

Before we start, please...

- Join Slack bit.ly/binfo_slack
 - You will find all of these links and code in the **#install** channel
- Download the data
https://github.com/BIGslu/2022_ASM_Microbe_RNAseq/blob/main/0_data/0_data.zip
- Install R cran.r-project.org
- Install RStudio www.rstudio.com/products/rstudio
- Install packages in RStudio console
 - `install.packages(c("tidyverse", "lme4", "lmerTest", "BiocManager", "devtools"))`
 - `BiocManager::install(c("edgeR", "biomaRt", "limma"))`
 - `devtools::install_github("BIGslu/RNAetc")`
 - `devtools::install_github("BIGslu/kimma")`

RNA-seq analysis in R

Kim Dill-McFarland (she/her)

Kelly Sovacool (she/her)

Holly Hartman (she/her)

Candace Williams (she/her)

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

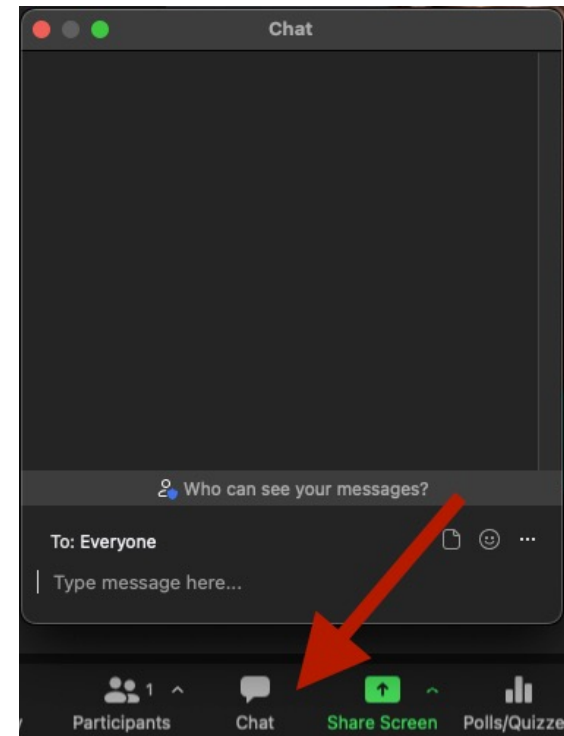
https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html

Outline

PST	MST	CST	EST	
8 AM	9 AM	10 AM	11 AM	Introduction and install check
8:20 AM	9:20 AM	10:20 AM	11:20 AM	Introduction to R and the tidyverse
11 AM	12 PM	1 PM	2 PM	Break
12 PM	1 PM	2 PM	3 PM	RNAseq data cleaning
1 PM	2 PM	3 PM	4 PM	Linear modeling
3 PM	4 PM	5 PM	6PM	End

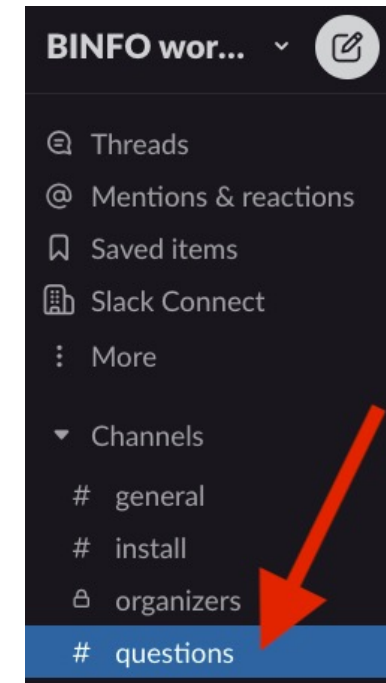
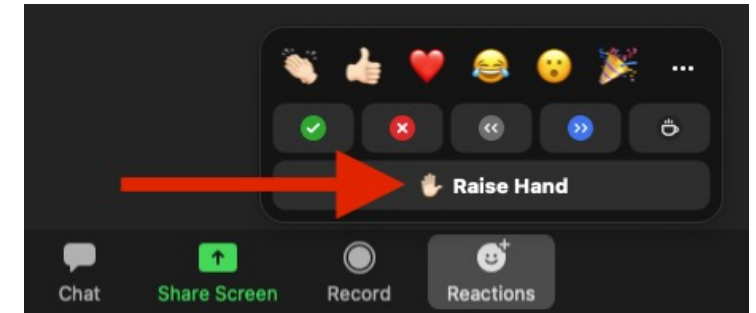
Zoom

