

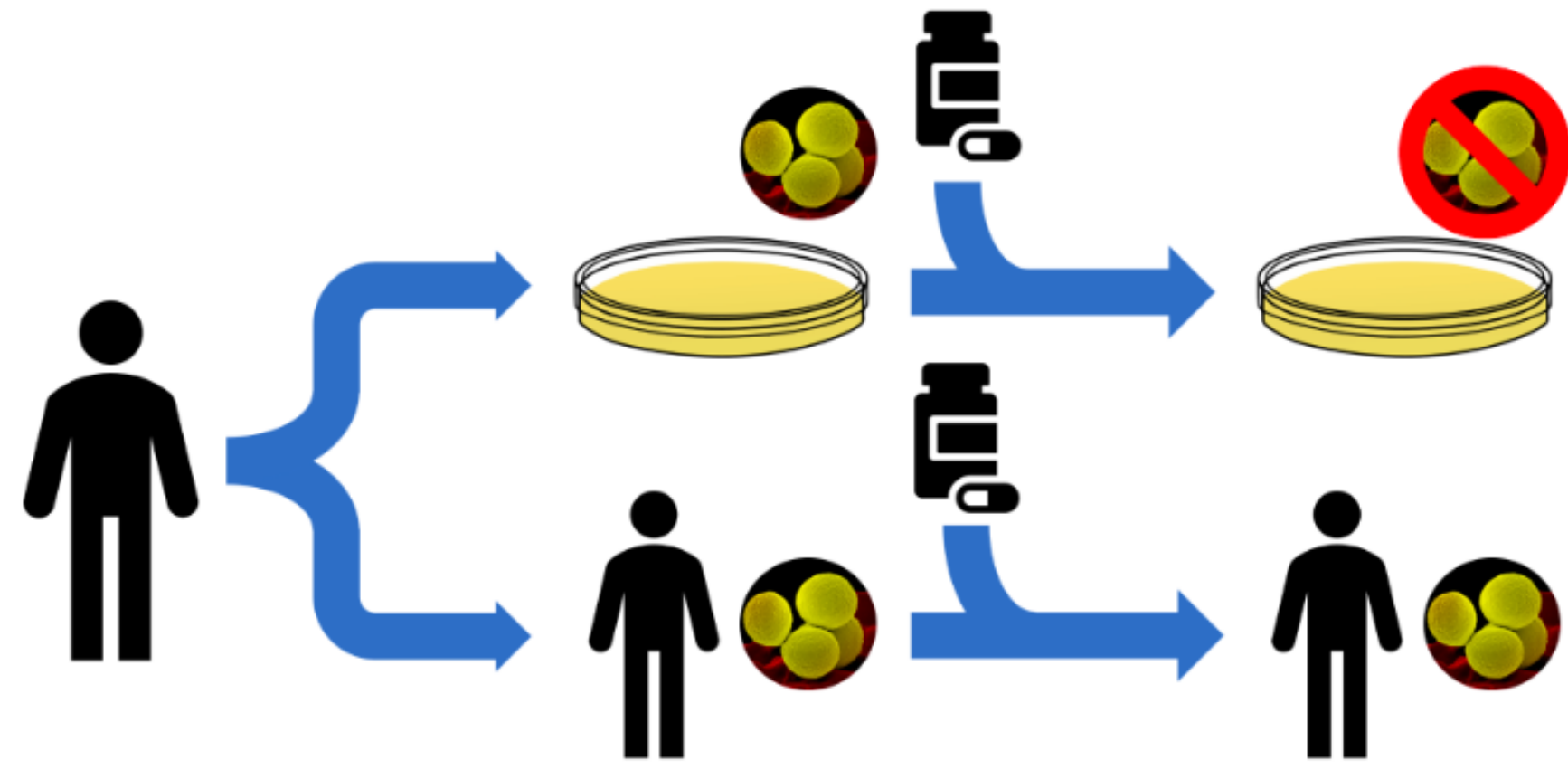


Unraveling Transcriptomic Mechanisms in Persistent MRSA Bacteremia Through Tensor-Based Data Integration



Hamza Ahmed¹, Jacqueline Durney¹, Jackson Chin², Aaron Meyer²
¹BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
²Department of Bioengineering, Samueli School Of Engineering, UCLA

Introduction



Methicillin-resistant *Staphylococcus aureus* (MRSA) bacteremia poses a frequent and life-threatening infection, and can originate from both community-acquired and healthcare-associated settings.

