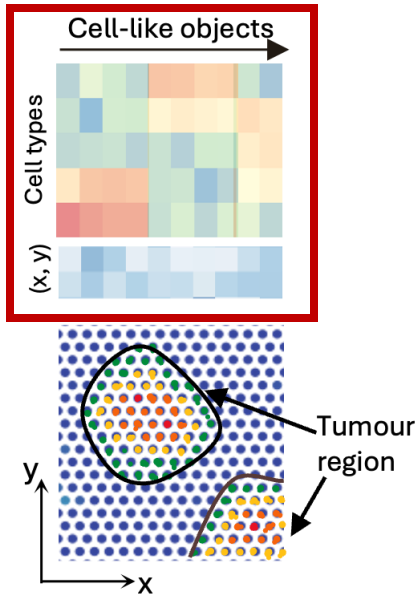
Tango dancer Jarny

Rock climber Kim-Anh

# How does it work




**Input**

- One or more annotated scRNA-seq datasets
- Query ST

**Training**

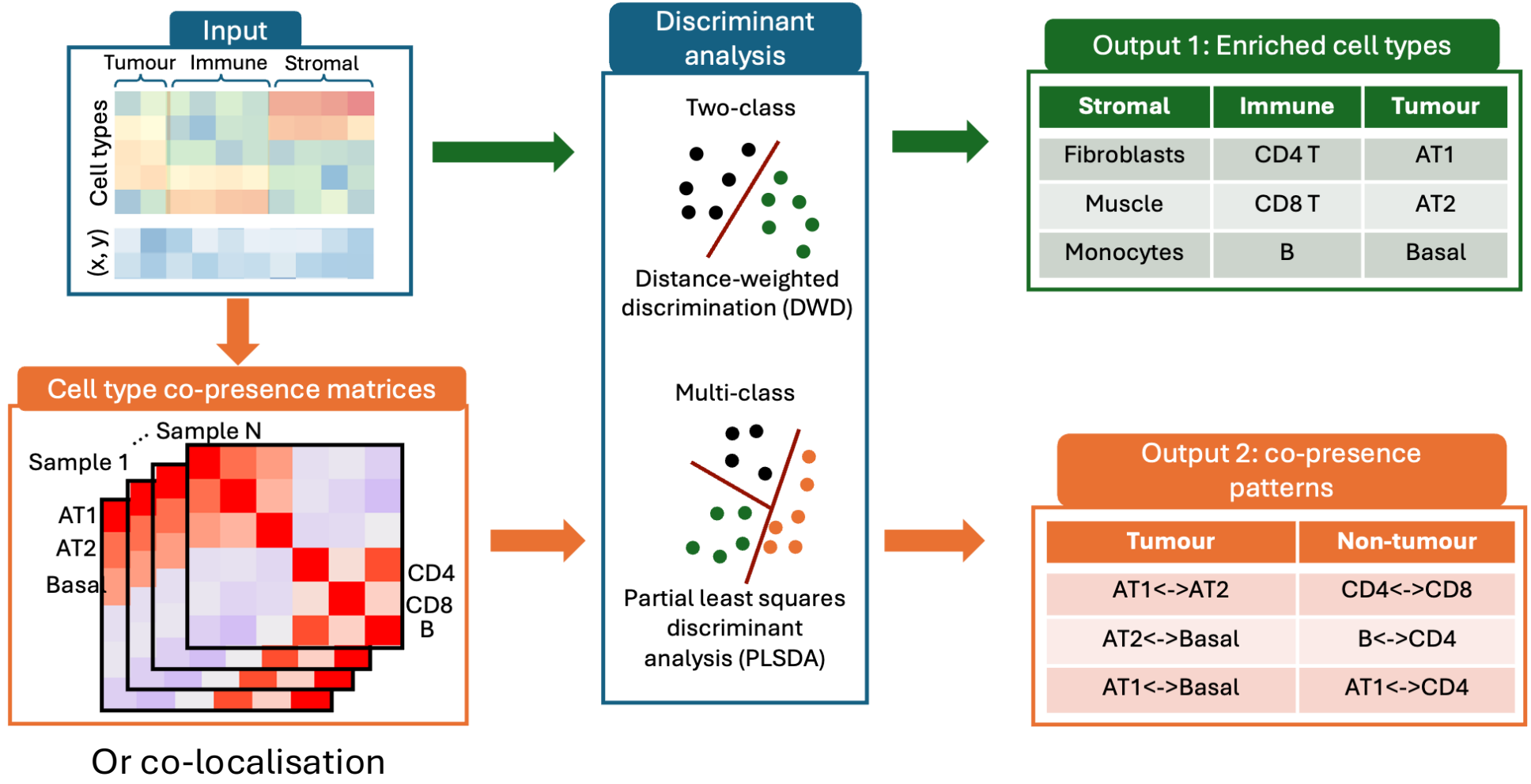
PLS regression (soft classification) model from each reference

