Single-cell and Spatial Multiome Data Analysis

Jing Zhang

12/xx/2025

Motivation

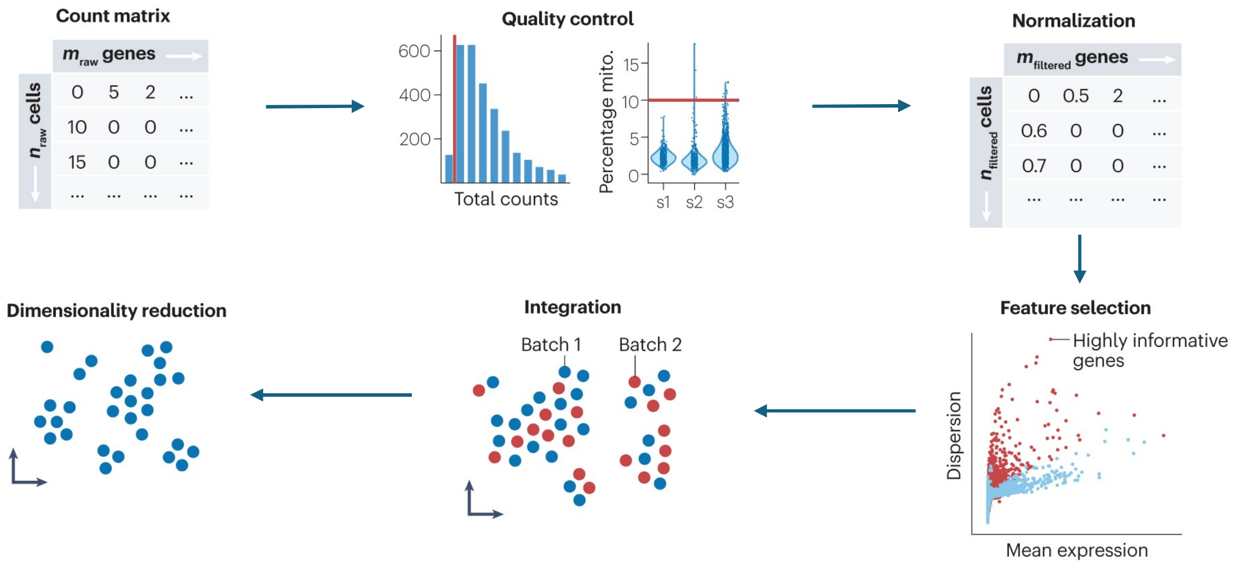
Single-cell Sequencing

- **Resolve cellular heterogeneity** beyond bulk averages
- **Identify cell types, states, and trajectories** in complex tissues
- **Link gene regulation to function** through RNA + chromatin modalities
- **Enable precision biology**: developmental mapping, disease mechanisms, therapeutic targets
- **Mechanistic insight at scale** — millions of cells across donors/conditions

Spatially-resolved Transcriptomics

- **Preserve spatial context** lost in dissociated single-cell assays
- **Map cell types to tissue architecture** (layers, niches, boundaries)
- **Identify spatial domains & microenvironments** driving function or disease
- **Resolve cell–cell interactions** via proximity, ligand–receptor patterns

scRNA-seq Pipeline



Figures adopted from Heumos *et al.*, Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023