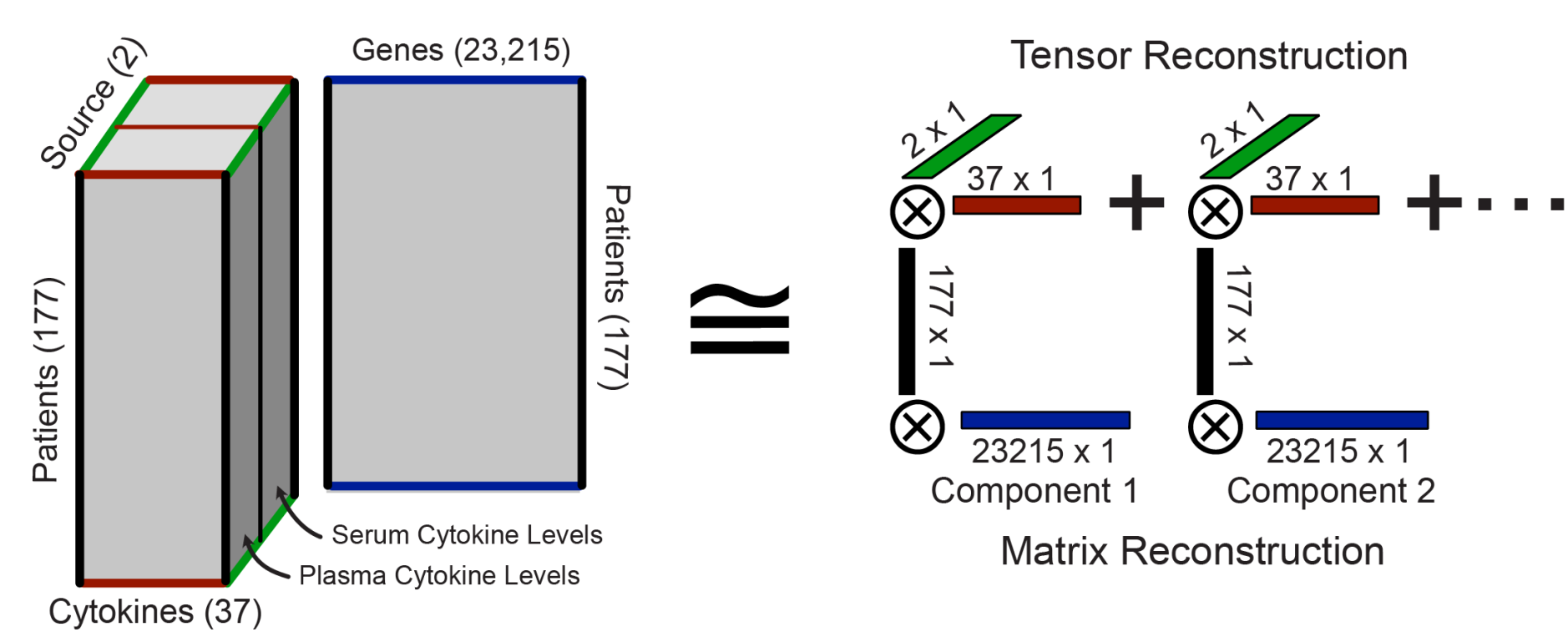
Appropriate antibiotic regimens often fail to resolve bacteremia *in vivo* despite being successful *in vitro*, suggesting that host-pathogen interactions may play an important role in MRSA persistence.

Resolving bacteremia (RB): resolves upon appropriate treatment of antibiotics.

