

Single-Cell Transcriptomics: From Bulk to Precision Biology

Resolving Cellular Heterogeneity, One Cell at a Time
*Jiekun (Jackie) Yang, Department of Genetics, Rutgers
University, 2/26/2025*

Why Single Cells?

Bulk RNA-seq Limitations

- Loss of Cellular Heterogeneity
- Difficulty in Identifying Rare Cell Types



Bulk

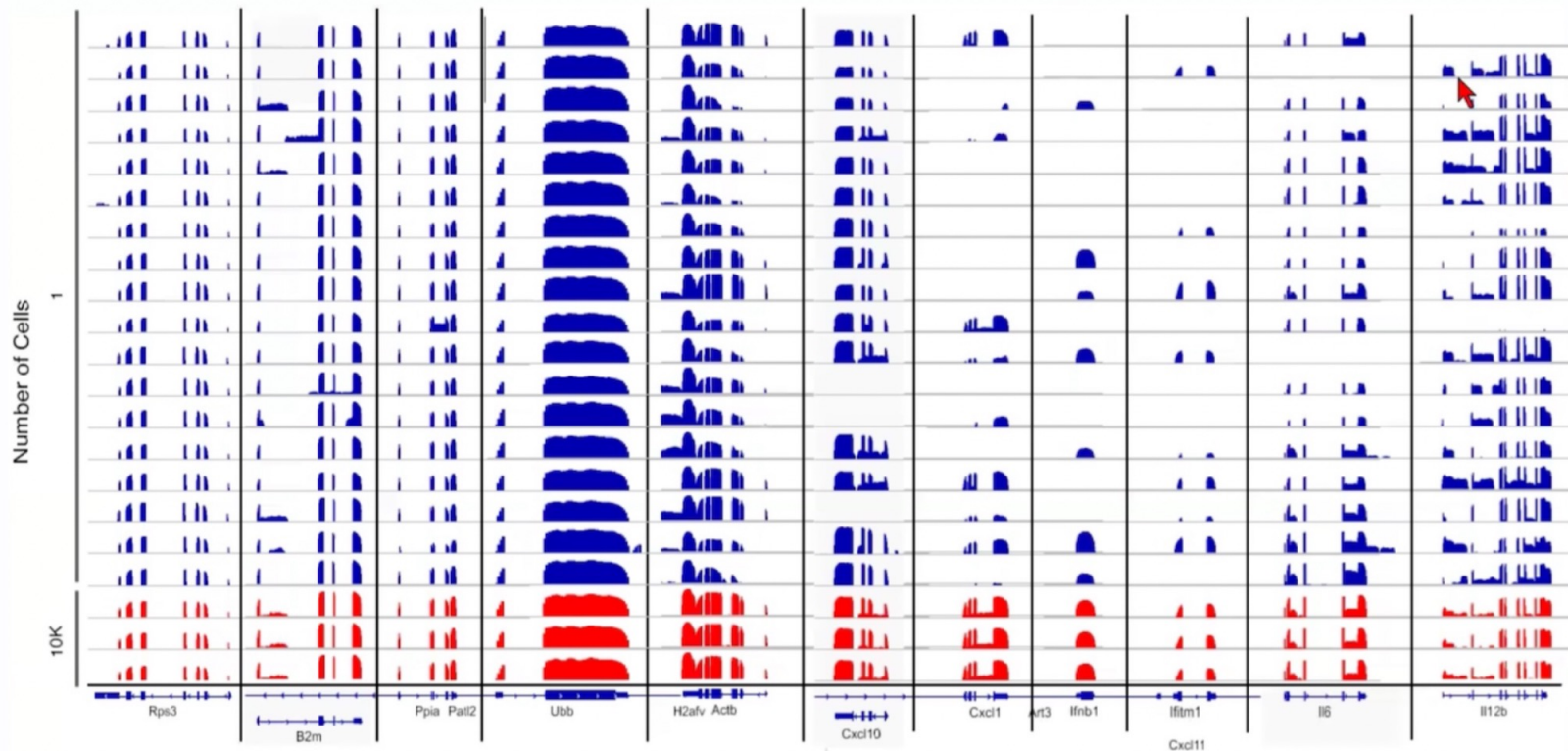


Single Cell

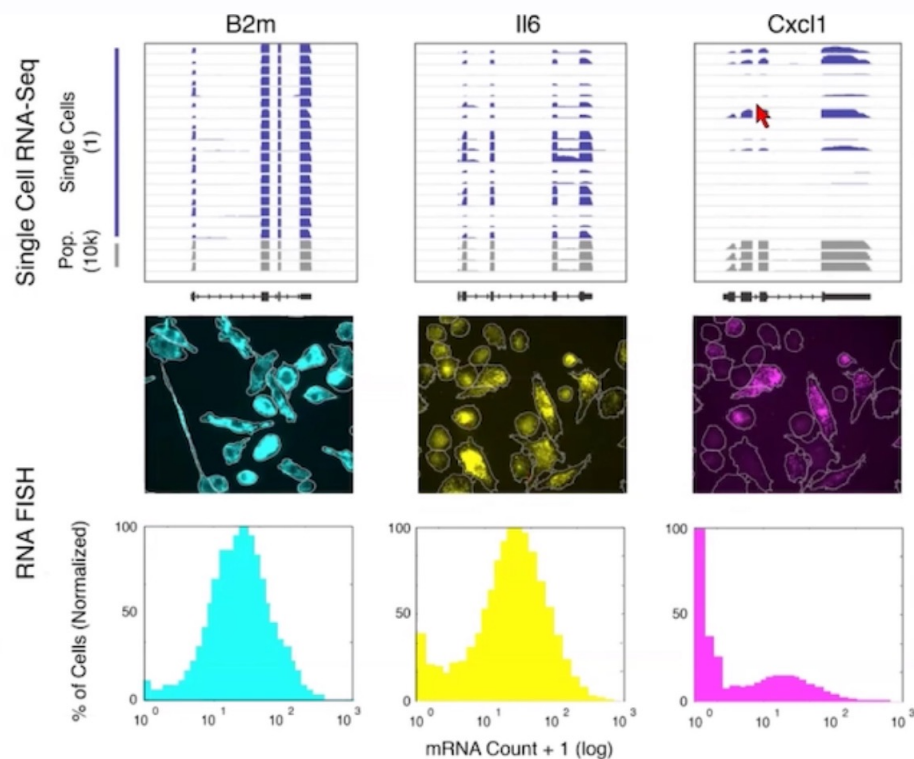


Spatial

- Inability to Study Cell-State Transitions
- Limited Resolution for Cell-Type-Specific Signatures
- Difficulty in Detecting Lowly Expressed Genes



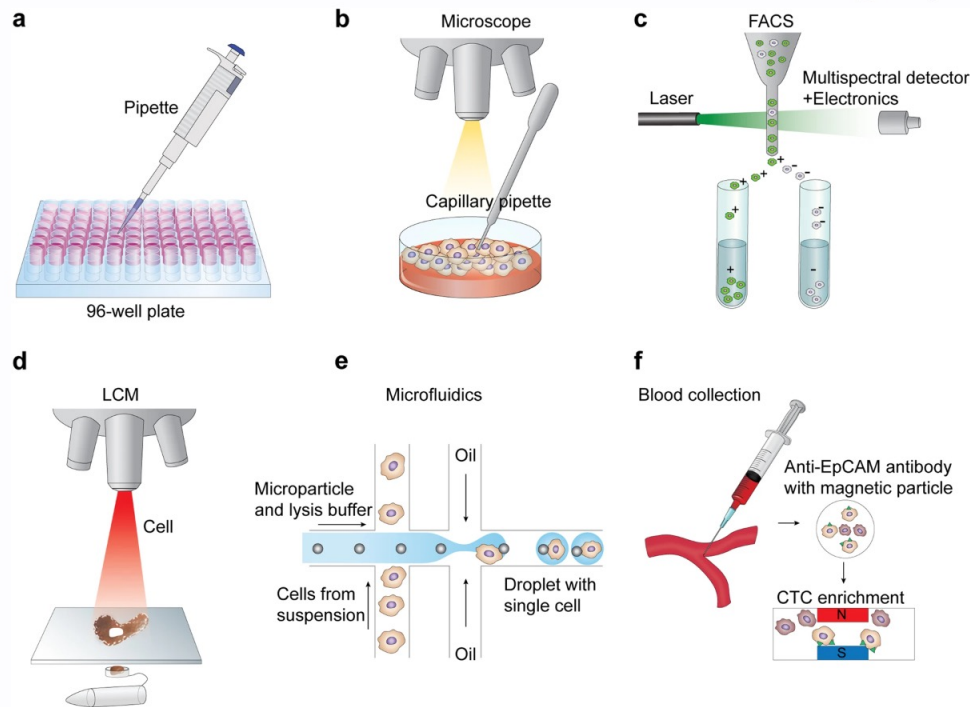
RNA FISH validates scRNA-seq cell to cell variability



