BF550: Fall 2024 Project 2 **Yuki Ito**

Papers used

Kedlian, V.R., Wang, Y., Liu, T. *et al.* Human skeletal muscle aging atlas. *Nat Aging* **4**, 727–744 (2024). https://doi.org/10.1038/s43587-024-00613-3

Introduction

The motivation of this study is to gain a comprehensive understanding of how human skeletal muscle ages at the cellular level. Skeletal muscle is essential for movement, metabolism, and immune regulation, making up 40% of our body mass. Skeletal muscle aging is characterized by the decline of both muscle mass and strength, often leading to sarcopenia. This can significantly affect the quality of life for older individuals, increasing their risk of falls, fractures, and other health problems. Despite its impact, the precise cellular and molecular mechanisms underlying this age-related decline remain incompletely understood.

To address this gap, researchers profiled the transcriptomics of 90,902 cells and 92,259 nuclei from intercostal muscle biopsies of 8 young (approximately 20-40 years old) and 9 aged (approximately 60-75 years old) donors, employing both single-cell RNA-seq and single-nucleus RNA-seq. Specifically, they used droplet-based 3′ sequencing. This combined approach allows us to capture a wider range of cell types and investigate both the cellular composition and the transcriptional changes within muscle tissue during aging.

In this report, I focused specisically on the analysis of mononuclear muscle stem cells (MuSCs), which surround multinucleated myofibers (MFs) in skeletal muscle and plays a key role to generate new MFs after damage. By profiling the transcriptomes of young and aged muscle tissue, this study aims to identify key cellular and molecular changes associated with human skeletal muscle aging and to uncover potential therapeutic targets for mitigating age-related muscle decline.

Method

First, as a pre-processing step, all skeletal muscle single-cell data were aligned to and quantified against the reference genome. For nuclei datasets, the pre-mRNA version of reference genomes was used. Then ambient RNA contamination, potential doublets, potential empty droplets, and doublet cells were identified and discarded. Furthermore, cells with more than 10% and nuclei with more than 5% mitochondrial genes expressed were removed as potential low-quality cells.

I started analysis on single-cell data pre-processed in this way. I downloaded pre-processed and annotated sc/snRNA-seq human muscle and MuSCs data from the Human Skeletal Muscle Aging Atlas database (https://www.muscleageingcellatlas.org/) in .h5ad format and loaded the data using Scanpy. From the human muscle dataset, a subset of MuSCs was extracted based on the indices of the annotated MuSC dataset. Next, the subset of MuSCs gene expression data was normalized so that the total counts for each cell summed to 10,000. This step makes cells comparable regardless of differences in sequencing depth. In addition, logarithm transformation was applied to all the gene expression values to compress large values and emphasize smaller changes. After the normalization, I selected the top 10,000 most variable genes across all cells since these cells capture the most biologically meaningful variations. Then the donor ID in the MuSCs data object was specified as the batch effect key for single-cell variational inference (scVI) to account for differences between donors. After that, the scVI model was built and trained with the following parameters: n_latent=30, n_layers=2 in order to correct batch effect and reduce the dimensionality of the data to a latent space representation while preserving biological variability. A nearest-neighbor graph was constructed using the scVI latent representation, and the Leiden algorithm was applied with a resolution parameter of 1.0 to partition cells into distinct clusters. UMAP was performed to visualize the clustering, with cells colored according to annotation_level2, representing different MuSC subtypes. Finally, a dot plot was generated to visualize the expression patterns of selected genes across different groups of cells. The plot was generated with the following parameters: ['PAX7', 'MYF5', 'SPRY1', 'TNFRSF12A', 'NOP16', 'NPTX2', 'FNDC4', 'MT1A', 'ICAM1', 'CXCL1', 'CXCL2', 'TNFAIP3', 'NFKBIZ', 'IRF1', 'IER3', 'MYOG', 'TNNI1', 'TNNT2', 'CKB', 'ACTC1', 'CLDN5', 'MYOD1'], groupby='annotation_level2', categories_order=['MuSC', 'TNFRSF12A+MuSCs', 'ICAM1+MuSCs', 'MYOG+MuSCs',], cmap='Blues', standard_scale='var'. Genes in the gene list are marker genes associated with specific cell types such as MuSCs, $MYOG^{+}$MuSCs, and two other lesser-known subtypes: TNFRSF12A and ICAMI MuSCs.

I executed all the code using 8 CPU cores.

Results & Conclusions

The generated UMAP visualization of MuSC cell subpopulations is shown in Fig. 1. Four distinct subpopulations were identified within the overall MuSC cluster. These include the well-characterized quiescent MuSCs (labeled as MuSC), a transient differentiating state (labeled as MYOG+MuSCs), and two less well-characterized subpopulations: TNFRSF12A+ and ICAM1+ MuSCs. Notably, previous studies[1, 2] suggest that MYOG^{+}MuSCs may represent a potential experimental artifact arising from isolation procedure, as differentiation markers can be upregulated during cell stress or manipulation.

