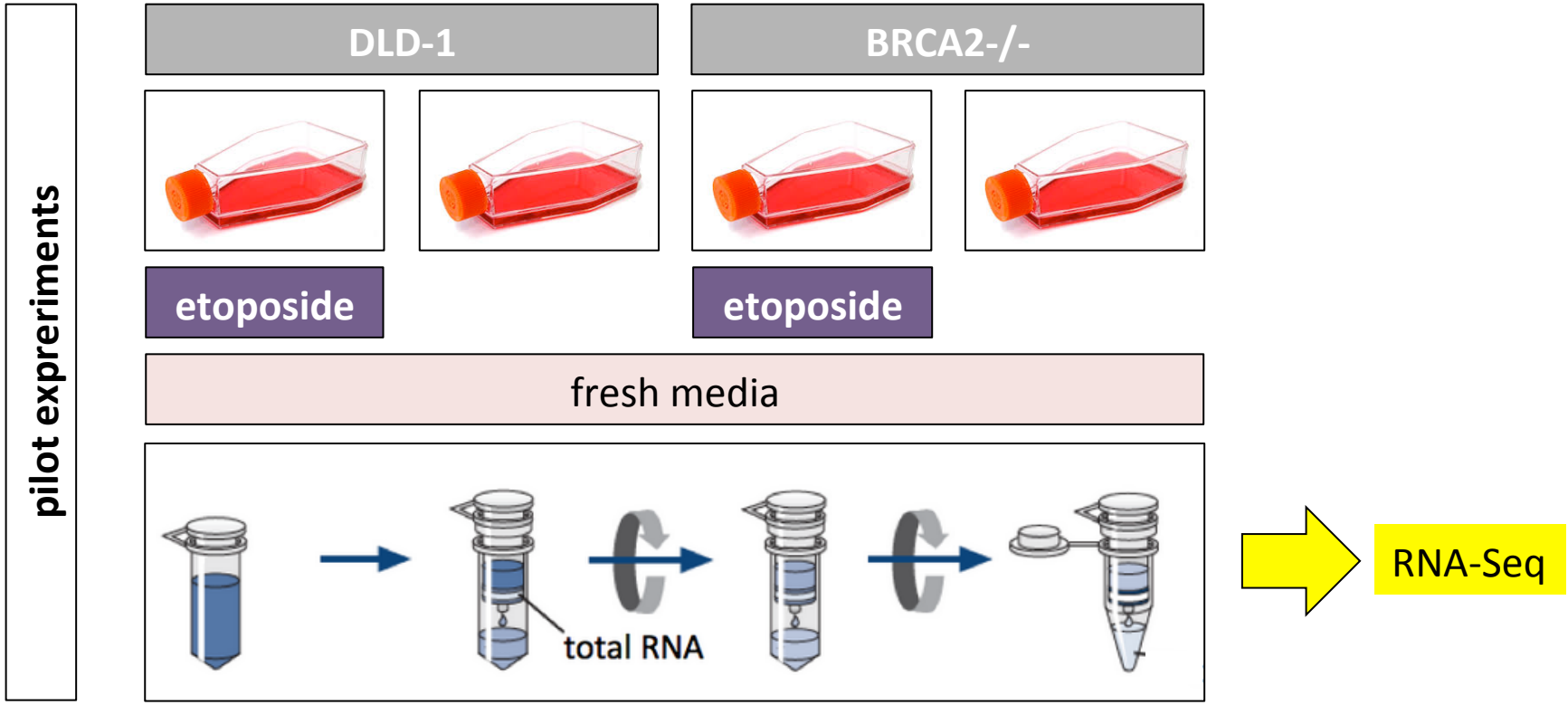
M2D7:
Analyze RNA-Seq data

04/06/2017

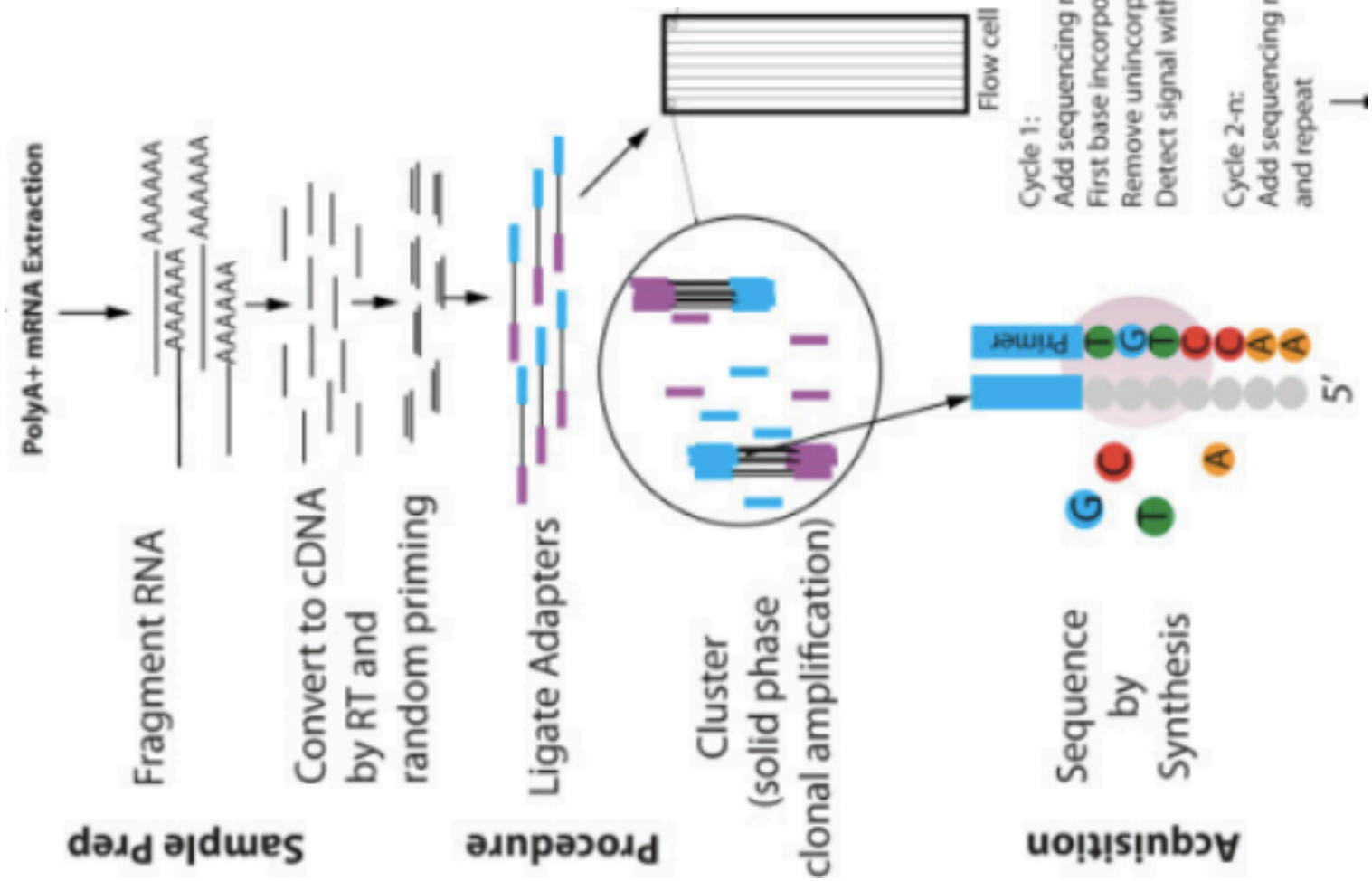
Let's analyze RNA-Seq data