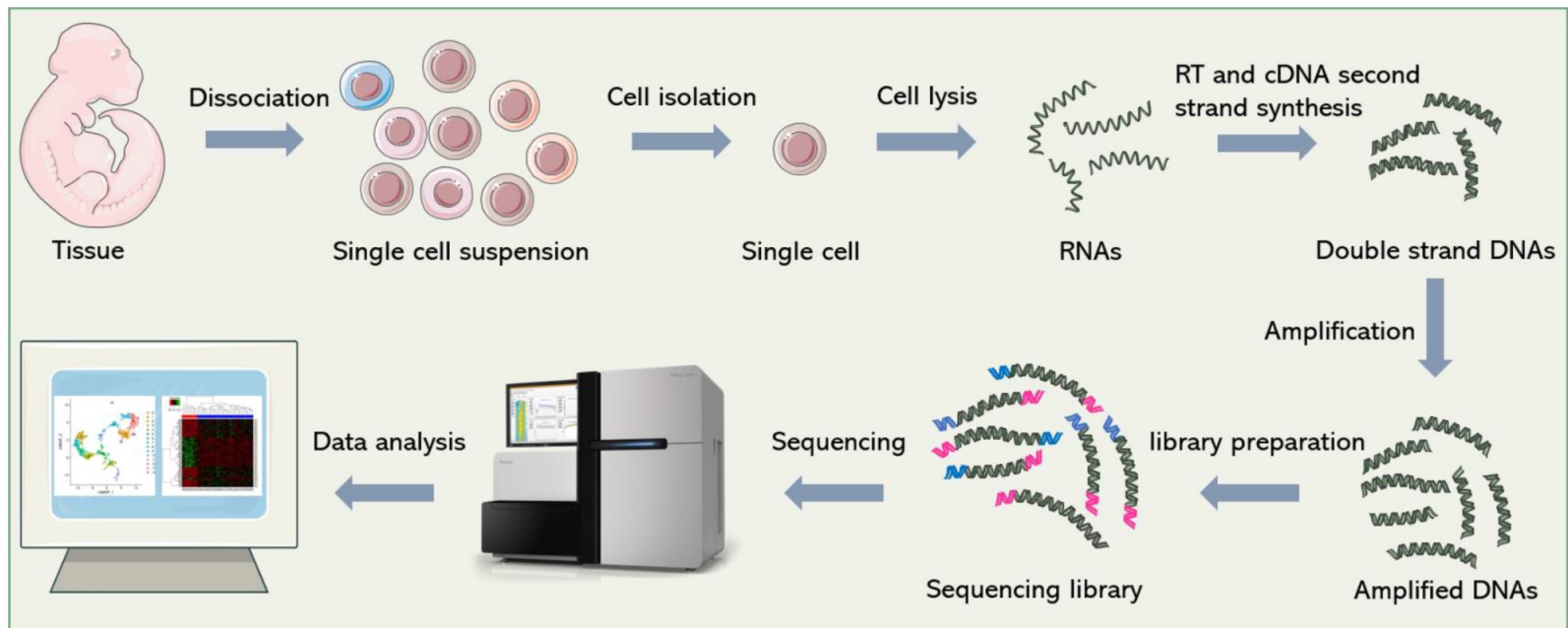
Traditional Approaches

FACS: Limited to known markers, heterogeneity exists

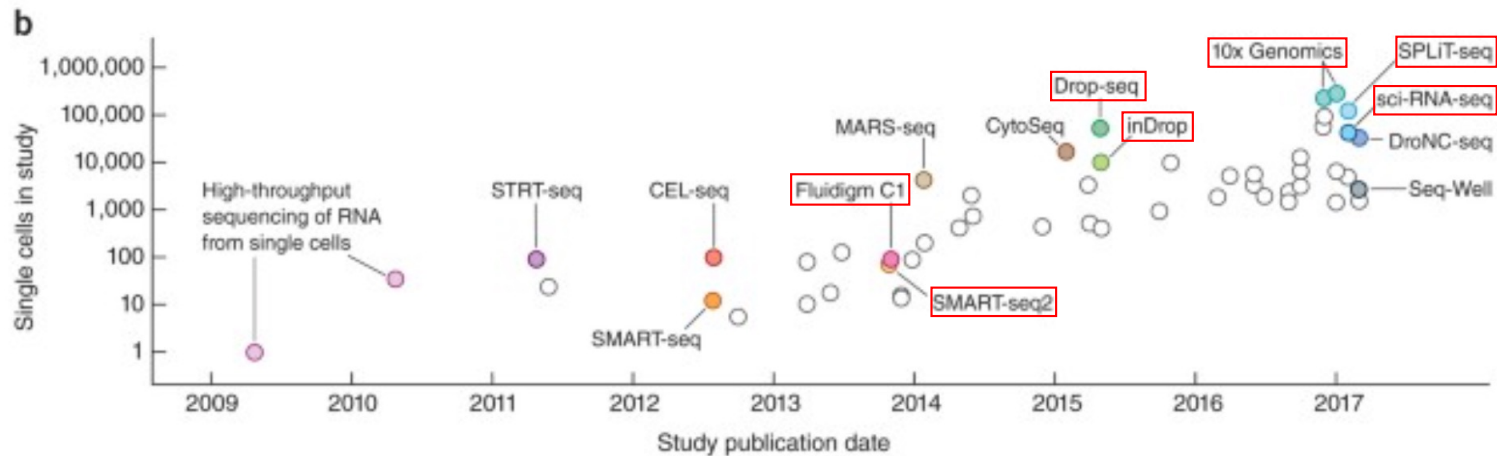
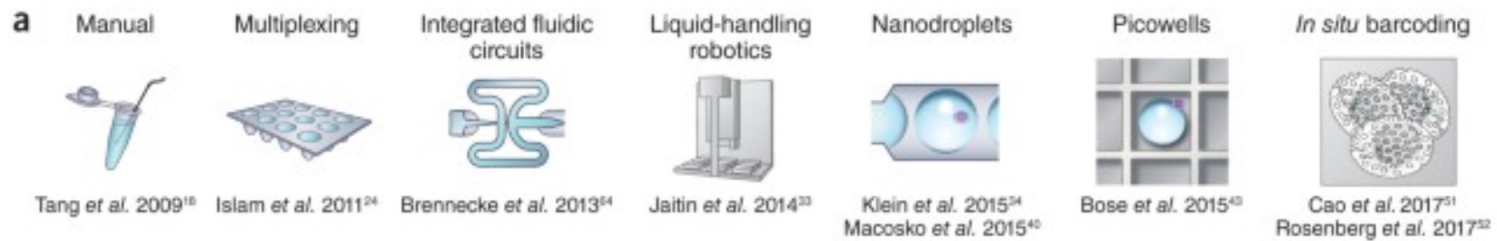
Microdissection: Low throughput



