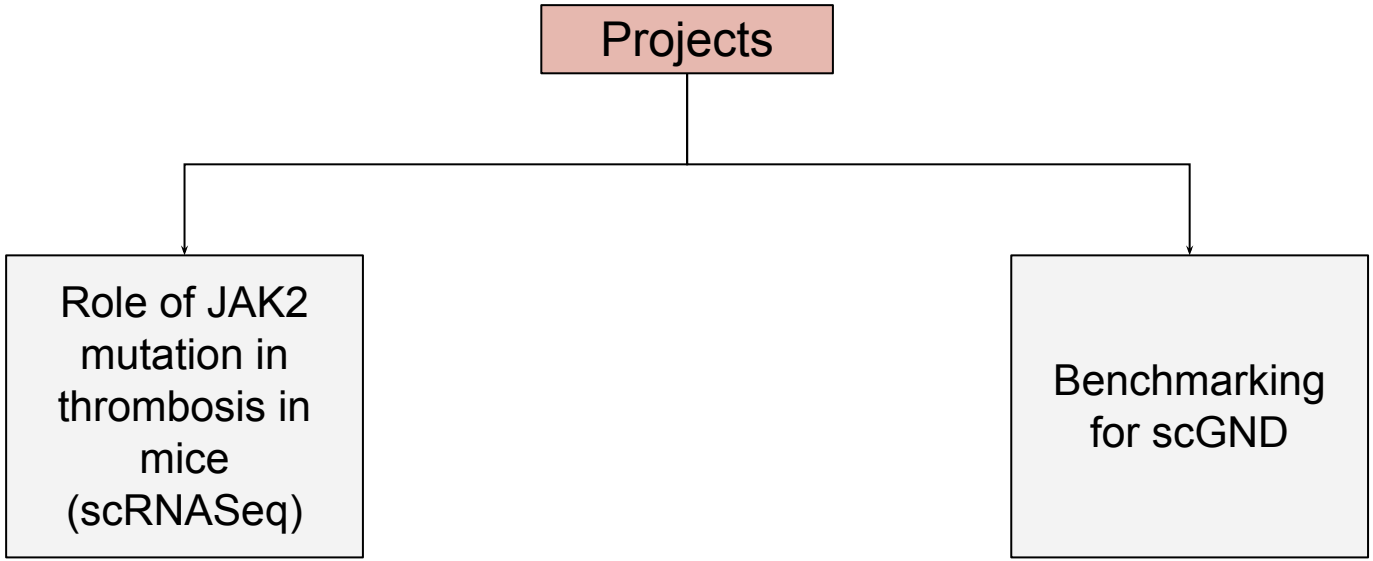
scRNASeq Analysis and Benchmarking

PI: Dr. Chao Zhang
Post Doc: Yuchen Liu

Overview


Projects
Role of JAK2 mutation in thrombosis in mice (scRNASeq)
Benchmarking for scGND

Role of JAK2 mutation in Thrombosis in Mice

Introduction:

1. C57BL/6J mice
2. Inbred strain: genetic uniformity and well documented characteristics
3. Not useful in studies involving hearing impairment, obesity, etc
4. JAK2V617F mutation was induced.
5. Mutation drives development of leukemia, and primary myelofibrosis (scar tissue build-up in bone marrow)
6. Promotes JAK-STAT pathway, which leads to uncontrolled cell proliferation + survival

Goals

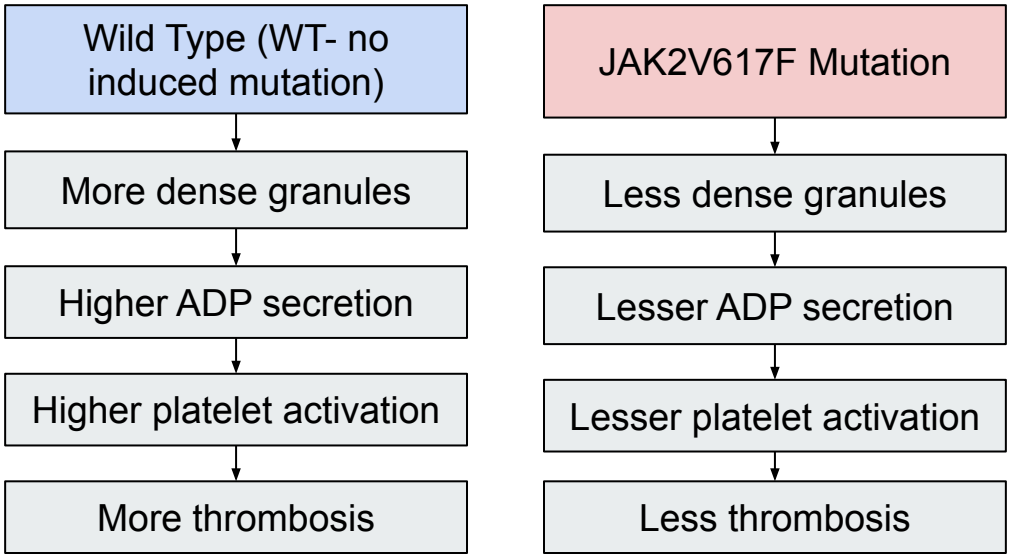
1. Investigating the risk of thrombosis (blood clots) in myeloproliferative neoplasms (MPNs: group of blood cancers) influenced by JAK2V617F.
2. Examine thrombosis tendency and platelet activation properties (due to its critical role in clotting) in mouse scRNASeq cells.
3. Compare IgG (Immunoglobulin G) and Hmb1 (Hemin-binding protein 1) responses.

