

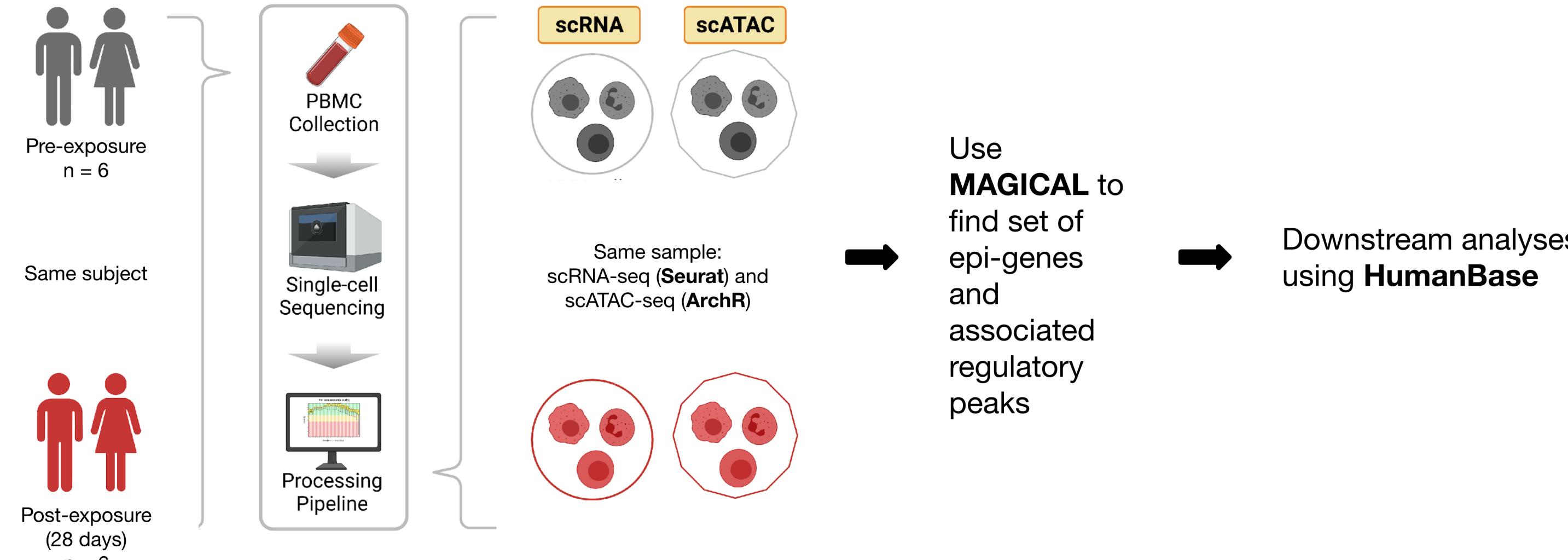
# Convalescent Epigenetic Changes Following Influenza Infection

William Thistlethwaite<sup>1</sup>, Xi Chen<sup>2</sup>, Yuan Wang<sup>1</sup>, Vincy Cheng<sup>3</sup>, Elena Zaslavsky<sup>3</sup>, Stuart Sealfon<sup>3</sup>, Olga Troyanskaya<sup>1,2</sup>

<sup>1</sup>Lewis-Sigler Institute at Princeton University, Princeton, NJ;