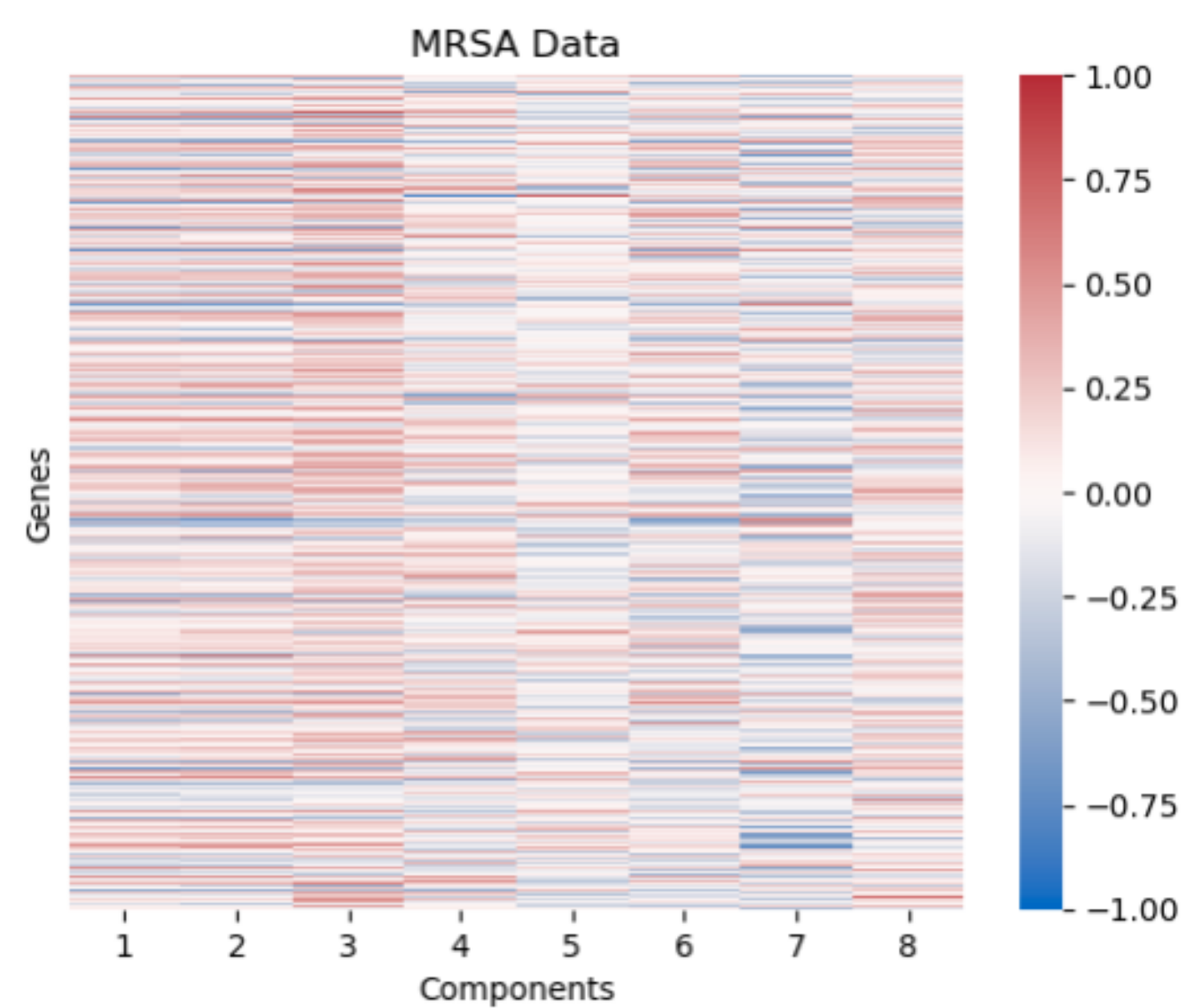
Persistent bacteremia (PB): does not resolve after 5-7 days following appropriate antibiotic treatment.

Data

Problem: Tensor reduced data requires additional processing to recognize transcriptomic patterns in gene factors.