scRNA-seq Workflow

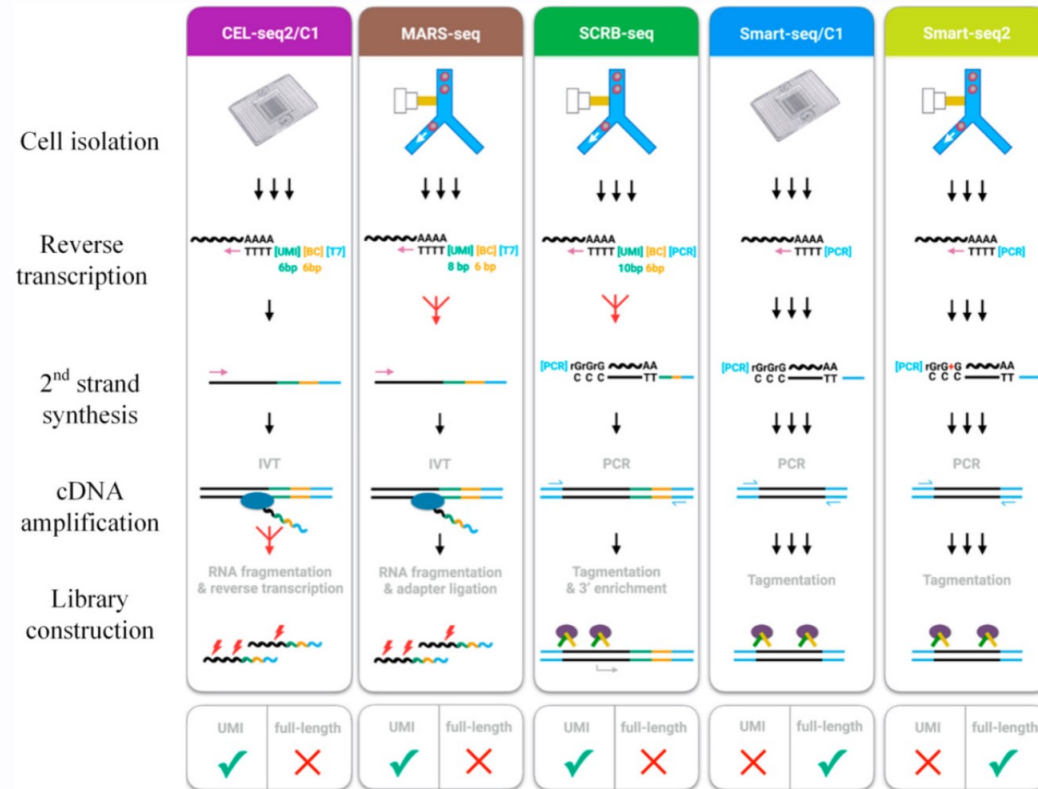


Technology Evolution



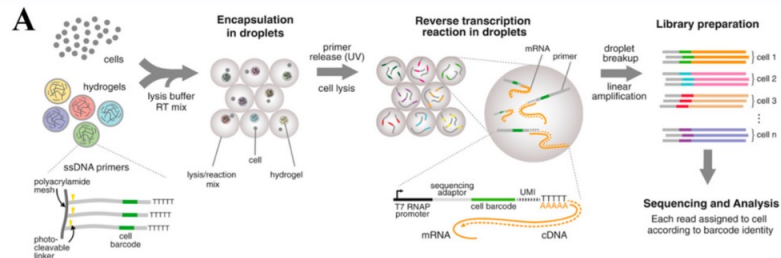
Conventional scRNA-seq

Plate-based platforms: low throughput

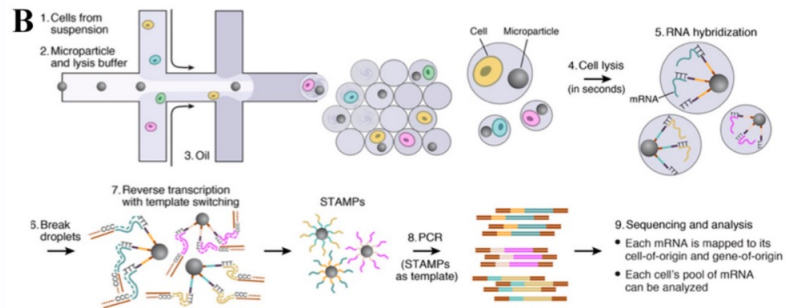


High-Throughput Revolution

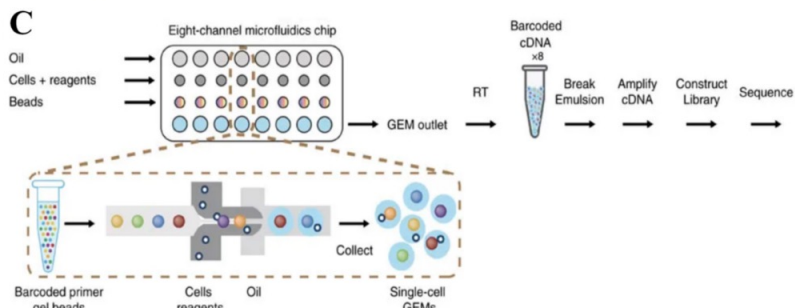
In-Drop



Drop-seq



10X Genomics

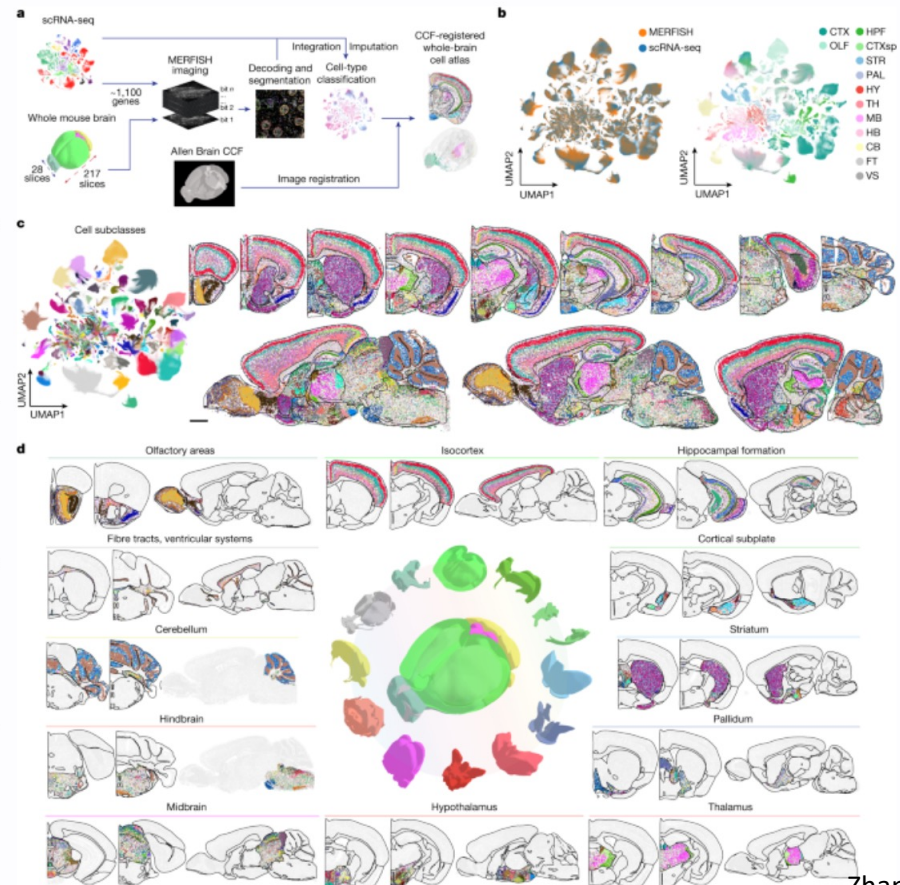
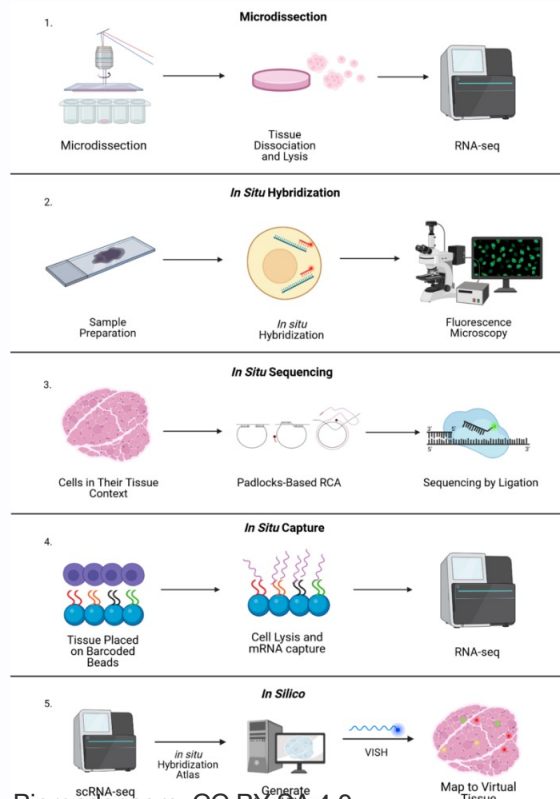


Sequential combinatorial barcoding

SPLIT-seq/sci-RNA-seq



Spatial Transcriptomics



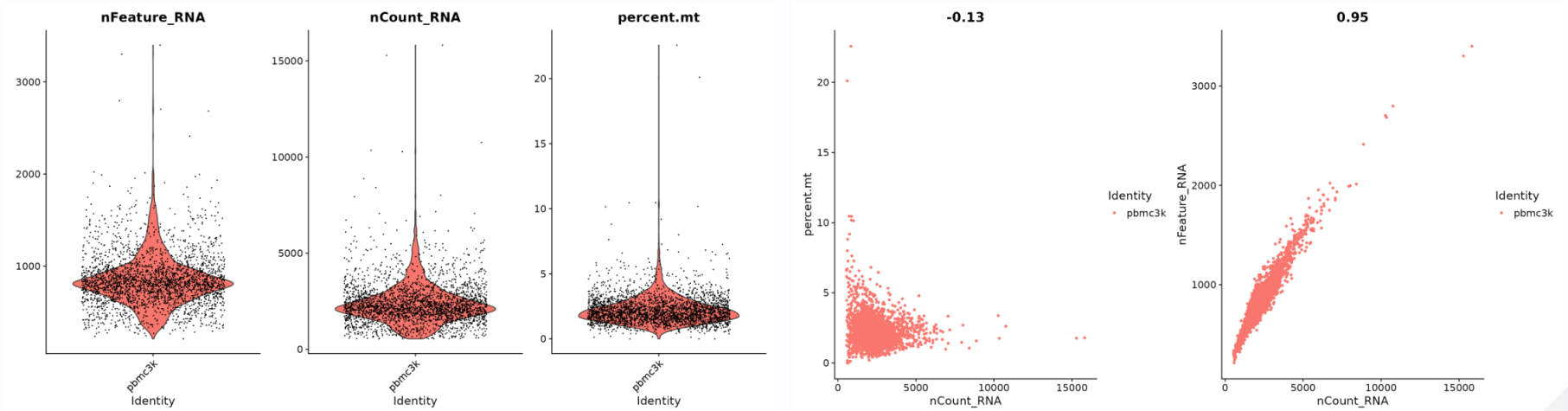
By SlifertheRyeDragon - Biorender.com, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=101273866>

Computational Challenges

Data Preprocessing

Quality control (Seurat, CellBender)

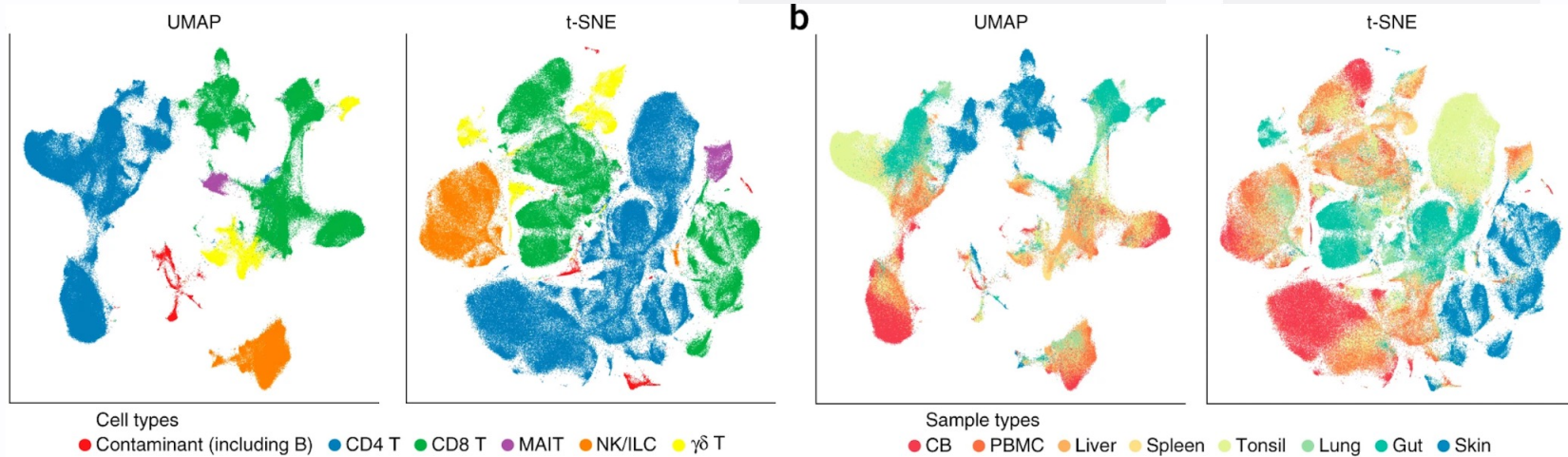
Normalization (Log-normalize, SCTransform).



Dimensionality Reduction

PCA: Linear projection, preserves global variance.

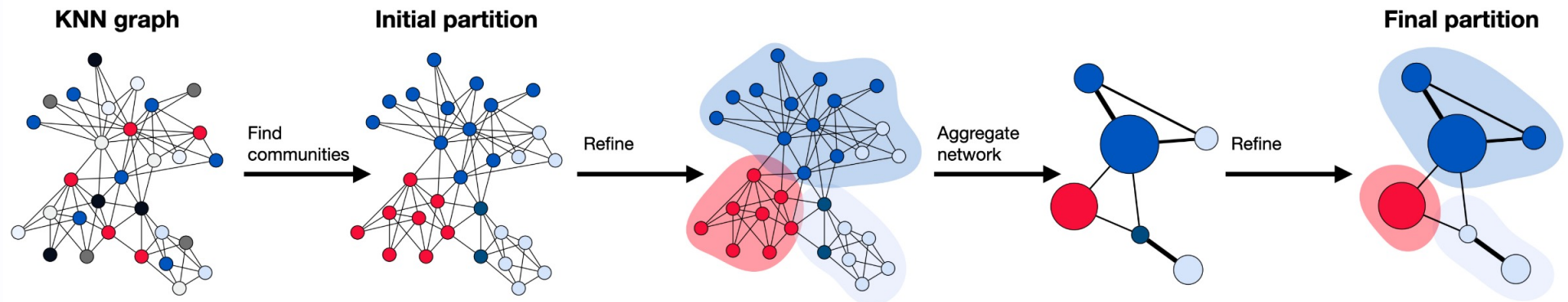
t-SNE/UMAP: Constructs neighbor graph → optimizes low-dimensional layout (`n_neighbors` , `min_dist`).