- Cameras encouraged but not required
- Pet, child, and roommate cameos welcome!
- Feel free to step away at any time as needed
- Use Zoom chat for reactions and comments,
NOT for questions





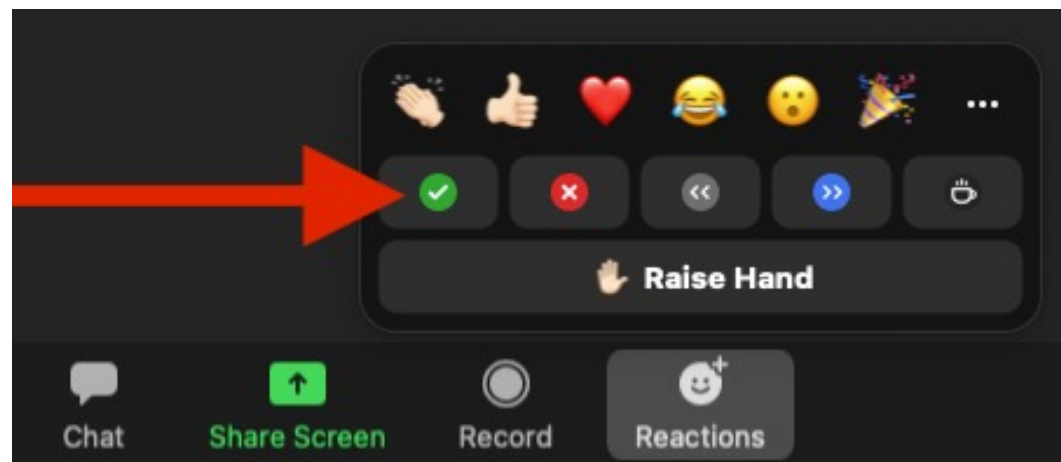
How to ask questions

- In Zoom, raise hand and once called upon, use microphone
- Write in the Slack channel **#questions**
 - Once an instructor is working on it, you will see 🖋️
 - Once it is complete, you will see ✅
- Please do not use the Zoom chat for questions as they are likely to get lost in the chatter!



Instructors checking in

- We will periodically check in to see if we can move to the next section
- Use Zoom reactions to let us know
 -  You are ready to move on
 -  You need more time or have an error
 - If you have an error or question, be sure to ask in the [#questions](#) Slack channel!



Recording

- This workshop will be recorded and made available to attendees through a private link
- 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 #questions

Materials


https://github.com/BIGslu/2022_ASM_Microbe_RNAseq

The screenshot shows the GitHub interface for the repository **BIGslu / 2022_ASM_Microbe_RNAseq**. The repository is public and has 1 watch, 0 forks, and 0 stars. The main branch is **main**. The repository contains 1 branch and 0 tags. The file list shows the following files and their last commit dates:

File	Commit Message	Commit Date
0_data	Update data	5 days ago
1_introR	Fix paths	5 days ago
2_tidyverse	setup dir structure	13 days ago
3_RNAseq_cleaning	setup dir structure	13 days ago
4_linear_models	setup dir structure	13 days ago
5_gene_set_analysis	setup dir structure	13 days ago
pantry	Create Kim_past_workshop_links	13 days ago
.gitignore	Add kinship data	13 days ago
2022_ASM_Microbe_R...	Create 2022_ASM_Microbe_RNAseq.Rproj	13 days ago
LICENSE	Initial commit	6 months ago
README.md	Update README.md	5 days ago
outline.md	Update outline.md	5 days ago

The repository has no description, website, or topics provided. It also has no releases or packages published. The contributors section shows 2 contributors.