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



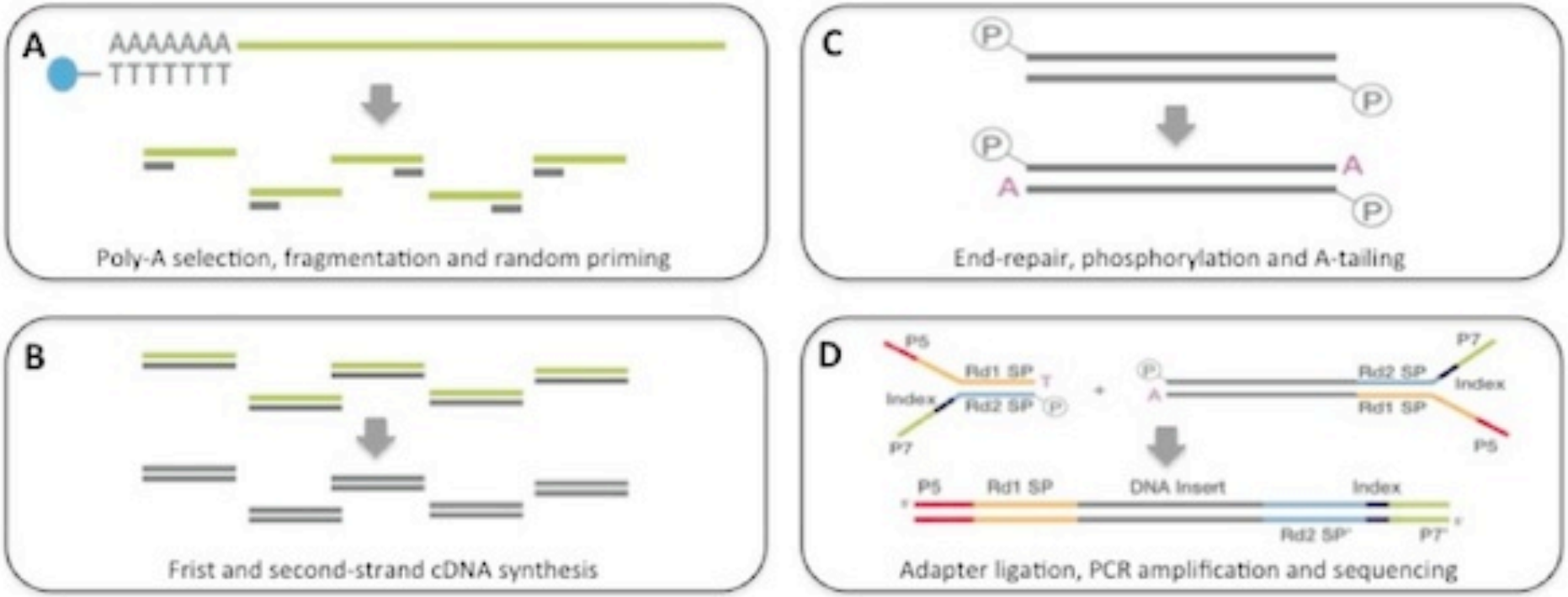
Next generation sequencing (NGS) by Illumina for RNA-Seq

