



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/>)

06/14/21



Who are we? The Data Analytics Core

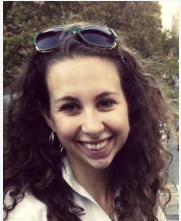


Mission statement

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



Tim Sullivan, BA
Bioinformatics Research Scientist



Owen Wilkins, PhD
Bioinformatics Research Scientist

➤ Genomic data analysis

- Analytical support for Dartmouth researchers
- Pipeline development & maintenance

➤ Bioinformatics consulting

➤ Publication & grant support

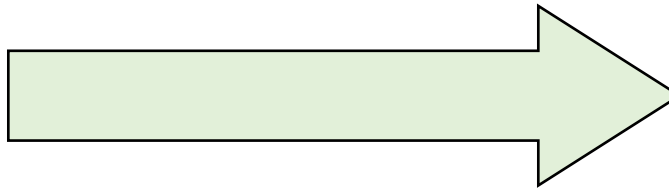
- Writing for methods & results sections
- Letters of support

➤ Training

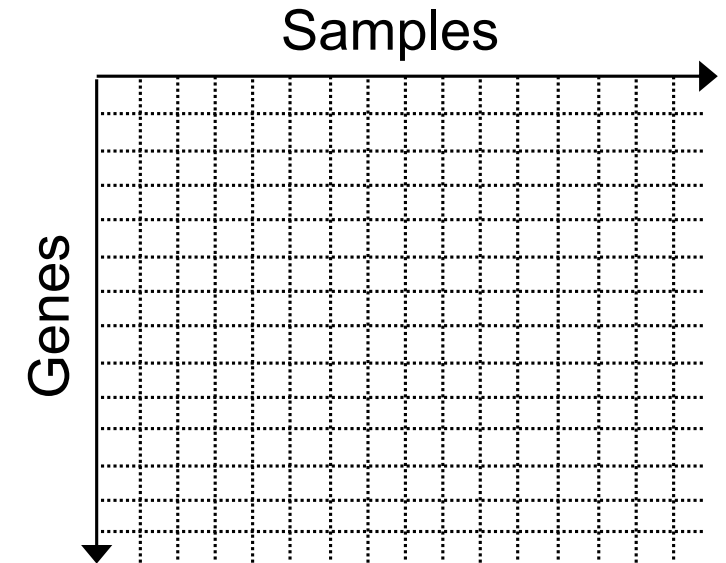
- Group workshops

Workshop outline

Millions/billions of reads from sequencing

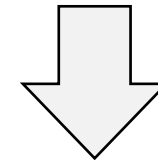


Raw & normalized counts



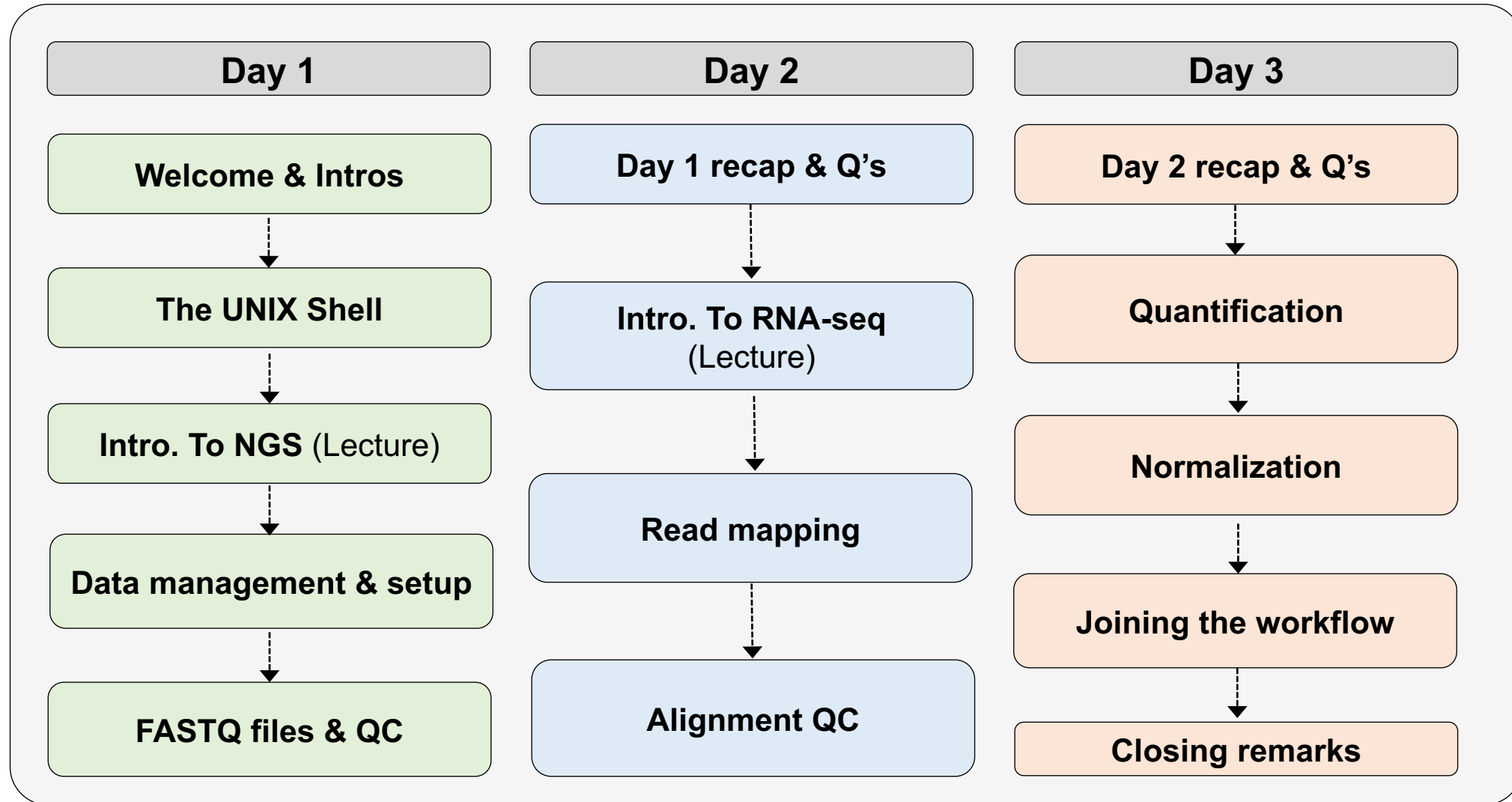
FASTQ file:

```
@SRR1039508.1 HWI-ST177:290:C0TECACXX:1:1101:1225:2130 length=63
CATTGCTGATACCAANNNNNNNNGCATTCTCAAGGTCTTCCTCCTTCCCTTACGGAATTACA
+
HJJJJJJJJJJJJJJJ#####00?GHIJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJHHHFFFFFD
```



Differential expression
workshop (July)

Workshop outline



Goals of the workshop

- Develop a working understanding of the analytical workflow for a modern bulk RNA-seq experiment
- Build a working knowledge of sample preparation considerations for RNA-seq experiments
- Learn how to process raw NGS data in FASTQ format to generate a gene expression matrix
- Learn how to perform a detailed quality control analysis

Schedule

- Can be found at: <https://github.com/Dartmouth-Data-Analytics-Core/RNA-seq-Primary-Data-Analysis-workshop-June-2021/blob/main/schedule.md>
- 12pm-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-Primary-Data-Analysis-workshop-June-2021>

- Lessons, lectures, cheat sheets, etc. are all here

- We will be running Bash code from markdowns (.md) into your terminal application