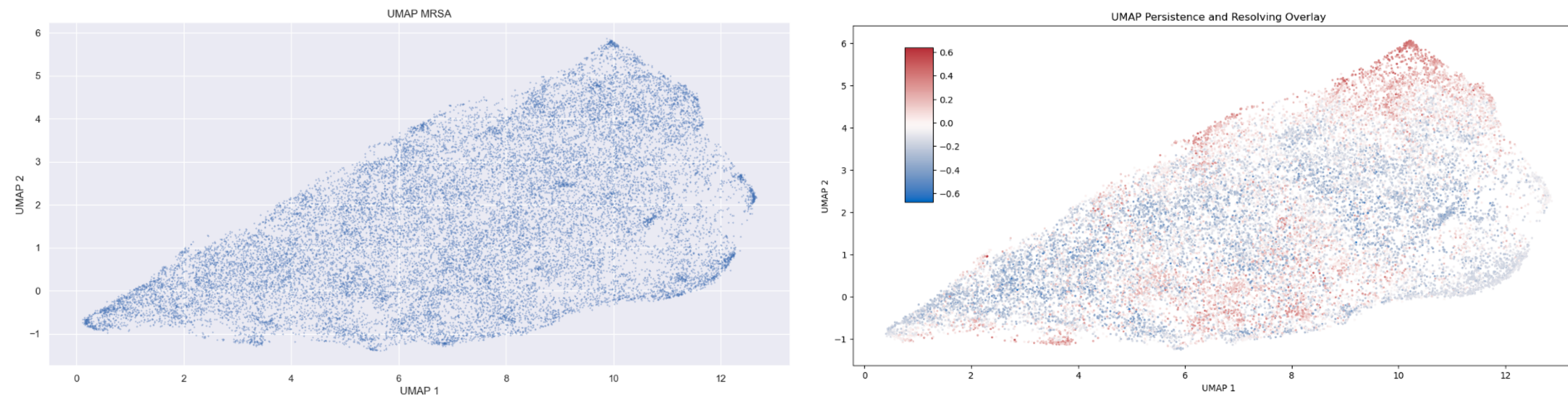


Tensor factorization enables multi-dimensional analysis of cytokine and gene expression from resolving and persistent patients.



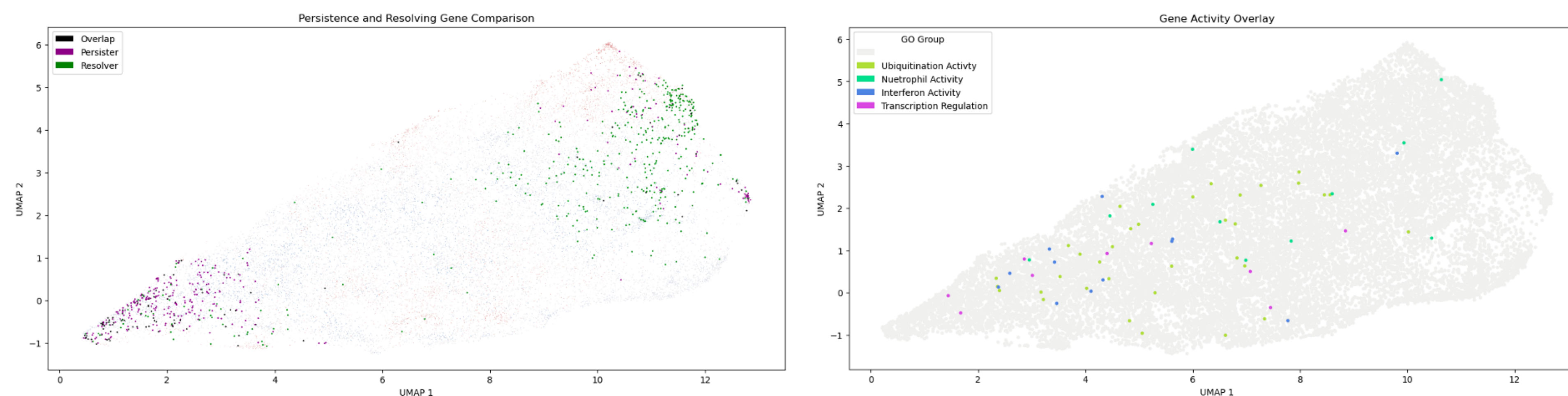
Tensor factorization reduces the data to 8 components. The above heat map relates gene expression to components. PB and RB were found to correspond to component 1 and 3 respectively, and were found in previous work.