at MIT's BioMicro Center



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

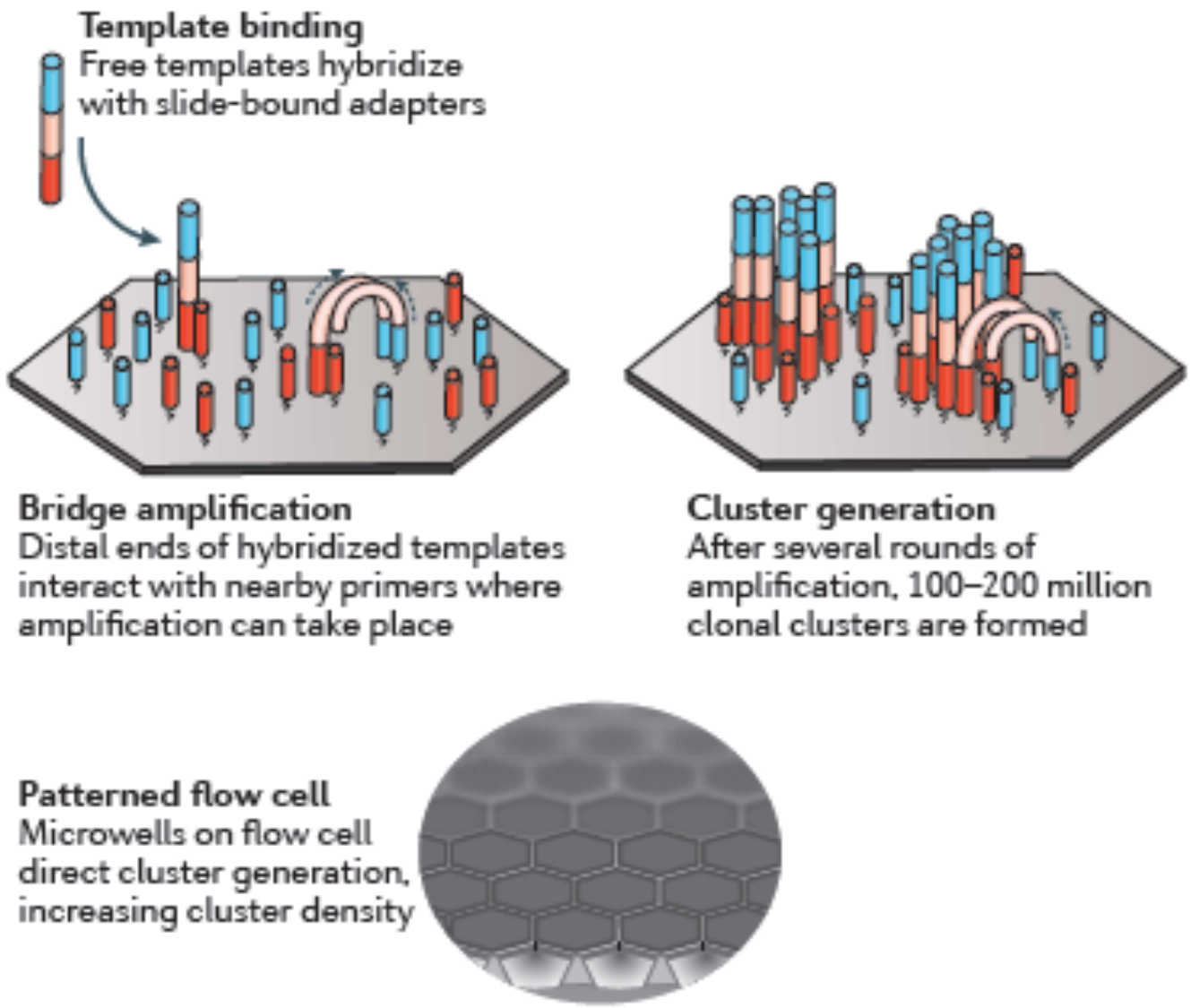
Illumina Tru-Seq RNA-seq protocol



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.

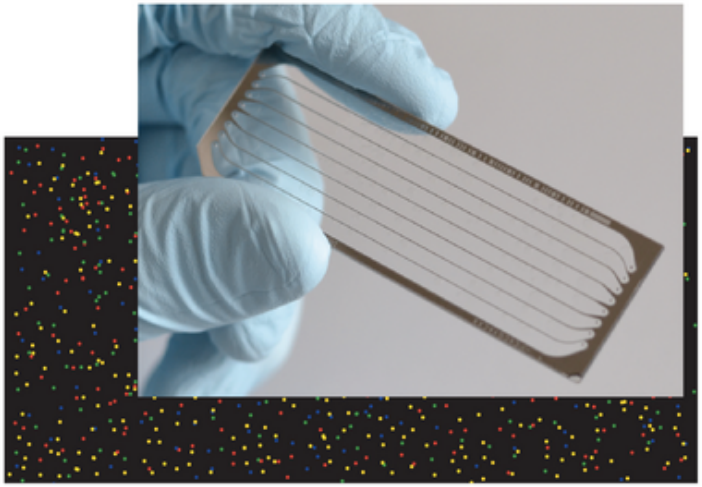
bitesizebio.com

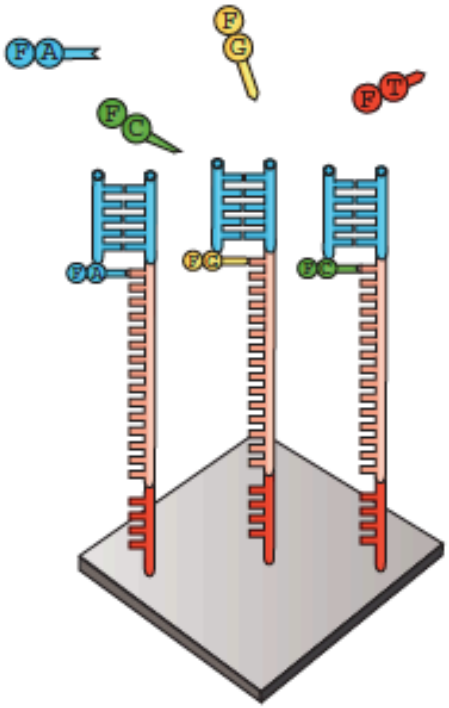
Solid-phase bridge amplification for clonal amplification of cDNA



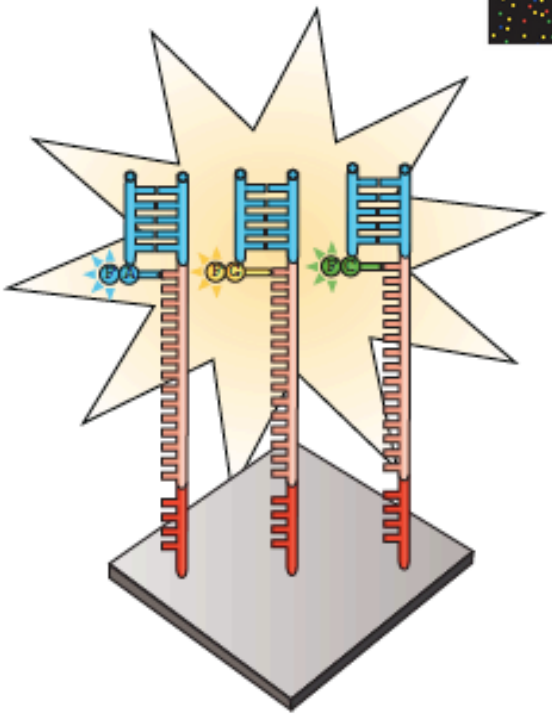
Review: http://www.nature.com/nrg/journal/v17/n6/full/nrg.2016.49.html

