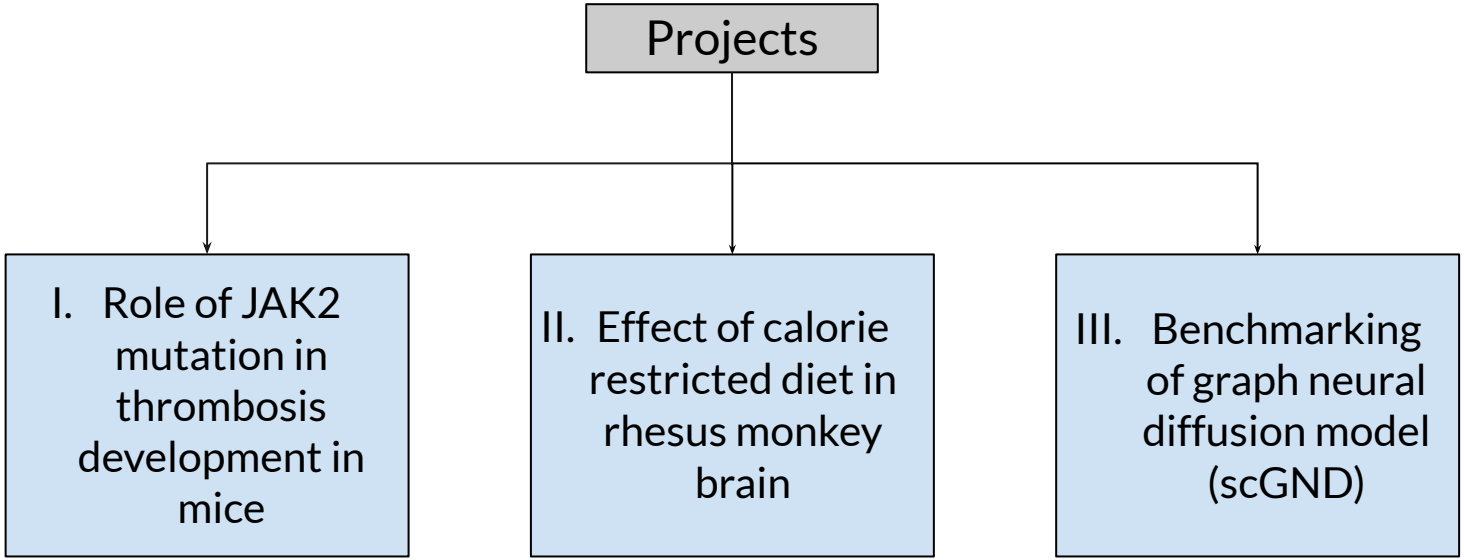
scRNA-seq Projects and Benchmarking

- Saumya Pothukuchi
- PI: Dr. Chao Zhang (chz2009@bu.edu)
- Post Doc: Yuchen Liu (ycliu137@bu.edu)

