M2D4:
Analyze RNA-seq data and prepare for qualitative PCR experiment

1. M2Q1
2. BE Comm Lab workshop
3. Prelab discussion
4. Design qPCR primers
5. Analyze RNA-seq data

Overview of Mod 2 experiments

Purify RNA for quantitative PCR analysis

DAY 1

seed cells

DAY 2

treat cells

DAY 3

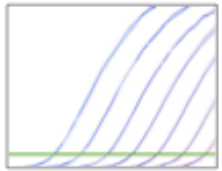
ACAUGGAUUA

↓

TGTACCTAAT

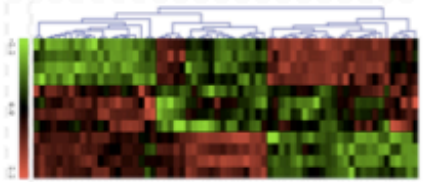
purify RNA and prepare cDNA

DAY 5



measure transcript levels

Analyze RNA-seq results

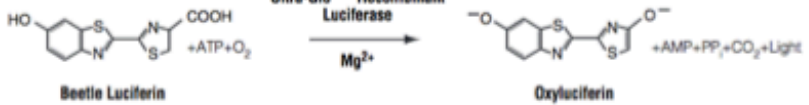


DAY 4: Evaluate altered gene expression
DAY 5: Explore public datasets
DAY 6: Complete RNA-seq data analysis

Design and perform cell viability experiment

DAYS 6 - 8

seed and treat cells

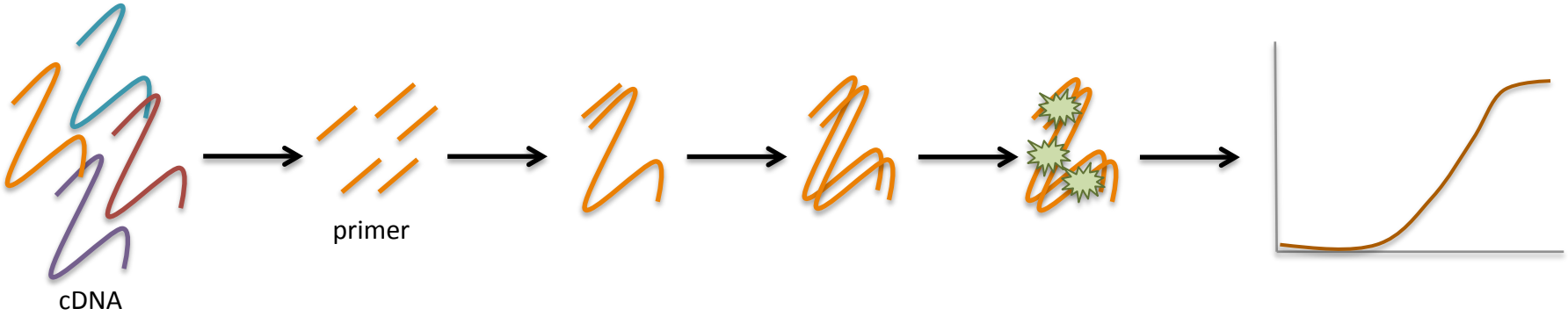


DAY 9

measure cell survival

qPCR measures transcript levels

- As in RNA-seq libraries, purified RNA is used to generate cDNA
- cDNA is probed using primers designed to specific genes of interest