UMAP Gene Visualization



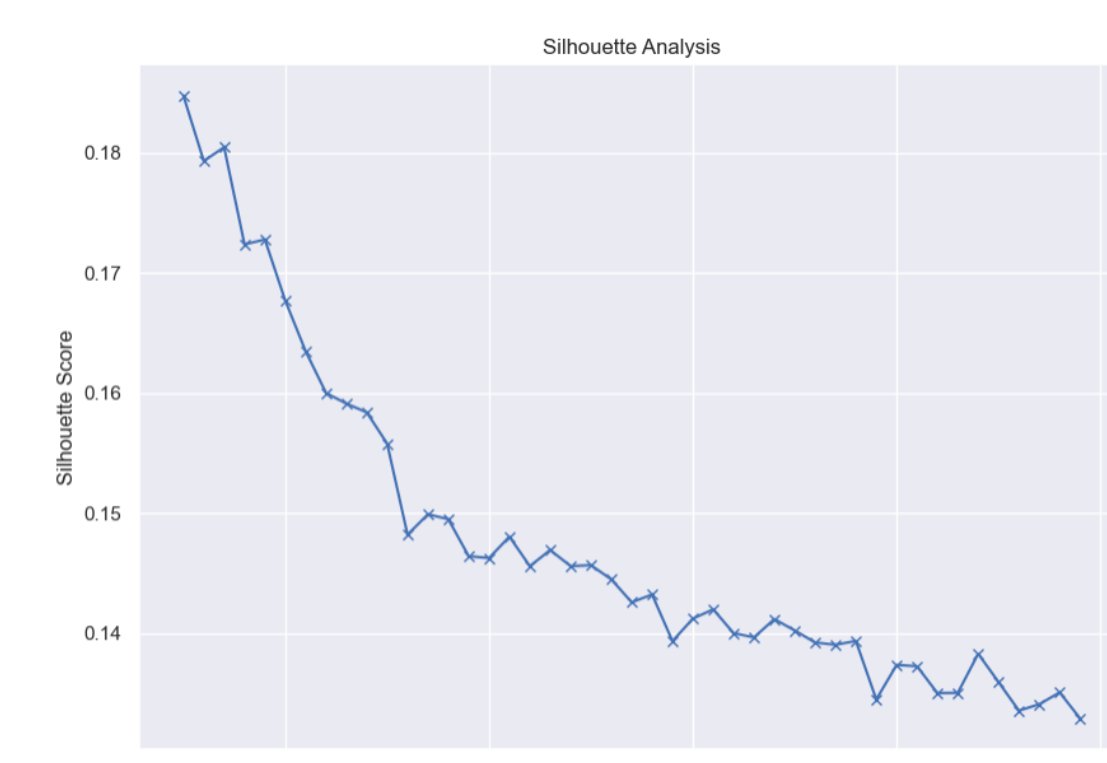
UMAP can be used to visualize the 23,000 genes in 2 dimensional space. Overlaying component specific expression data shows relationships between genes.