**Output**

'Phenotype space embeddings'

Comparable across references and samples after *rescaling*

# Spatial biology in phenotype space

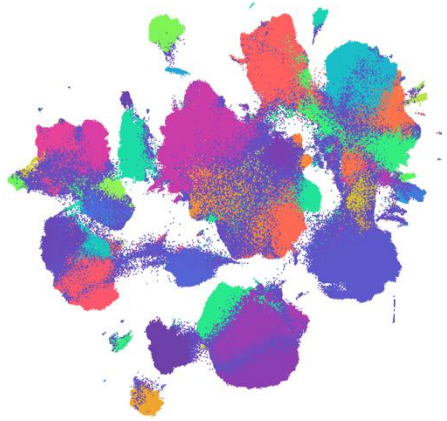


# Motivating example: Visium

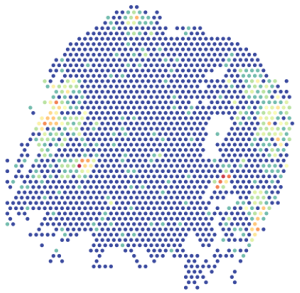
- *Platform*: 10x Visium
- *Cell like object*: spot with diameter 55µm
- *Sample info*: 18 healthy and cancerous lungs (NSCLC)
- *Reference*: integrated Human Lung Cell Atlas (HLCA)

*Challenges*
- Interpretable deconvolution;
- Healthy reference, cancerous query

**Reference**
HLCA with 61 healthy cell types



KRT19 (cancer)



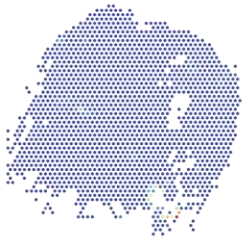
CD79A (B cell)