Data

1. 5 samples: 2 with induced JAK2V617F mutation, 3 WT mouse stem cells.
2. Further divided into IgG and Hmb1 samples, which gives 4 main categories (JAK2V617F_IgG, JAK2V617F_Hmb1, C57BL6J_IgG and C57BL6J_Hmb1).

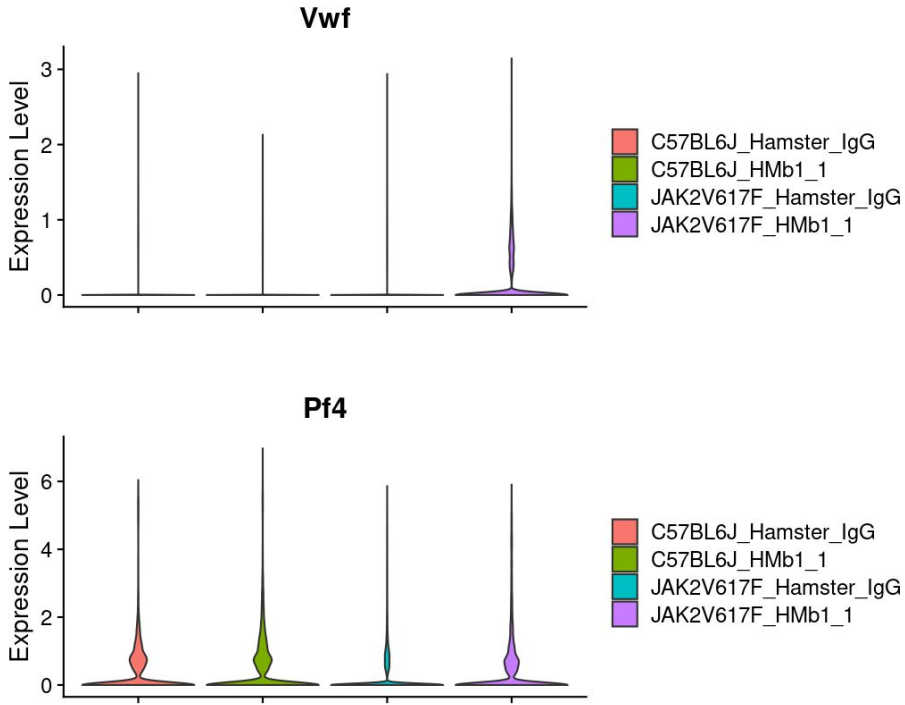
Methods and Observations

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



Current Focus

1. Expression of megakaryocytes (MKs: large platelet precursors) markers in scRNASeq samples.
2. Characteristic of primary myelofibrosis (mimicked by JAK2 mutation), hence should observe higher expression levels in JAK2 samples.



Benchmarking for scGND (Graph Neural Diffusion Model)

Introduction:

1. DL models for scRNASeq lack biological interpretability.
2. scGND aims to address this by balancing local and global equilibrium in the cell-cell interaction graph by using a diffusion model.
3. Model does not require training so is immune to training bias.
4. Tool built by lab and needs to be optimised + benchmarked.

Goals

1. Comparative study between how different tools handle biological interpretability.
2. Performing workflow steps until data integration to study how well the tools handle local/global equilibrium.

Methods

1. Benchmarking involves performing comparative study using ~15 different packages on ~15 industry standard scRNASeq datasets.
2. Performing workflow steps until data integration to study how well the tools handle local/global equilibrium.
3. Introduced cell type imbalance in datasets to compare.
4. Using KBet which overlays embeddings and scores them.