Clustering

Leiden: Avoids disconnected communities.

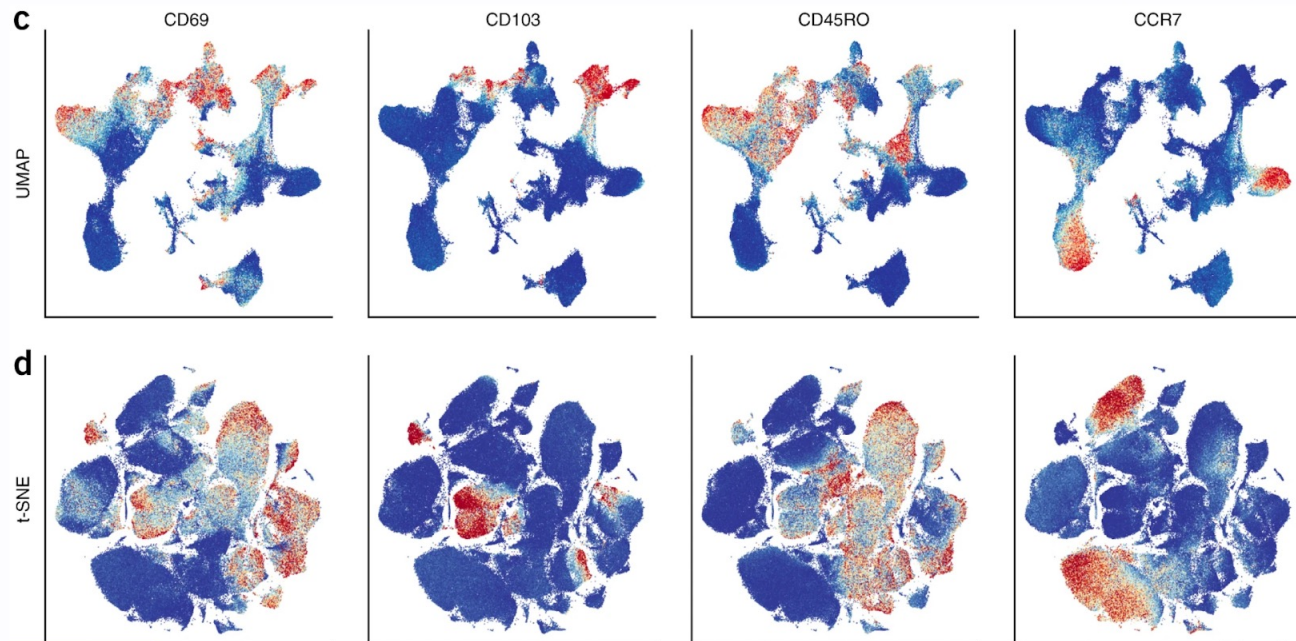
Louvain: Maximizes modularity.



Resolution: 0.2 (broad) vs. 1.0 (fine-grained).

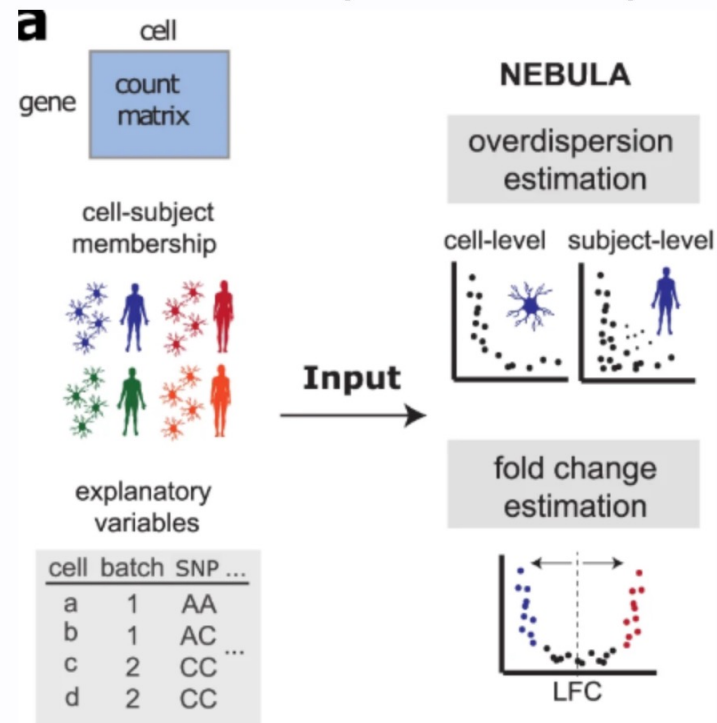
Cell Type Annotation

Marker genes, reference databases (CellMarker, PanglaoDB), tools (SingleR, CellTypist)



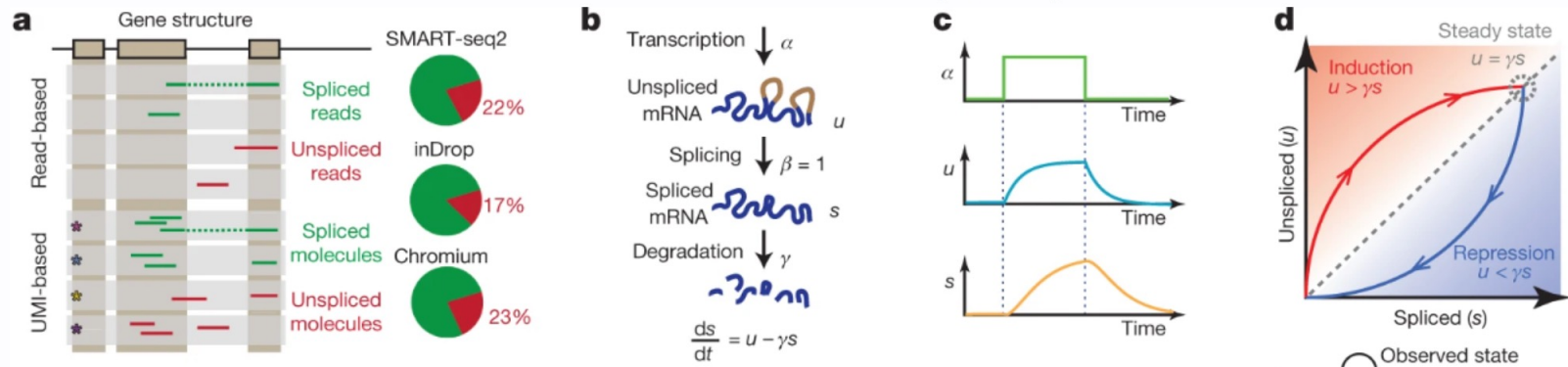
Differential expression

Cell-based (MAST), Pseudobulk (DESeq2), Linear-Mixed Models (NEBULA)

