

# Comparing Models of Subject-Clustered Single-Cell data

Lee Panter

Audrey Hendricks<sup>1</sup>, PhD-Committee Chair and Advisor  
Stephanie Santorico<sup>1</sup>, PhD-Committee Member  
Rhonda Bacher<sup>2</sup>, PhD-Committee Member

<sup>1</sup>The University of Colorado-Denver

<sup>2</sup>The University of Florida

# Presentation Overview (1/2)

## Analysis Summary

- ▶ What problem is being addressed?
  - ▶ Single-Cell (SC) data increasing prevalence
  - ▶ SC data with multiple subjects emerging for analysis
  - ▶ Not clear how to analyze subject level relationships
- ▶ What needs to be done?
  - ▶ Demonstrate: existing statistical methods can account for subject level relationships (SLRs) (in SC data)
  - ▶ Compare: how each method differs/resembles the others
- ▶ What do I do to address the problem?
  - ▶ Outline five modeling methods
    - ▶ Demonstrate: how (if) the models account for SLRs in SC data
    - ▶ Compare: how model frameworks compare in theory
  - ▶ Apply modeling methods to a motivating SC data example
    - ▶ Demonstrate:

# Presentation Overview (2/2)

## **Presentation Goals**

- ▶ Demonstrate a need for the analysis conducted
- ▶ Describe case study, and why it is relevant to the analysis needed
- ▶ Outline methods and how they address established needs, and why they are relevant to case study
- ▶ Highlight important results from applying methods to case analysis data
- ▶ Show that results are relevant to the demonstrated need

# Presentation Outline (1/2)

- ▶ Introduction

- ▶ Motivation

- ▶ *Demonstrate a need for the analysis conducted*

- ▶ Single-Cell Data

- ▶ Descriptions of Applied Data

- ▶ *Describe why case study is relevant to needed analysis*

- ▶ Model Descriptions

- ▶ Overview of selected models, notation

- ▶ Individual model descriptions

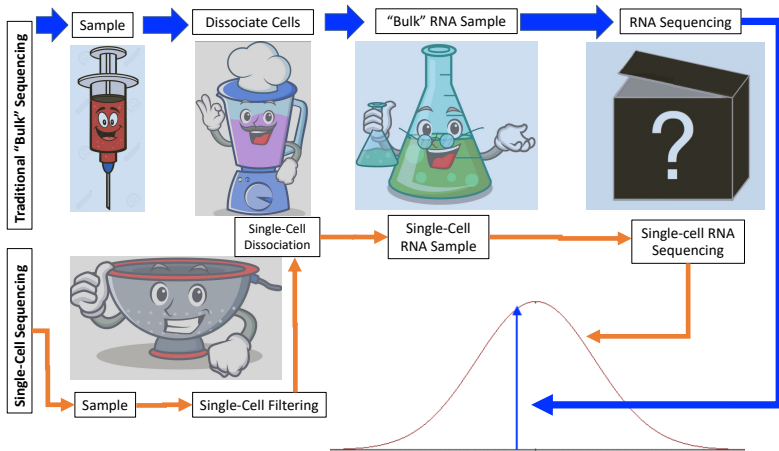
- ▶ Model comparisons & parameter interpretations

- ▶ *Demonstrate how model comparisons address established analysis needs, and why they are relevant to the case study*

# Presentation Outline (2/2)

- ▶ Case Study Analysis Results
  - ▶ Comparing models by directly using estimated parameter values
  - ▶ Comparing models using nested model statistics
    - ▶ *Highlight important results from applying methods to case analysis data*
- ▶ Discussion
  - ▶ Interpretation of important results
  - ▶ Justification of importance
    - ▶ *Show that results are relevant to the demonstrated need*

# Introduction



# Introduction: Motivation

- ▶ Bulk Sequencing Methods:
  - ▶ Analyze combined expression of thousand/millions of cells
  - ▶ Often fail to capture underlying variability within sample profiles
- ▶ Single Cell (SC) Sequencing Methods:
  - ▶ Analyze expression measurements specific to individual cells
  - ▶ Hundred/thousands of SC measurements used to estimate expression variability within samples
- ▶ Benefits of SC Methods
  - ▶ Suited for detecting expressions differentially expressed across conditions[1]
  - ▶ Useful for identifying rare cellular subpopulations [2]
- ▶ Technological Improvements in Producing SC data
  - ▶ Makes SC data easier, and more economical to generate than ever before
  - ▶

# References

1. Bacher R, Kendzierski C (2016) Design and computational analysis of single-cell rna-sequencing experiments. *Genome biology* 17: 63.
2. Ståhlberg A, Andersson D, Aurelius J, et al. (2010) Defining cell populations with single-cell gene expression profiling: Correlations and identification of astrocyte subpopulations. *Nucleic acids research* 39: e24–e24.
3. Arazi A, Rao DA, Berthier CC, et al. (2018) The immune cell landscape in kidneys of lupus nephritis patients. *bioRxiv* 363051.