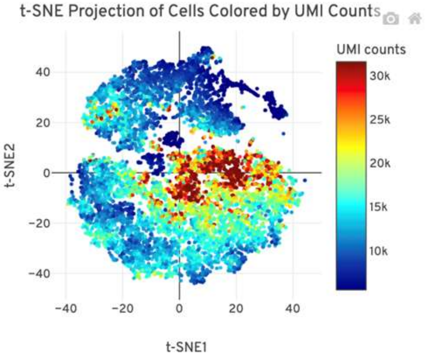
scRNA-seq Pipeline

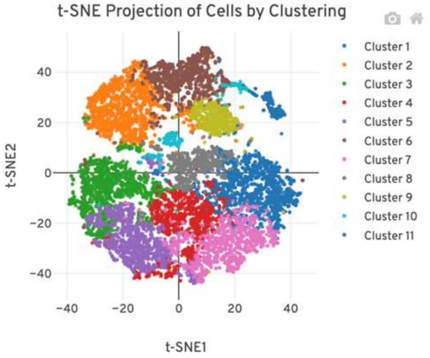
Count matrix generation

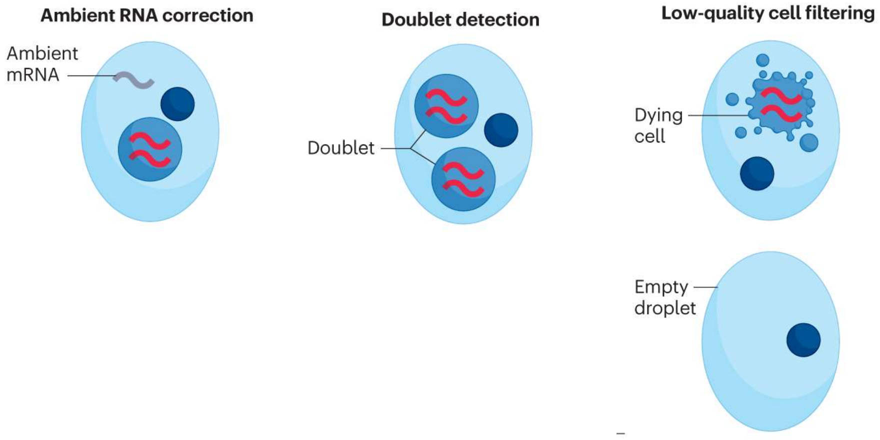
- Align sequenced reads to reference genome
- Tools: Cellranger (10X Genomics)

Quality control

- Remove low-quality cells
- Correct for biological & technical variations
- Tools: Seurat/Scanpy


t-SNE Projection of Cells Colored by UMI Counts


t-SNE Projection of Cells by Clustering



https://www.10xgenomics.com/support/software/cell-ranger/7.2/analysis/outputs/cr-flex-outputs-web-summary-multi
Figures adopted from Heumos *et al.*, Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023

scRNA-seq Pipeline

Normalization

- Adjust for sequencing depth and technical variability
- Make gene expression comparable across cells
- Tools/Algorithms: log-normalization, SCTransform (Scanpy, Seurat)

Feature Selection

- Identify informative genes to improve downstream dimensionality reduction and clustering
- Tools: Highly Variable Gene in Seurat and Scanpy

Batch Integration

- Correct for donor, batch, and experimental effects while preserving true biological signals and enabling joint analysis
- Tools: Harmony, LIGER, scVI

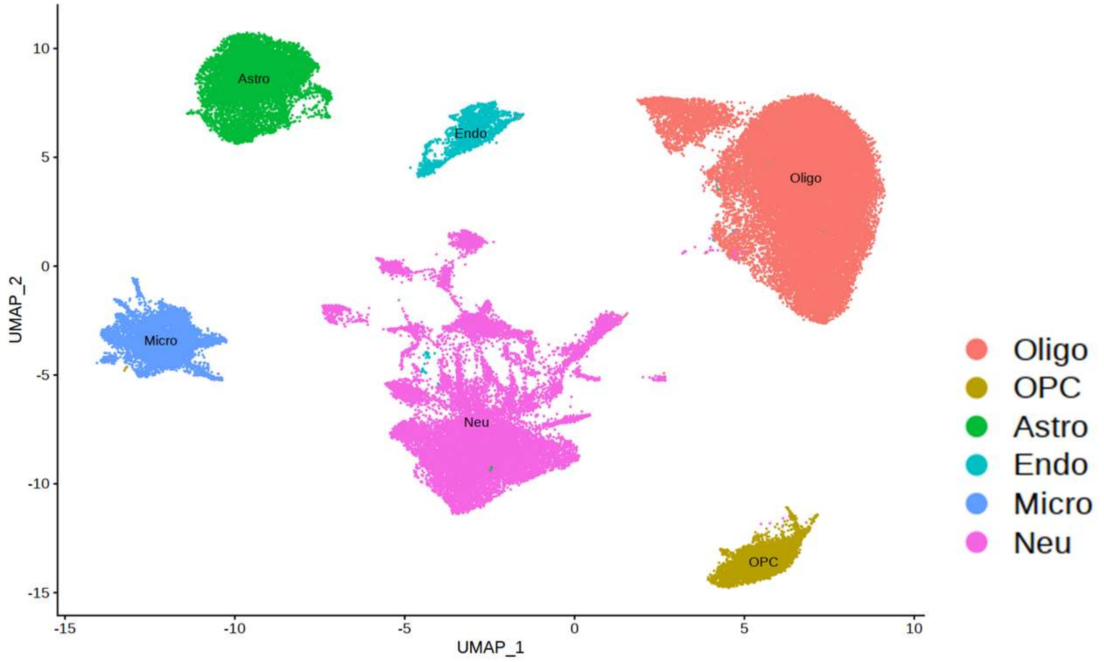
scRNA-seq Pipeline

Dimensionality Reduction

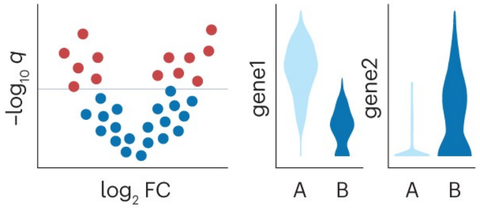
- Compress high-dimensional data for efficient analysis & visualization
- Tools: PCA, UMAP/t-SNE