Predicted B cell abundance

PhiSpace

RCTD

cell2location



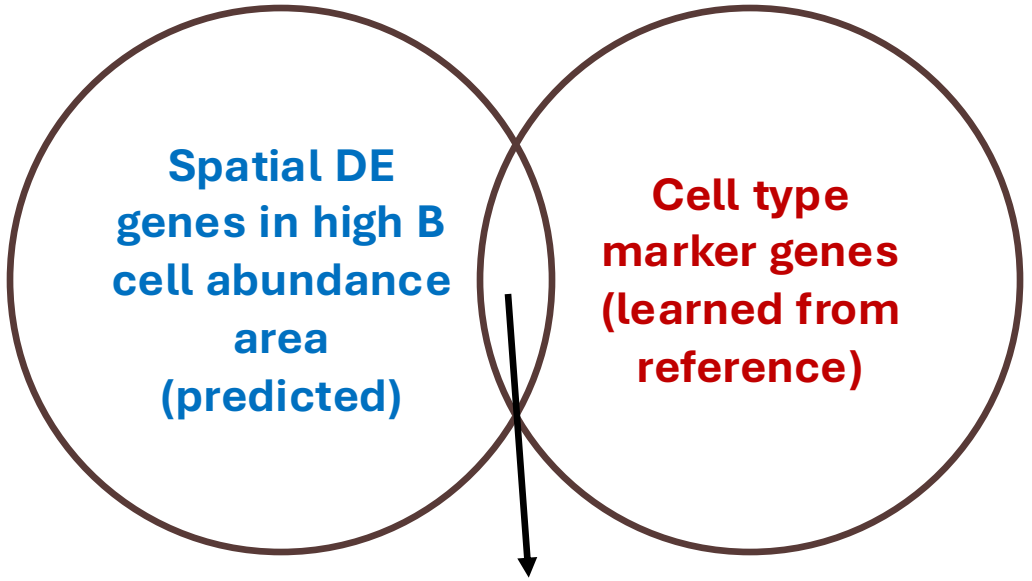
TACCO

**Which one should I trust?**

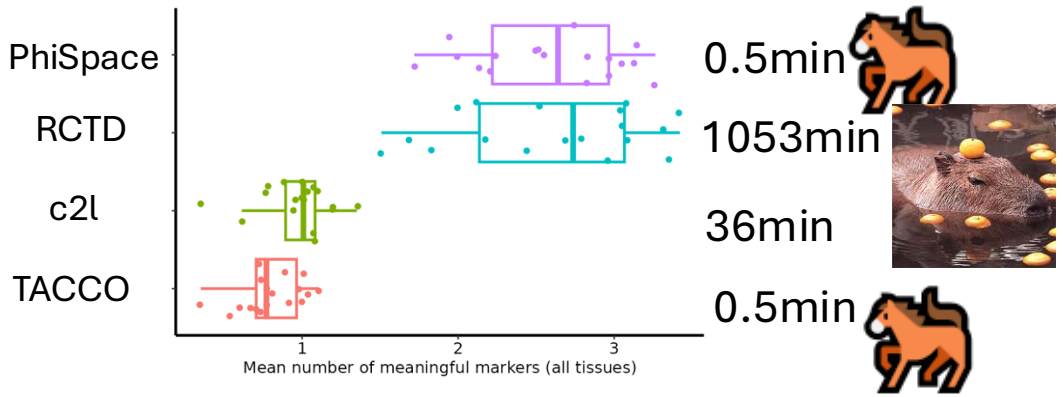
NSCLC: non-small cell lung cancer

# Meaningful markers

- No ground truth
- A proxy of truth:
  - If you say a spot has high B cell identity, spot should express B cell marker genes
  - ‘Meaningful markers’