Persistent and Resolving Gene Analysis

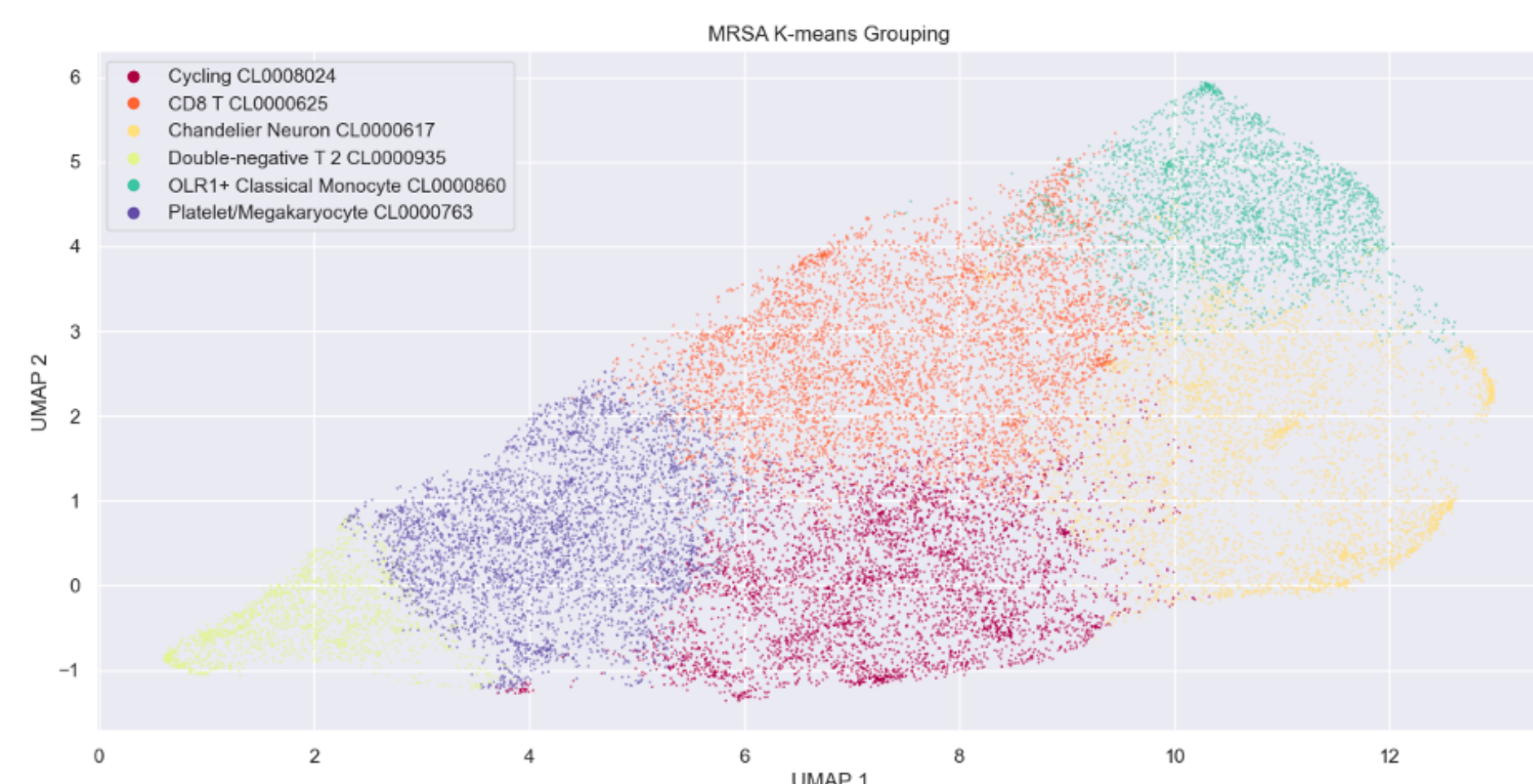


Gene enrichment analysis indicates upregulation of T-Cell activity and gene transcription in the RB related component. PB component genes were found to have increased superoxide formations and neutrophil activity.