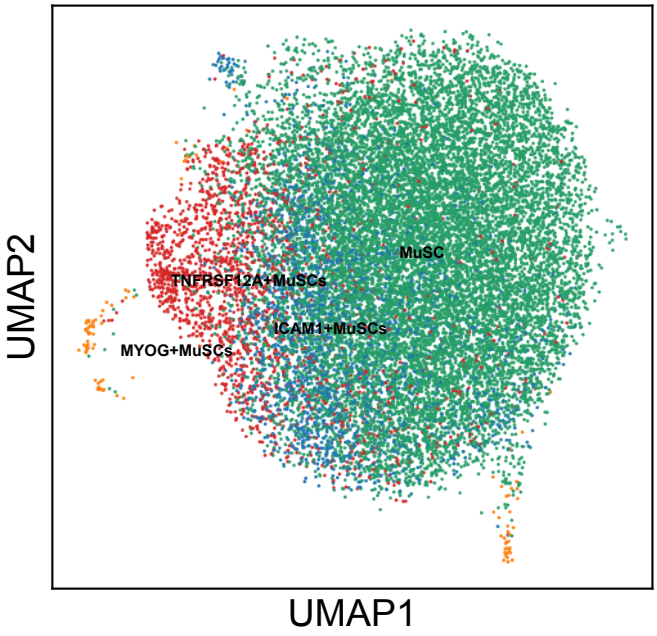

Fig. 1: UMAP visualization of MuSC subpopulations

The generated dot plot of marker genes associated with each cell subtypes found through clustering is shown in Fig. 2. The size of each circle represents the fraction of cells in each subtype, and the color intensity reflects the mean expression of the gene in that group. Marker genes on the x-axis include those associated with MuSCs (e.g., PAX7, MYF5, SPRY1), MYOG^{+}MuSCs (e.g., MYOG, MYOD1), and the lesser-known subtypes TNFRSF12A^{+}and ICAM1^{+}MuSCs. PAX7, MYF5, and SPRY1, which are known markers of quiescent muscle stem cells expressed strongly in MuSC group. In TNFRSF12A^{+}MuSCs subtype, the expression of TNFRSF12A, which is related to inflammatory responses and MT1A, which is involved in oxidative stress protection was enriched. In ICAM1^{+} MuSCs subtype, ICAM1 strongly expressed with co-expression of other signaling-related or adhesion molecules such as CXCL1 and CXCL2. Finally, Differentiation markers, including MYOG, MYOD1, TNNI1, and TNNT2, are highly expressed.

These results suggest that functional diffrences between MuSC cell subpopulations. The MuSC subtypes include quiescent and maintenance-related genes; the TNFRSF12A^{+}MuSCs subtype is likely involved in stress or injury processes; the ICAM1^{+} MuSCs subtype is associated with cell adhesion and immune signaling; and the MYOG^{+}MuSCs subtype plays a crucial role in the differentiation step.

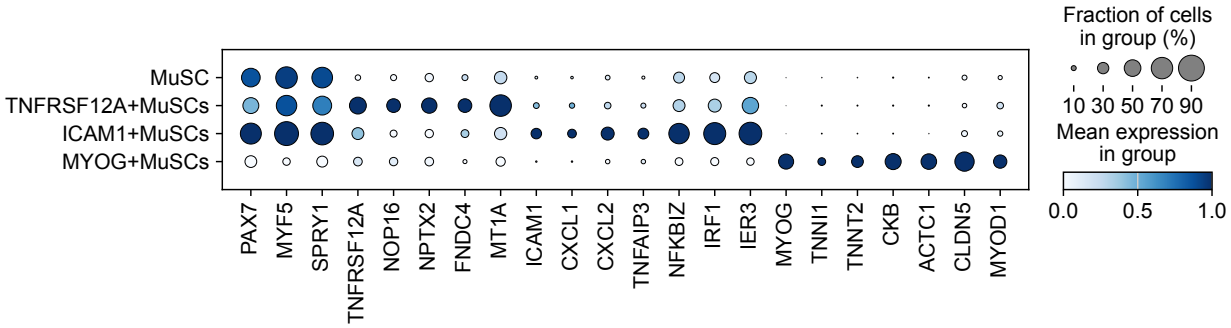

Fig. 2: Dot plot of marker genes associated with cell subtypes

References:

1. Dell'Orso, S. et al. Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. *Development* **146**, dev174177 (2019).
2. Machado, L. et al. In situ fixation redefines quiescence and early activation of skeletal muscle stem cells. *Cell Rep.* **21**, 1982–1993 (2017).