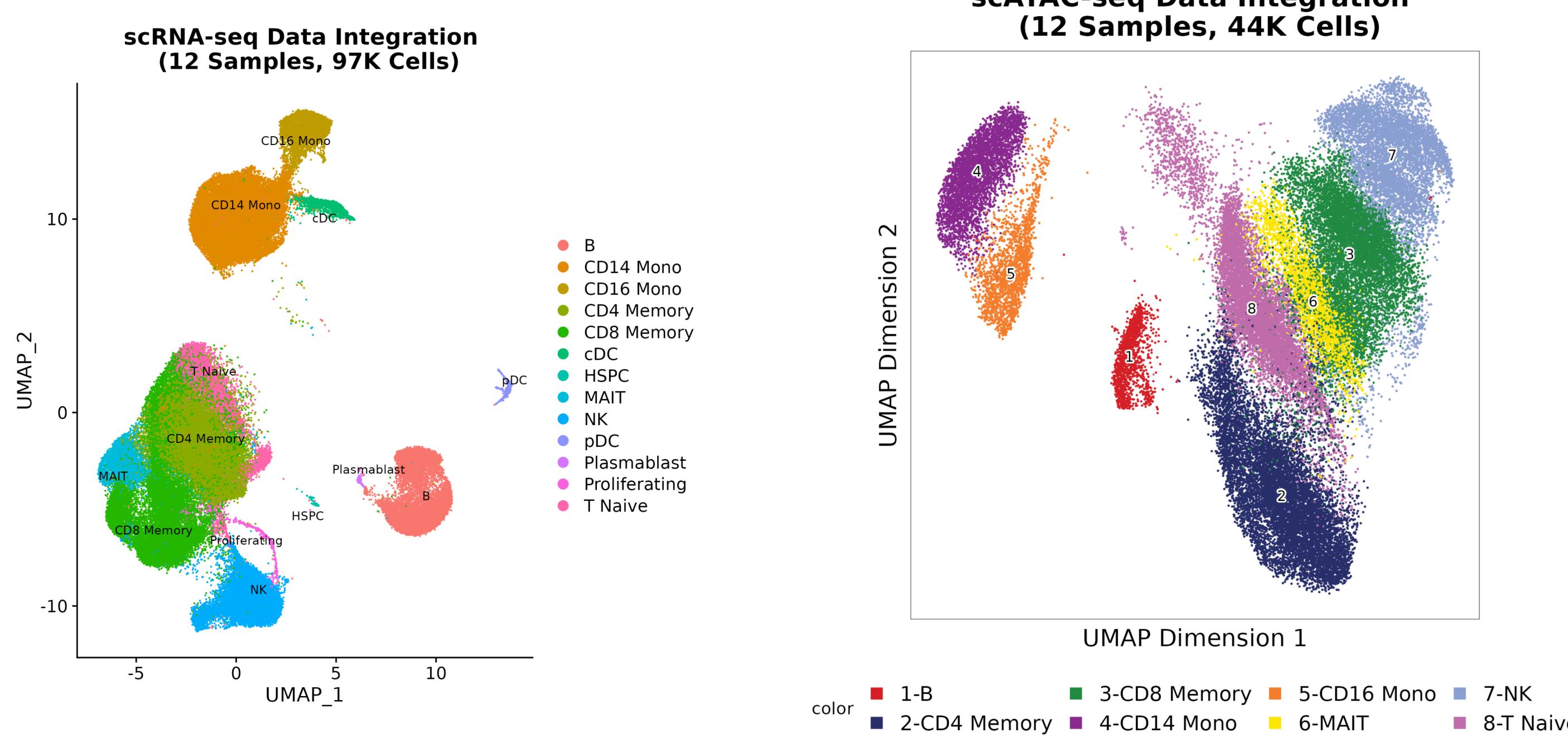
<sup>2</sup>Flatiron Institute, Simons Foundation, New York, NY; <sup>3</sup>Icahn School of Medicine at Mount Sinai, New York, NY