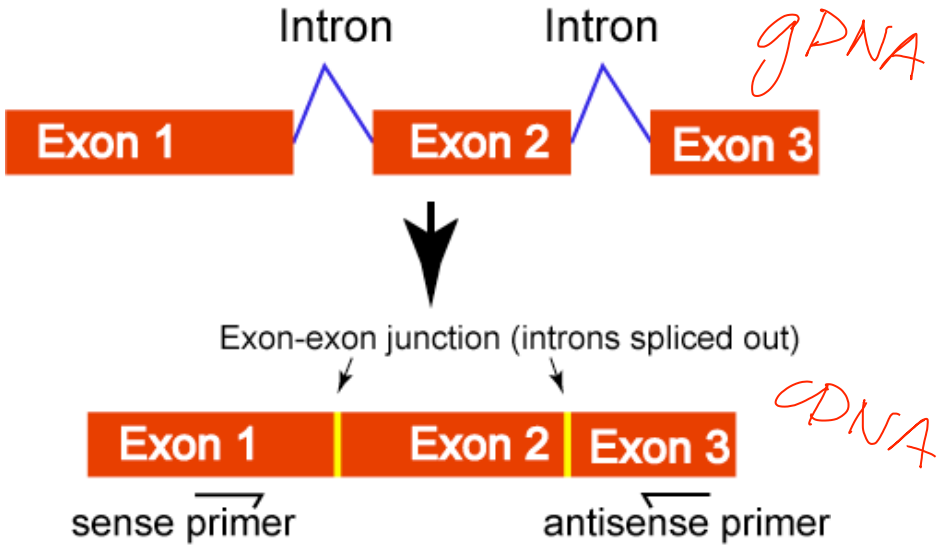
You will measure p21 and GAPDH expression

- p21 (CDKN1A)
  - Regulator of cell cycle progression at G1
  - Arrests cell cycle in response to stress
  - Controlled by tumor suppressor p53
- GAPDH (glyceraldehyde 3-phosphate dehydrogenase)
  - Involved in glycolysis
  - Constitutively expressed
  - ‘Housekeeping’ gene

Primer design guidelines for qPCR

- Must span exon-exon junction

- Length = 100 bp
- GC content = 50 – 60%
- Tm = 60 °C
  - forward ≅ reverse
- G / C at 3’ end



Why are you using qPCR to measure p21 expression? GAPDH expression?

How does your qPCR data relate to the RNA-seq data?

RNA-seq data was pre-processed

- Data from sequencer provided reads
- Reads were aligned to reference genome
- Aligned reads were counted

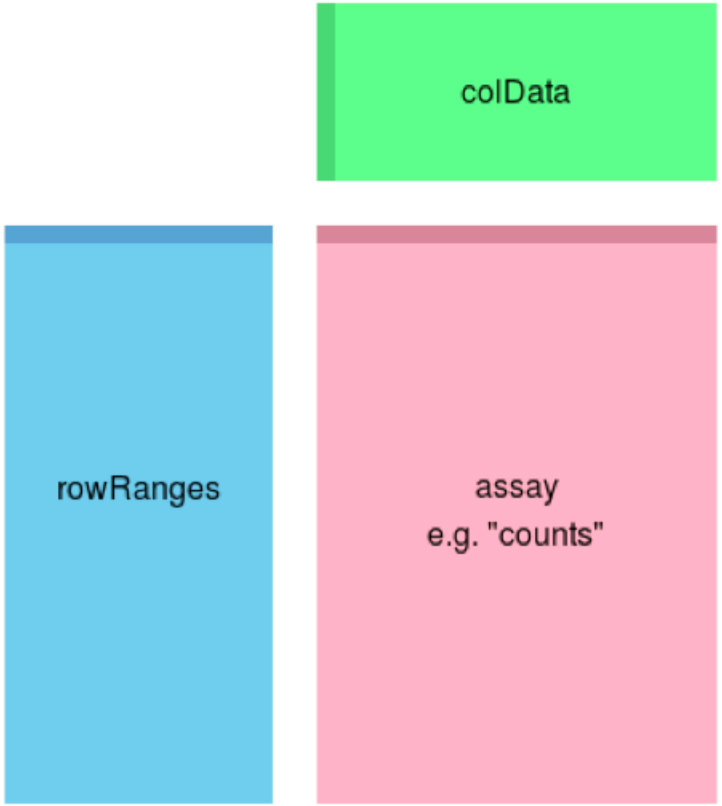
length of gene
number of total reads] RPKM

You will use count information in your data analysis!