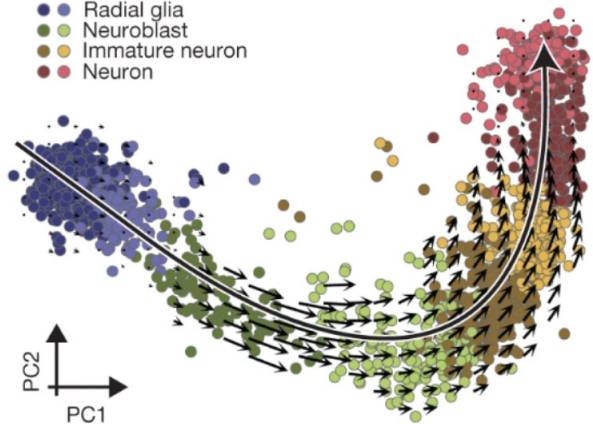


Trajectory inference

Monocle3, RNA velocity, CytoTRACE

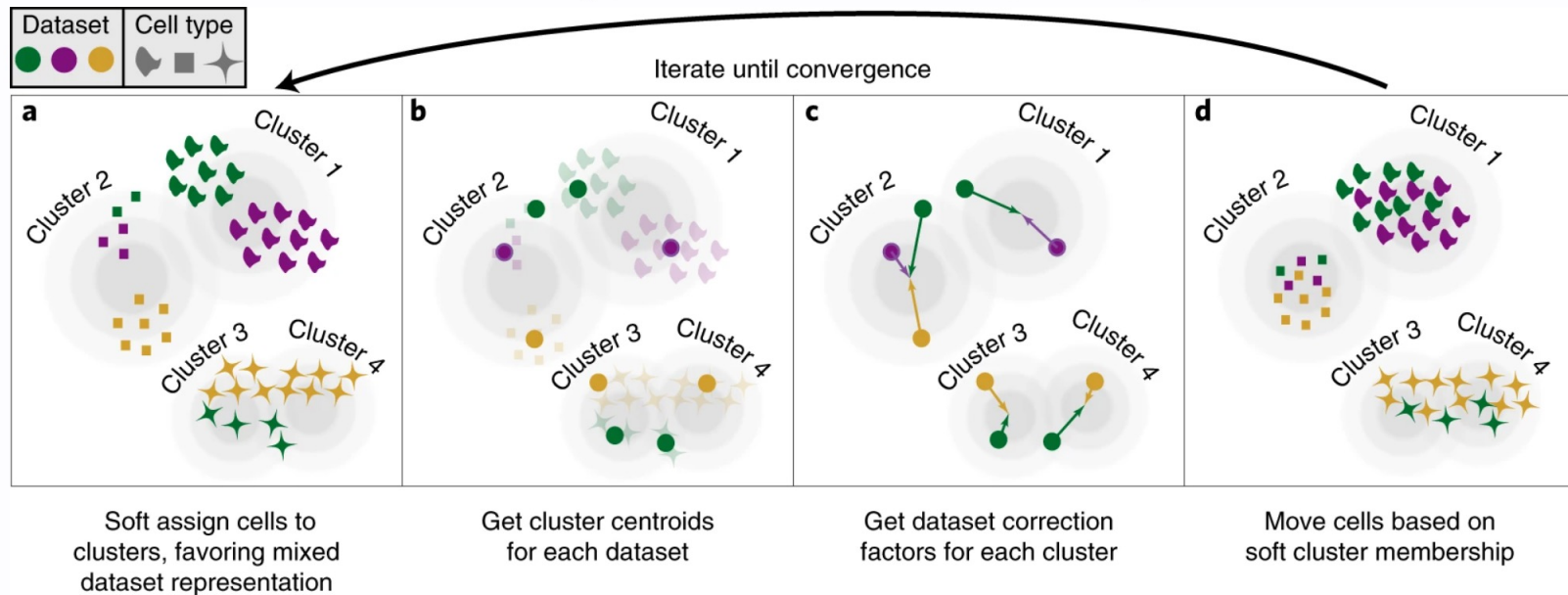


- a**
- Radial glia
 - Neuroblast
 - Immature neuron
 - Neuron



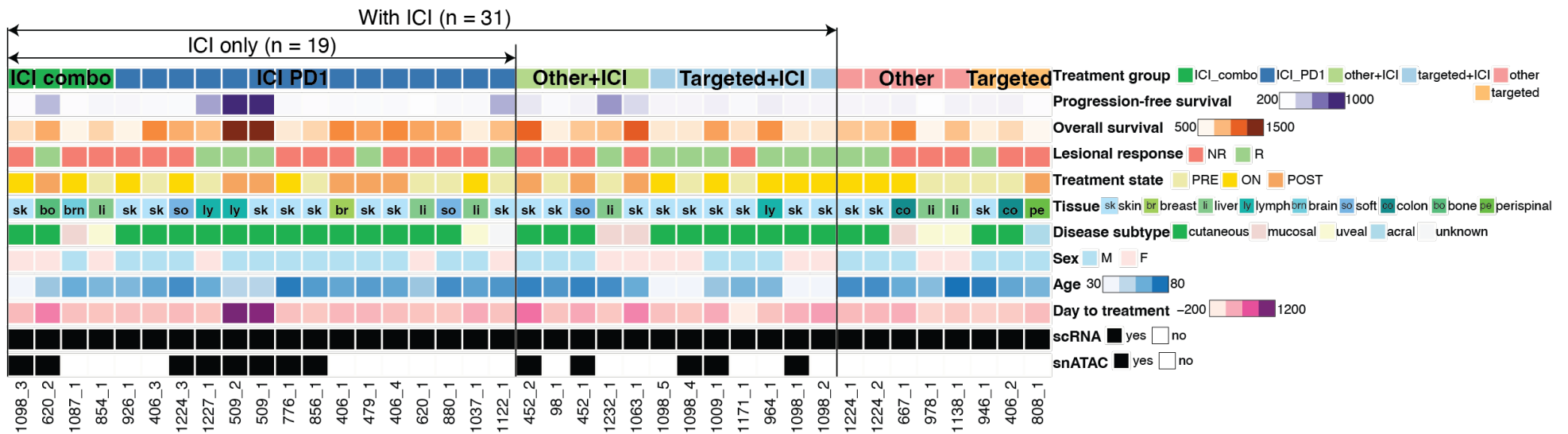
Multi-Dataset Integration

Harmony, Seurat CCA, ML/DL

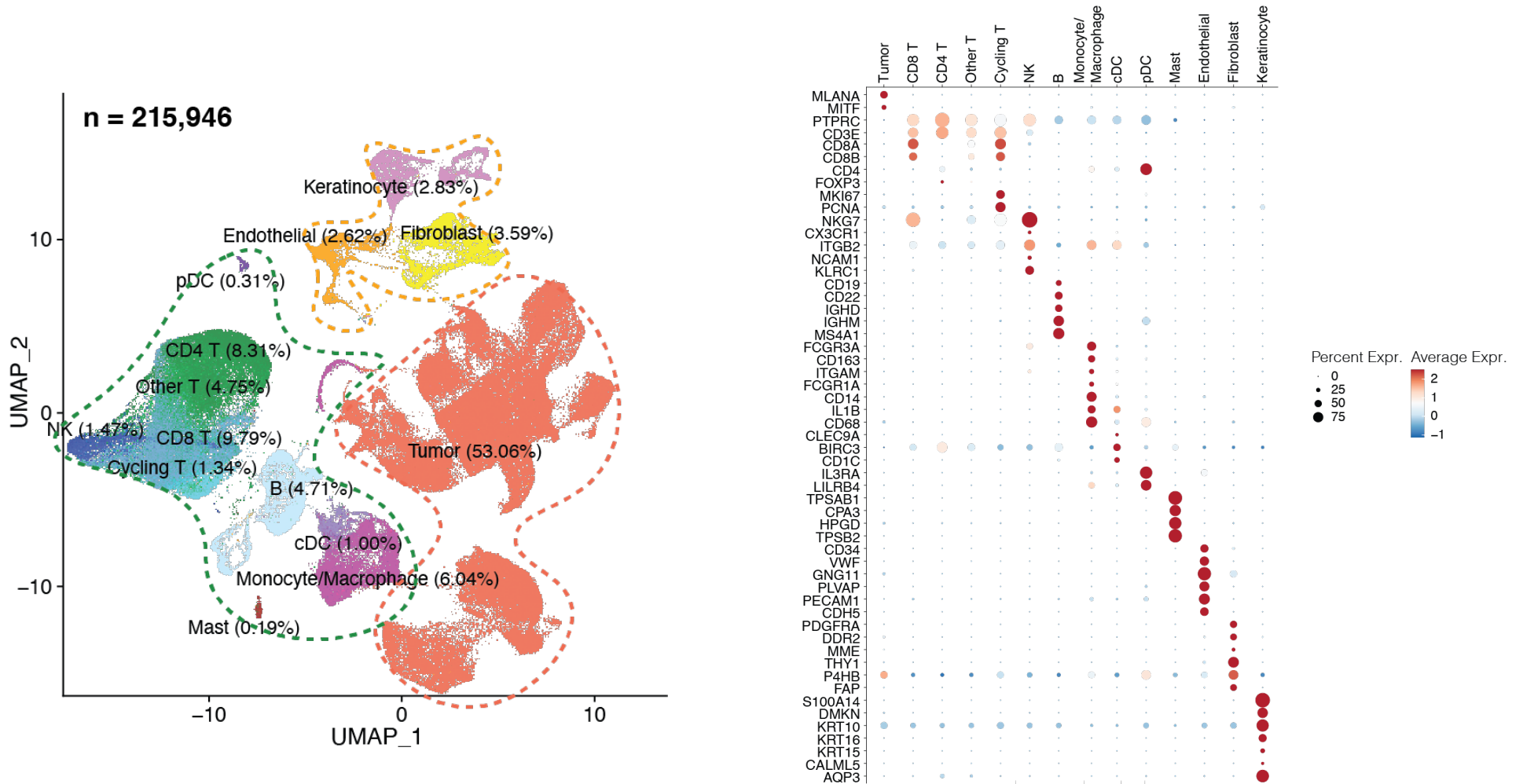


Case Studies: Single-cell Dissection of Immunotherapy Response in Metastatic Melanoma

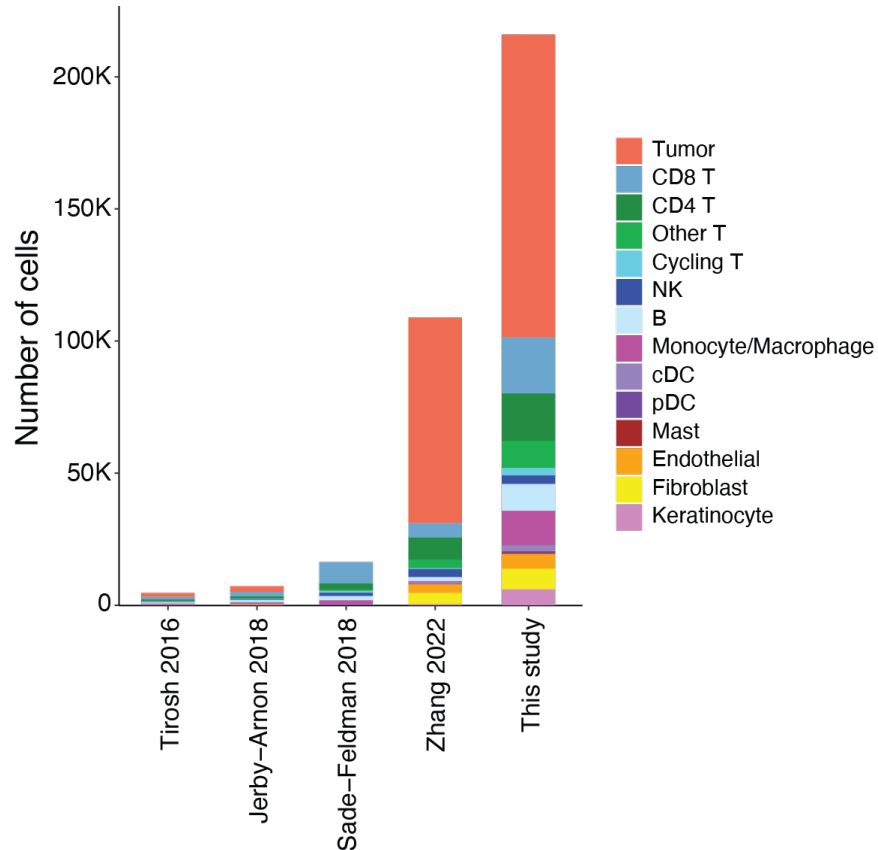
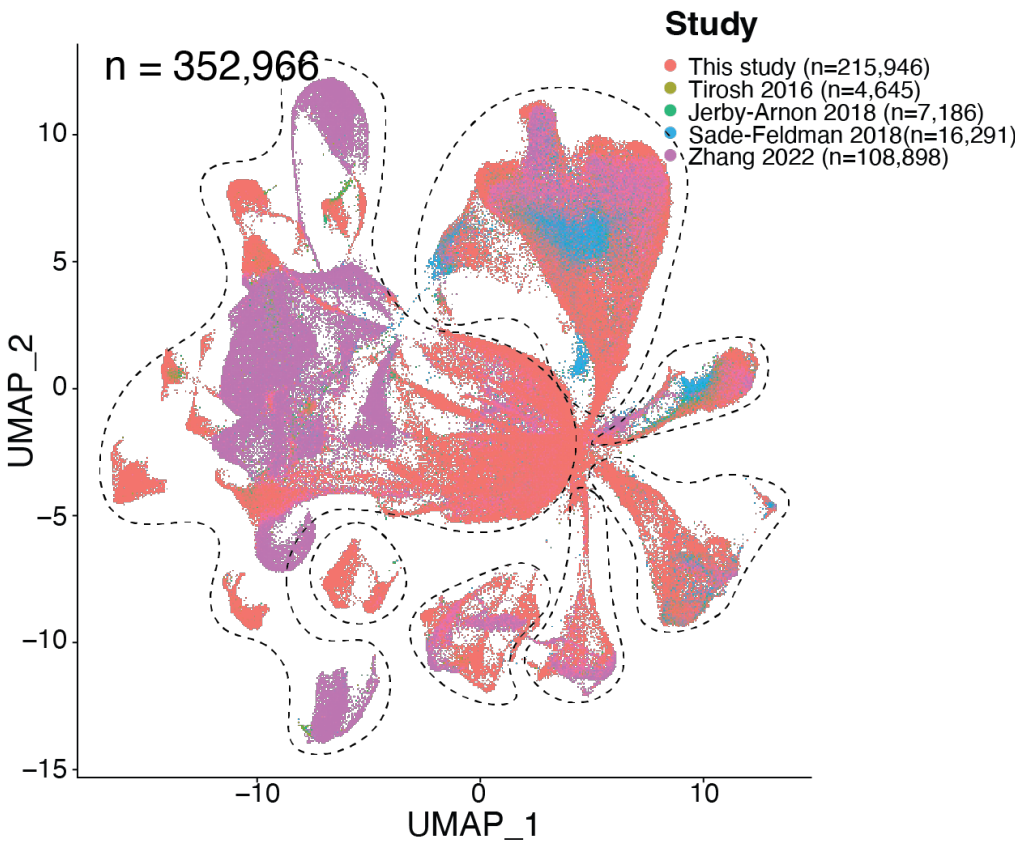
We used 39 scRNA-seq and 15 paired snATAC-seq samples