## OVERVIEW OF STUDY



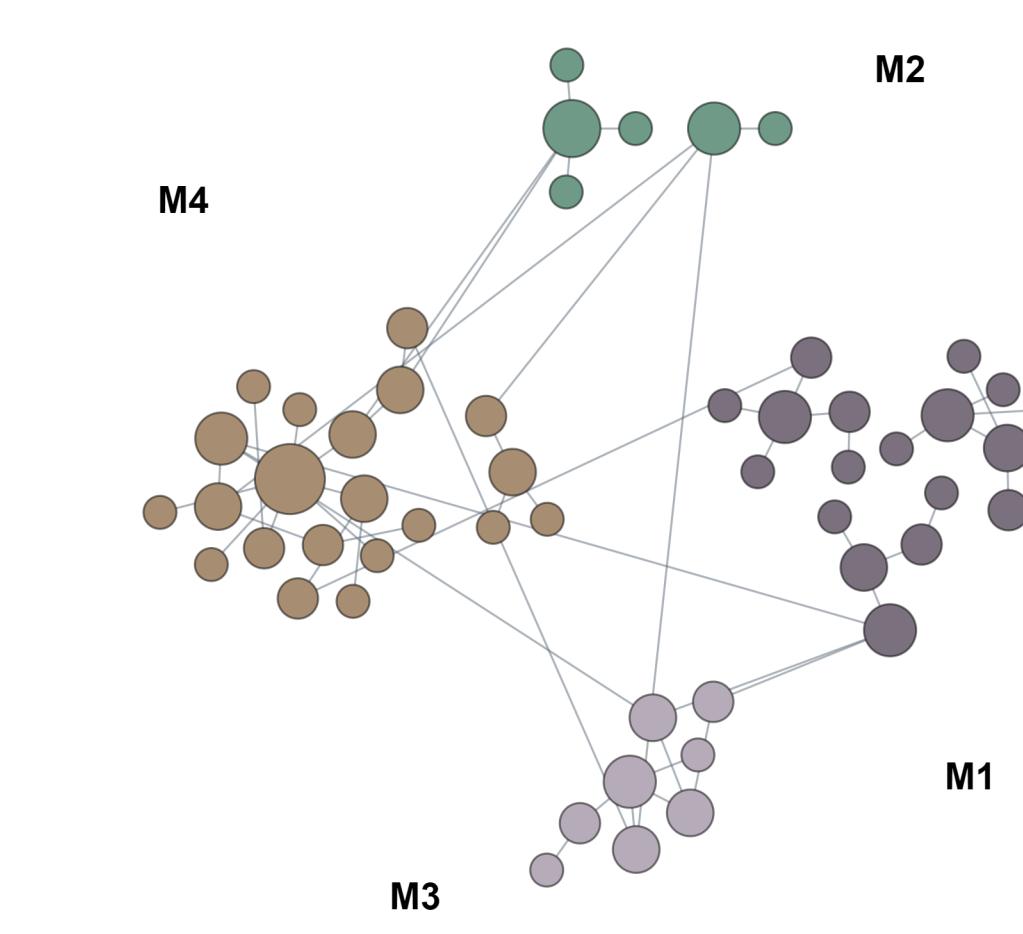
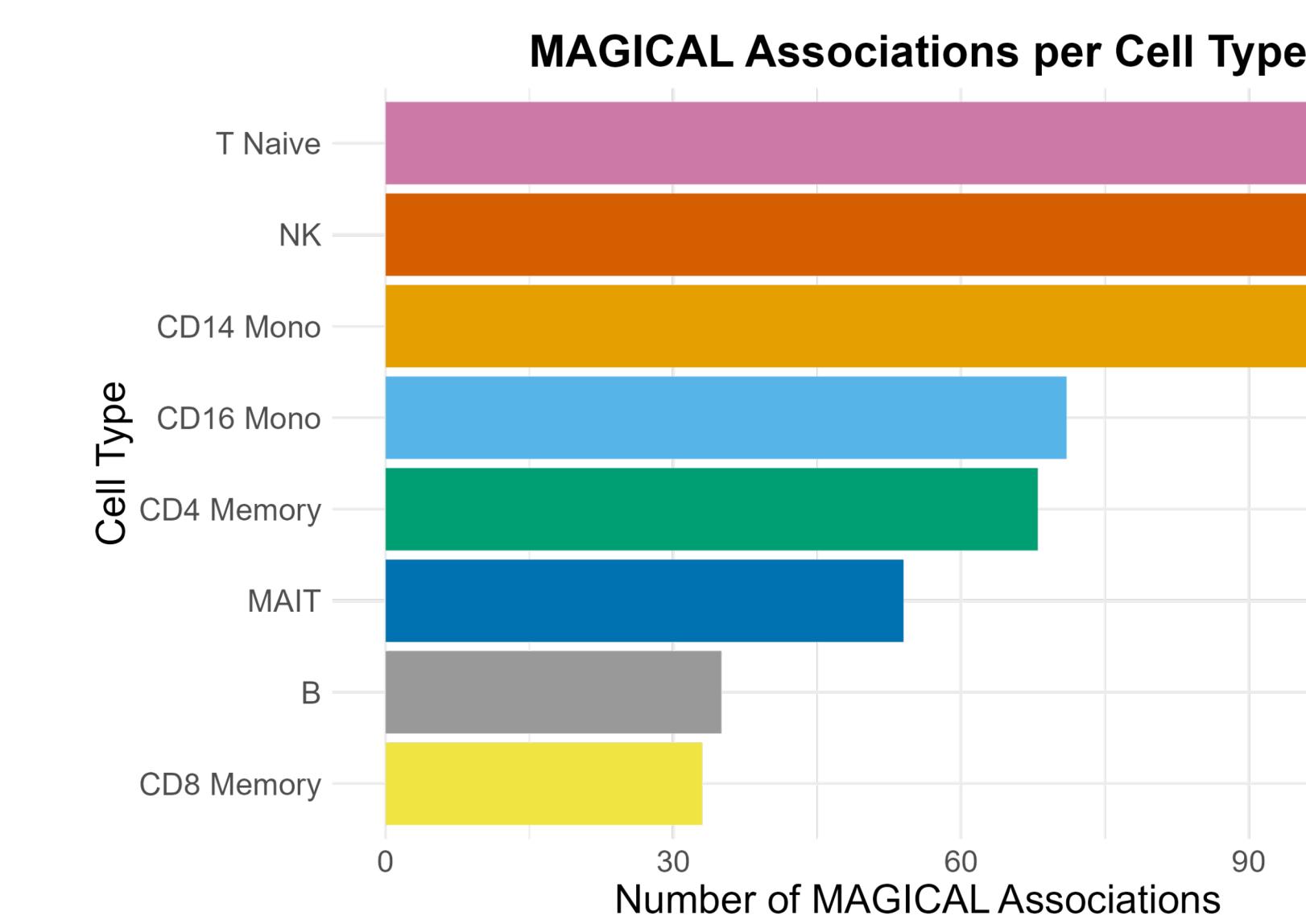
Peripheral blood mononuclear cell (PBMC) samples were collected from six patients one day before and 28 days after exposure to the influenza (H3N2) virus. Single-cell RNA-sequencing (scRNA-seq) and single-cell ATAC-sequencing (scATAC-seq) were performed on each sample. Resulting raw data were aligned to the human reference genome GRCh38 (hg38) using Cell Ranger<sup>1</sup> and then further processed using Seurat<sup>2</sup> (scRNA-seq) and ArchR<sup>3</sup> (scATAC-seq). Differentially expressed genes (DEGs) and differentially accessible peaks (DAPs) from the data were used by MAGICAL to derive a set of epi-genes and associated regulatory peaks. Finally, downstream functional module analyses were performed using HumanBase.