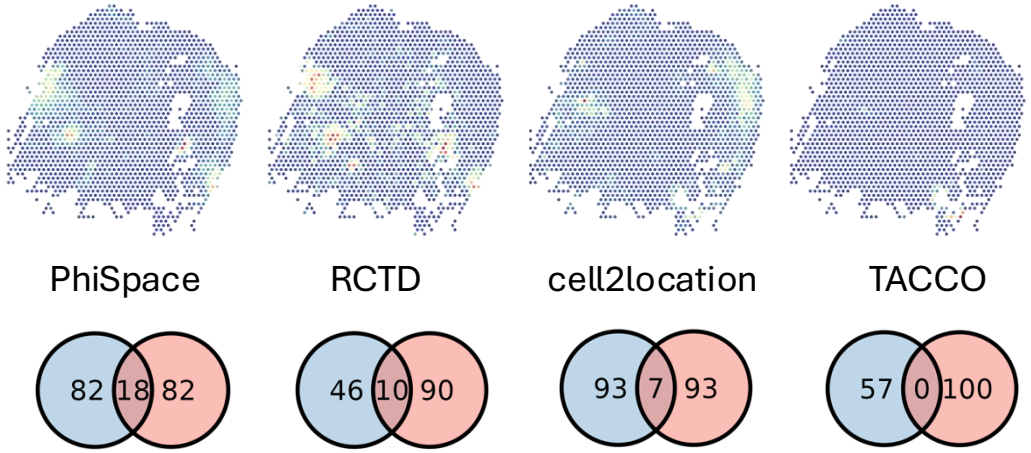



Φ-Space ST annotation is **meaningful** and **fast**




All 18 tissues

Predicted B cell abundance



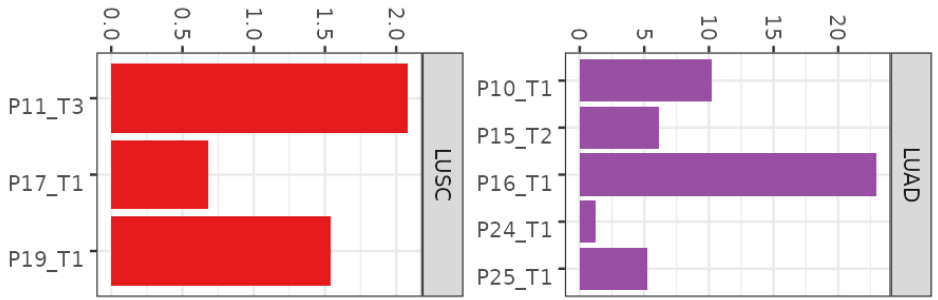
# Co-presence

18 tissues from

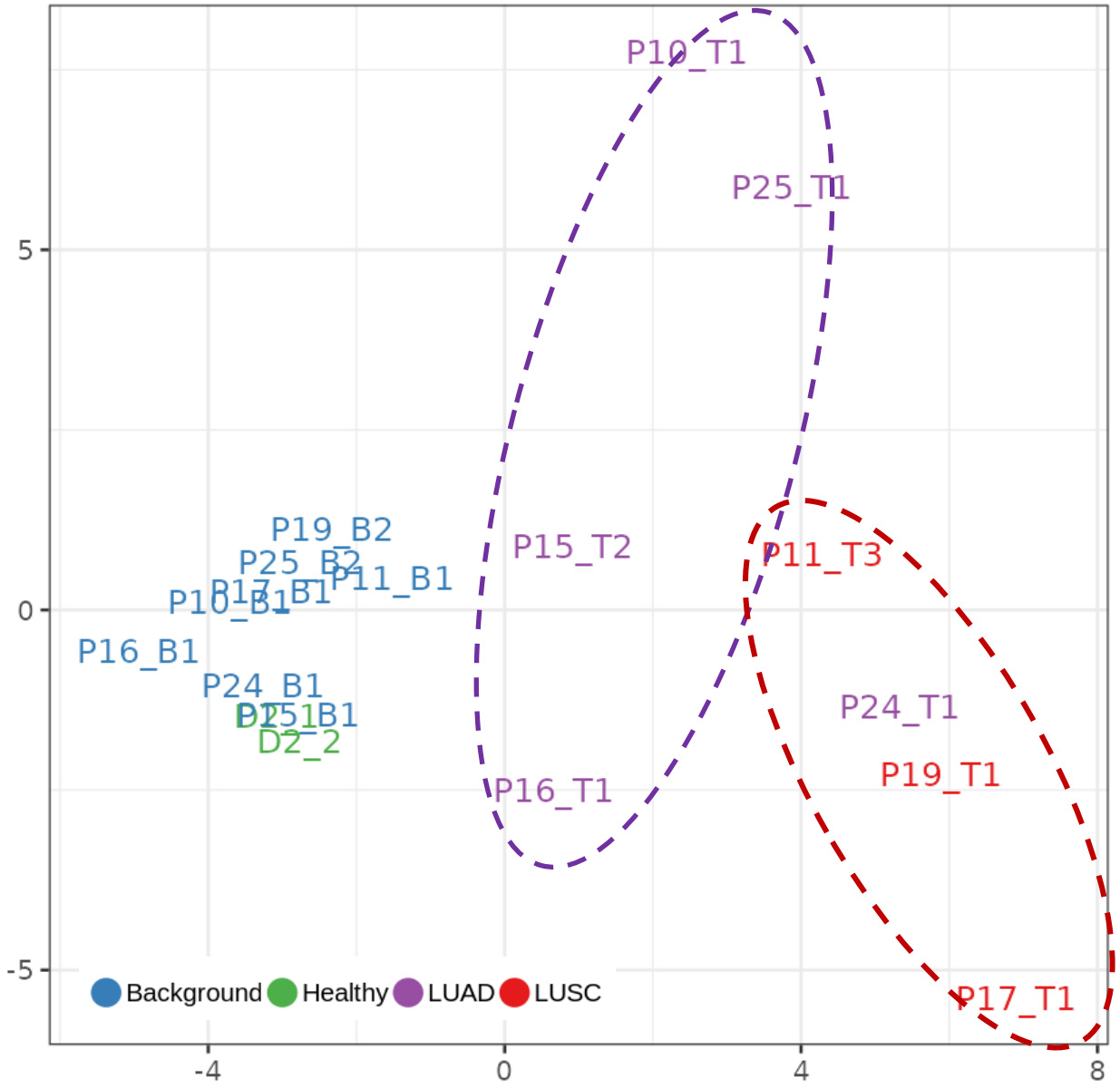
- Healthy: healthy donor
- Background: non-cancerous tissue from cancer patients
- LUAD: lung adenocarcinoma (better prog)
- LUSC: lung squamous cell carcinoma (worse prog)

