RNA-seq data analysis workshop

Owen M. Wilkins, PhD

Bioinformatics scientist

Center for Quantitative Biology, Geisel School of Medicine at Dartmouth

Email: DataAnalyticsCore@groups.dartmouth.edu

Website: (https://sites.dartmouth.edu/cqb/projects-and-cores/data-analytics-core/

07/08/20

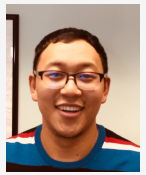
Who are we? The Data Analytics Core

Mission statement

Facilitate advanced genomic & bioinformatic data analysis solutions to CQB faculty & the Dartmouth research community

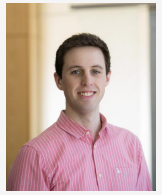
James O'Malley, PhD
Director

Shannon Soucy, PhD
Senior Research Scientist



Yue (Frank) Wang, PhD
Bioinformatics Research Scientist

Carol Ringelberg
Data analyst



Owen Wilkins, PhD
Bioinformatics Research Scientist

- **Genomic data analysis**
  - Analytical support for Dartmouth researchers
  - Pipeline development & maintenance
- **Bioinformatics consulting**
- **Publication & grant support**
  - Writing for methods & results sections
  - Letters of support
- **Submission of raw & processed data to public databases**
- **Training**
  - One-on-one analysis
  - Group workshops

Goals of the workshop

- Build an appreciation of key concepts for experimental design and generation of a typical RNA-seq dataset
- Develop understanding of pre-processing and QC steps for preparing a gene expression matrix for a typical RNA-seq dataset
- Gain familiarity working with RNA-seq data at the command-line (Bash)
- Learn the fundamentals of differential expression analysis and build foundational skills for performing differential expression in your own datasets (R/Bioconductor)

Schedule

- **Can be found at:** https://github.com/Dartmouth-Data-Analytics-Core/RNA-seq_workshop_July2020/blob/master/schedule.md

- **9am-5pm each day**

- **Schedule is best guess, and we may deviate from it based on time**

- **If you will be absent for a session, just let us know**

Logistics I

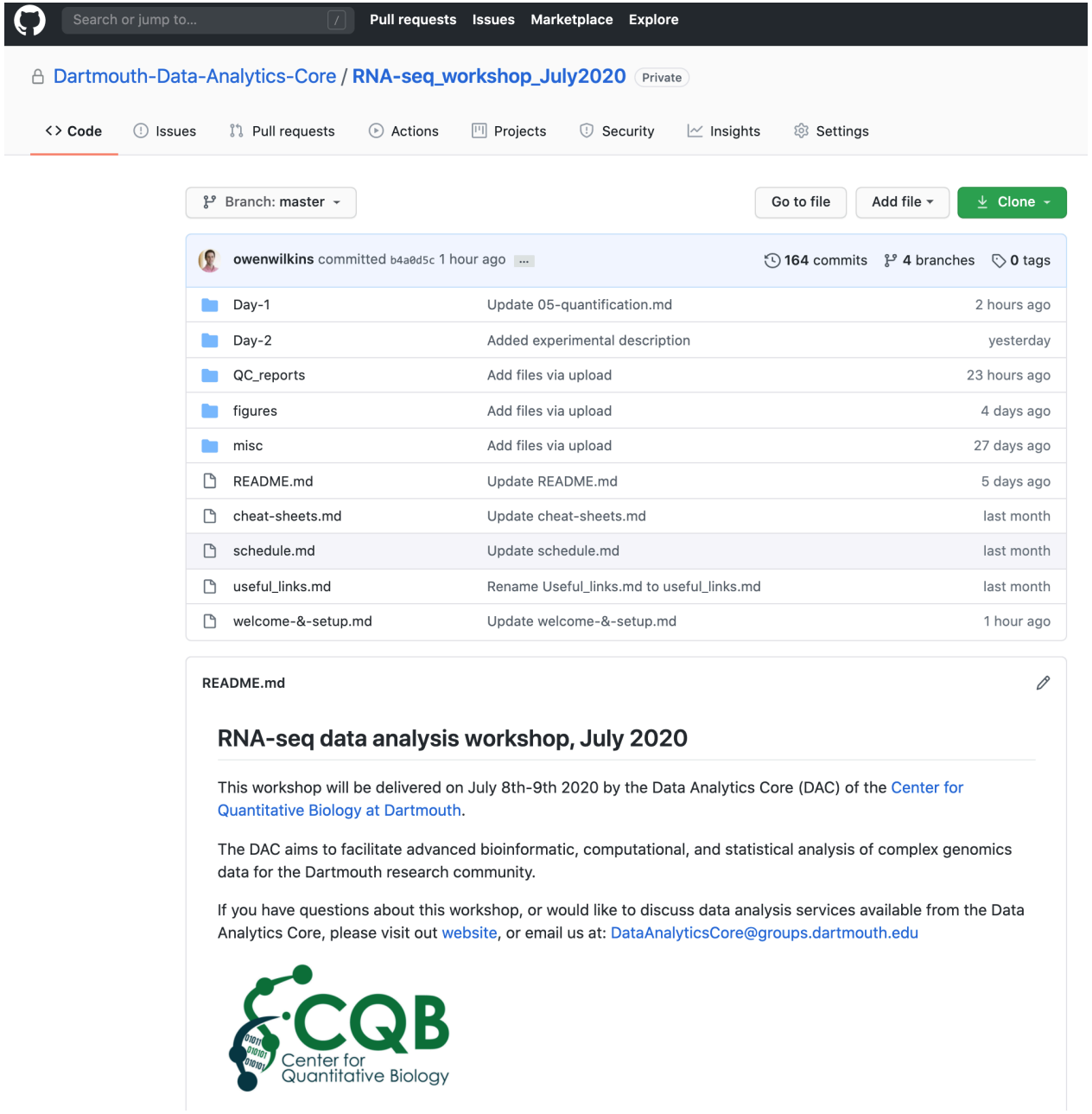
- **Course materials are all online** (and will stay there):

