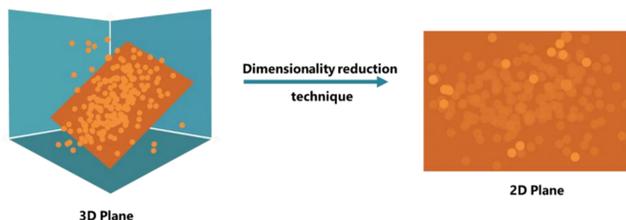


# Principal Component Analysis: Interpreting the Mechanisms of Candida Albicans

Jalissa Emmens & Carly Gordon

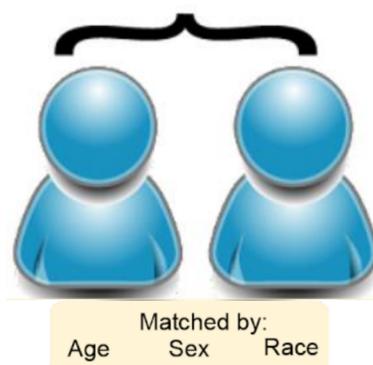
## Principal Component Analysis

- Principal Component Analysis (PCA) transforms large datasets of variables into a smaller set while retaining most of the information in the larger set.
- PCA helps to capture patterns in the data with fewer variables



## Data

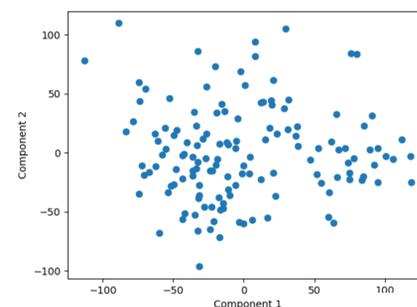
- Contains a collection of mRNA extracted from the whole blood of 138 hospital patients that presented an infection
- All 138 patients had infections that fell under the following:
  - Bacterial
  - Viral infection
  - Candidemia
  - Systemic inflammatory response syndrome (SIRS)



## Introduction

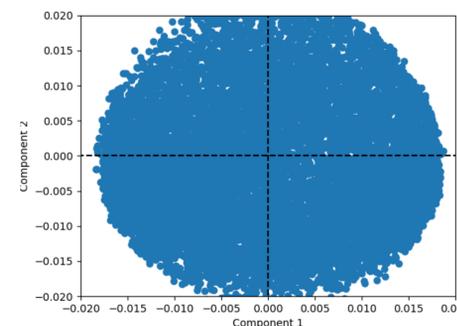
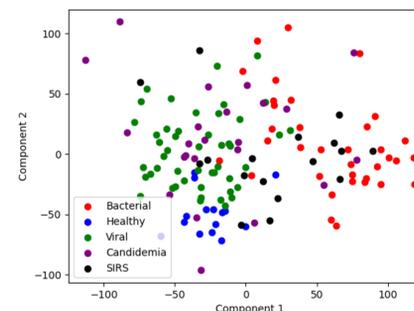
- Candida albicans are microorganisms that are commonly found in the GI tract, mouth, and vagina
- *C. albicans* rapidly transitions to a pathogen causing a variety of infections

## Results



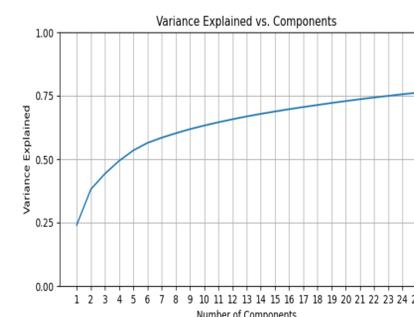
Scores plot; all 138-patient information shown; variables containing similar information cluster together as they are positively correlated.

The same figure as above, instead color-coded by the specified infection.



Loadings plot; 11,111 variables shown denoting every gene across all patient data

Variance explained by PCA components; 25 components explain around 75% of the total variance across all data sources.



## Conclusions

- PCA reduces the dimensionality of large data sets by transforming large sets of variables into a smaller set
- This reduces the data to a much smaller set of variables while minimizing information loss
- Component 1 is higher for patients with bacterial infections
- Component 2 is low for healthy patients
- Some overlap in viral and candidemia patients in, who have low component 1 and high components 2
- As the number of patients increases the variances explained drastically increases
- Component 1 shows the most variation in data, component 2 shows the second most

## References

- <https://vitalflux.com/pca-explained-variance-concept-python-example/#What is Explained Variance>
- <https://www.sartorius.com/en/knowledge/science-snippets/what-is-principal-component-analysis-pca-and-how-it-is-used-507186#:~:text=The%20PCA%20score%20plot%20of%20the%20respective%20capital%20city>

## Acknowledgments

Thanks to Dr. Meyer & Jackson Chin