The screenshot shows a GitHub repository page for 'Dartmouth-Data-Analytics-Core / RNA-seq-Primary-Data-Analysis-workshop-June-2021'. The repository is private and has 49 commits. The file list includes folders for 'Day-1', 'Day-2', 'Day-3', 'QC_reports', 'figures', 'lectures', and 'misc', as well as files like '.DS_Store', 'README.md', 'cheat-sheets.md', 'schedule.md', 'useful_links.md', and 'welcome-&-setup.md'. The 'README.md' file is selected, showing the title 'RNA-seq Primary Data Analysis Workshop - June 2021' and the text: 'This workshop will be delivered on June 14th, 16th, and 18th 2021 by the Data Analytics Core (DAC) of the Center for Quantitative Biology at Dartmouth. All the teaching materials required for the workshop can be downloaded from this repository. Datasets used in the workshop will be available through the Dartmouth Discovery cluster, which we will use to perform analysis during the workshop. You must obtain a discovery account before the workshop to participate.'

Logistics II

Multiple tabs open

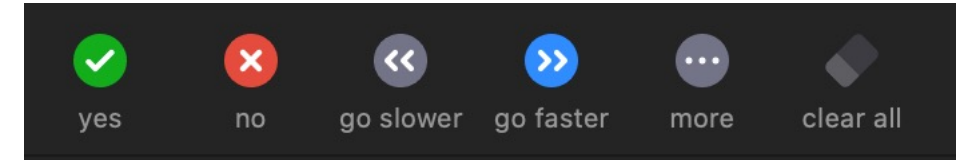
- Terminal window
(ssh to discovery7)
- Web browser to GitHub repo

- - If you finish: edit the code, try different options, generate scripts
- Use the ***cheat sheets!***

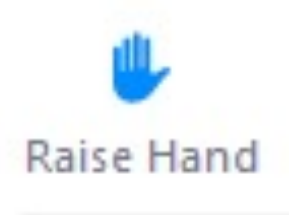
[illegible]

Logistics III

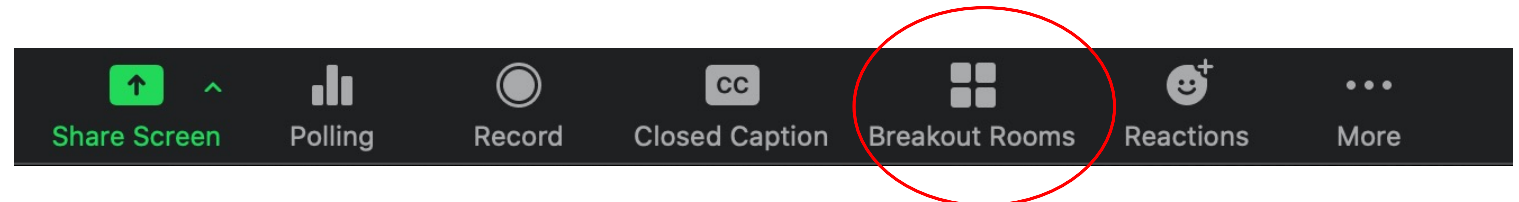
- Use buttons in Participants tab in zoom



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



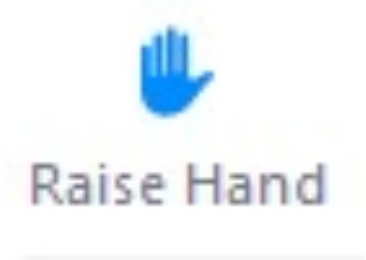
- We'll be using *breakout rooms (BRs)*
 - We will use these when we split up to run code independently
 - We've tried to pair everyone based on experience
 - If your stuck, message us, and we will come help you in your (BR)
 - When we are going to move on, breakout rooms will close



- 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)
 - Tim Sullivan (Timothy.J.Sullivan@dartmouth.edu)

Questions after the workshop?

Bioinformatics office hours

Friday 1-2pm (every other week, check calendar:
<https://sites.dartmouth.edu/cqb/upcoming-events/calendar/>

Zoom link: <https://dartmouth.zoom.us/s/96998379866>

Passcode: bioinfo

- Also, please give us feedback about this workshop, there will be a survey!
- And please ask lots of questions!

Questions?

...then.. Introductions!

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