RNA-seq data analysis

Reads aligned to genes were loaded into data structure called “DESeqDataSet”

- colData: sample information
- rowRanges: gene information
- assay: matrix of counts assigned to each gene for each sample



Workflow for RNA-seq data analysis

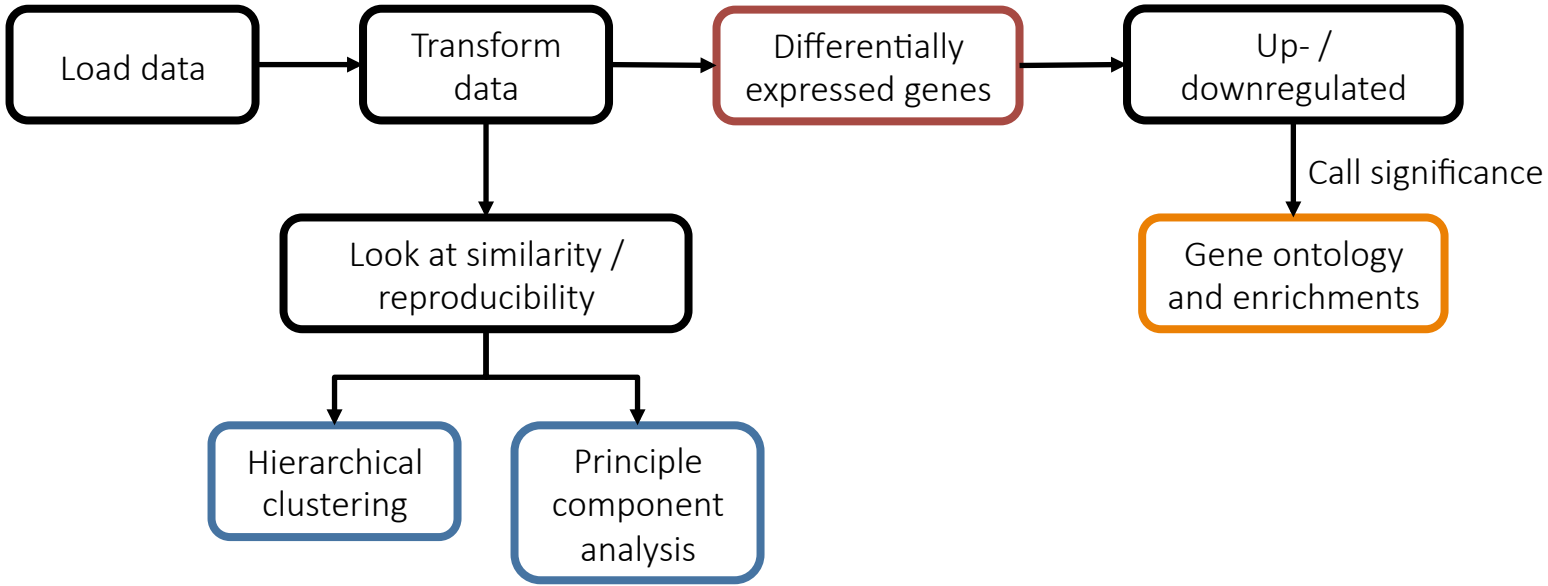
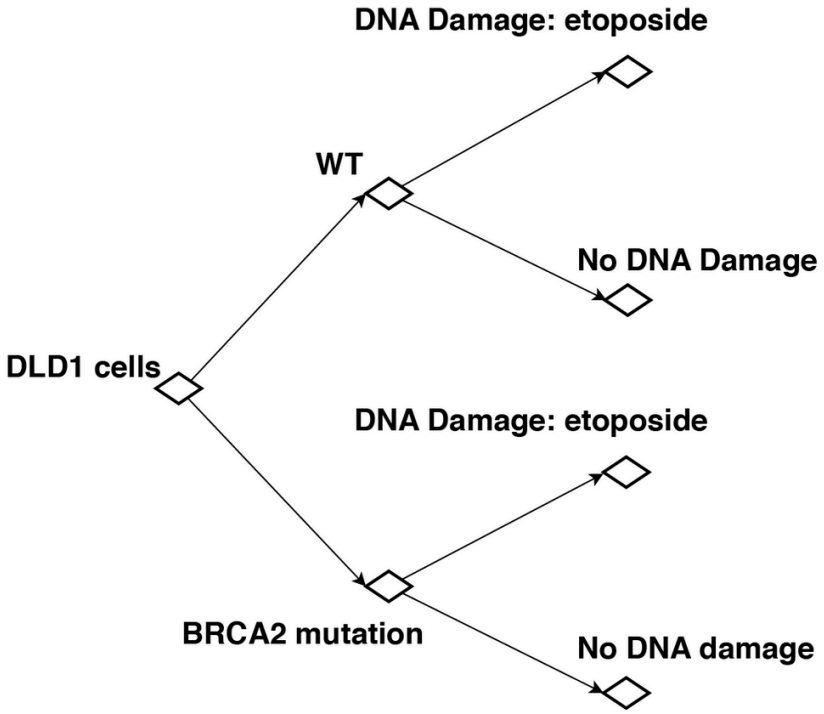