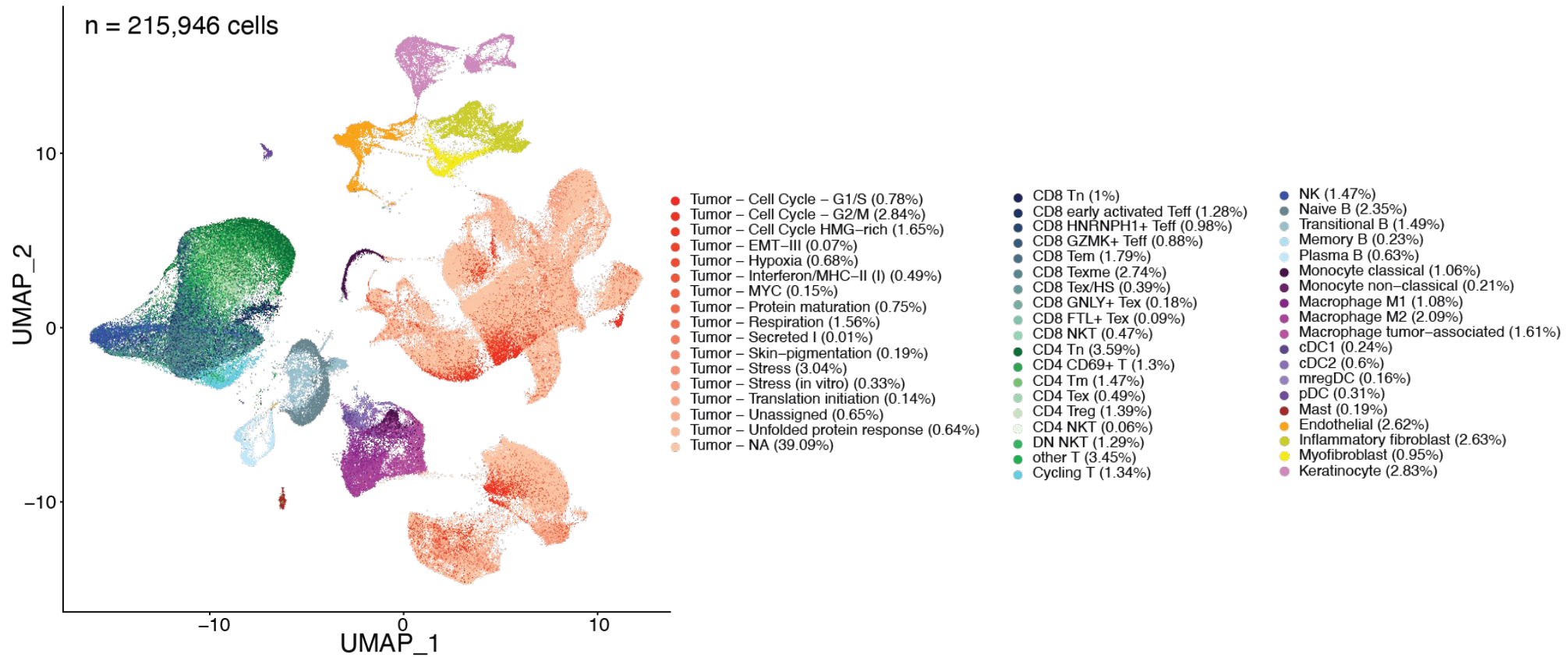
We annotated 14 cell types for ~216K cells



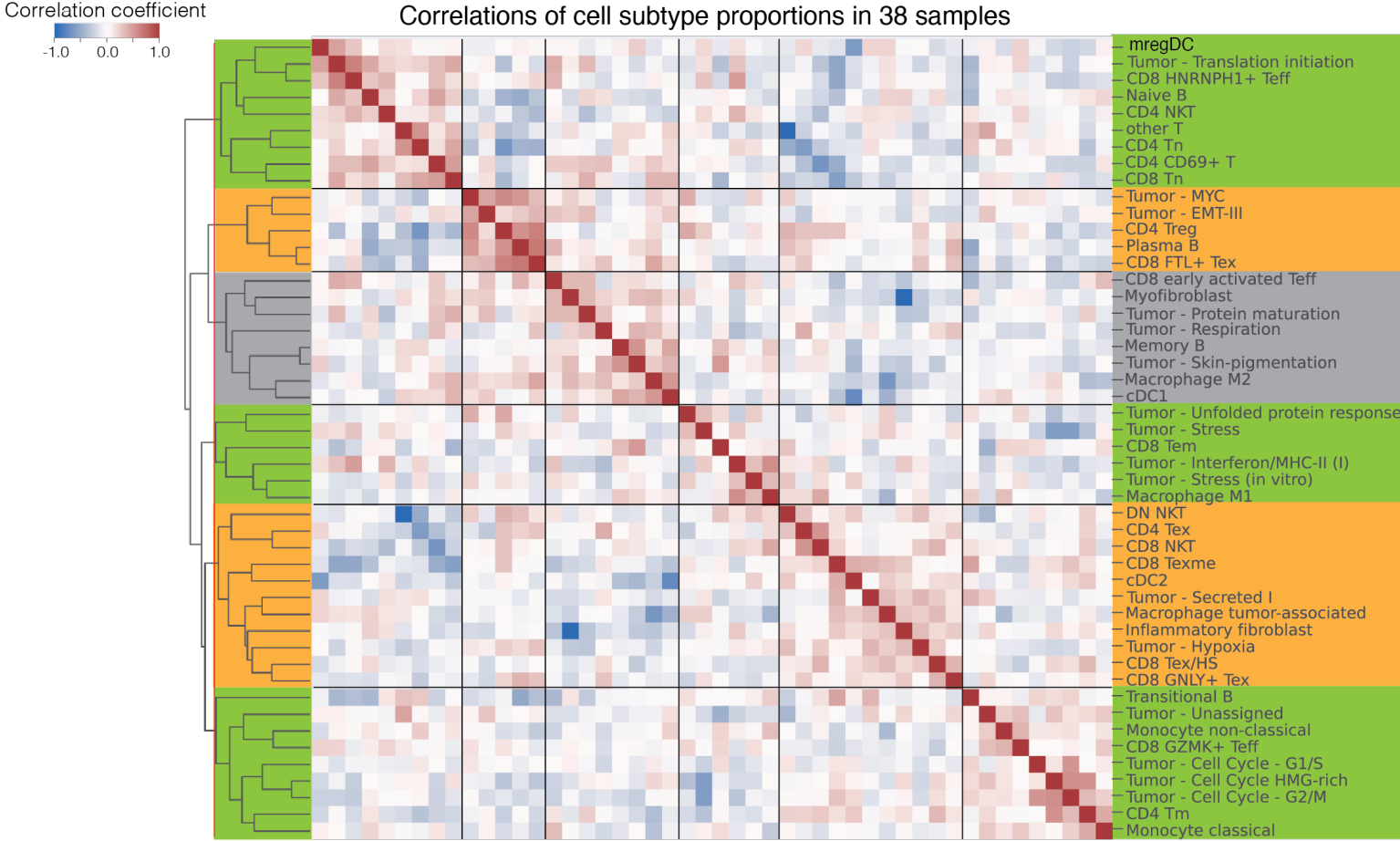
We contributed more cells than all previous studies combined