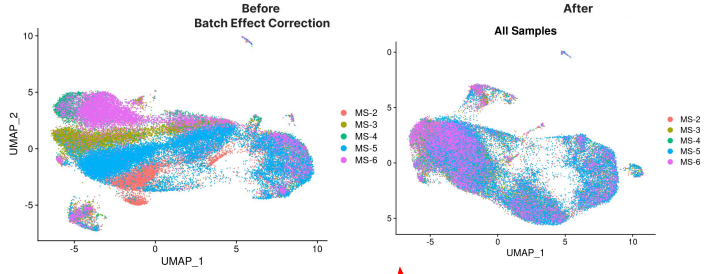
Internship Projects Overview

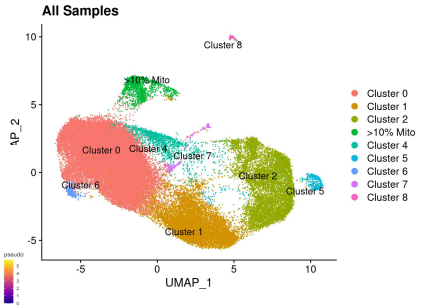


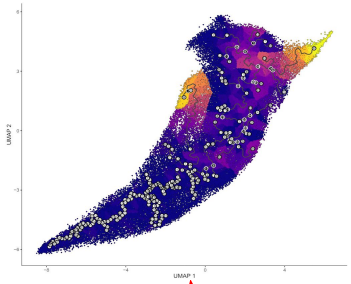
scRNA-Seq Workflow Using Seurat

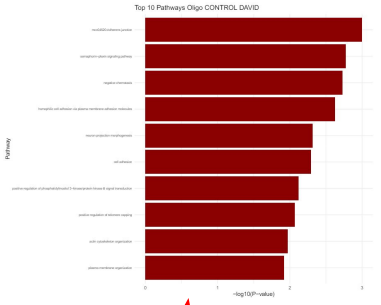
10X Genomics

Sequence reads → Counts Matrix → QC: Filter Cells → Normalise Data → Data Integration → Clustering → Marker Identification → Trajectory Analysis / DE Genes b/w Cell Types/ Groups / Pathway Analysis





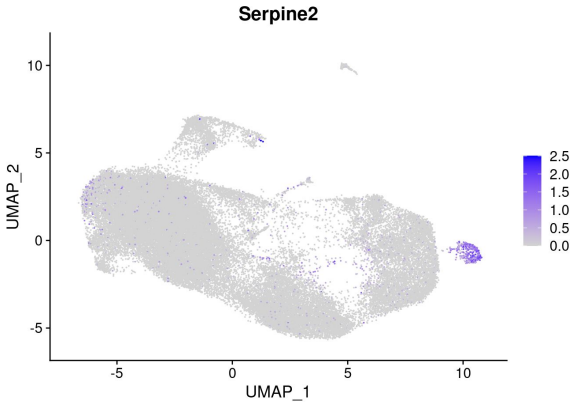




scRNA-seq Workflow Details

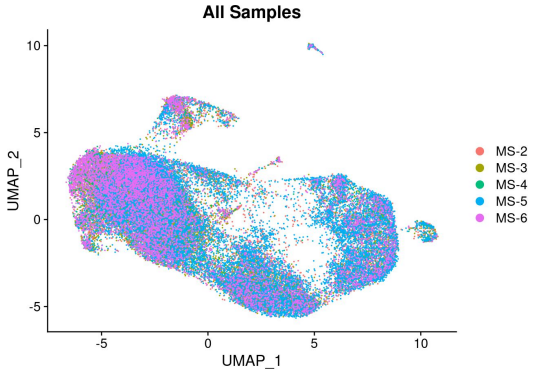
Marker Identification

Subset of genes that can distinguish between cell subpopulations in the data.



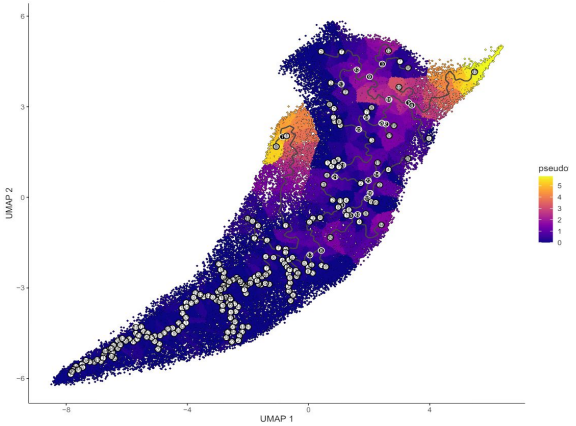
Batch Effect

Technical artifacts in individual group of samples that are processed differently relative to other samples.



Trajectory Analysis

Reconstructs a path that describes how cells transition through different cell states.



