M2D7: Analyze RNA-Seq data

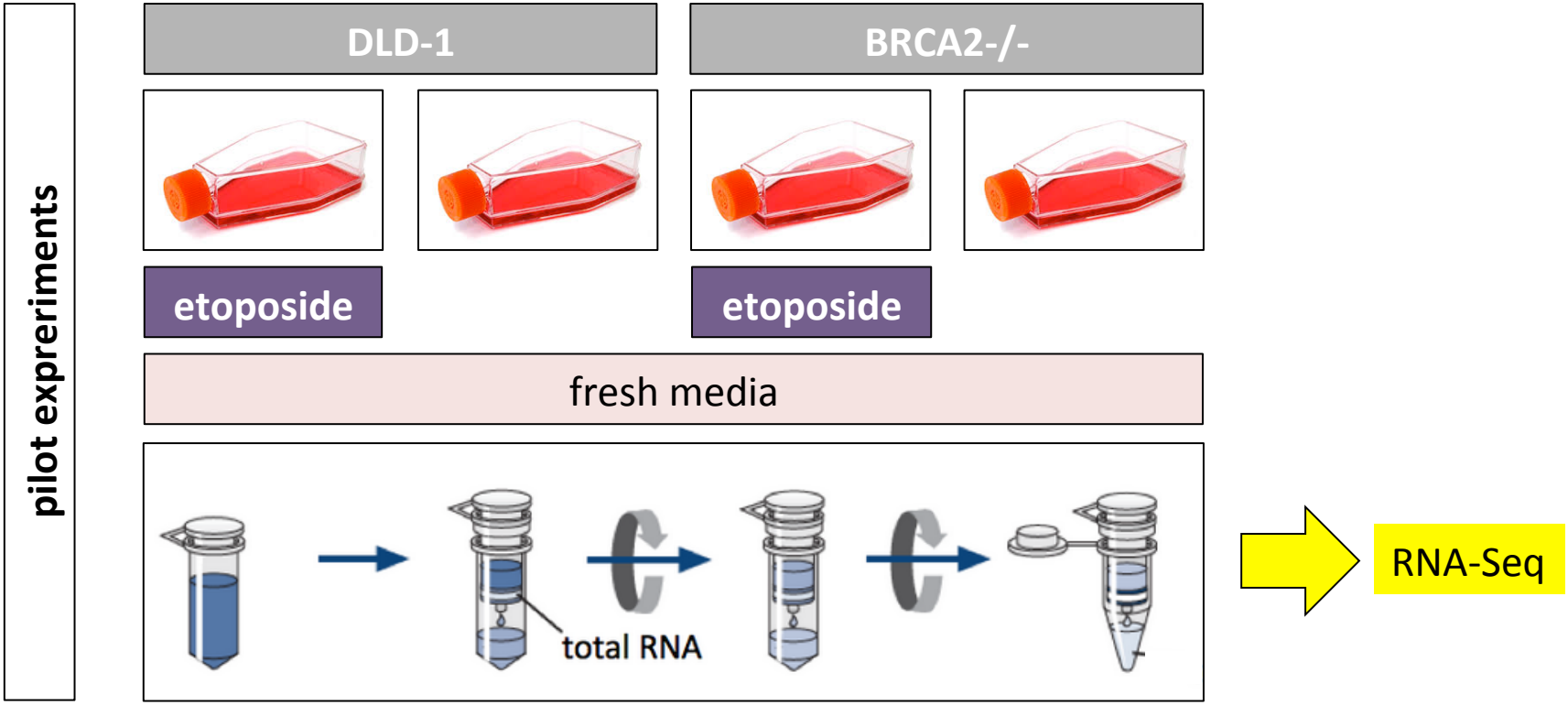
04/07/2017

1. Prelab Discussion
2. Intro to R Exercise
3. Start Mod2 Report early!

Let's analyze RNA-Seq data

DNA damage reponse

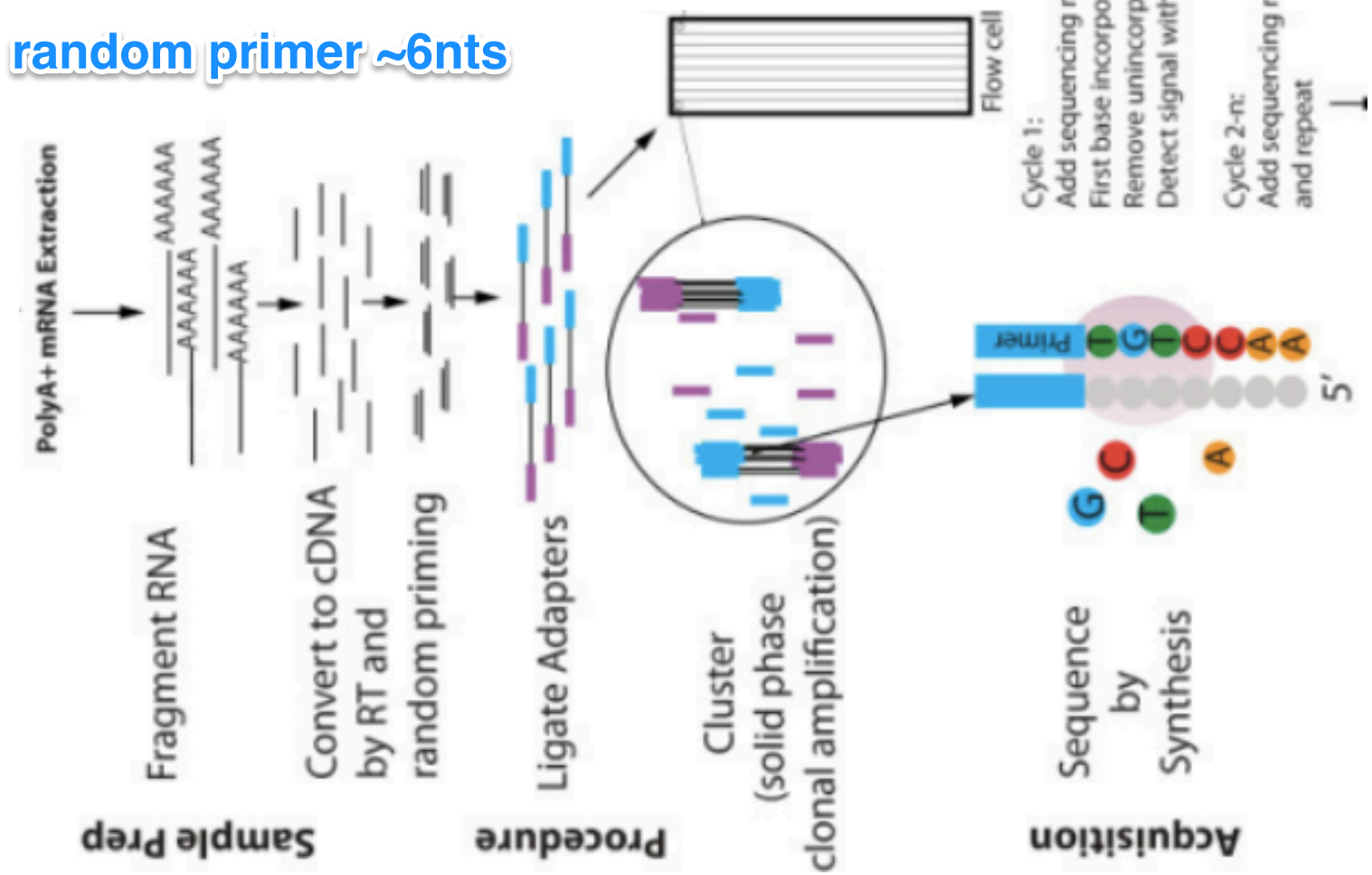
- With qPCR, you looked at p21 transcript levels (norm. to GAPDH)
- With RNA-Seq, we'll peek into the entire transcriptome