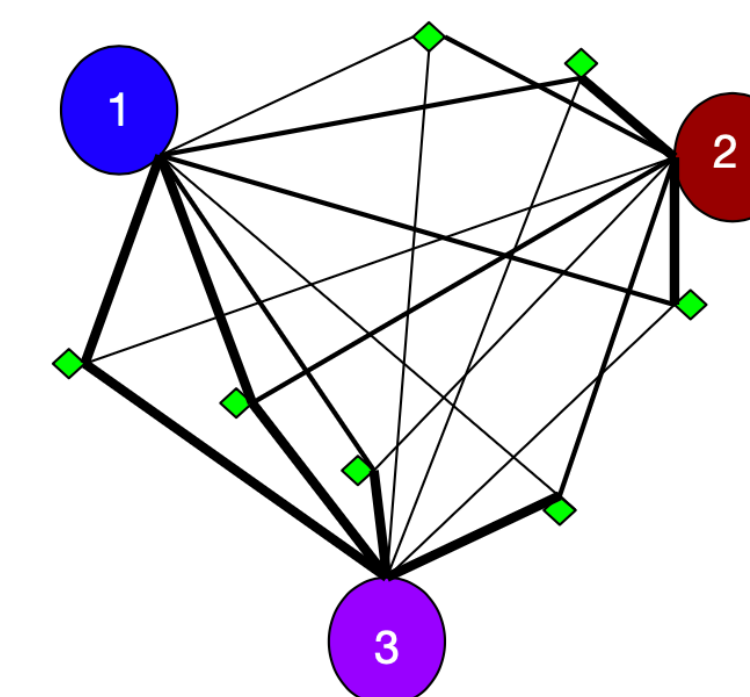
K-means Clustering Enrichment



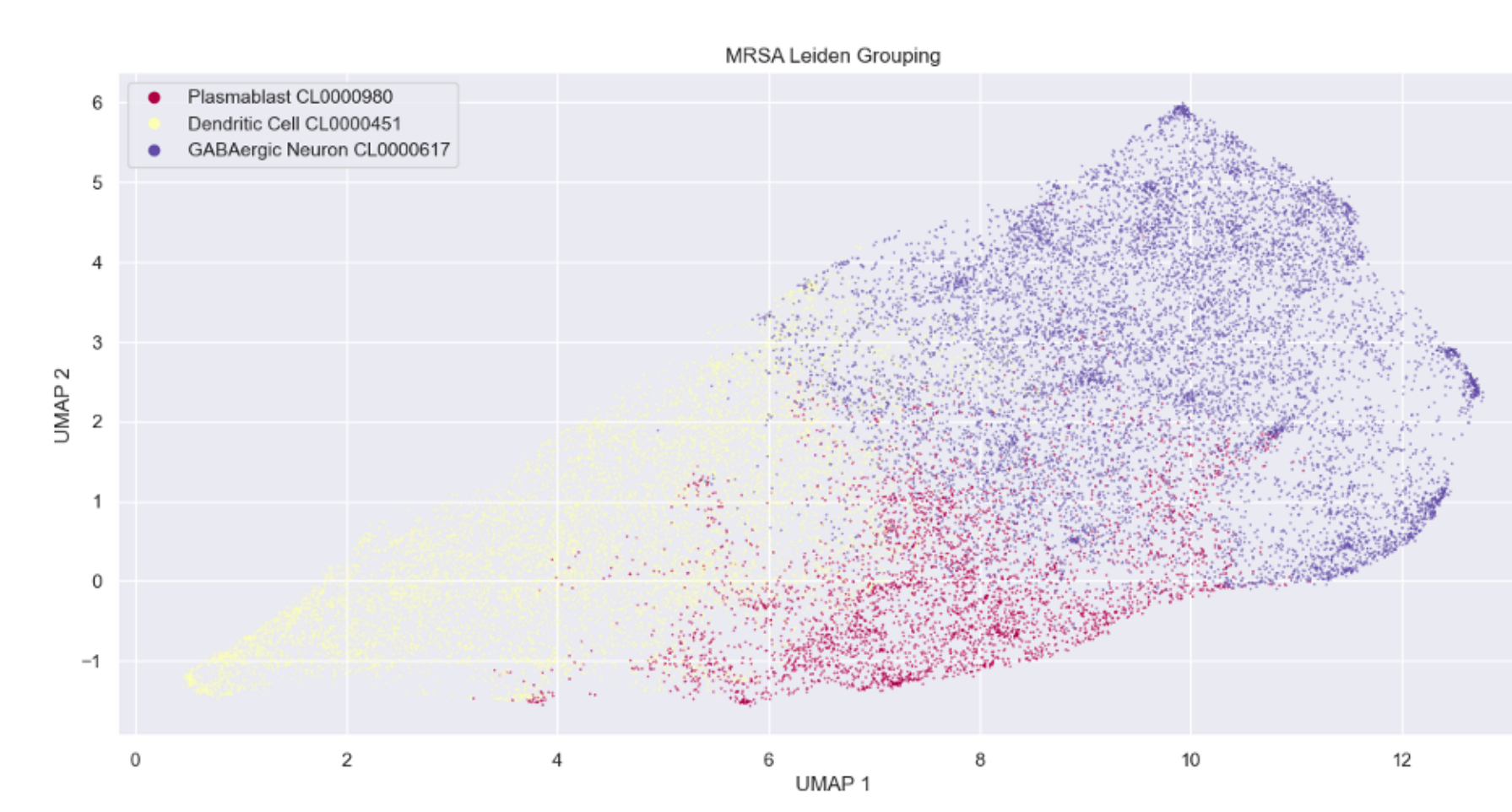
K-means clustering on component data created 6 clusters after silhouette analysis. Gene enrichment found relevant cell types and gene ontology groups for clusters on the UMAP.