Compute a co-presence matrix from each tissue.

**Favourable prognosis score**




**PCA of co-presence matrices**



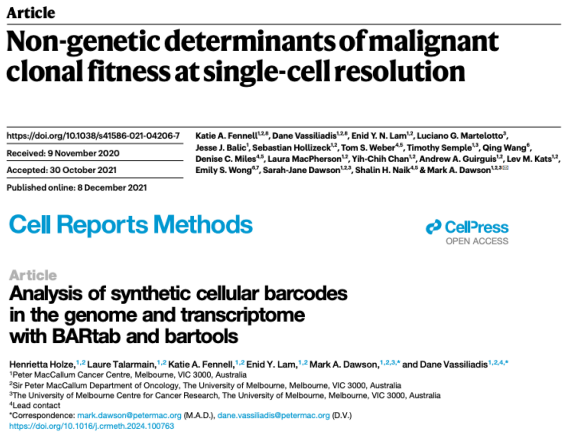
# Subcellular ST with cancer cell lineage tracing

- *Platform*: BGI Stereo-seq
- *Cell like object*: spatial bins side lengths ~25µm
- *Sample info*
  - 1 mouse spleen sample with AML cells
  - SPLINTR lineage tracing, same AML **clone** same **barcode**
- *References*: Spleen, BM, CITE, Neutro

Henrietta Holze

Dane Vassiliadis

Mark Dawson



Article

Non-genetic determinants of malignant clonal fitness at single-cell resolution

https://doi.org/10.1038/s41586-021-04206-7

Received: 9 November 2020

Accepted: 30 October 2021

Published online: 8 December 2021

Katie A. Fennell[1,2,8], Dane Vassiliadis[1,2,8], Enid Y. N. Lam[1,2], Luciano G. Martelotto[3], Jesse J. Balic[1], Sebastian Hollizeck[1,2], Tom S. Weber[4,5], Timothy Semple[1,2], Qing Wang[6], Denise C. Miles[4,5], Laura MacPherson[1,2], Yih-Chih Chan[1,2], Andrew A. Guirguis[1,2], Lev M. Kats[1,2], Emily S. Wong[4,5], Sarah-Jane Dawson[1,2,3], Shalin H. Naik[4,5] & Mark A. Dawson[1,2,3]

Cell Reports Methods

CellPress OPEN ACCESS

Article

Analysis of synthetic cellular barcodes in the genome and transcriptome with BARtab and bartools

Henrietta Holze,[1,2] Laure Talarmain,[1,2] Katie A. Fennell,[1,2] Enid Y. Lam,[1,2] Mark A. Dawson,[1,2,3,*] and Dane Vassiliadis[1,2,4,*]

[1]Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia

[2]Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, VIC 3000, Australia

[3]The University of Melbourne Centre for Cancer Research, The University of Melbourne, Melbourne, VIC 3000, Australia