OVERALL

Next generation sequencing (NGS) by Illumina for RNA-Seq

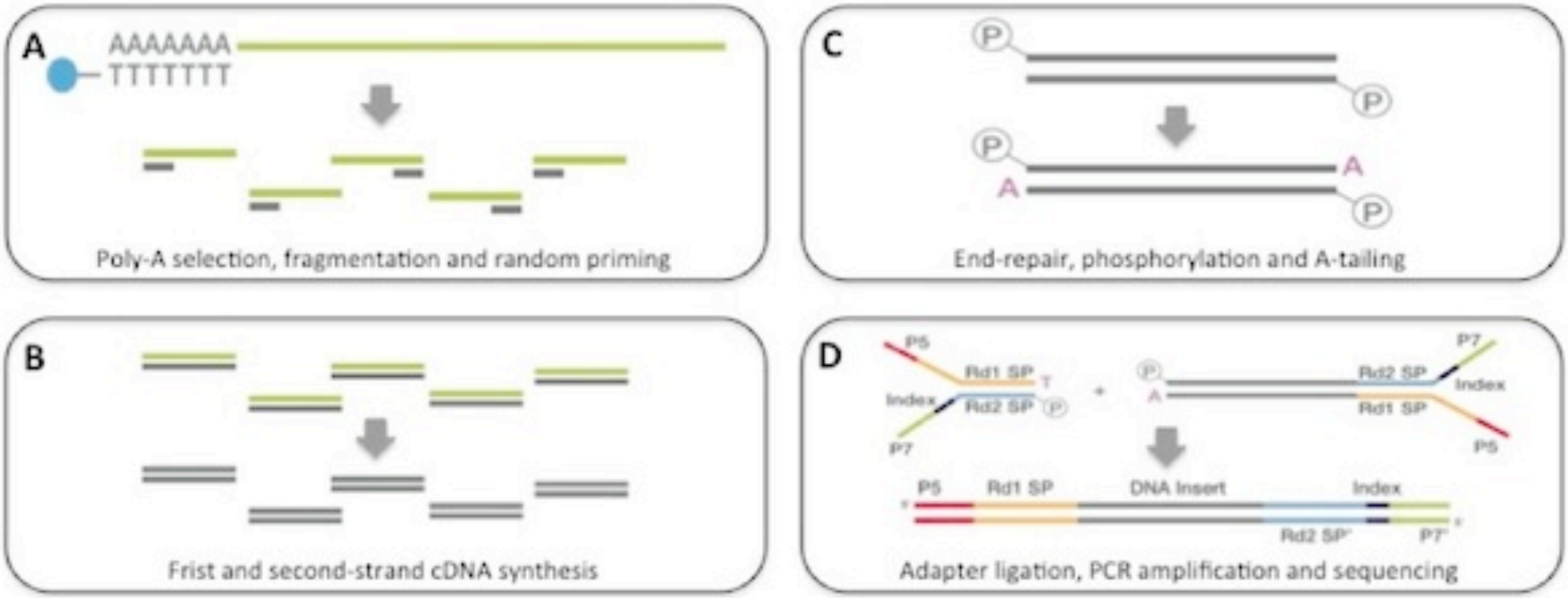
at MIT's BioMicro Center



Malone JH, Oliver B, BMC Biol. 2011, https://openi.nlm.nih.gov/

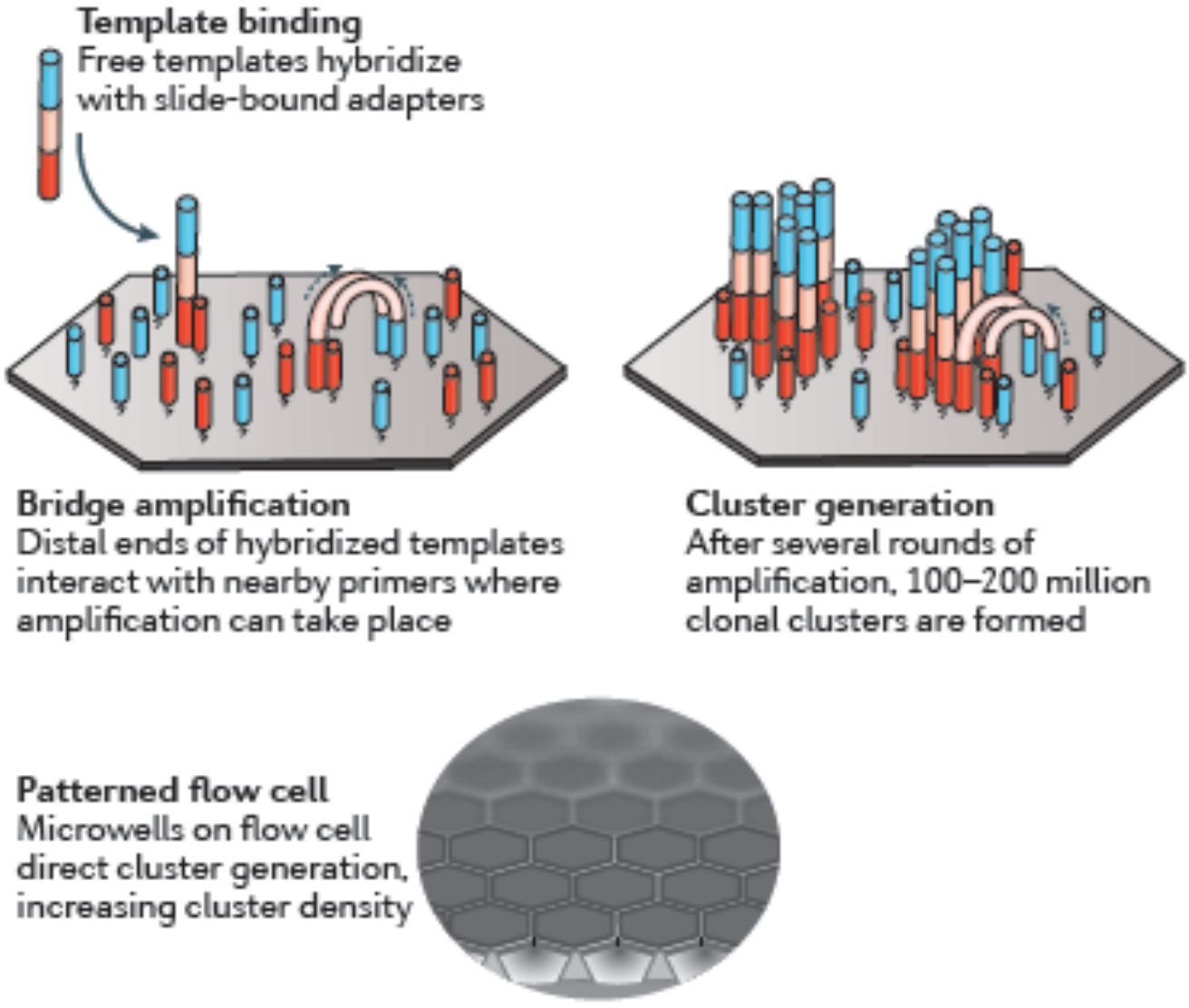
1ug RNA for library prep

Illumina Tru-Seq RNA-seq protocol