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

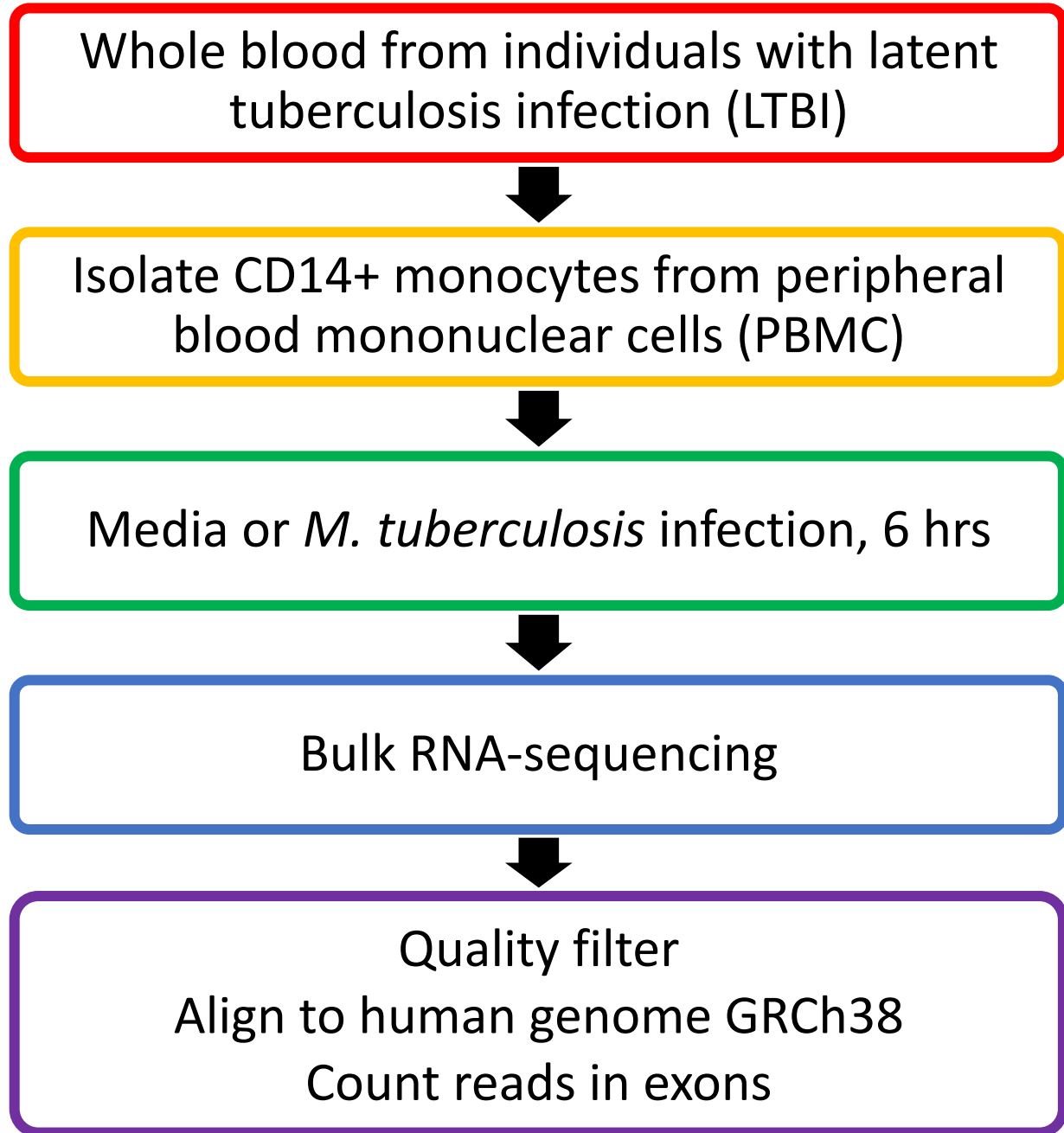
Simmons JD, Dill-McFarland KA, *et al.* 2022. Monocyte transcriptional responses to *Mycobacterium tuberculosis* associate with resistance to tuberculin skin test and interferon gamma release assay conversion. mSphere. In press.

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



Raw counts data

hgnc_symbol	pt01_Media	pt01_Mtb	pt02_Media	pt02_Mtb	...
5_8S_rRNA	20.05	20.01	7.11	6.06	
5S_rRNA	28.07	28.18	22.81	22.29	
7SK	2.28	1.28	3.1	1.04	
A1BG	25	18.3	16.68	10.79	
A1BG-AS1	46.71	46.97	40.93	27.34	
A1CF	0	0	0	0	
A2M	1271	585.93	495.99	210	
A2M-AS1	19.99	16.06	6	8	
...					

Sample metadata

libID	ptID	condition	age_dys	sex	ptID_old	RNAseq	methylation	total_seq
pt01_Media	pt01	Media	12410	M	pt00001	TRUE	FALSE	9114402.2
pt01_Mtb	pt01	Mtb	12410	M	pt00001	TRUE	FALSE	8918699.42
pt02_Media	pt02	Media	12775	M	pt00002	TRUE	FALSE	9221554.97
pt02_Mtb	pt02	Mtb	12775	M	pt00002	TRUE	FALSE	7733260.22
...								