

# T-SNE Clustering and C5.0 Analysis

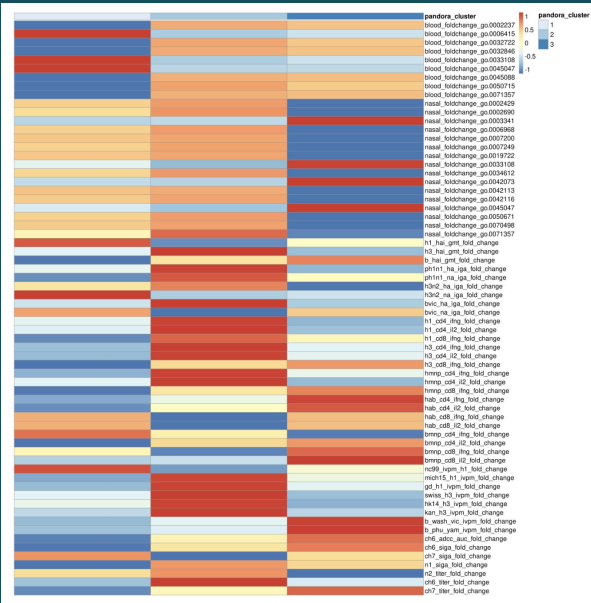
**Research Question:** By clustering patients with similar profiles together, can baseline measurements predict how a patient responds to the LAIV vaccine?

## Method:

1. Cluster patient response via t-SNE
  - a. Use fold change data only to prevent overfitting
2. Apply predictive C5.0 analysis on clusters using baseline data
  - a. Decisions trees seemed to give best AUCs

## Results:

- Training AUC: 0.8678
- Predictive AUC: 0.8758
- Model allows us to predict response to LAIV based on patient baselines → useful in understanding what kind of immunity the patient will have



**Cluster 1:** high gene response in the blood, drop in flu antibody response

**Cluster 2:** increase in CD4+ and CD8+ IFN- $\gamma$  and IL-2 for h1/h3 strains, increased antibody response for h1/h3

**Cluster 3:** high gene response in nasal, increase in IFN- $\gamma$  and IL-2 for hab/bmnp strains, increased ADCC and secretory IgA response