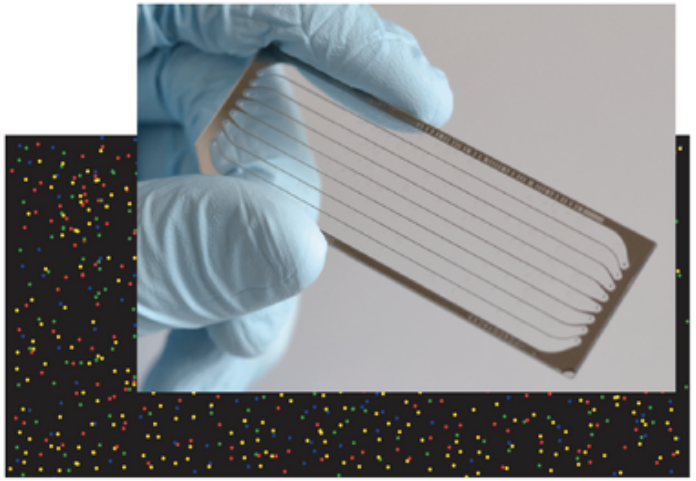
Library prep begins from 100ng-1ug of Total RNA which is poly-A selected (A) with magnetic beads. Double-stranded cDNA (B) is phosphorylated and A-tailed (C) ready for adapter ligation. The library is PCR amplified (D) ready for clustering and sequencing.