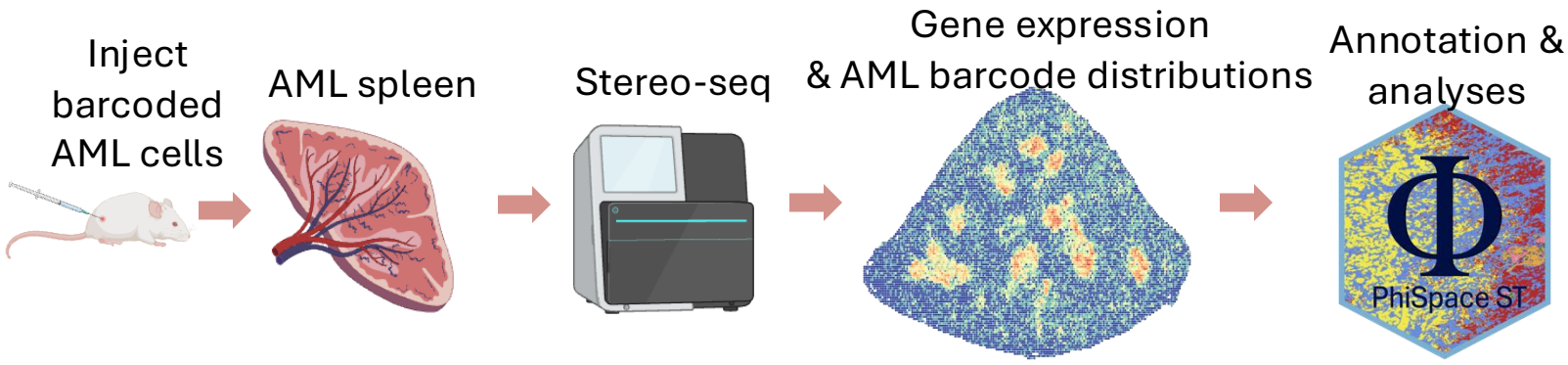
[4]Lead contact

*Correspondence: mark.dawson@petermac.org (M.A.D.), dane.vassiliadis@petermac.org (D.V.)

https://doi.org/10.1016/j.crmeth.2024.100763

*Challenges*

- Lack of cell segmentation;
- Lack of reference atlas;
- Liquid tumour lacking structures.