Project 1: Role of JAK2 Mutation

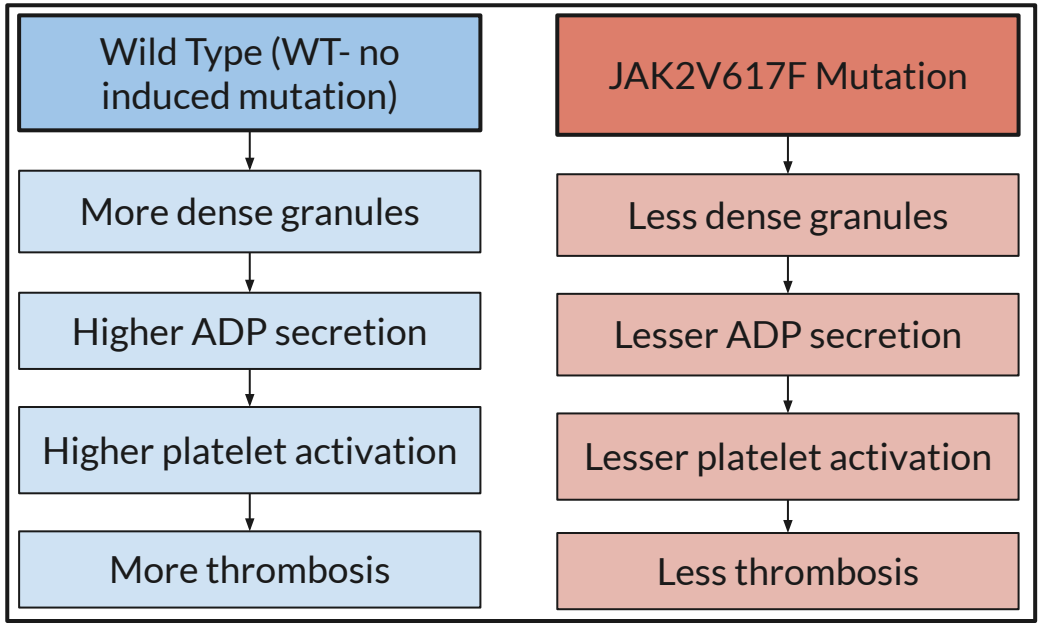
1. C57BL/6J mice stem cells.
2. JAK2V617F mutation was induced.
3. Mutation drives development of leukemia, and primary myelofibrosis (scar tissue build-up in bone marrow).
4. Promotes JAK-STAT signaling pathway, which leads to uncontrolled cell proliferation and survival.

Goals

1. Investigating the risk of thrombosis (blood clots) in myeloproliferative neoplasms (MPNs: group of blood cancers) influenced by JAK2V617F mutation.
2. Examine thrombosis tendency and platelet activation properties (due to its critical role in clotting) in mouse scRNASeq cells.

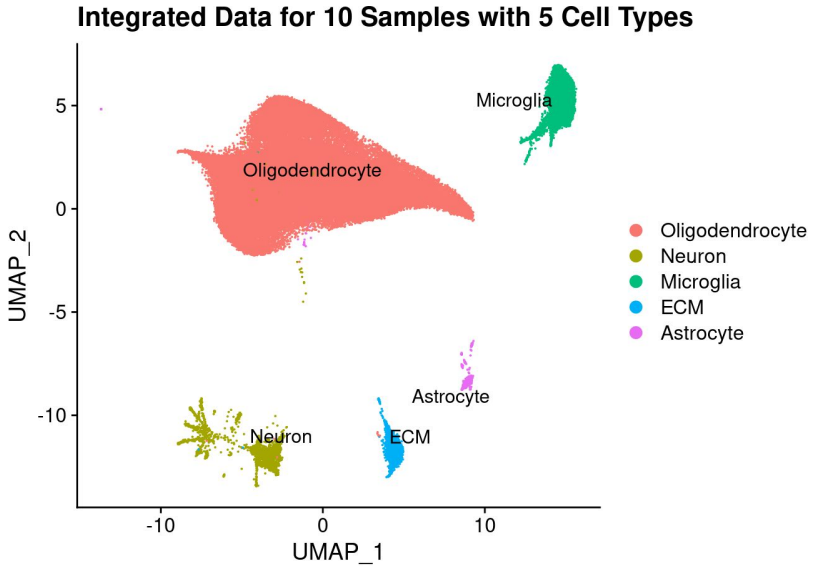
Methods and Observations

1. End to end sample analysis using Seurat for 5 samples: 2 with mutation.
2. Platelets from mice with JAK2 mutation show reduced aggregation, less thrombus formation and longer bleeding time.