Sequencing by synthesis (SBS)

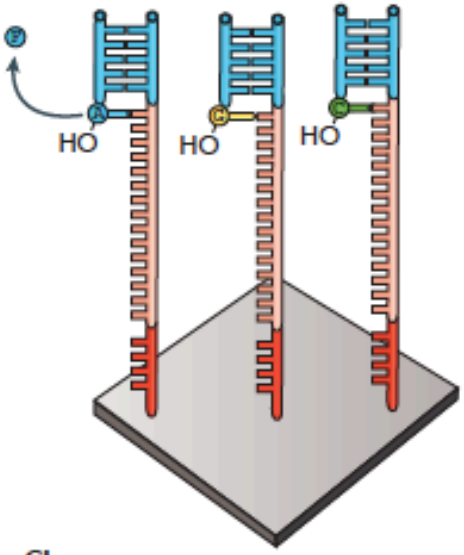




Nucleotide addition
Fluorophore-labelled, terminally blocked nucleotides hybridize to complementary base. Each cluster on a slide can incorporate a different base.

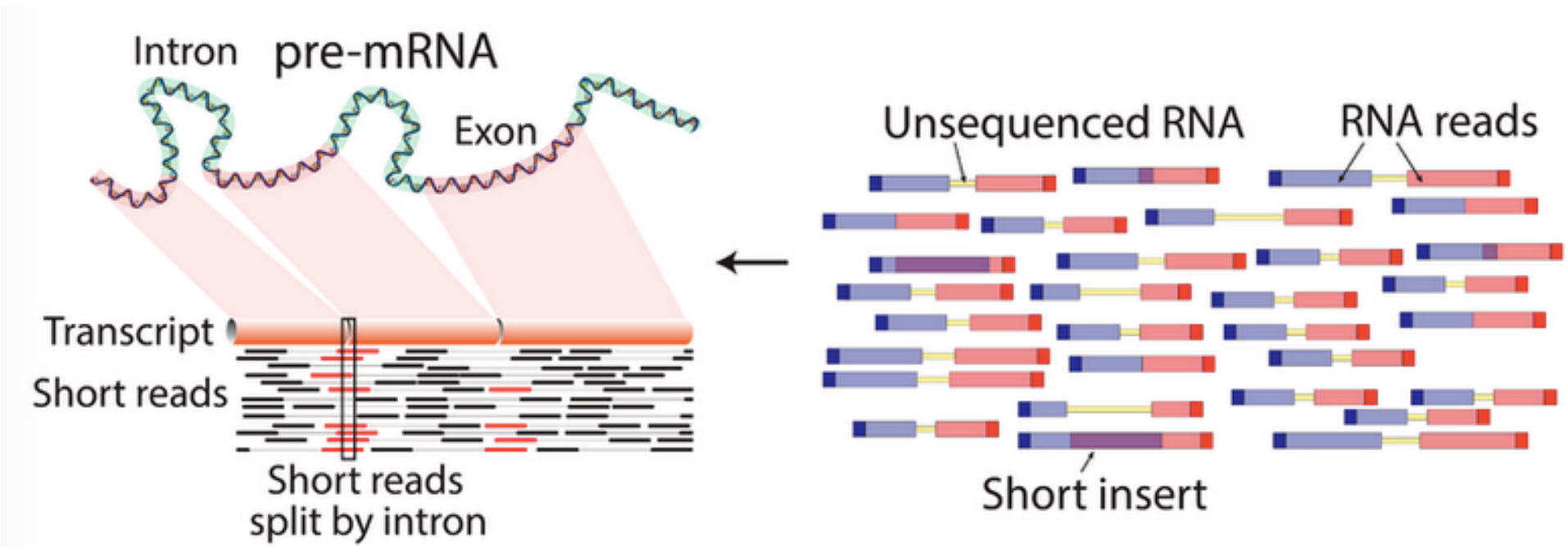


Imaging
Slides are imaged with either two or four laser channels. Each cluster emits a colour corresponding to the base incorporated during this cycle.



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.