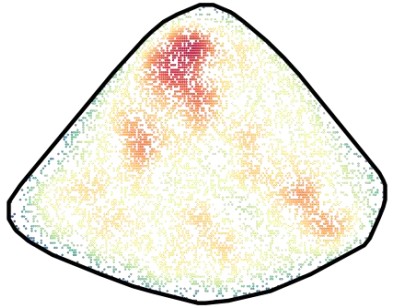


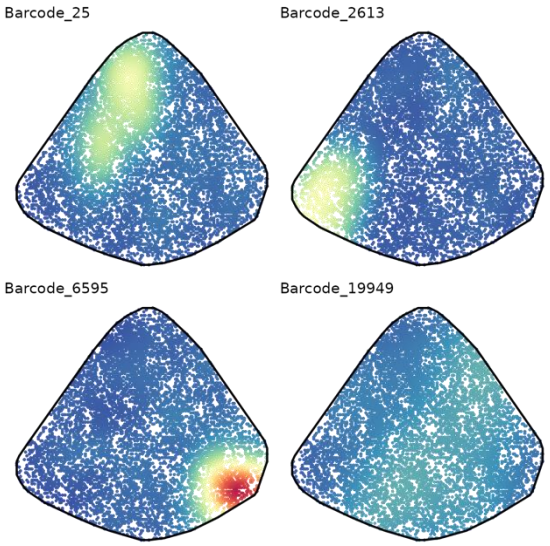

AML: acute myeloid leukemia

SPLINTR: single-cell profiling and lineage tracing

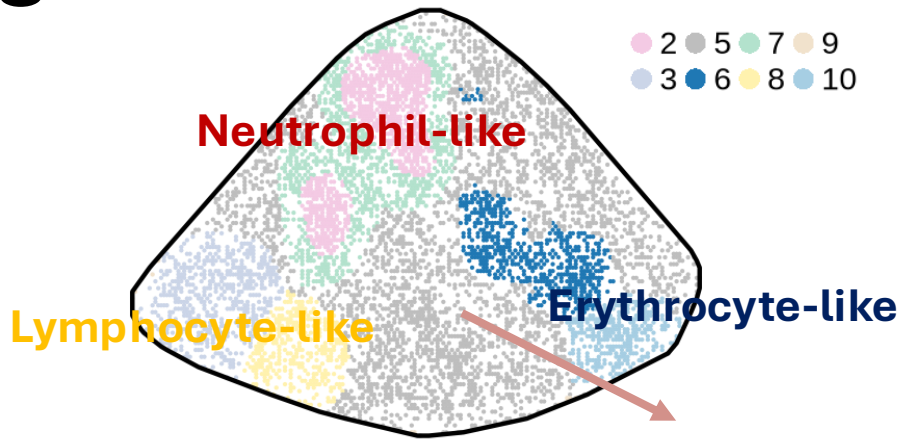
# Cell states of ‘meta-clones’

Density of all AML barcodes






AML Meta-clones



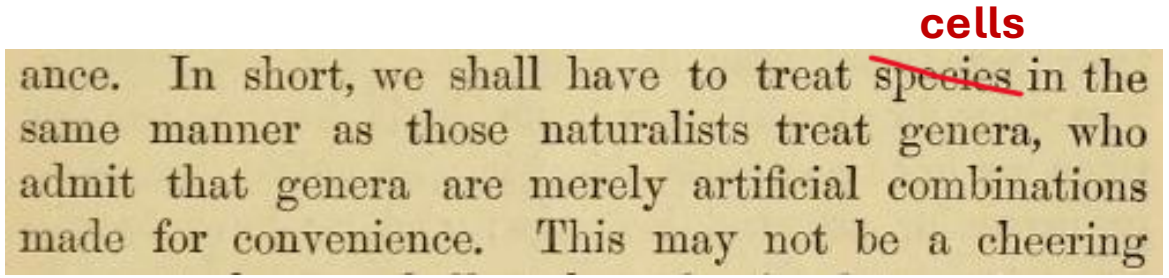

Uncoloured: nonbarcoded background

Meta-clone-specific enriched cell types

| | Meta-clone 2 | Meta-clone 3 | Meta-clone 6 | Meta-clone 7 | Meta-clone 8 | Meta-clone 10 |
|---|---|---|---|---|---|---|
| 1 | Granulo (BM) | Imm NK (BM) | RBC (CITE) | HPC (BM) | HPC (BM) | Naive B (BM) |
| 2 | Neutro (Spleen) | Trans B (Spleen) | RBC (Spleen) | IMM 1 (Neutro) | T (BM) | ProEryThBla (BM) |
| 3 | HPC (BM) | T (BM) | ErythBla (BM) | Neutro (Spleen) | Pre-B cycl (Spleen) | Pre-B cycl (Spleen) |
| 4 | T1 (Neutro) | MAT 3 (Neutro) | ProEryThBla (BM) | Pre-B cycl (Spleen) | Imm NK (BM) | ICOS+ Tregs (CITE) |
| 5 | Imm NK (BM) | CD8 T (Spleen) | IMM 2 (Neutro) | Pre-B (BM) | Pre-B (BM) | CD4 T (CITE) |