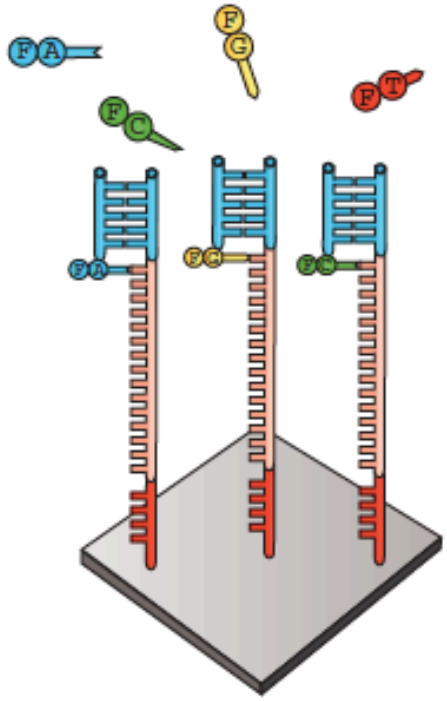
Solid-phase bridge amplification for clonal amplification of cDNA



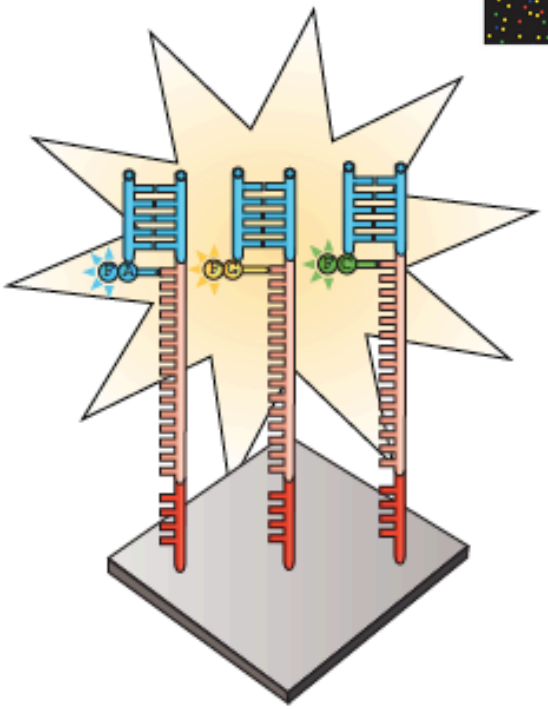
Review: http://www.nature.com/nrg/journal/v17/n6/full/nrg.2016.49.html

Sequencing by synthesis (SBS)

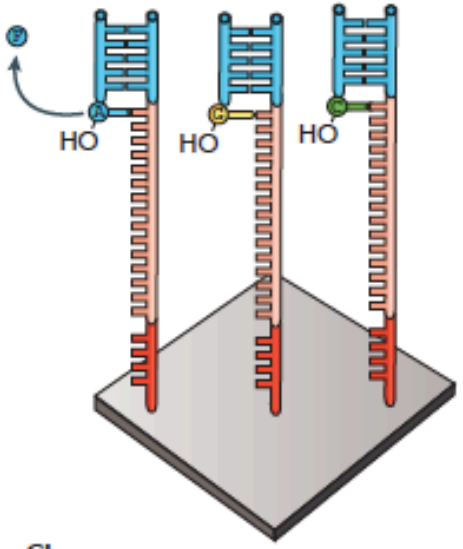




Nucleotide addition
Fluorophore-labelled, terminally blocked nucleotides hybridize to complementary base. Each cluster on a slide can incorporate a different base.