Image from Casper Enghuus, Sp17 TA 20.109

RNA-seq sample overview

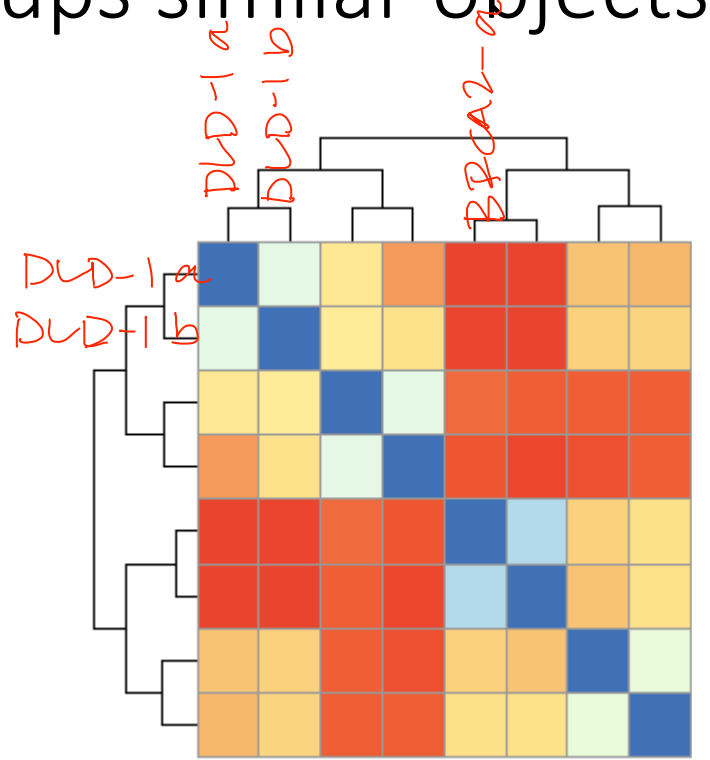
- 2 cell lines used
- 2 variables tested
- Completed in duplicate

How many samples will you be comparing / analyzing?

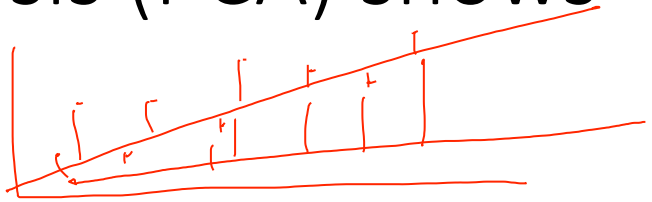


Hierarchical clustering groups similar objects

- How are samples clustered (ie how is similarity quantified)?
- What do colors represent?