Project 2: Effect of Calorie-Restricted Diets in Monkey Brains

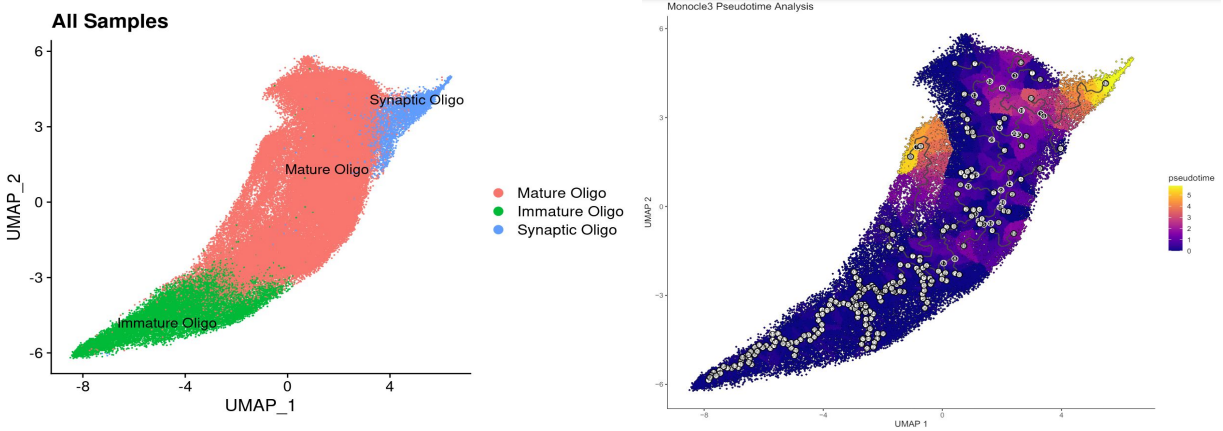
1. CR is known to increase lifespan.
2. Reduces oxidative stress and inflammation.
3. Potentially delay the onset of age-related white matter degeneration.
4. S100b calcium-binding cytoplasmic protein: neurological conditions like atrophy, neurofibrillary tangles and plaques
5. Protein expressed in 5 major cell types.



Current Focus

1. Volcano Plot and Dot Plot of DEGs
2. DAVID Pathway Analysis. Bar plots of upregulated/downregulated pathways.
3. GSEA Analysis using ClusterProfiler.

Trajectory Analysis for Oligodendrocyte Cluster



Project 3: Benchmarking for scGND (Graph Neural Diffusion Model)

- Benchmarking involves performing comparative study using ~10 different packages on 3 industry standard scRNA-seq datasets.
- Performing workflow steps until data integration to study how well the tools handle local/global equilibrium.
- Introduced cell type imbalance (0-100%) in datasets to compare.
- Using KBet which overlays embeddings and scores them.

References

1. Matsuura, Shinobu, et al. "Platelet dysfunction and thrombosis in JAK2V617F-mutated primary myelofibrotic mice." Arteriosclerosis, thrombosis, and vascular biology 40.10 (2020): e262-e272.
2. Alamente Y. “The effect of calorie restriction on age-related white matter degeneration in rhesus monkeys” https://hdl.handle.net/2144/15068.
3. Liu, Y. C., Zou, A., Lu, S. L., Lee, J. H., Wang, J., & Zhang, C. (2024). scGND: Graph neural diffusion model enhances single-cell RNA-seq analysis. bioRxiv, 2024-01.

Thank you! Questions?