Clustering and Cell Type Annotation

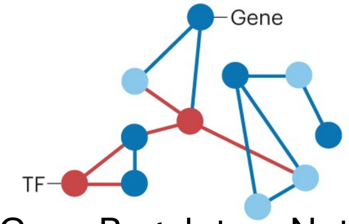
- Identify groups of cells with shared transcriptional profiles
- Use biologically meaningful markers to assign cell types
- Tools: LEIDEN/Louvain (clustering), Azimuth (annotation)



scRNA-seq Downstream Tasks

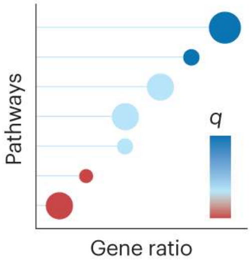


Differentially Expressed Genes

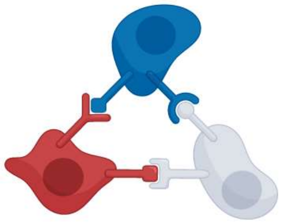


Gene Regulatory Networks

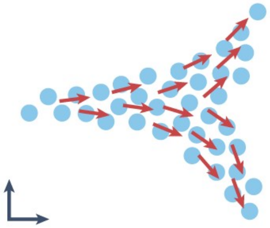
scRNA-seq



Functional Genomics



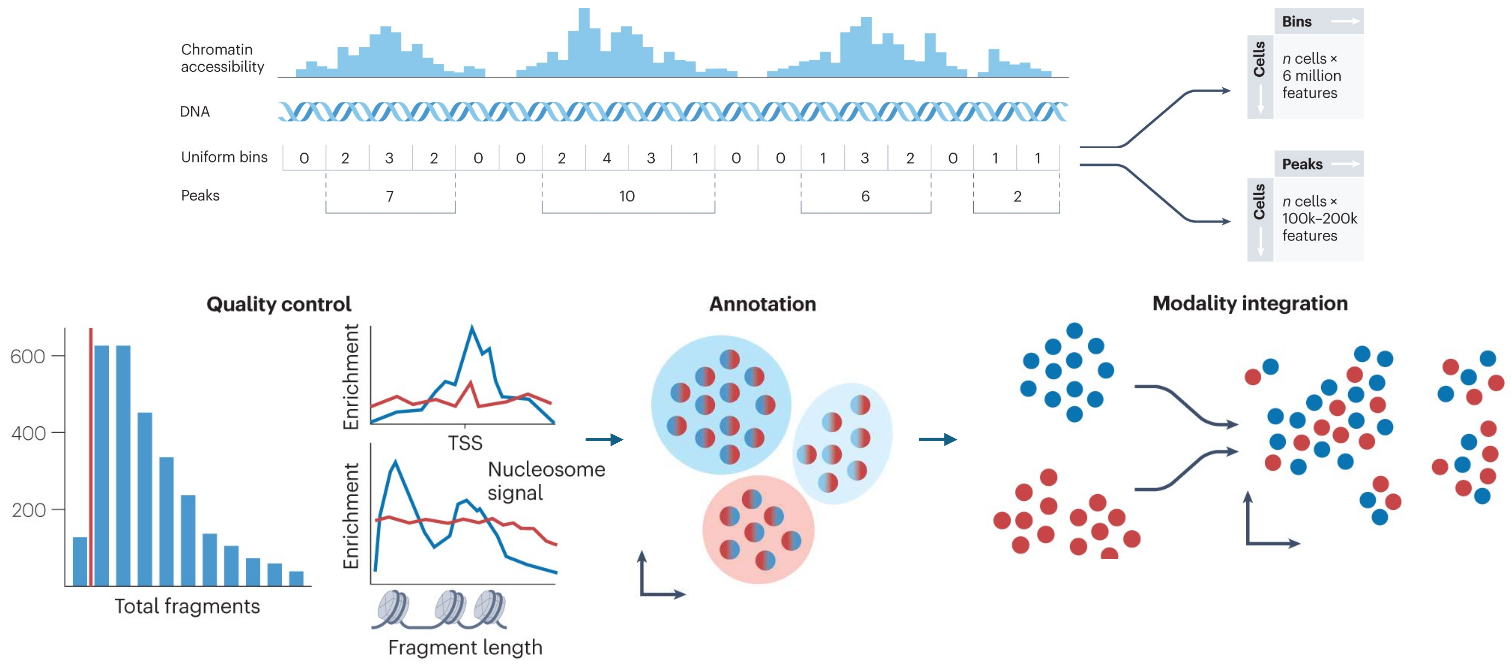
Cell-to-Cell Communications



Trajectory Inference