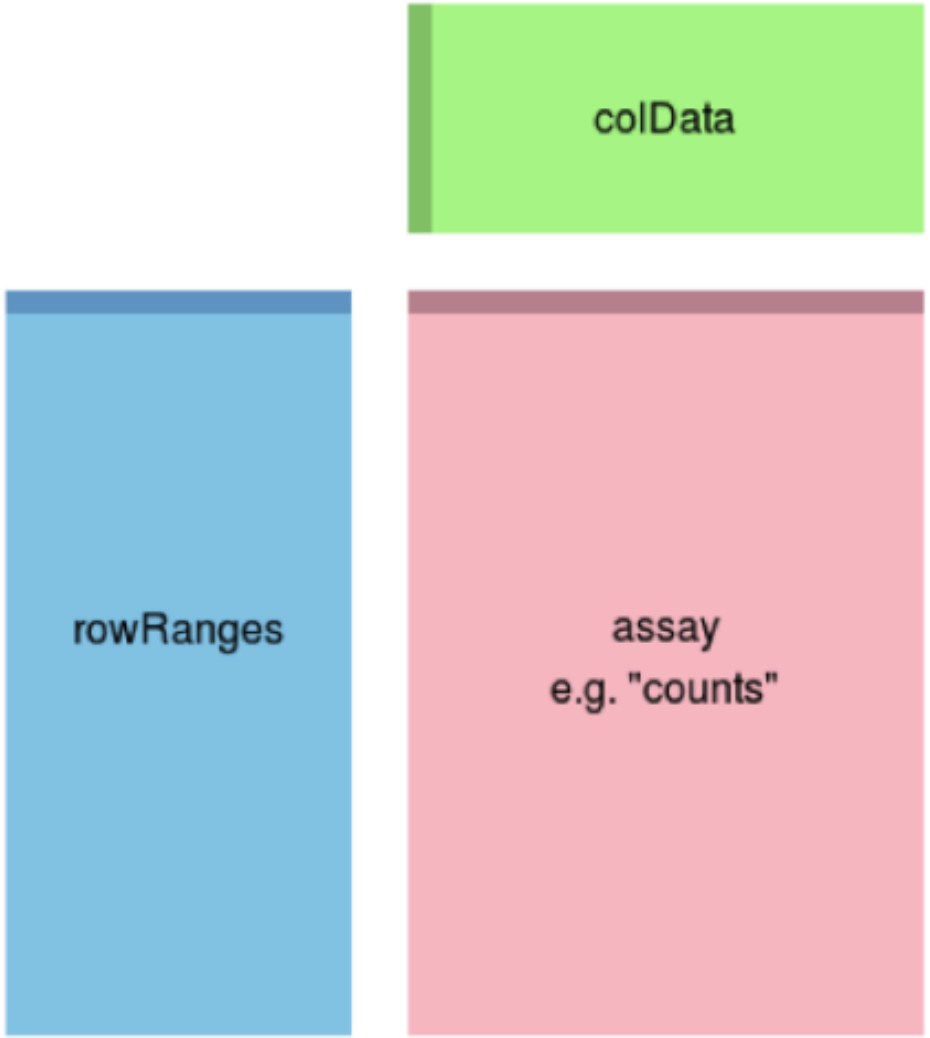
Finally, map to genome, transcriptome, and predict exon junctions



en.wikipedia.org

Data structure

(preprocessed by Amanda)



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

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

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?

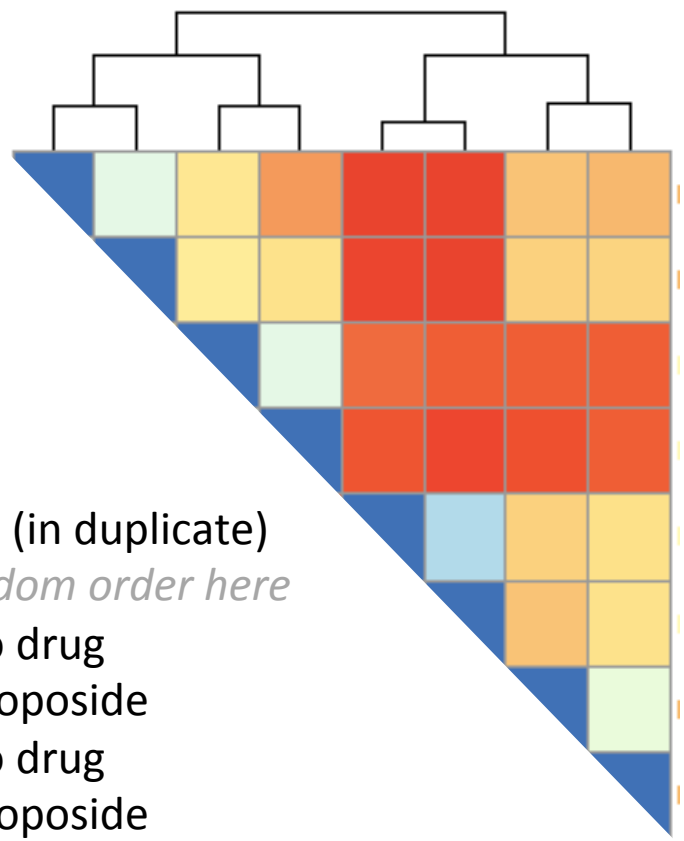
This raises the following questions

- 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



columns are (in duplicate)