## CELL TYPES CLUSTER ACROSS INTEGRATED DATA



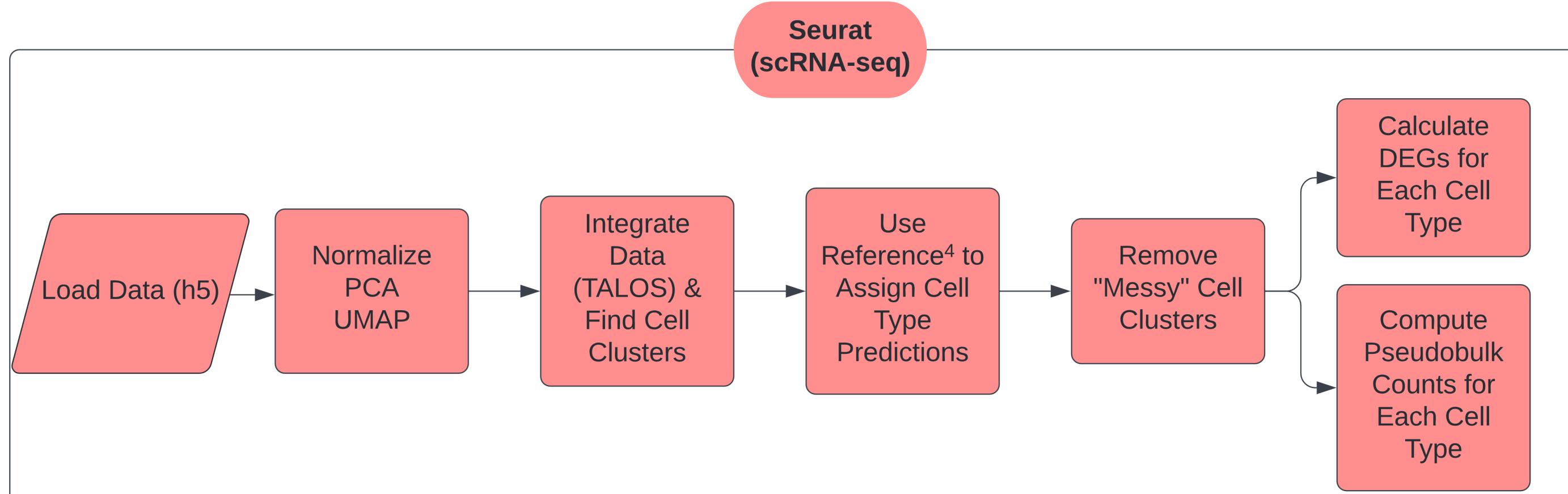
After filtering and removal of clusters containing many different cell types ("messy clusters"), cells across all samples clustered by cell type in both scRNA-seq and scATAC-seq data.