Figures adopted from Heumos *et al*., Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023

scATAC-seq Pipeline



Figures adopted from Heumos *et al.*, Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023

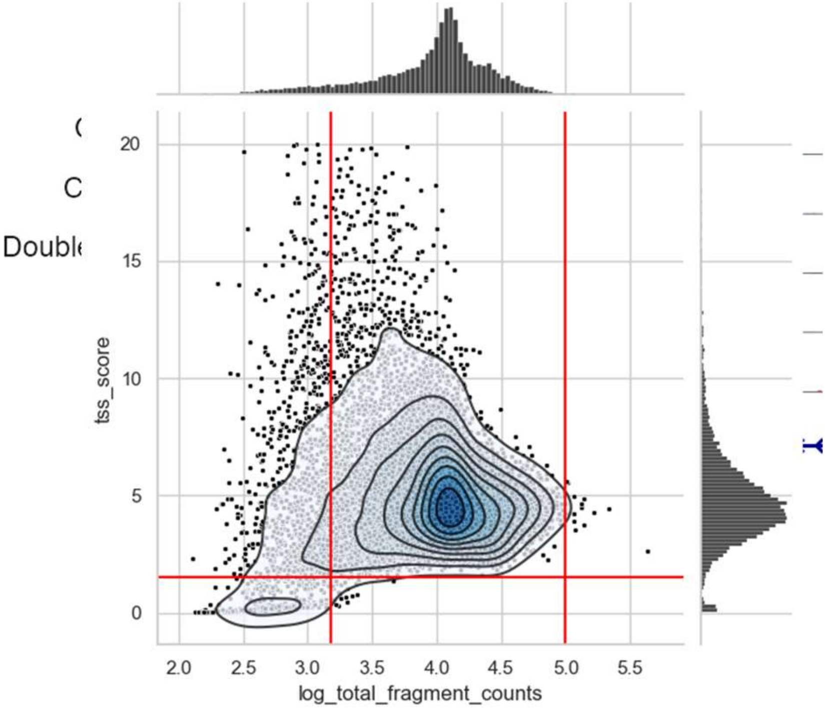
scATAC-seq Pipeline

Peak Calling

- Identify **informative** regions by aggregating reads across cells to detect statistically enriched "**peaks**"
- Tools: MACS3

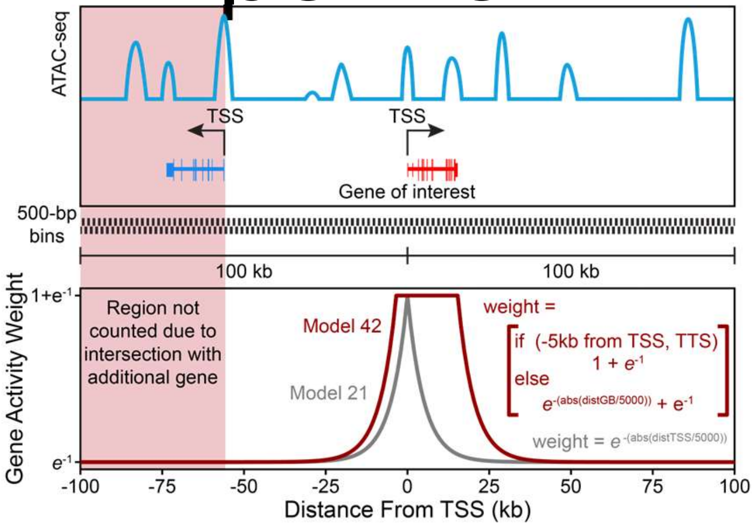
Quality Control

- Filter out **low-quality nuclei** (low fragment counts, low TSS enrichment, low fragments in peaks)
- Tools: ArchR, Signac, SnapATAC