listed in random order here

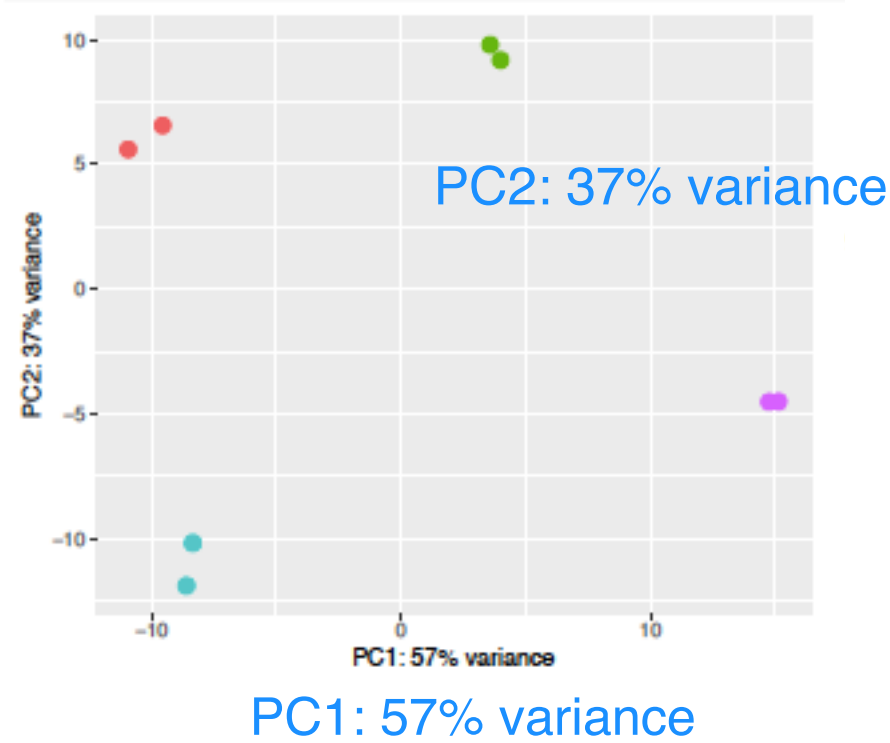
DLD-1 no drug

DLD-1 etoposide

BRCA2-/- no drug

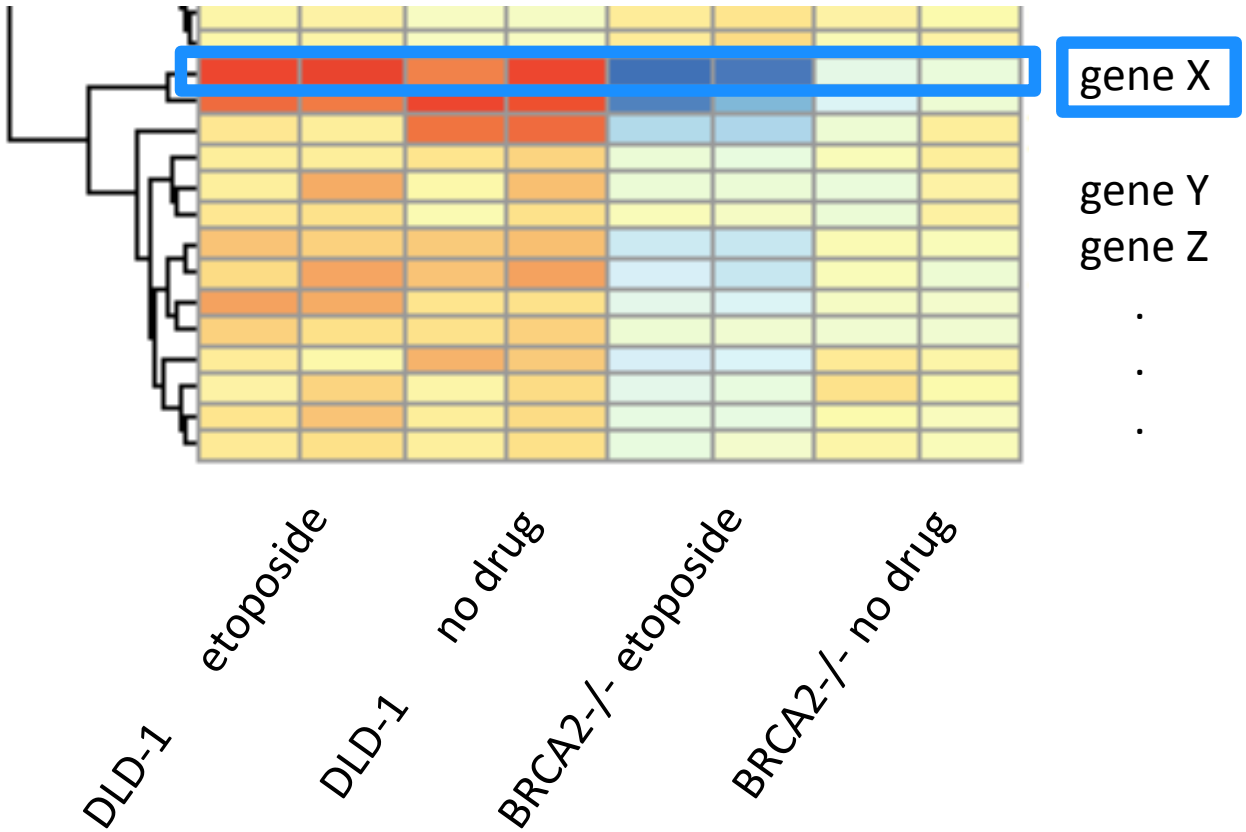
BRCA2-/- etoposide

Principal component analysis (PCA)



We ask:

Are specific genes differentially expressed?