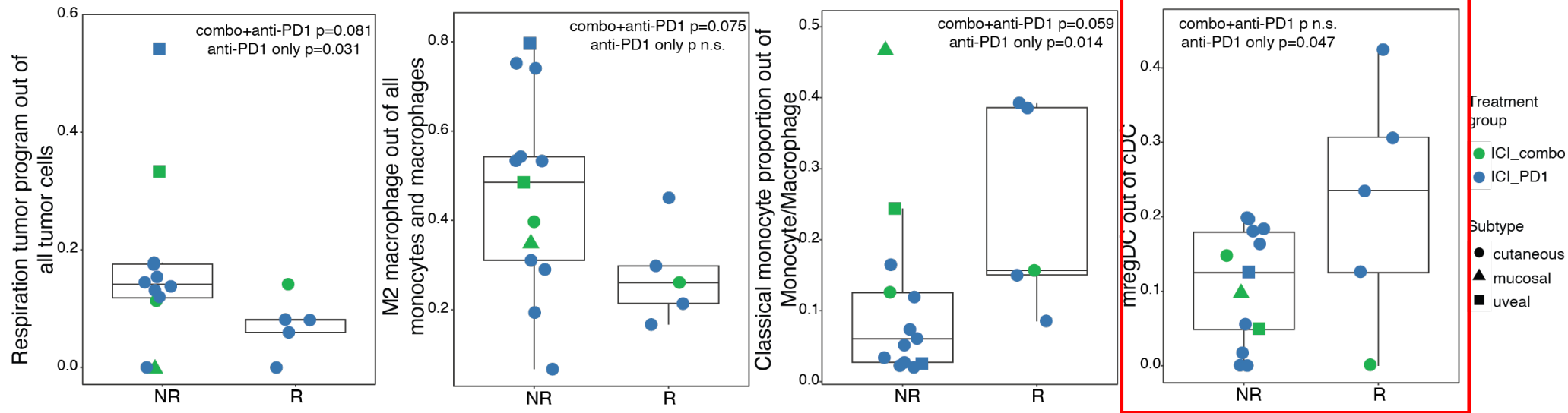
We identified 55 cell subtypes for ~216K cells



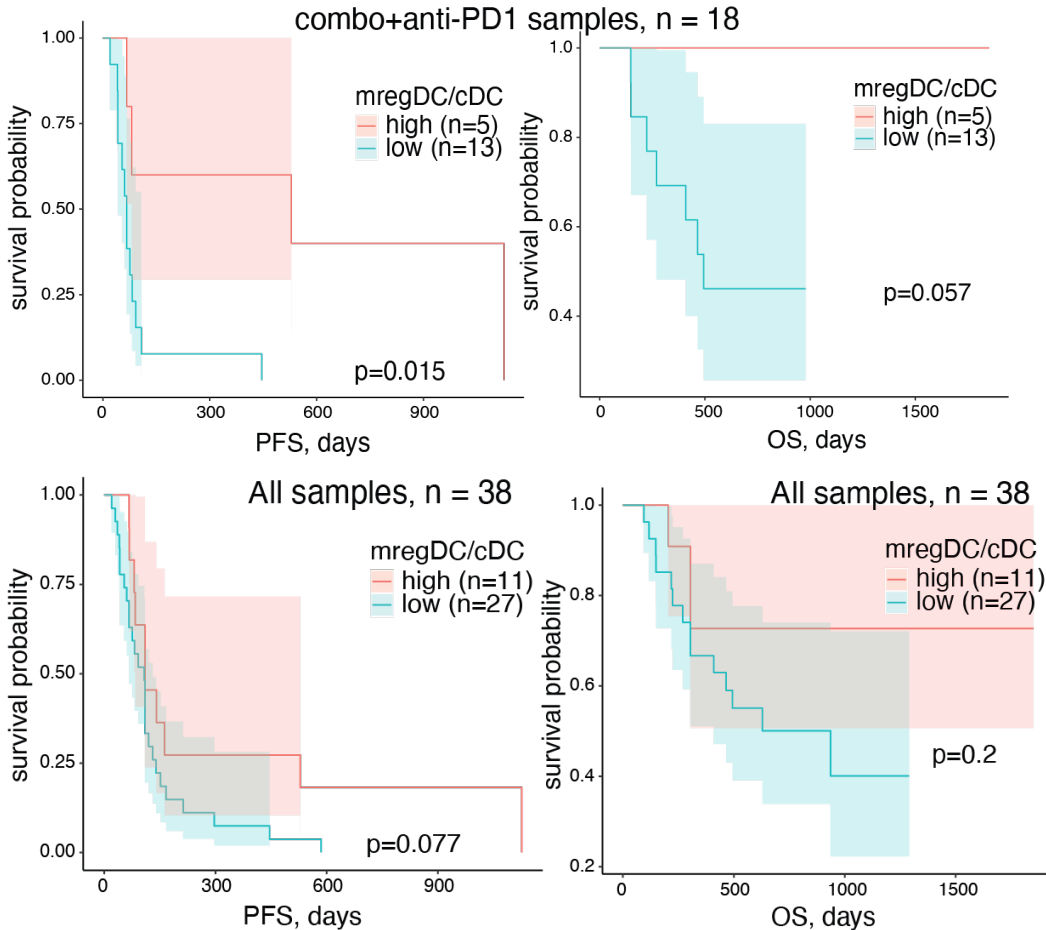
We revealed 6 cell subtype correlation modules with varied tumor responses



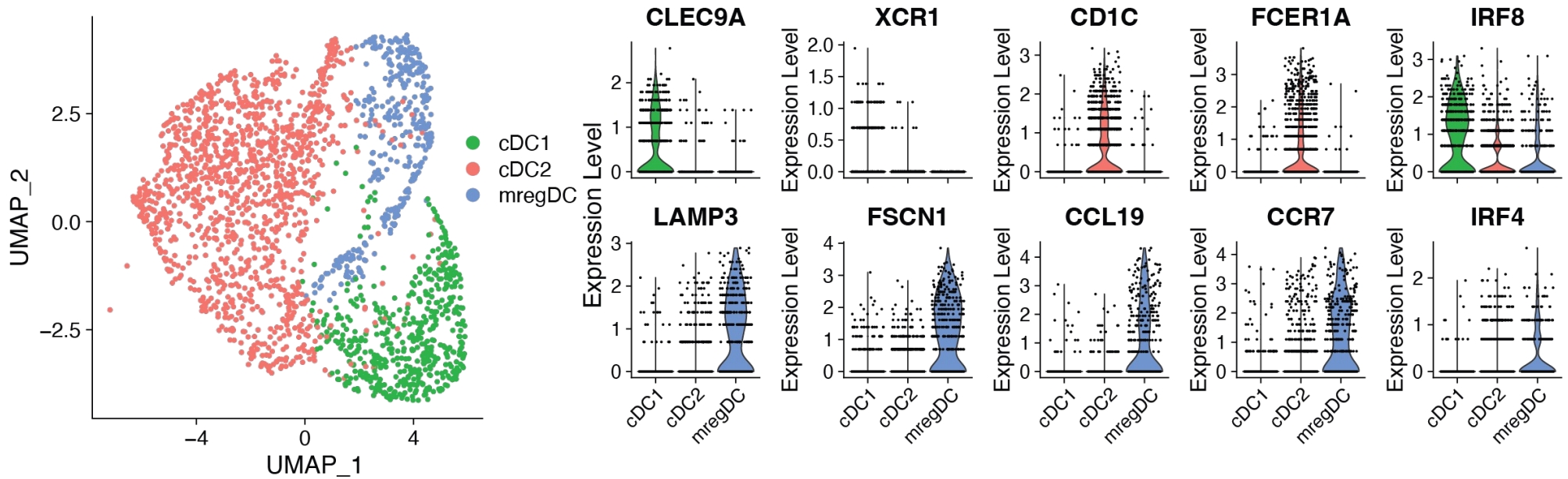
We detected cell subtypes associated with ICI response



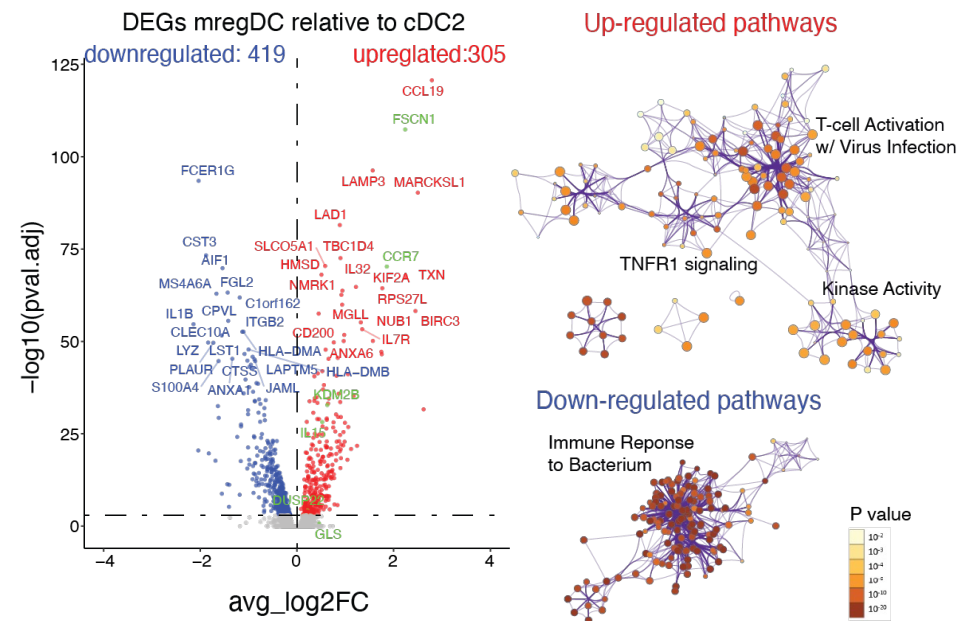
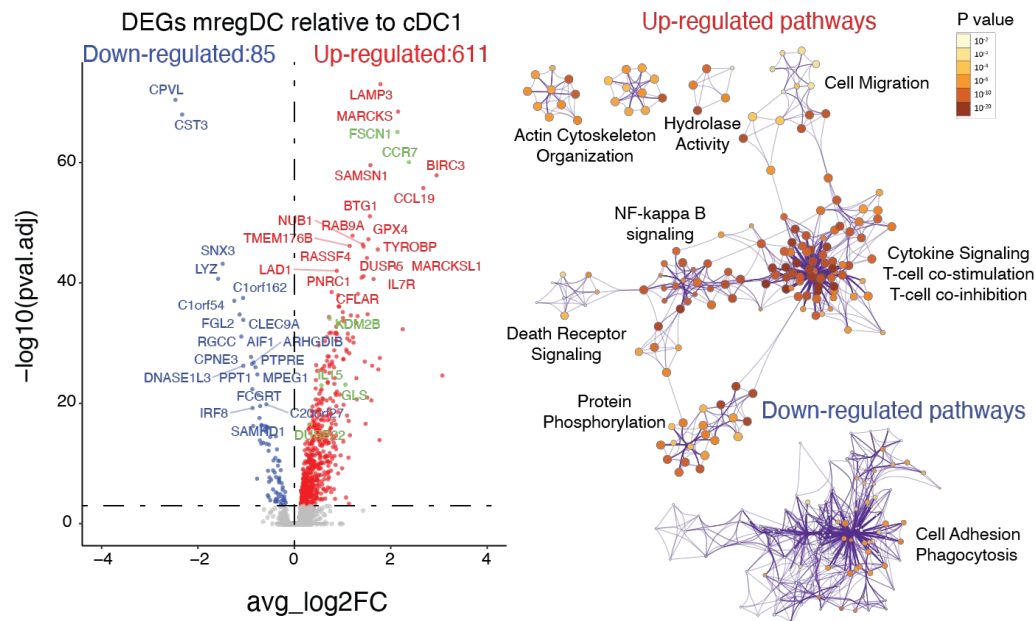
We observed PFS benefit for ICI-treated mregDC-high samples