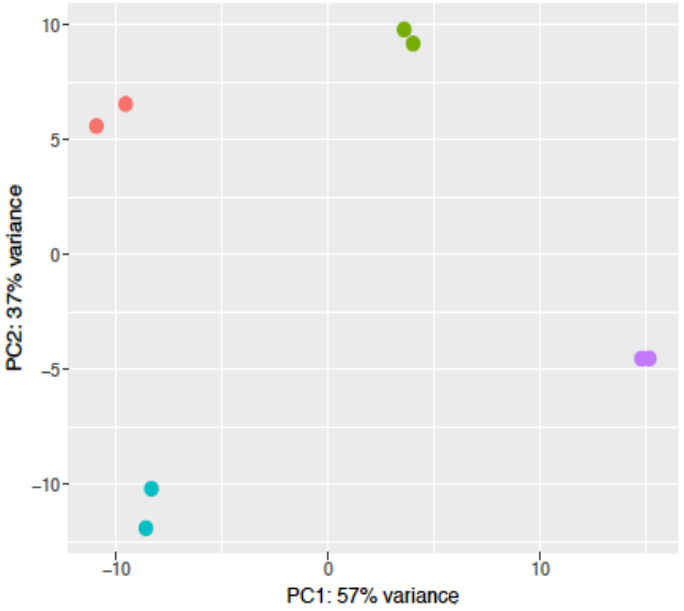


Principle component analysis (PCA) shows relatedness of objects



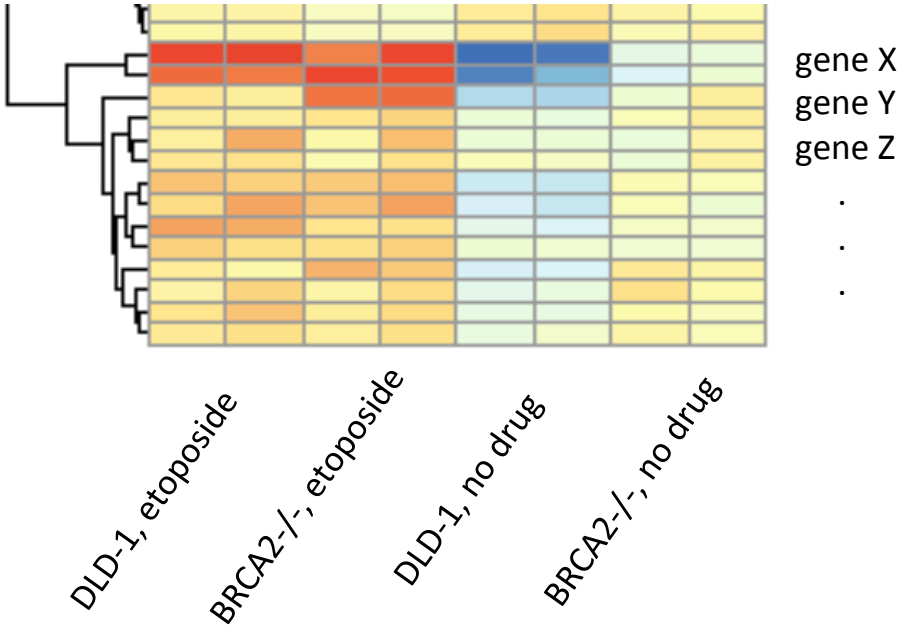
- How are principle components defined?

- How much of the variance between samples is represented in this plot?