https://stuartlab.org/signac/articles/visualization
https://www.sc-best-practices.org/chromatin_accessibility/quality_control.html

scATAC-seq Pipeline



Dimensionality Reduction

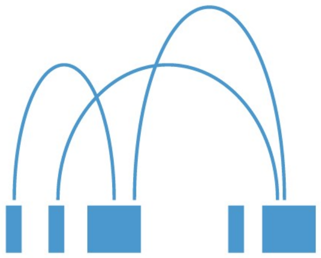
- Compress high-dimensional (often > 300k) **binarized** data for efficient analysis & visualization (similar to scRNA-seq)
- Tools: Latent Semantic Indexing (LSI)

Clustering, Gene Activity Scores, and Cell Type Annotation

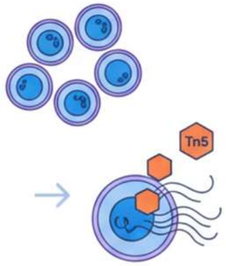
- **Aggregate** similar cells with Louvain/LEIDEN clustering
- **Calculate** gene activity scores (chromatin accessibilities around genes)
- **Assign** cell types together with biologically meaningful markers
- Tools: ArchR, Signac

https://www.archrproject.com/bookdown/calculating-gene-scores-in-archr.html

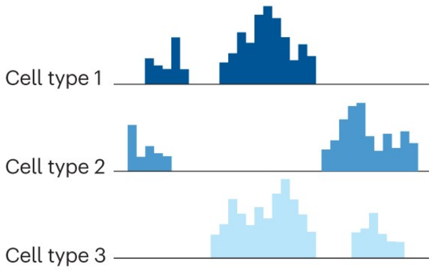
scATAC-seq & Multiome Downstream Tasks



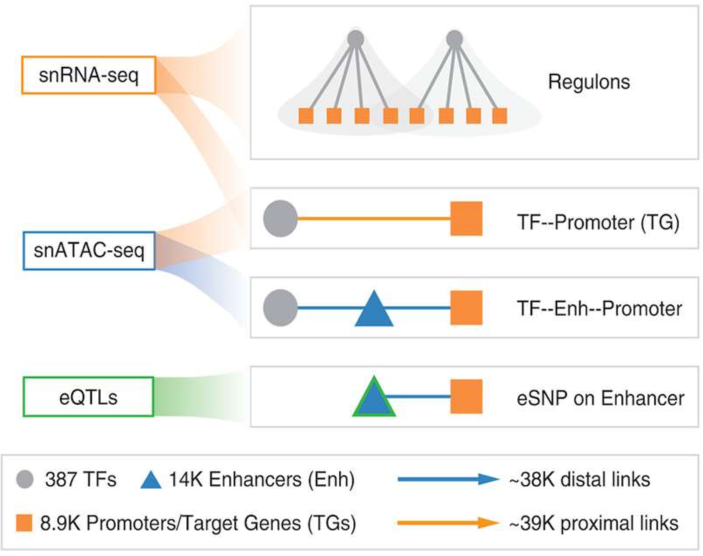
co-accessibility



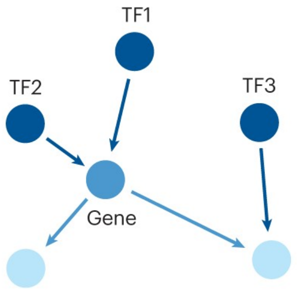
scATAC-seq



Differentially accessible peaks



TF-CRE-Motif Linkages

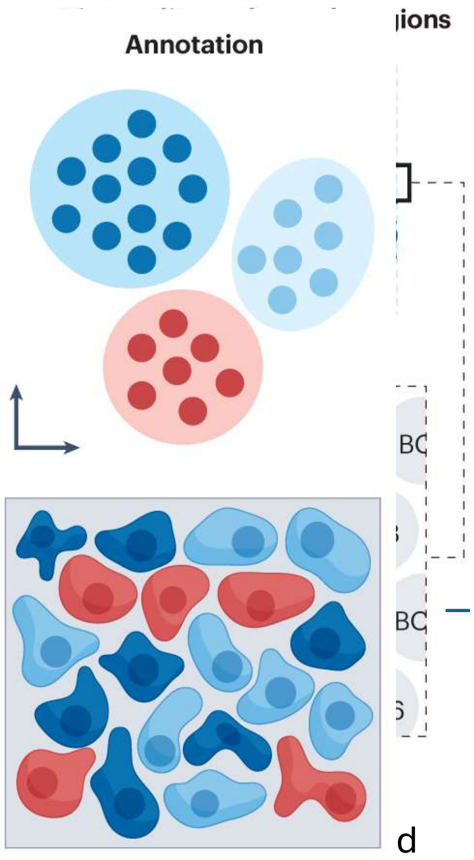


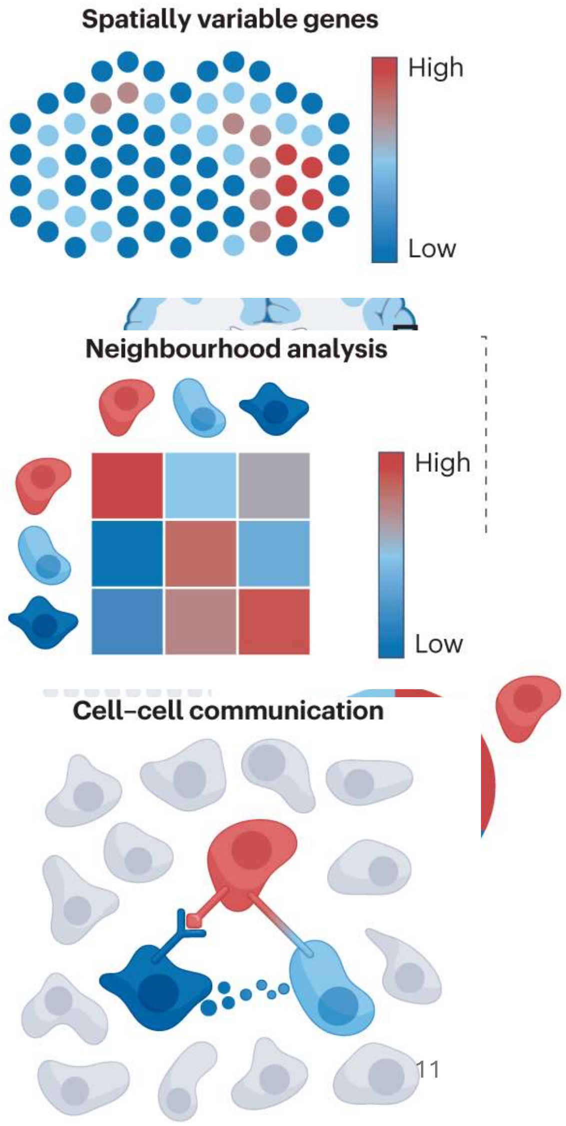
Gene Regulatory Networks

Figures adopted from Heumos *et al.*, Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023
Emani *et al.*, Single-cell genomics and regulatory networks for 388 human brains, *Science*, 2024

Spatial Transcriptomics and Multi-omics

- **Very similar to single-cell sequencing, but with spatial coordinates**
  - Two types: sequencing-based, imaging-based