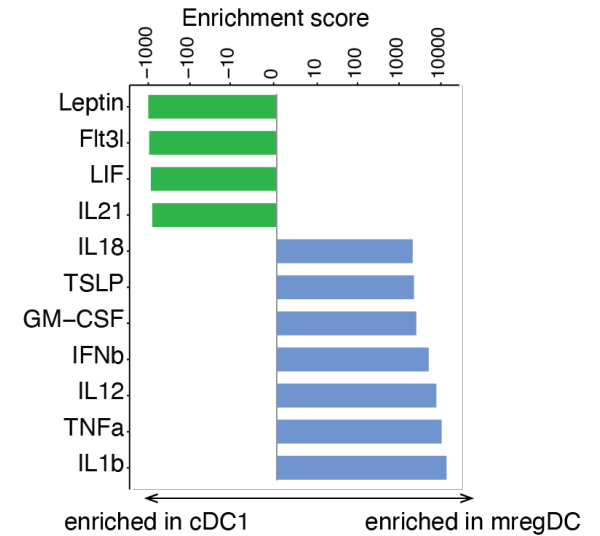
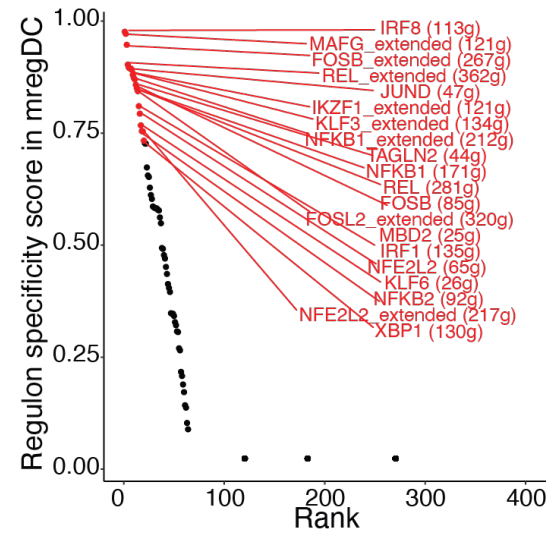
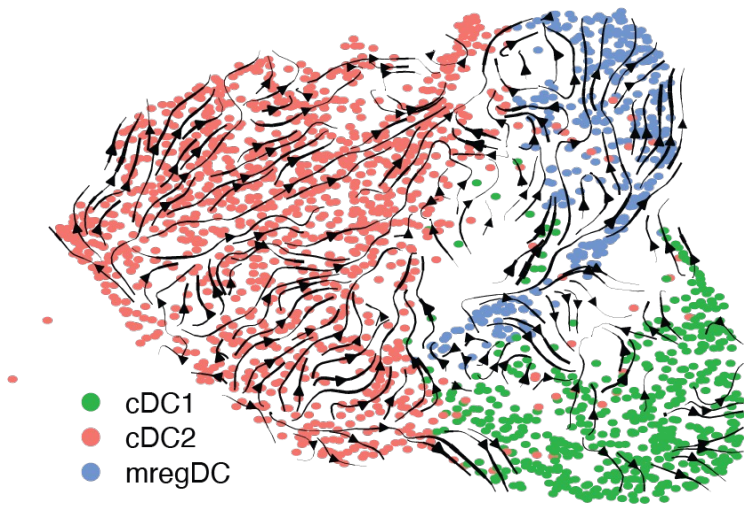
Conventional Dendritic Cell subtypes and markers



mregDC likely derives from cDC1, with certain gene programs activated

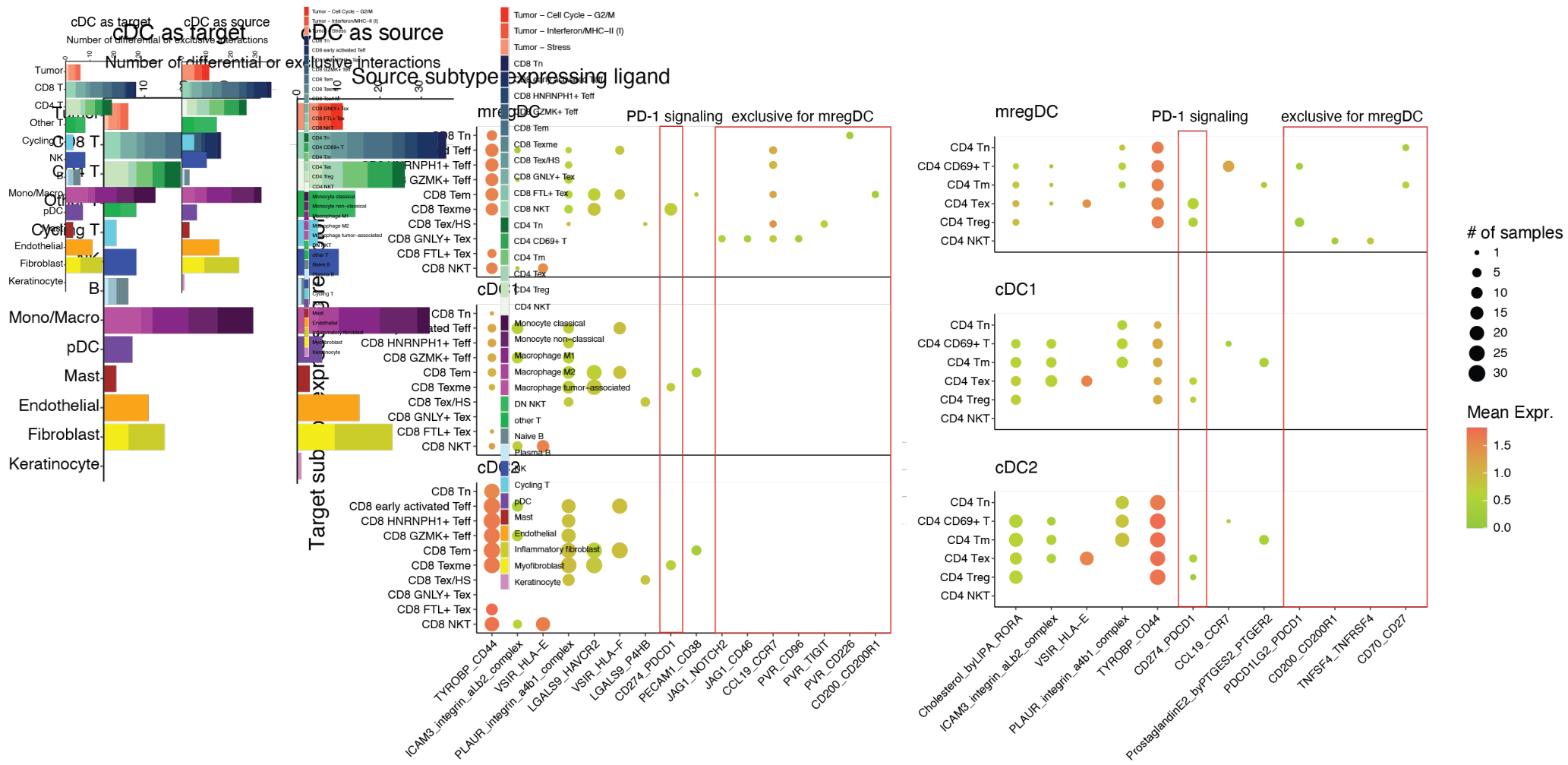


Trajectory, gene-regulatory network and immune response inference



Cui *et al.* (2024) Nature

Increased immunoregulatory and immune recruiting signals from mregDCs



Deep Learning in scRNA-seq

- **Why DL?**
 - Handling high-dimensional data, non-linear patterns.
- **Key Methods:**
 - Autoencoders (scVI, scANVI) for denoising.
 - Graph Neural Networks (GNNs) for cell-cell interactions.
 - Transformers for gene-gene relationships.
- **Applications:**
 - Cell type prediction, data integration, synthetic data generation.

Beyond scRNA-seq: Multi-Omics & scATAC-seq

- **scATAC-seq:**
 - Chromatin accessibility, regulatory element mapping.
- **Multi-Omics Integration:**
 - CITE-seq (RNA + protein), SNARE-seq (RNA + ATAC).
 - Challenges: Technical noise, data harmonization.
- **Future Directions:**
 - Spatial multi-omics, live-cell imaging integration.

Conclusion & Future Directions

- **Key Takeaways:**
 - Single-cell tech enables unprecedented resolution.
 - Computational methods are critical for interpretation.
 - Multi-omics and DL are shaping the future.
- **Challenges:** Scalability, benchmarking.

Thank You!

Questions?

Contact: jackie.yang@rutgers.edu

Website: yangcompbio.org