

Bootcamp Last Day!

BMS Bootcamp 2019

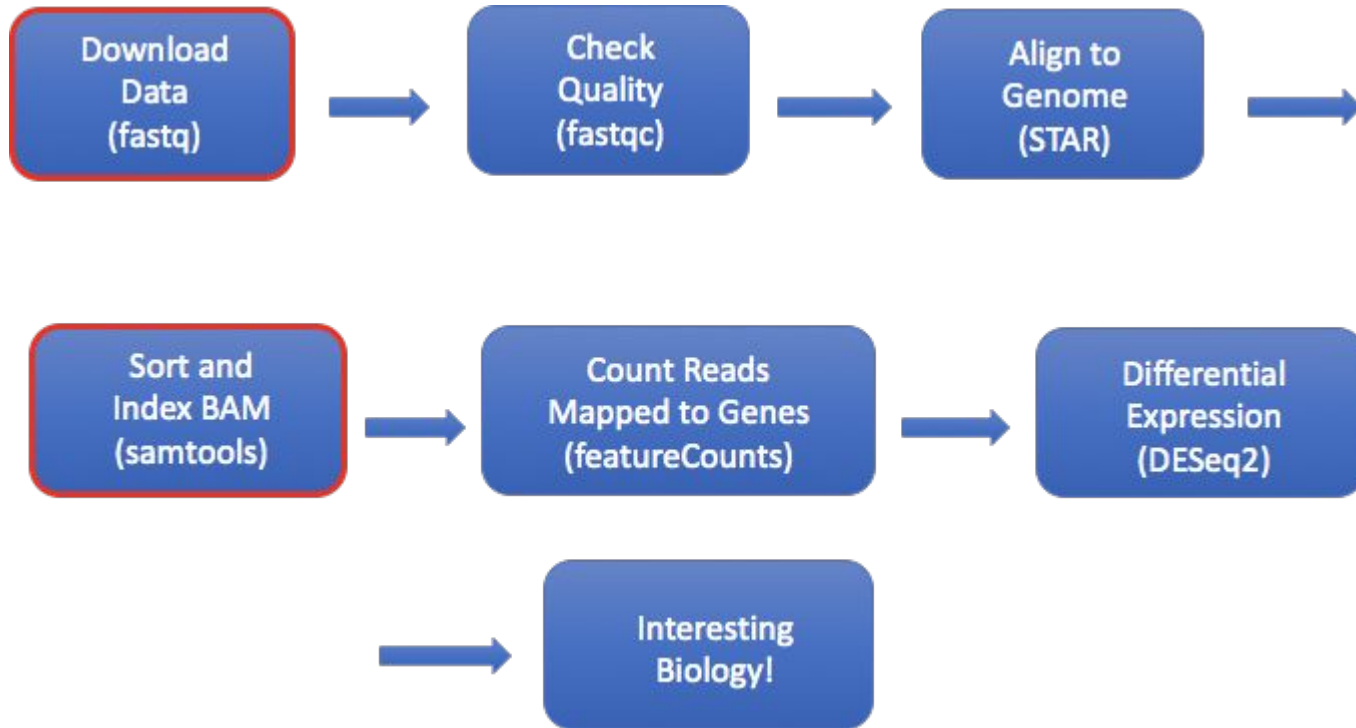
Schedule

- More data analysis
- Finish presentations
- Present!

Learning Objectives

- Basic coding in Python/R
- Visualize data
- Draw biological conclusions from data analysis

RNA-Seq Pipeline

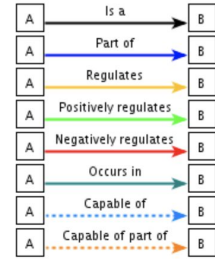
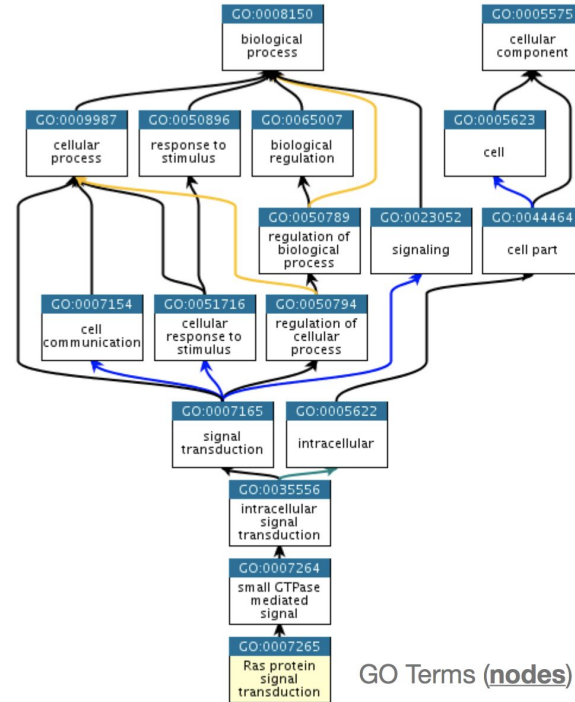


Why pathway analysis?

- From these huge lists of genes, we want to find trends in these lists
- Are genes that we find differentially expressed also found to be enriched in certain pathways?
- Can this lead to biological conclusions about treatment vs control?

GO Term Analysis

- Vocabulary of terms used to describe biological processes that are controlled and searchable
- We can use these to annotate the gene lists we get of up and downregulated genes



Parent terms are more general & child terms more specific

Many open-source ontology databases

PANTHER - <http://pantherdb.org/>

GOrilla - <http://cbl-gorilla.cs.technion.ac.il/>

DAVID - <https://david.ncifcrf.gov/>

KEGG Pathways - <https://www.genome.jp/kegg/>

Reactome - <https://reactome.org/PathwayBrowser/#/>

Limitations of this analysis

- These annotations are biased to certain pathways that are more annotated/studied
- Also biased to the size of a pathway
- Tissue specific annotations are lost
- Post transcriptional regulation is ignored

Big-Picture Objectives



Learn how to code in UNIX



Read documentation for software packages



Use Python to analyze and present data



Answer your own research questions with computational techniques

Please fill out our survey!

Every year we have altered the class based on the feedback from last year's class.

Will be sent out in an email after class.