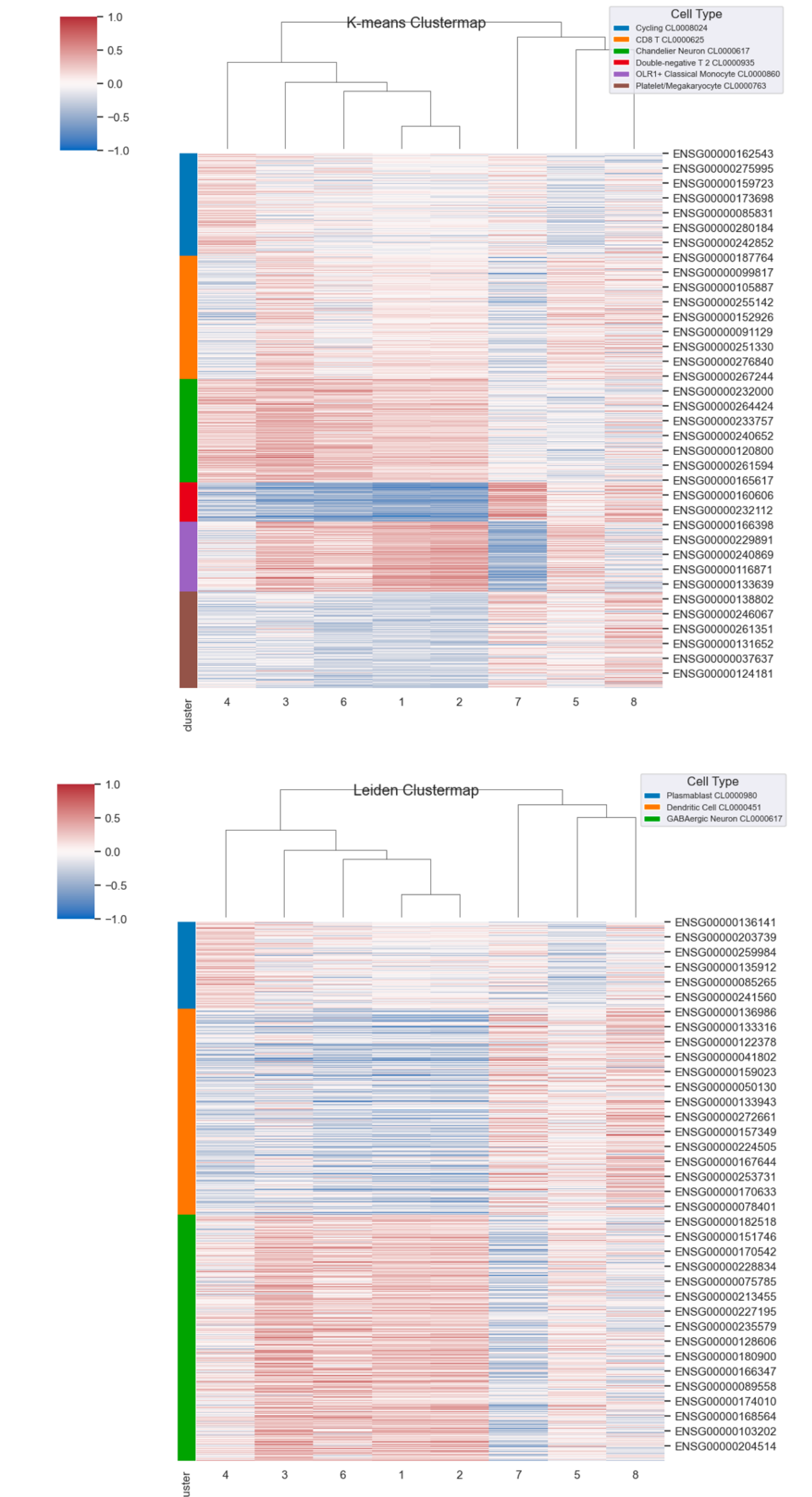
Leiden Community Finding



Component data was transformed into a network treating components as hubs and genes as nodes. Leiden community finding converged to 3 groups of components and genes. Separation between MRSA contributing components and batch effects were found.



Component Groupings



K-means and Leiden groupings can show component related patterns relative to cell types and gene ontology groups. CD8 T Cells are consistently found in the groupings.

Conclusion

PB-associated and RB-associated components have separate markers:

- PB-associated component shows an upregulation of gene transcription and superoxide and neutrophil activity.
- RB-associated component shows an upregulation of T cell activity and ubiquitination and a downregulation in transcription.

Different Clustering Algorithms:

- Leiden: Able to separate out groups based on components.
- K-means: Creates better separation on the UMAP.
- Cell types can be found and seem cohesive across groups.

Acknowledgements

We would like to thank Dr. Aaron Meyer and Jackson Chin, along with the Meyer lab for support through the project. We are also grateful for the Bruins In Genomics program at the QCB. Funded by: NIH PHS grant number 5 R25 DE 30117-3