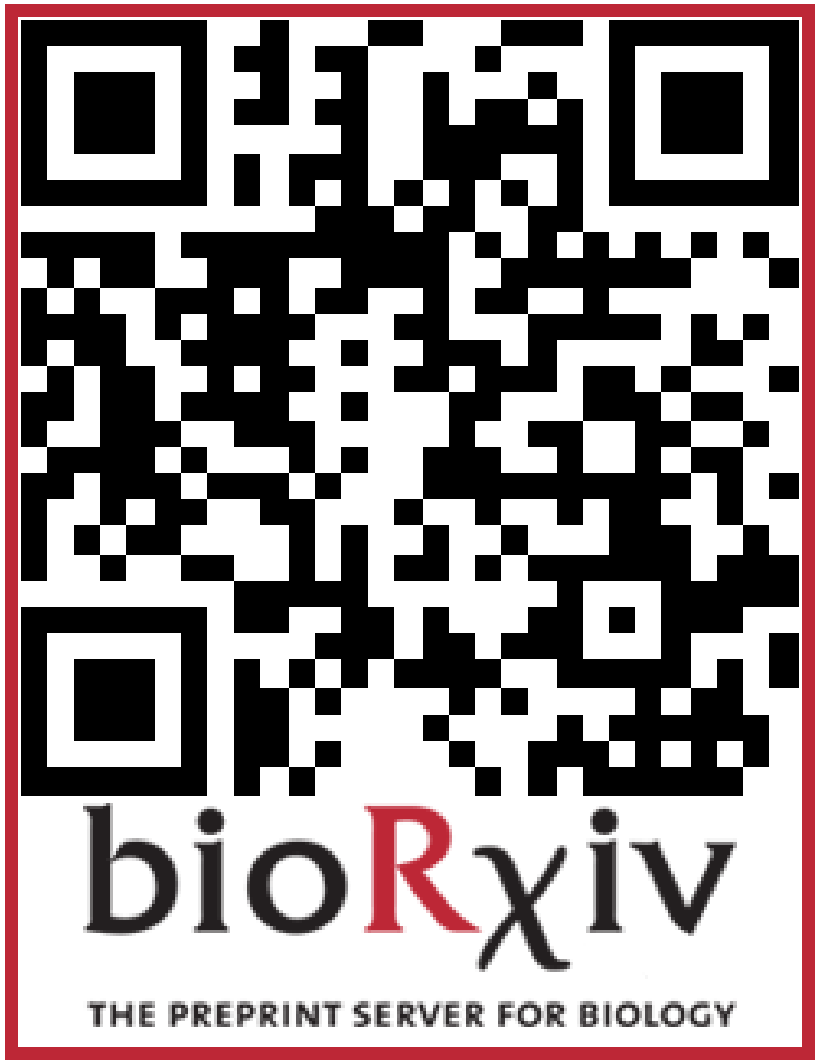
# Discussion

- Φ-Space ST goes beyond single sample
- Cell segmentation can be skipped for some questions: try segmentation-free methods!
- Φ-Space is a simple and robust method for evolving technologies
  - Bulk RNA-seq
  - scRNA-seq
  - Single-cell multiomics
  - Spatial transcriptomics
  - Spatial multiomics?
  - Spatial developmental trajectories?
  - …




Darwin, *Origin of Species*, 1859, p 485

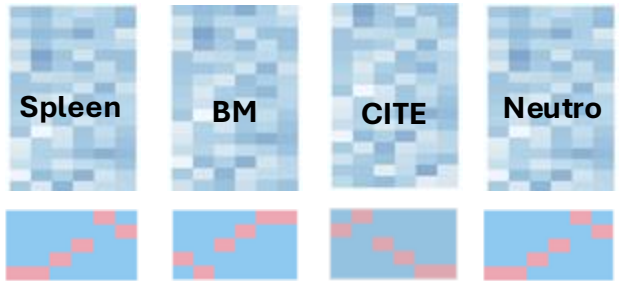
# Questions




Mao, J., Choi, J., & Le Cao, K.-A. (2025). *Φ-Space ST: a platform-agnostic method to identify cell states in spatial transcriptomics studies*. *bioRxiv*.

# Supp: Multi-reference bridging

References
4 scRNA-seq datasets

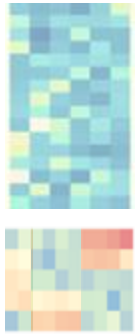



(Transfer discrete ann)

Bridge
scRNA-seq


(Transfer continuous ann)

Query
Stereo-seq from AML mouse



- Spleen: mouse spleen
- BM: mouse bone marrow
- CITE: CITE-seq mouse spleen
- Neutro: mouse neutrophils

From same AML mouse spleen

# Supp: Co-presence results for Visium

Role of fibroblasts highlighted