  - Resolutions vary from spots with multiple cells to subcellular resolution
  - Technologies: 10X Visium, Xenium, Visium HD, Stereo-seq, ……
- **Downstream Tasks**
  - Domain Identification, Spatially Variable Genes, Neighborhood Analysis, Cell-to-Cell Communications, Histology Image Integration





Figures adopted from Heumos *et al*., Best practices for single-cell analysis across modalities, *Nature Reviews Genetics*, 2023

Biological Insights

- From scRNA-seq:
  - **Cellular heterogeneity revealed** — identification of rare cell types, activation states, differentiation intermediates
  - **State-specific transcriptional programs** become visible, enabling discovery of disease-associated or treatment-responsive subpopulations
- From scATAC-seq:
  - **Regulatory landscape mapping** — active enhancers, promoters, and TF motifs driving specific cell states can be inferred
  - **Chromatin dynamics** reveal early regulatory shifts that precede gene expression changes
- From scMultiome:
  - **Cell-level enhancer–gene linkage** reveals regulatory elements controlling lineage programs
- From spatial transcriptomics:
  - **Cellular organization in situ** — mapping cell types to tissue domains, identifying microenvironments and niche-specific signals

Fiture Directions

- Multi-omic Integration (RNA + ATAC + Spatial)
- Unified Foundation Models for Single-Cell & Spatial Data
- Generative Modeling & *in-silico* Perturbations
- Regulatory Sequence–to–Function Modeling
- Multi-resolution & Cross-omics Alignment
- Integrative Regulatory Network Inference
- ……