Before we start, please…

- Join Slack bit.ly/BIGslu and the channel #2022-08_r-workshop

- Download the data https://github.com/BIGslu/workshops/raw/main/2022.08.15_R.tidyverse.workshop/data/data.zip

- Install R cran.r-project.org

- Install RStudio www.rstudio.com/products/rstudio

- Install packages in RStudio console
  - `install.packages(c("tidyverse", "BiocManager", "ggrepel"))`
  - `BiocManager::install(c("limma"))`

Introduction to R and the tidyverse:
RNA-seq edition

Kim Dill-McFarland (she/her)

Madison Cox (she/her)

Tomasz Janczyk (he/him)

Basilin Benson (he/him)

Code of conduct

We are dedicated to providing a welcoming and supportive environment for all people, regardless of background or identity. By participating in this community, participants accept to abide by The Carpentries' Code of Conduct. Any form of behavior to exclude, intimidate, or cause discomfort is a violation of the Code of Conduct. In order to foster a positive and professional learning environment, we encourage the following kinds of behaviors in all platforms and events:

- Use welcoming and inclusive language

- Be respectful of different viewpoints and experiences

- Gracefully accept constructive criticism

- Focus on what is best for the community

- Show courtesy and respect towards other community members

https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html

Outline

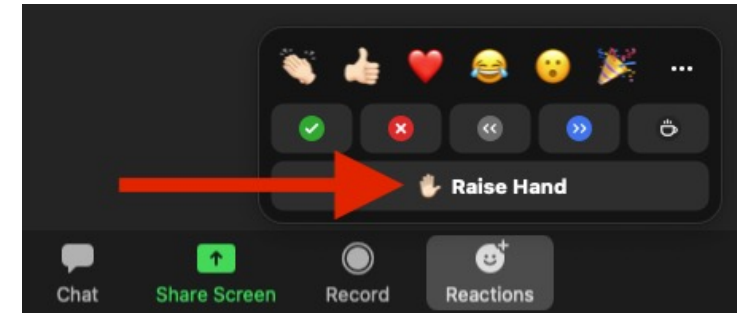| Date	PST	
Mon 8/15	1 - 2 PM	Introduction to R and RStudio, session 1
	3 - 4 PM	Introduction to R and RStudio, session 2
Wed 8/17	1 - 4:30 PM	Data manipulation in dplyr and tidyr
Thu 8/18	1 - 4:30 PM	Data visualization in ggplot

Questions

In-person
- Raise hand at any time
- Use Pink Post-It

Zoom
- In Zoom, raise hand and once called upon, use microphone

- Please do not use the chat for questions

- Write in the Slack channel #2022-08_r-workshop
  - Once it is complete, you will see ✅
  - If relevant to the larger group, instructor will repeat aloud and address it
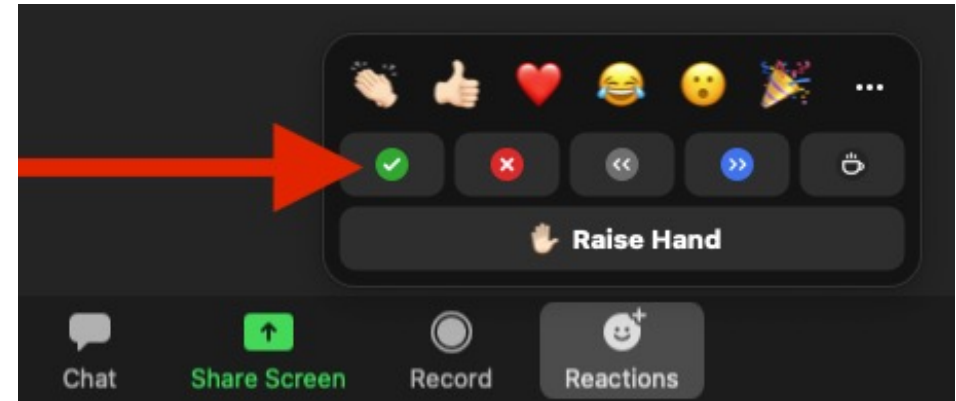
Checking in

In-person

- Use Post-Its
  - Green: Ready to move on
  - Pink: You need more time or have an error

Zoom

- Use Zoom reactions
  - ✅ You are ready to move on
  - ❌ You need more time or have an error

Recording

- This workshop will be recorded and shared on YouTube

- We will pause recording periodically to allow questions from those who do not wish to be recorded

- You can also always post questions in Slack to avoid being recorded

Materials

https://bigslu.github.io/workshops/2022.08.15_R.tidyverse.workshop/index.html

Introduction to R and the tidyverse

Welcome! In this 3-session workshop, we use real-world human RNA-seq data to introduce R data manipulation and visualization. Participants will get hands-on experience with coding in R and be introduced to tools in the R tidyverse. You **do not** need any prior experience in R or coding!

Materials

- Entire workshop

Each lesson

- 0: Setup and installation
- 1: Intro to R and RStudio
- 2: Data manipulation in dplyr and tidyr
- 3: Data visualization in ggplot

Motivation

- Why write code?
  - More reproducible (scripts!)
  - Customization and control
  - Access high-performance and high-throughput resources

- Why R?
  - Open-source and free
  - Supportive community
  - Statistical foundation
  - Well-documented and popular packages for RNAseq
  - We could go on… We 💜 R

Introduction to the data

Simmons JD, Dill-McFarland KA, *et al*. 2022. Monocyte transcriptional responses to *Mycobacterium tuberculosis* associate with resistance to tuberculin skin test and interferon gamma release assay conversion. mSphere. In press.

https://github.com/hawn-lab/RSTR_RNAseq_Mtb_public

Introduction to the data

- *Mycobacterium tuberculosis* (Mtb) is the causative agent of tuberculosis (TB)

- TB is among the top infectious killers worldwide... and had been for centuries

- Mtb predominantly infects lung macrophages

- Exposure to Mtb results in infection that is cleared, contained, or progresses to disease

- Mechanisms that distinguish these outcomes are unknown

Experimental pipeline

Whole blood from individuals with latent tuberculosis infection (LTBI)

↓

Isolate CD14+ monocytes from peripheral blood mononuclear cells (PBMC)

↓

Media or *M. tuberculosis* infection, 6 hrs

↓

Bulk RNA-sequencing

↓

Quality filter
Align to human genome GRCh38
Count reads in exons