Introduction to RNA Seq Course

We will meet at FNT 1.104 from Monday - Friday 1-4pm for our classes.

You are strongly encouraged to attend in person, but if you would like to attend virtually, use this zoom link:

https://utexas.zoom.us/j/91424696609

Zoom Instructions:

Please make sure your ***zoom version is updated (and that you have a zoom account)***-- this is required to join a UT-sponsored Zoom session. See this link for more details about zoom requirements: https://zoom.its.utexas.edu/home

Other setup:

Please make sure you have an ssh client installed on your computer. All macs come with terminal, so no installation is required. For windows laptops, install putty and winSCP. You will also need a TACC account with multi factor authentication set up.

Use our summer school reservation when submitting jobs to get higher priority.

sbatch --reservation=RNASeq-Mon <batchfile>.slurm

sbatch --reservation=RNASeq-Tue <batchfile>.slurm

sbatch --reservation=RNASeq-Wed <batchfile>.slurm

sbatch --reservation=RNASeq-Thu <batchfile>.slurm

sbatch --reservation=RNASeq-Fri <batchfile>.slurm

```
#If the RNAday reservation does not work for you, try the command without the --reservation=RNAday
```

Course Overview

This is a course designed to give you an overview of RNA-Sequencing in a hands-on manner. It will comprise of lectures and guided tutorials. For the tutorials, we will be used a canned dataset, but a large part of the last class day will be devoted to practicing the techniques learned on your own dataset. This course has the following objectives:

1. To teach you about the different options that are available to you when setting up a RNA-Seq study.
2. To teach you about the different options that are available to you when analyzing a RNA-Seq dataset.
3. To familiarize you with how the Texas Advanced Computing Center (TACC) can be used to simplify and speed up your data analysis.
4. To familiarize you with some of the typically used RNA-Seq analyses methods.
5. To provide a vocabulary to understand NGS and RNA-Seq terminology and to provide give you a starting point of where to begin you own data analysis, and enough experience that you can begin that analysis on your own.

Your Instructors

Name	Affiliation	Expertise	How to contact?
Dhivya Arasappan (Instructor)	CBRS	Unix, TACC, RNA-Seq, Single cell RNA-Seq, hybrid genome assembly, pacbio	darasappan@austin.utexas.edu or come to FNT1. 206D
Michael Keist (Teaching Assistant)	CNS	Bulk RNA-Seq, R, Unix	mwkeist@utexas.edu

Day 1: Introduction to RNA-Seq

- Logistics-Day 1
- Part 1: Lecture: Introduction to RNA-Seq
- Part 2: Unix and TACC Refresher

        - The Quickest Unix Refresher ever
        - TACC Overview
            - Using The Stampede2 cluster
            - Stampede2 directory structure
            - Submitting jobs to Stampede2
            - Stampede2 TACC/Unix Exercise

- Part 3: Getting NGS data ready
    - The Big Picture
    - Evaluating your raw data
    - Dhivya's suggestions for QC cutoffs
- Day 1 take aways
- Overview of TACC Submission

Day 2: Read Mapping and Pseudomapping

- Logistics-Day 2
- Read Mapping
    - Lecture: Introduction to read mapping
    - Mapping with BWA
    - (OPTIONAL) Mapping with HISAT2
    - Mapping with STAR
    - Assessing mapping results
    - Assessing mapping results II
    - Pseudo mapping and quantification using kallisto
- Dhivya's suggestions for Mapping
- Day 2 take aways

Day 3: Testing for Differential Expression and Visualization

- Logistics-Day 3
- The Big Picture
- Part A: Differential expression of annotated genes/transcripts
    - Lecture: Gene Counting and Differential Expression
    - Gene counting
    - Testing for differential expression
    - Lecture: Visualization
    - Plotting DESEQ2 Results
    - (OPTIONAL) Visualization using Integrated Genome Viewer (IGV)
- Part B: (OPTIONAL) The new tuxedo suite for finding novel transcripts
    - Running the new tuxedo suite and Results
- Day 3 take aways

Day 4: Downstream Analysis of Differentially Expressed Genes and 3' Targeted RNA-Seq

- Logistics-Day4
- The Big Picture

- Part 1: GO Enrichment and Pathway Analysis
    - (OPTIONAL): Advanced GO/KOG Analysis
    - Pathway Analysis
    - (OPTIONAL) Gene Set Enrichment Analysis
- Part 2: Clustering using WGCNA
- (OPTIONAL) Lecture: Example project downstream analysis
- Part 3: Lecture: Introduction to tag-seq (3' targeted sequencing)
- Day 4 take aways

Day 5: Single Cell RNA-Seq And Practice

- Introduction to Single Cell RNA-Seq
- Single Cell RNA-Seq R script and Data
- Introduction to CBRS offered pods
- Practice time/Bring your own RNA-Seq Data

Extras

- Resources
- Installing tools on a server

MANY THANKS

Many thanks to **Dr. Scott Hunicke-Smith** and **Dr. Jeff Barrick** who taught the first Introduction to NGS course at UT. Many slides and wiki pages have been borrowed and adapted from their initial course.

Also, thanks to **Dennis Wylie, Anna Battenhouse, Benni Goetz** who have also provided some material for this course