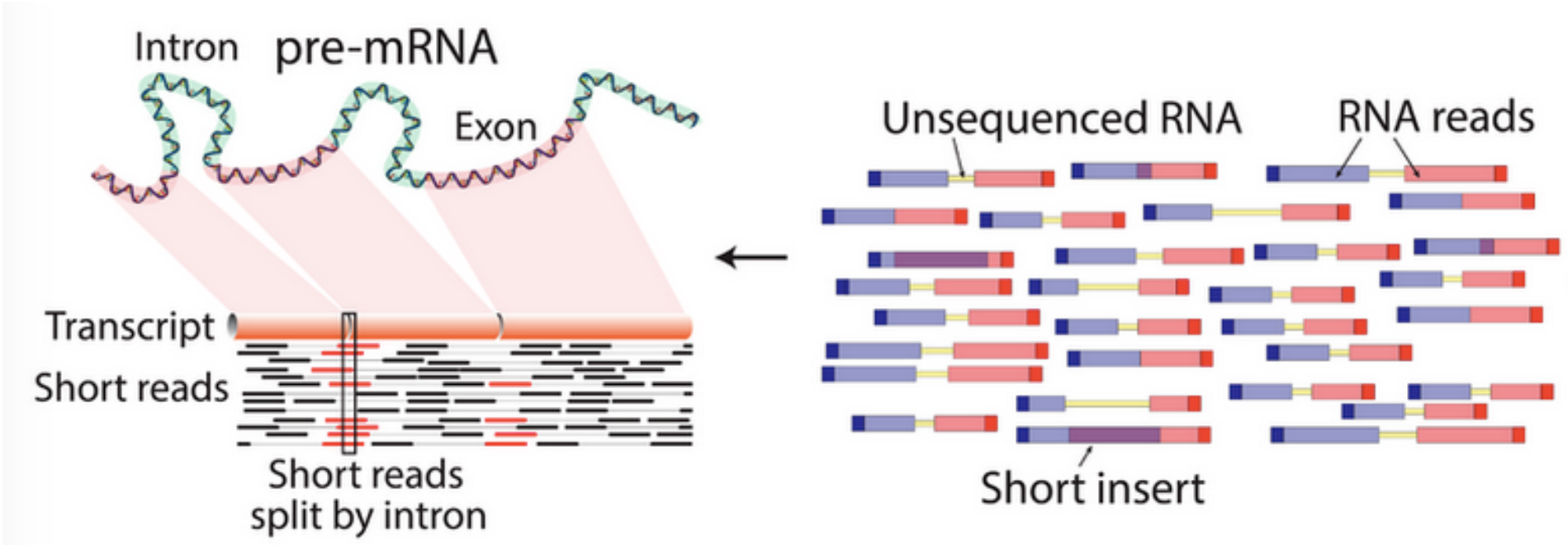
Imaging
Slides are imaged with either two or four laser channels. Each cluster emits a colour corresponding to the base incorporated during this cycle.



Cleavage
Fluorophores are cleaved and washed from flow cells and the 3′-OH group is regenerated. A new cycle begins with the addition of new nucleotides.

Finally, map to genome, transcriptome, and predict exon junctions

reference based analysis



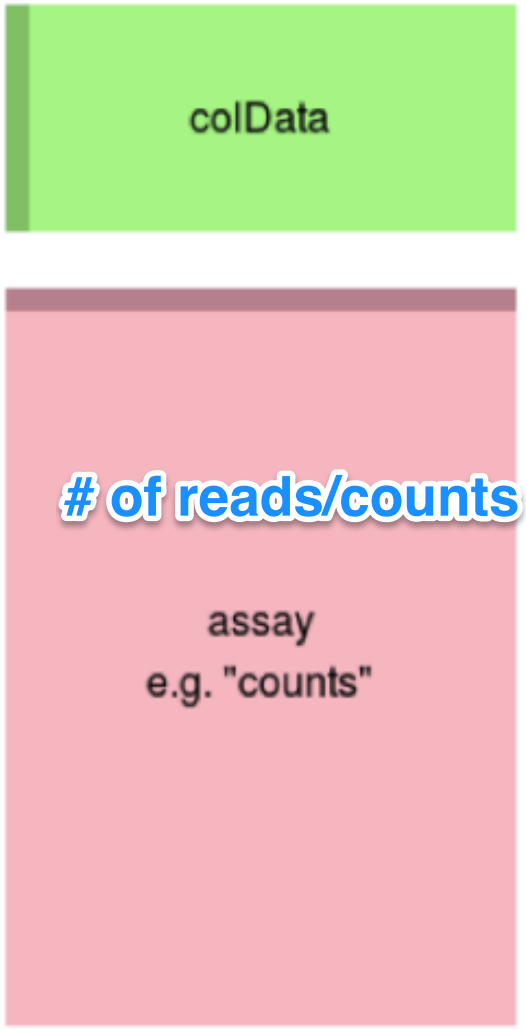
en.wikipedia.org

Bioconductor, genomic alignments package

Data structure

(preprocessed by Amanda)

experimental conditions



samples
(DLD-1 or BRCA2-/-,
+/- etoposide)

genes
(chromosome,
number of exons,
ID, etc)

genomic region

# of reads/counts

From Prof. Samson's lecture 03/09/17:

What experimental question will you ask in Module 2?

How does DNA repair affect the ability of cancer chemotherapy drugs to kill cancer cells?

How does cancer chemotherapy affect gene expression?