## DOWNSTREAM ANALYSIS RESULTS

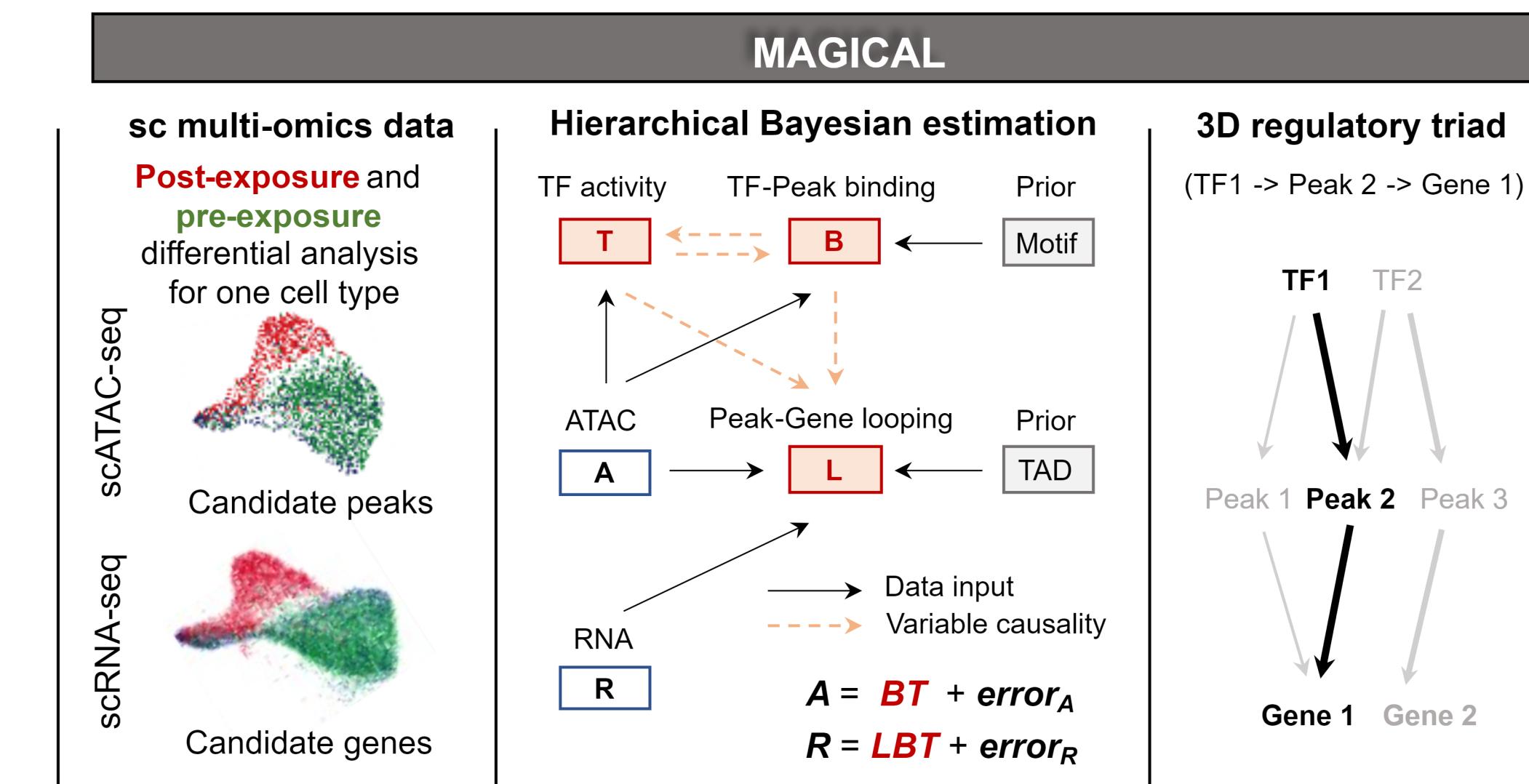