Assessing differential gene expression

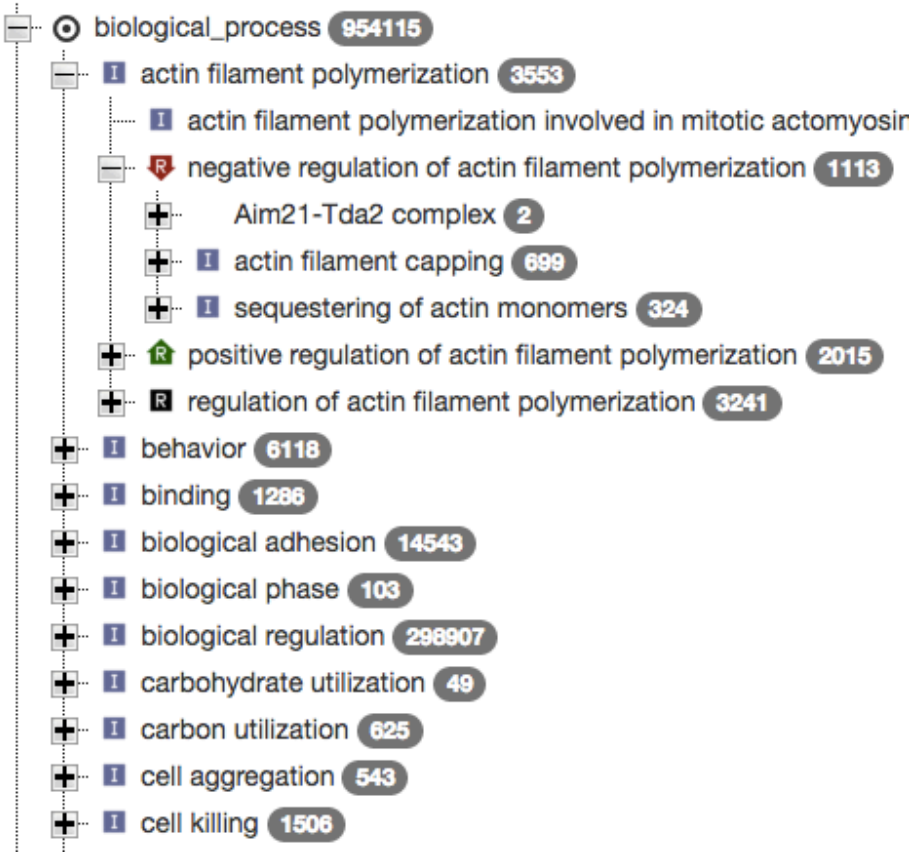
- How is ‘up / down’ expression defined?
- What information is provided by clustering (left of heatmap)?



Gene ontology (GO) terms based on gene product properties

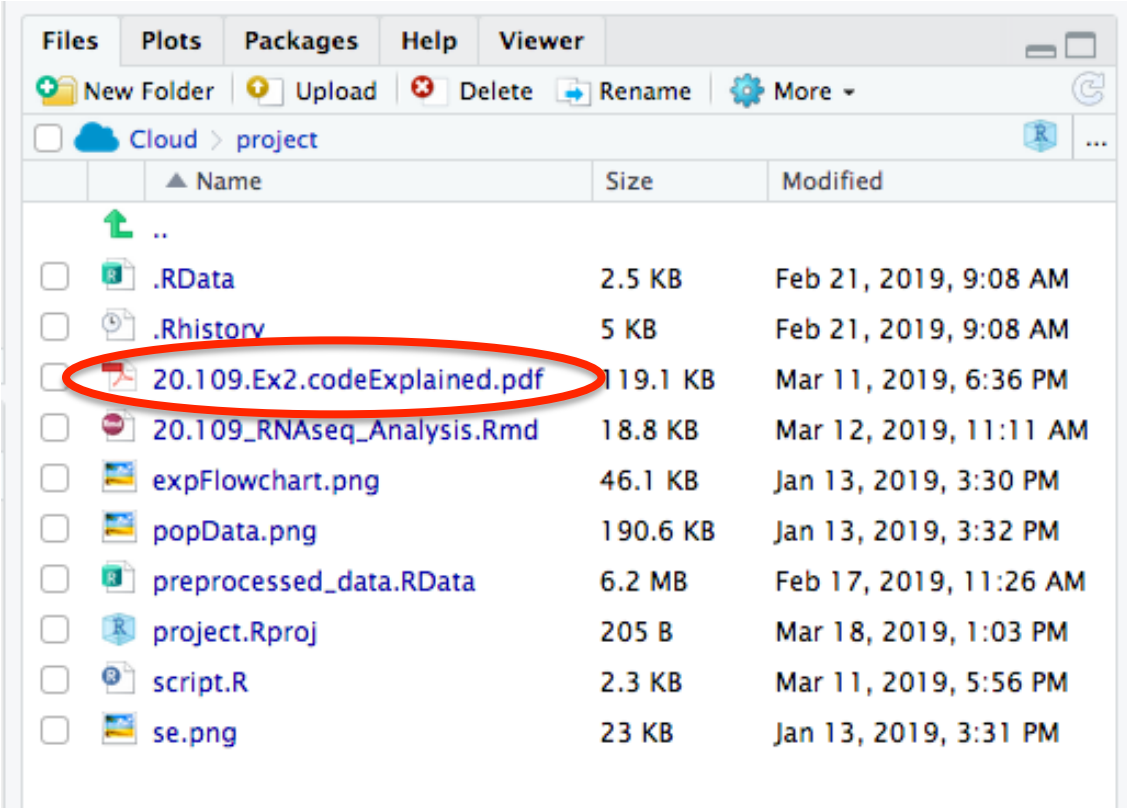
Gene classifications / characterizations based on:

- Molecular function
- Cellular location
- Biological processes



Getting help with R: ‘ask’ Amanda

Several resources available in bottom right window



Getting help with R: ask R

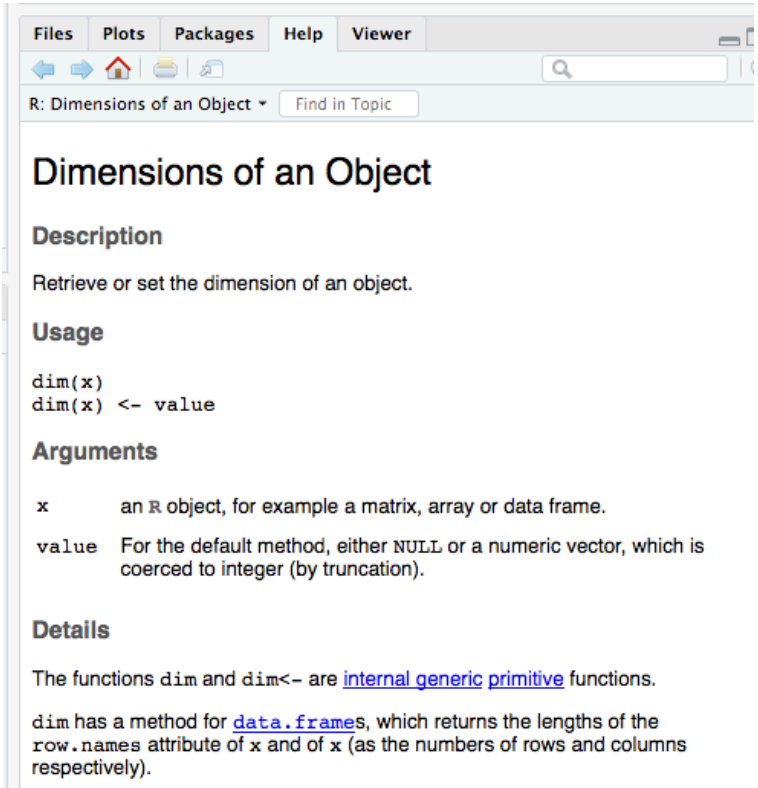
Example: Type `?dim` or `help(dim)`

Inquire about functions:

Method 1
`?function`

or

Method 2
`help(function)`



For today…

- Submit primer sequences before you leave!!
- Take time to understand the RNA-seq data analysis

For M2D5…

- Peer review classmate's M2 methods section
  - Provide numbered comments on a separate document