in the top left editor window,

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)`

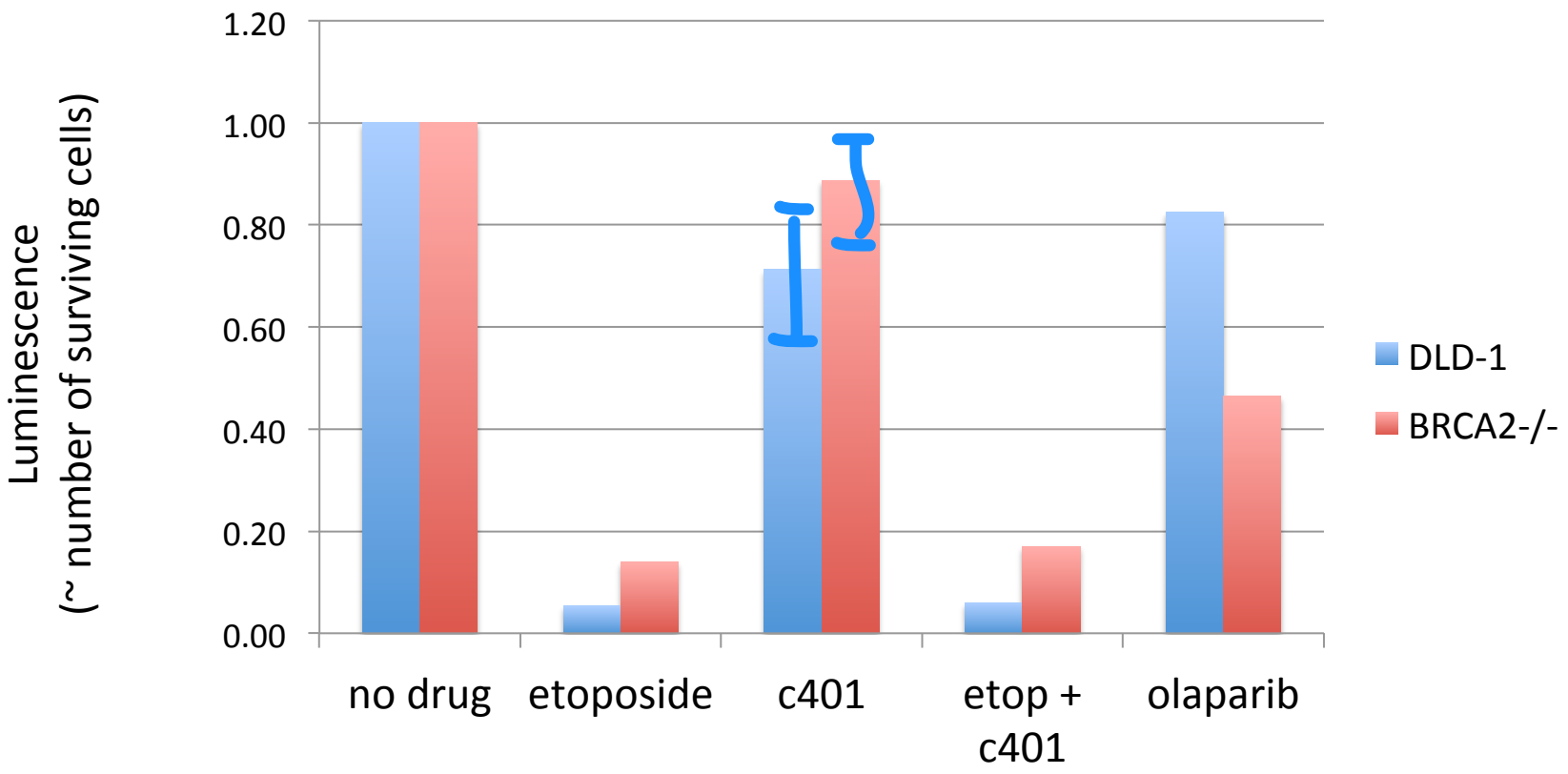
should be replaced by

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

If you're done early,

analyze your CellTiter Glo cell viability data

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



M2D3: you treated cells to examine viability

- etoposide:

creates double-stranded breaks

- compound 401:

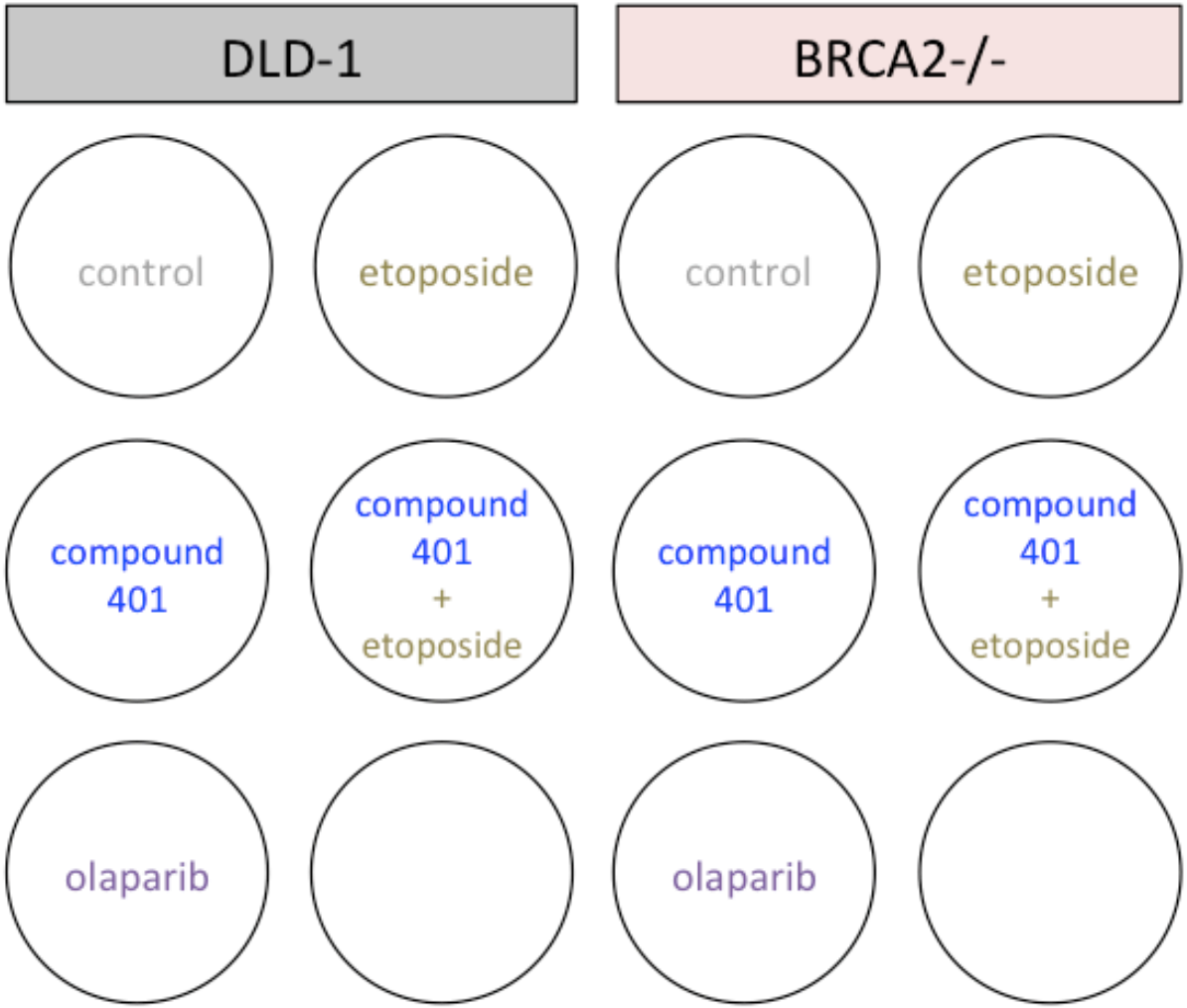
inhibitor of DNA-PK

NHEJ

- olaparib:

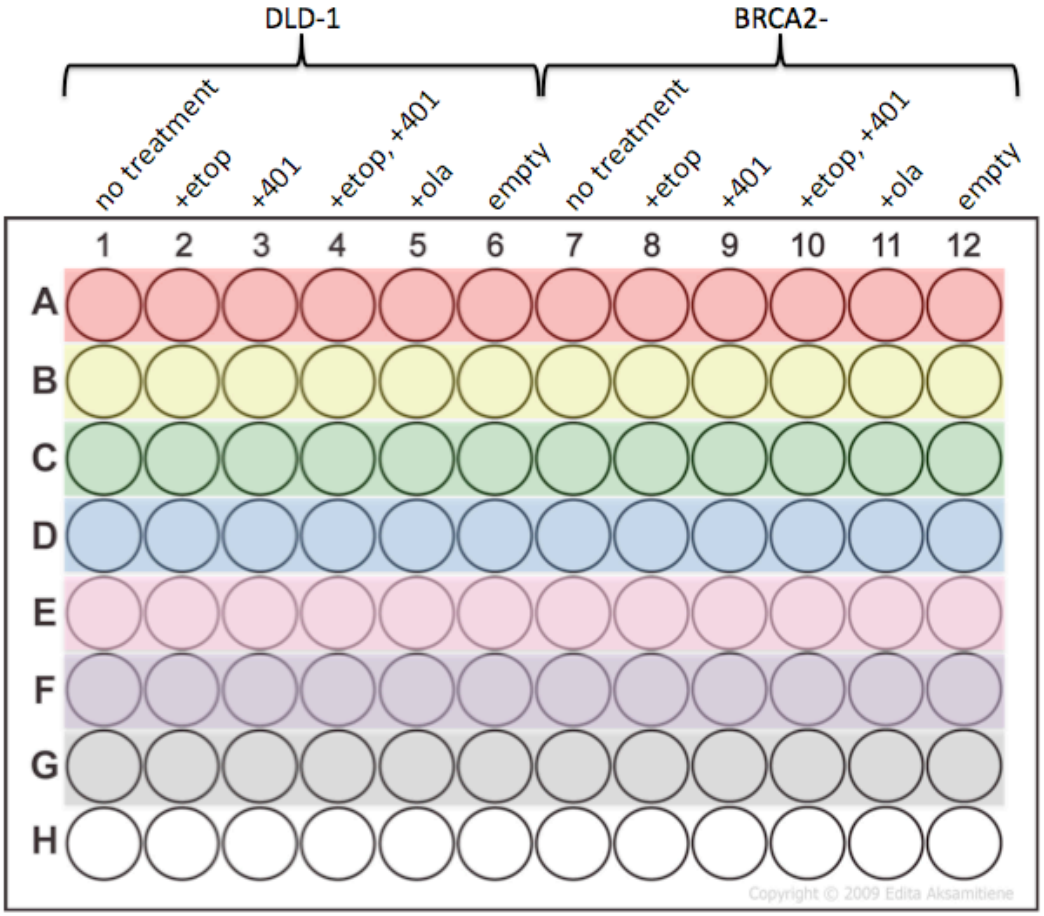
inhibitor of PARP

BER



On M2D4,

- In 12-well plate,
  - fresh media
  - CellTiter Glo reagent
  - shake & incubate
- Transfer to 96-well plate row
- Read luminescence at BioMicro Center



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			DLD-1						BRCA2-/-					
2			no drug	etoposide	c401	etop + c401	olaparib	no cells	no drug	etoposide	c401	etop + c401	olaparib	no cells
3	raw luminescence data		1	2	3	4	5	6	7	8	9	10	11	12
4	red	A	49654.81	1937.50	39366.52	2743.60	29555.24	314.92	38575.78	6354.88	41477.49	12075.84	18331.80	281.03
5	yellow	B	9551.49	420.72	10027.70	536.57	14732.55	308.67	20807.56	1052.96	28315.97	1016.54	15635.14	274.92
6	green	C	43007.98	2113.78	35111.33	4055.97	37168.98	356.34	31815.51	9306.73	34109.37	9106.13	17685.88	295.62
7	blue	D	46397.64	6578.69	45830.47	5609.64	46960.78	366.61	46395.73	10516.19	45732.32	10678.46	24473.21	287.94
8	pink	E	28199.15	2782.81	13052.43	2210.26	30470.25	361.31	39795.76	5506.20	38090.64	7337.06	22774.35	295.73
9	purple	F	24099.03	1499.50	6277.39	1615.06	19474.94	334.56	40117.90	4990.21	14923.51	5837.19	13012.29	294.85
10	white	G	63223.66	938.40	39206.80	1136.77	40042.56	358.81	56134.29	2187.20	40381.04	2050.56	16498.60	275.13

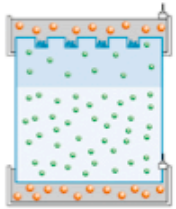
M2: Experimental overview

Confirm cell lines

DAY 1

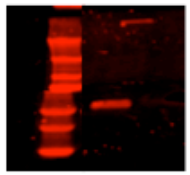
prepare cell lysates

DAY 2



SDS-PAGE and transfer

DAY 3



Western blot analysis

Purify RNA for quantitativ PCR analysis

DAY 1

seed cells

DAY 2

induce DNA damage

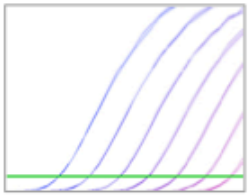
DAY 3

ACAUGGAUUA

TGTACCTAAT

purify RNA and prepare cDNA

DAYS 4 & 9

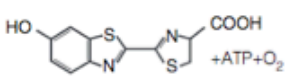


measure transcript levels

Examine effect of drug treatment on cell survival

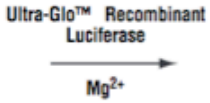
DAY 3

induce DNA damage and treat with drug

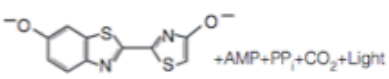


Beetle Luciferin

DAY 4



measure cell survival



Oxyluciferin

Analyze RNA-seq results

DAYS 6 & 7