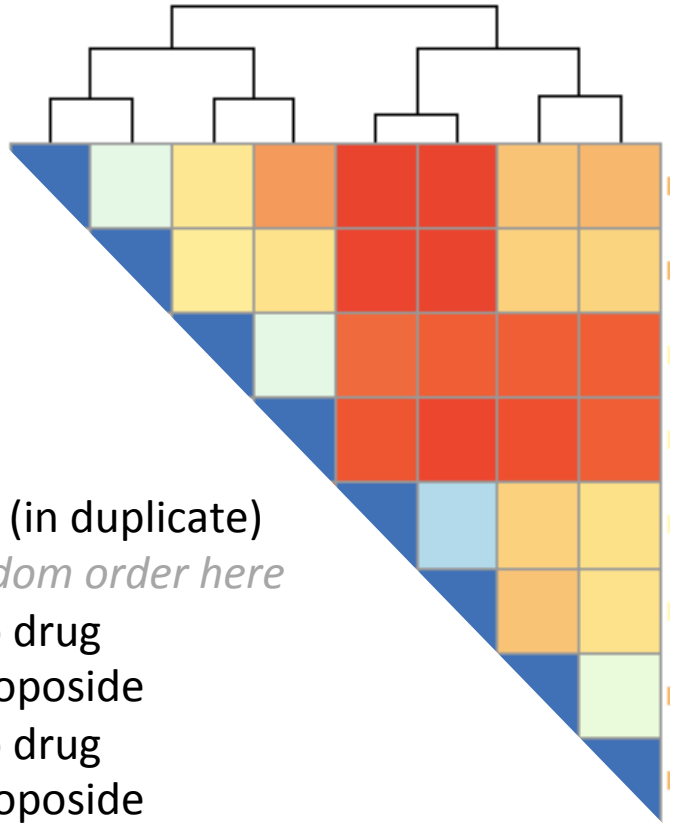
This raises the following questions

- How does DNA get damaged?
- What is DNA repair?
- Why does DNA repair exist?

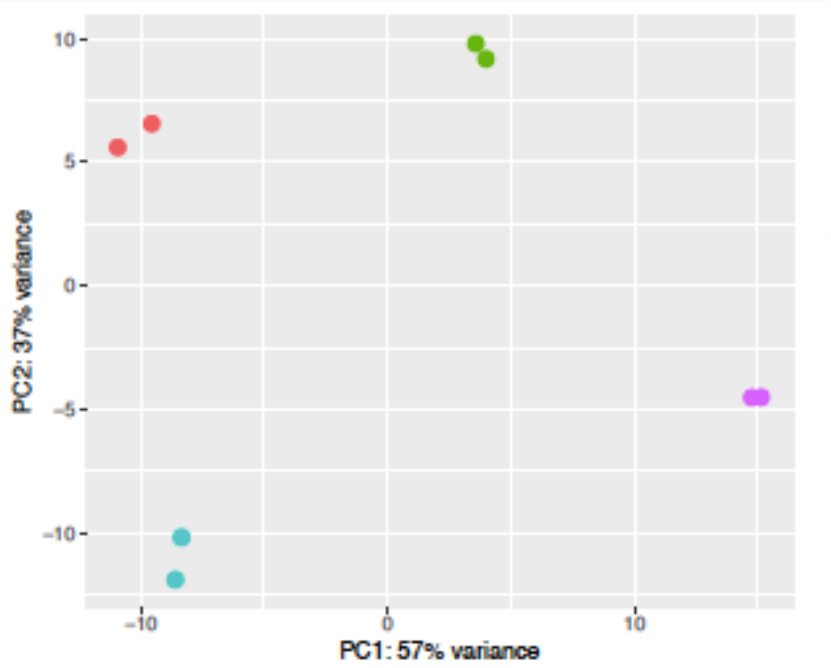
We ask:

Which experimental conditions are overall most similar (least similar)? Clustering:

Hierarchical clustering



Principal component analysis (PCA)



columns are (in duplicate)
listed in random order here

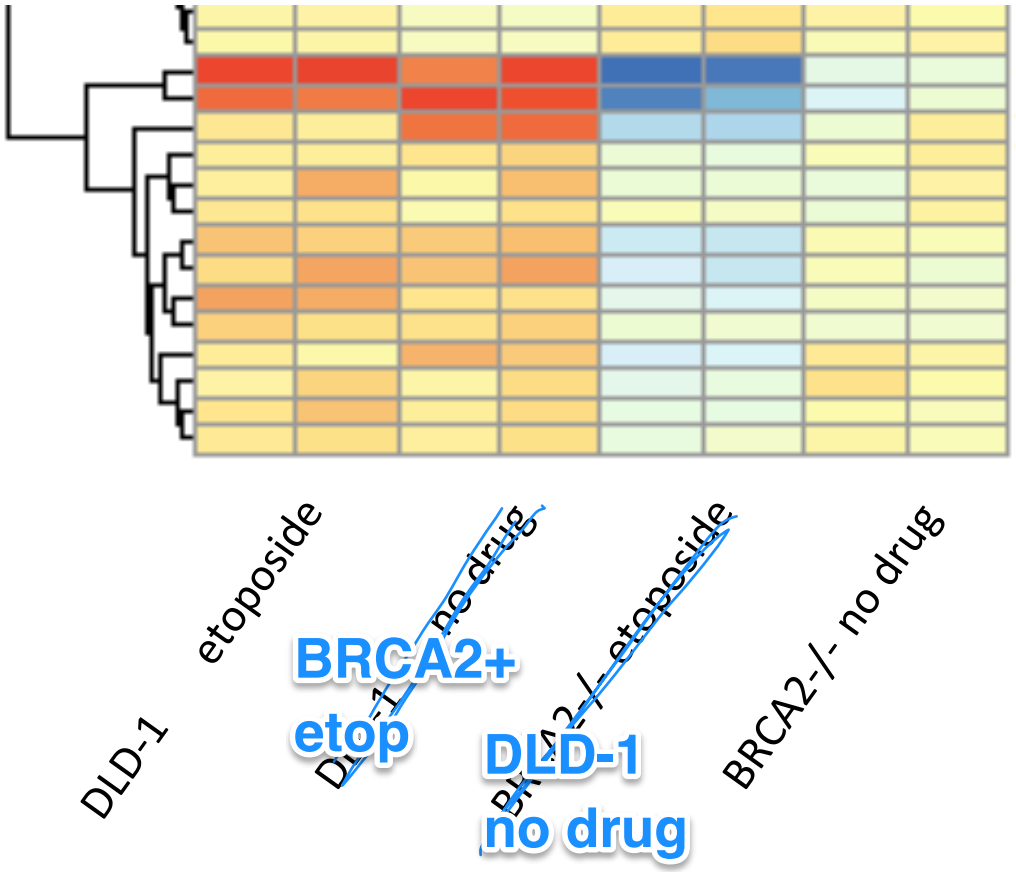
DLD-1 no drug
DLD-1 etoposide
BRCA2-/- no drug
BRCA2-/- etoposide