  https://github.com/Dartmouth-Data-Analytics-Core/RNA-seq_workshop_July2020

- **You should have the repo downloaded locally**

- **Day 1 we will be copying Bash code from markdowns (.md) into the terminal**

- **Day 2 we will be running R markdowns (.Rmd)**



Logistics II

- **For day 1:**

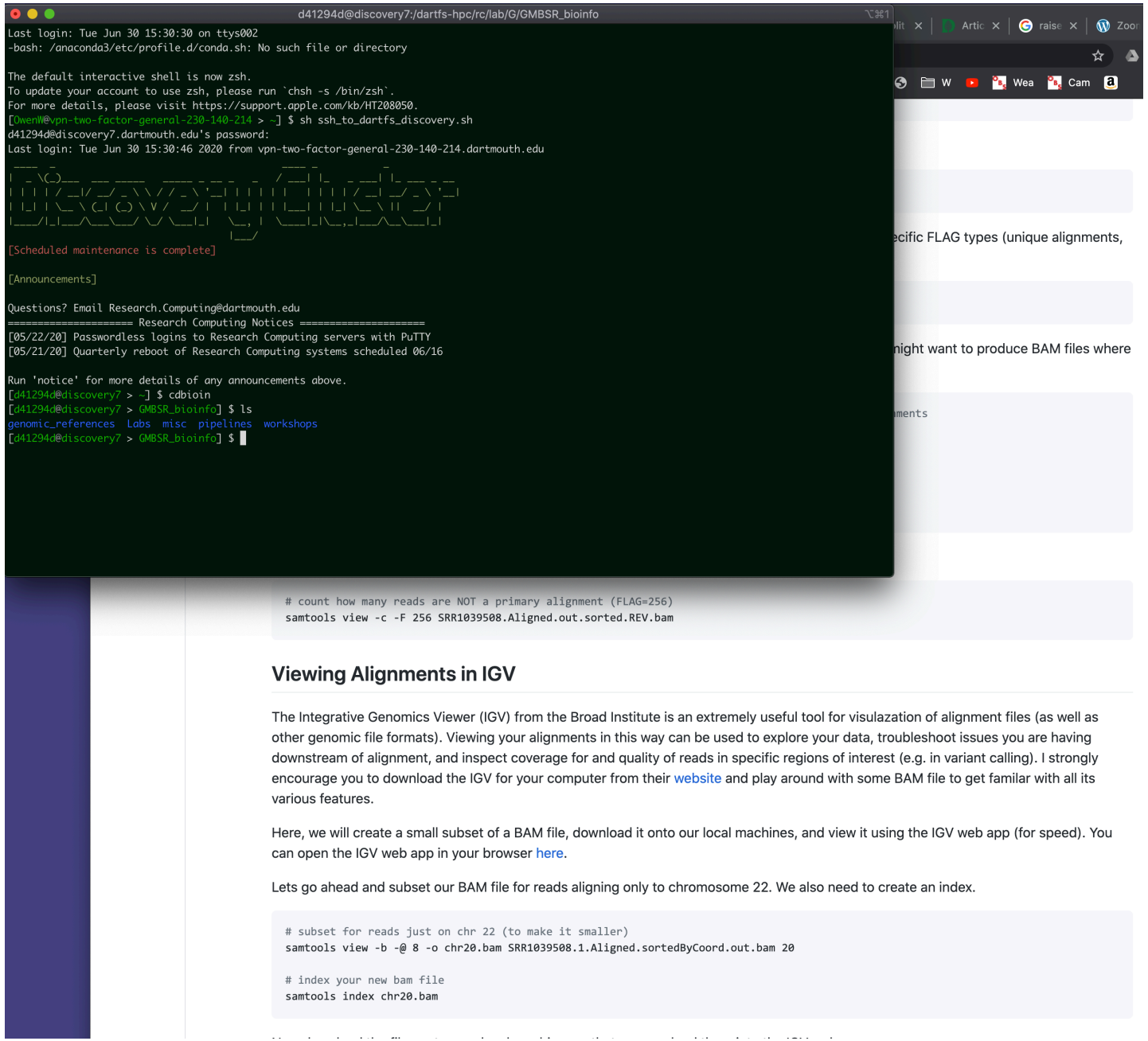
- Multiple tabs open

- Terminal window
  (ssh to discovery7)