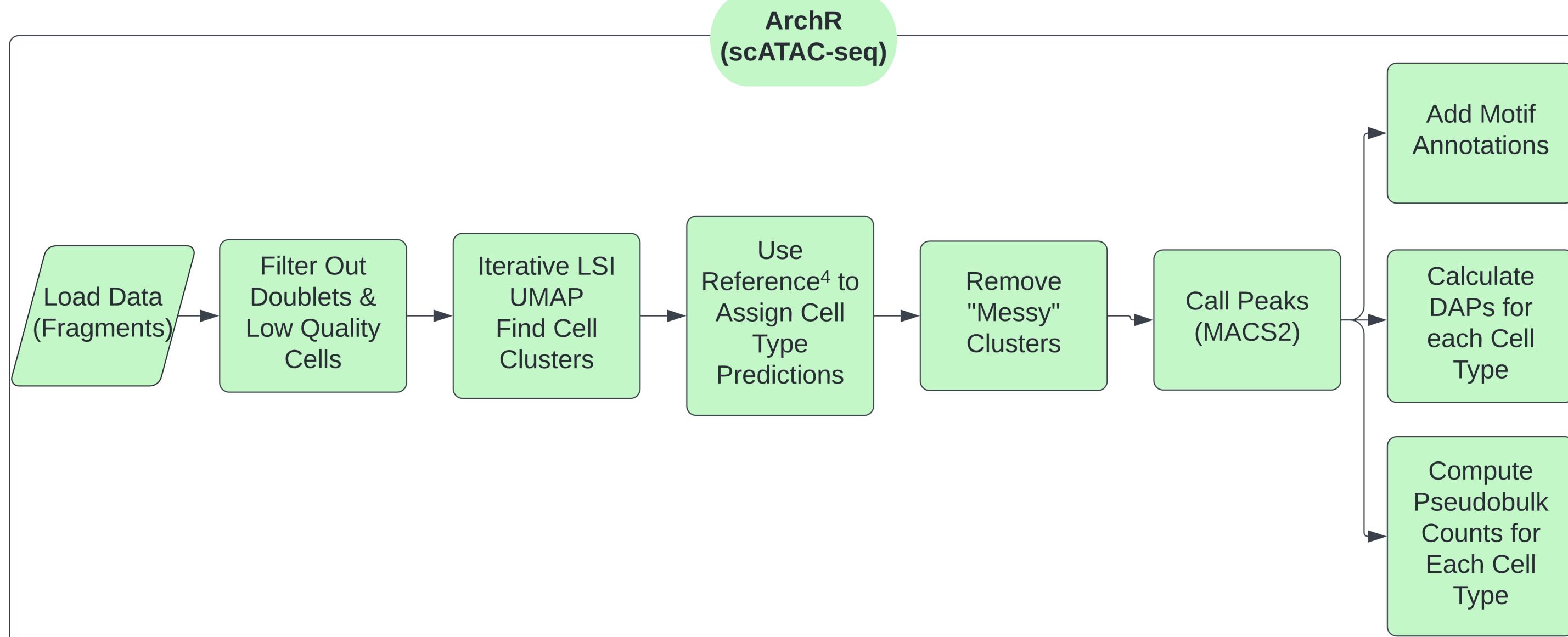
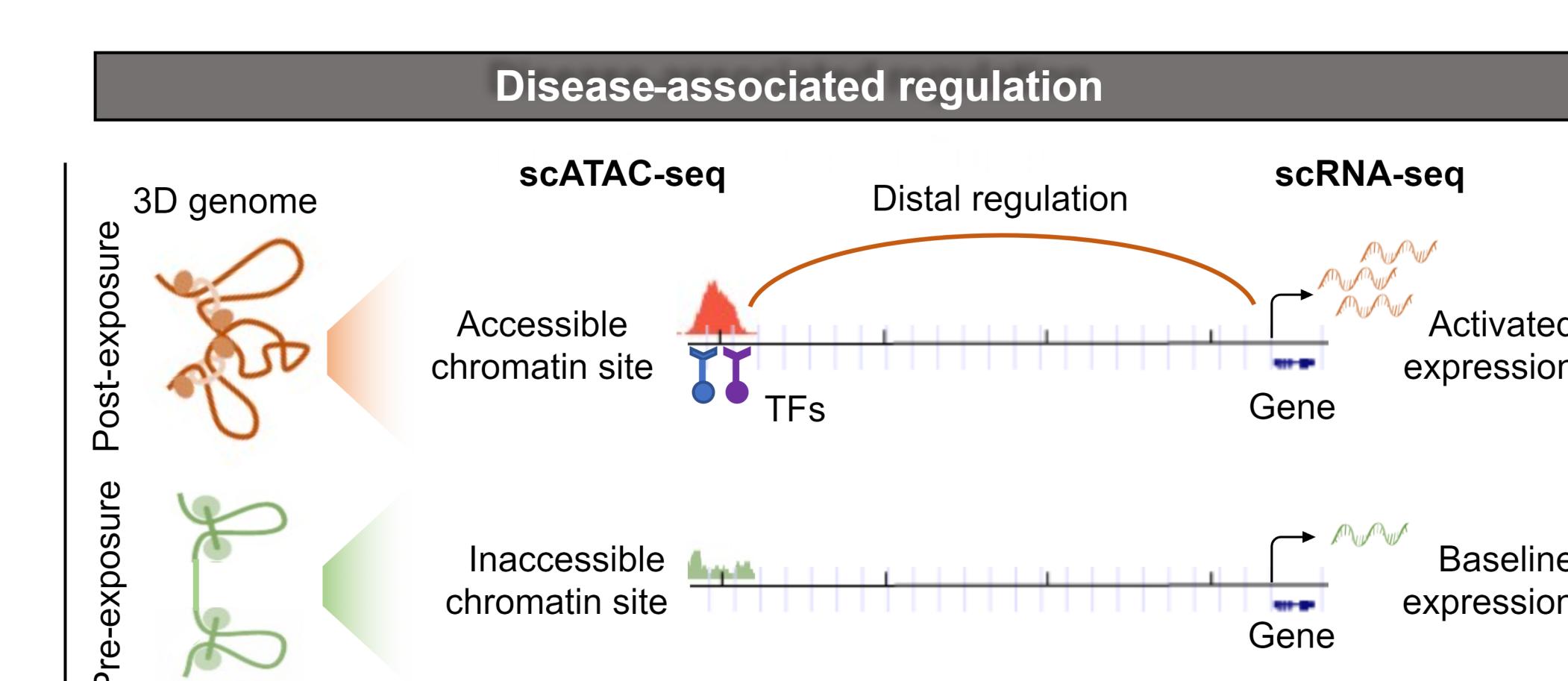