DNA damage genes

We ask:

Are specific genes differentially expressed?

up or down (or unchanged) compared to average of entire data set



command + return saves so much time!

```
p53_targets = c('CDKN1A','BTG2','FBXW7','GADD45A','SFN','GTSE1','ZNF385A',
                'PCBP4','GPX1','GPX2','SESN2','ALDH4A1','SOD2','CFLAR',
                'PTGS2','CCNG1','DDR1','HBEGF','PPM1D','MYO6','TNFRSF10D',
                'TNFRSF10B','APAF1','BAX','FAS','PMAIP1','PERP','TP53AIP1',
                'TP53I3','BBC3','SIVA1','PTP4A3','PML','PTPRVP','PIDD1',
                'DDB2','ERCC5','FANCC','XRCC5','MGMT','MLH1','MSH2',
                'RRM2B','POLK','XPC')
mat = assay(rld)[p53_targets, ]
mat = mat - rowMeans(mat)
pheatmap(mat, annotation_col=df)
```

However, note the **erratum**: On page 4,

`pheatmap (sampleDists, labels_row=filenames)`

should be replaced by

`pheatmap(sampleDists, labels_row=rownames(colData(dds)))`

on M2D3: you treated cells to examine viability

- etoposide:

creates double-stranded breaks

- compound 401:

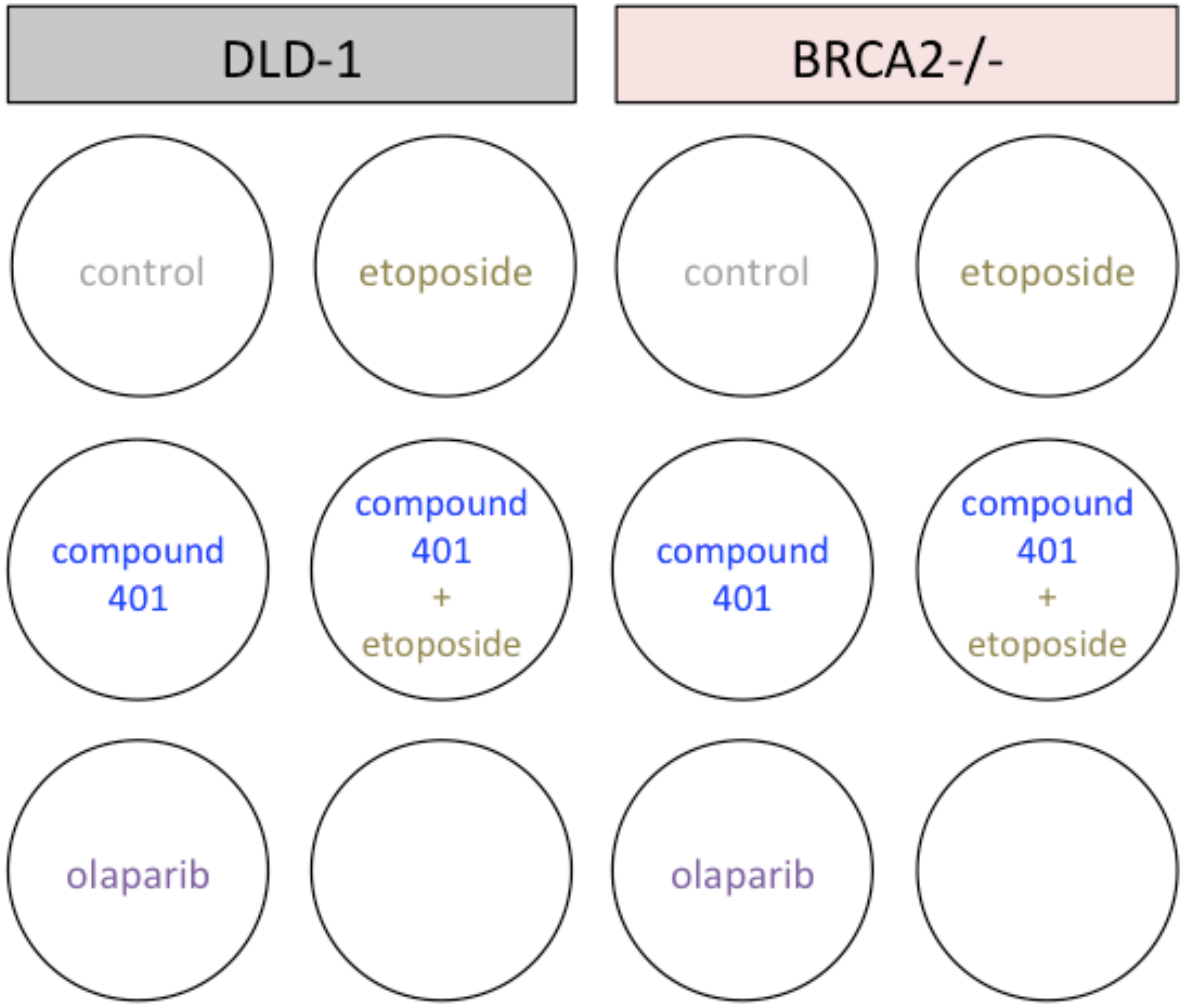
inhibitor of DNA-PK

NHEJ

- olaparib:

inhibitor of PARP

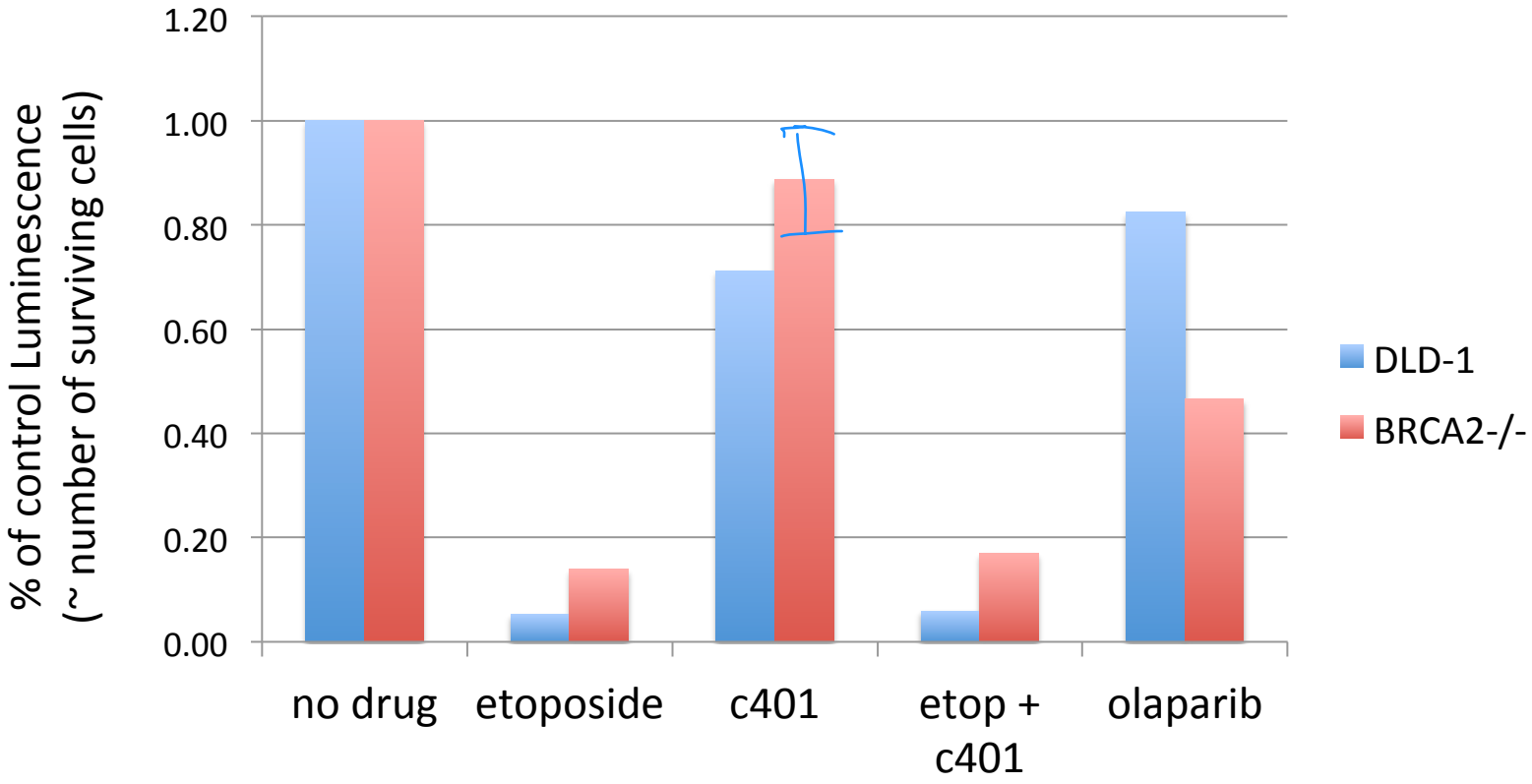
BER



If you're done early,

analyze your CellTiter Glo cell viability data

http://engineerbiology.org/wiki/Talk:20.109(S17):Module_2



Today in lab:

1. Get lab computer and analyze RNA seq data set
2. Make sure to take notes and save images where noted.
3. You should have enough time in lab today to analyze Cell titer glo data and start your Mod2 report figures!

- No homework due Wednesday! Journal Club 2
- Don't forget to Journal Club blog (April 15)
- HW for M2D9 (Friday) could take time, start now!