- Web browser to GitHub repo

- If you finish: edit the code, try different options, generate scripts

- Use the ***cheat sheets!***

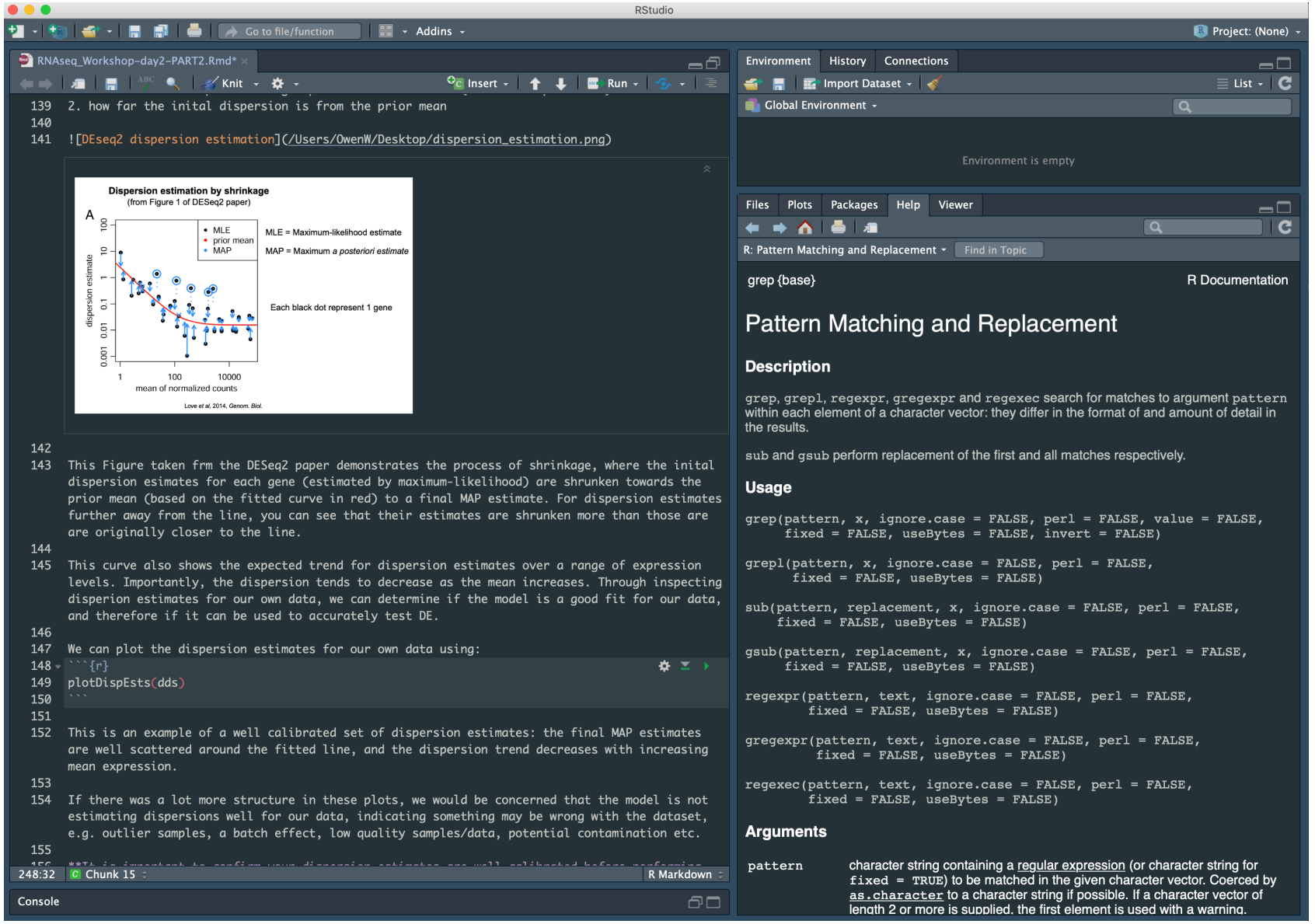


Logistics III

- **For day 2:**

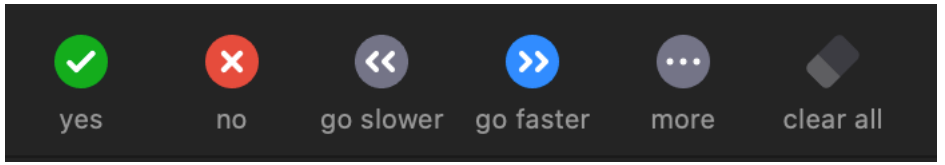
- Run Rmd files locally

- .Rmd files are in the repo you downloaded

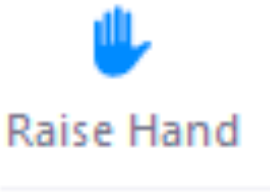


Logistics IV

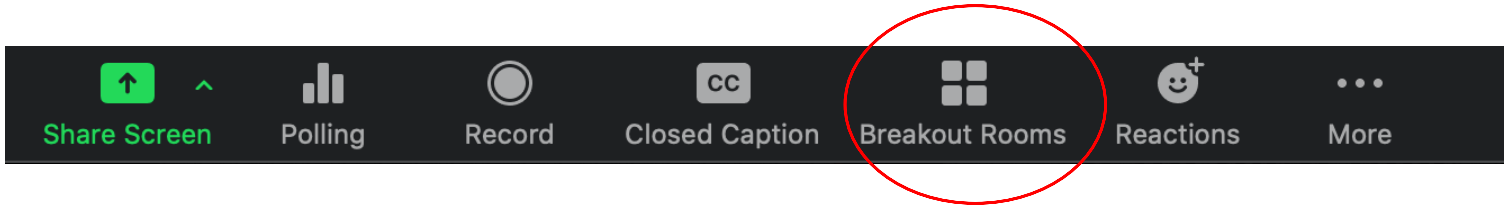
- **Use buttons in *Participants* tab in zoom**



- **You'll be muted, but if you want to ask a question, just raise your hand**



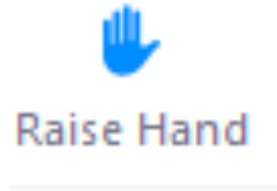
- **We'll be using *breakout rooms (BRs)***
  - We will use these when we split up to run code independently
  - We've tried to pair everyone based on experience
  - If your stuck, message us, and we will come help you in your (BR)
  - When we are going to move on, breakout rooms will close



- **Please be courteous on zoom..**

How to get help?

- **Raise your hand in zoom** (bottom right, participants tab)



- **Use the slack channel to message one of us**
  - Use the ***general*** channel if it might benefit everyone
  - Message us directly if its specific

- **If all else fails, email us:**
  - DAC: DataAnalyticsCore@groups.dartmouth.edu
  - Shannon Soucy (Shannon.Margaret.Soucy@Dartmouth.edu)
  - Owen Wilkins (omw@Dartmouth.edu)
  - Yue Wang (yue.wang@Dartmouth.edu)

Questions?

Questions later?

Bioinformatics office hours - Fridays 1-2pm. https://dartmouth.zoom.us/s/96998379866

Also, give us feedback about this workshop & other workshops you would like to see!

Bring questions for discussion at end of Day 2

…then.. Introductions!

Name, department/program, research interests, why are you taking the workshop?