MODULE	TOP TERMS (Max 10)	Q VAL	GENES	TERMS
M1	histone H3 acetylation histone acetylation internal protein amino acid acetylation internal peptidyl-lysine acetylation peptidyl-lysine acetylation protein acetylation protein acetylation peptidyl-lysine modification peptidyl-lysine modification histone modification covalent chromatin modification	0.01080515 0.01452371 0.01452371 0.01452371 0.01452371 0.01452371 0.01452371 0.02123665 0.03267650 0.03967589 0.03967589	19	11
M2	response to endoplasmic reticulum stress	0.01080515	6	1
M3	membrane organization	0.01452371	8	1
M4	peptidyl-serine modification positive regulation of protein kinase activity positive regulation of kinase activity	0.01766864 0.03287650	21	3

## PROCESSING DATA - SEURAT AND ARCHR



## DERIVING EPI-GENES USING MAGICAL



The primary outputs from the workflows include differentially expressed genes (DEGs), differentially accessible peaks (DAPs), and pseudobulk counts used by MAGICAL.

MAGICAL uses hierarchical Bayesian estimation to derive epi-genes and associated regulatory peaks using candidate genes and peaks generated by Seurat and ArchR, respectively.

MAGICAL associations (epi-gene peak pairs) for each cell type are displayed above. Epi-genes from each cell type were analyzed using HumanBase's functional module discovery tool within the "blood" network. CD14 Mono and CD16 Mono modules are displayed above. Acetylation, methylation, and stress-related terms are found throughout the modules.

## REFERENCES

- 10x Genomics Cell Ranger 6.1.0 and 10x Genomics Cell Ranger ATAC 2.0.0
2. Cell, 184 (13), 3573-3587 (2021)
3. Nature Genetics, 53, 403–411 (2021)
4. <https://azimuth.hubmapconsortium.org/>

