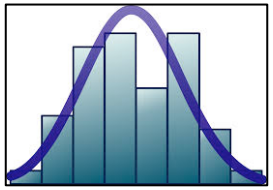
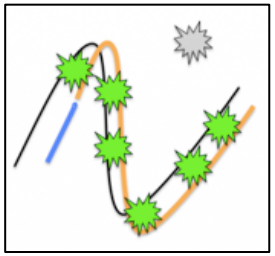
Congratulations on excellent journal club presentations!

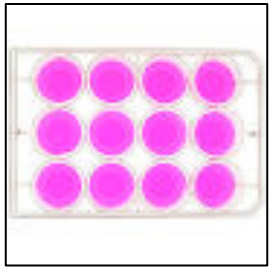
M2D9:
data analysis

04/13/2017

Today in lab







- Quiz
- Practice statistics
- Analyze qPCR data
- Analyze cell viability data

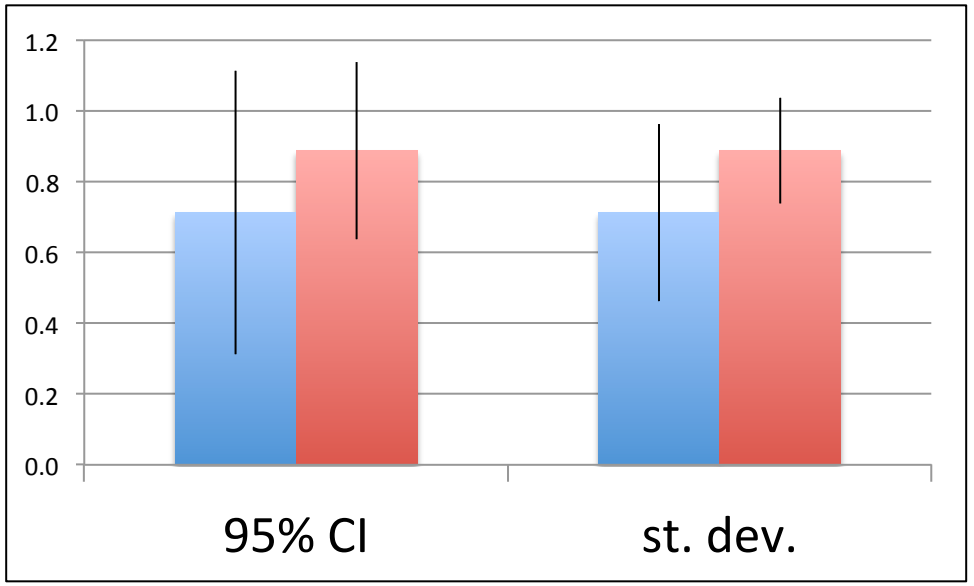
Wrapping up M2!

- Lab notebook:
  - M2**D7** (RNA-Seq data analysis) graded by Rob at 10pm tonight
- Blog post**s**
  - on journal club due 10pm on Saturday, April 15
  - on M2 in general due 10pm on Sunday, April 23
- M2 research article 20%
  - due 10pm on Saturday, April 22
  - extra office hours: (56-302)

Monday	no office hours	Noreen
Tuesday, April 18	9:30-11:30am	Leslie
	3:00-5:00pm	Noreen
Wednesday, April 19	9:30-11:30am	Leslie
	12:00-2:00pm	Maxine
	2:00-5:00pm	Noreen
Thursday, April 20	9:30-11:00pm	Maxine

Plot error bars as 95% confidence intervals (CI)

- 95% CI: the true value of the parameter will be within this interval 95% of the time if the experiment is repeated



- In Excel, find lower and upper bounds:

$$\bar{x} \pm \frac{t_{table} * stdev}{\sqrt{n}}$$

$$t_{table} = TINV(0.05, n-1)$$

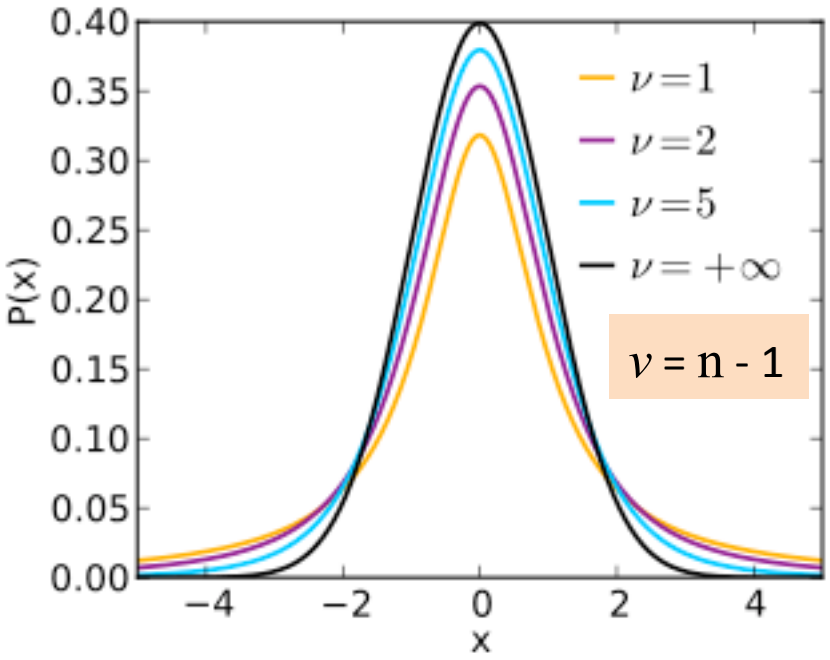
0.05 is the complement of 0.95 (of 95%)

n - 1 is the number of degrees of freedom

Student's *t*-distribution and *t*-test

- "Student"
  - pen name of William Sealy Gosset (Guiness brewery in Dublin !)

- *t*-distribution
  - symmetric and bell-shaped probability distribution, with heavier tails than the normal (Gaussian) distribution
  - to estimate the mean when sample size (n) is small and standard deviation unknown

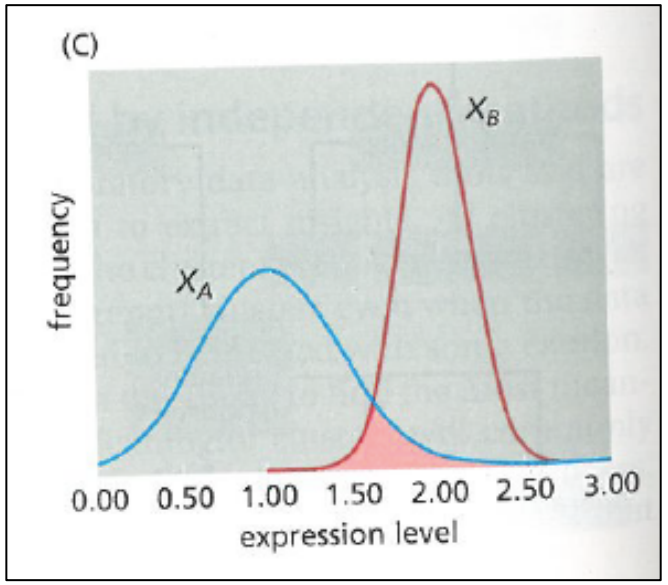


- *t*-test
  - the test statistic follows a *t*-distribution under the null hypothesis
  - used to determine if two sets of data are significantly different

t-test in Excel

- How certain are you that two populations are different?

$$p = TTEST(array1, array2, 2, 3)$$



2-tailed

unequal variance (heteroscedastic ☺)

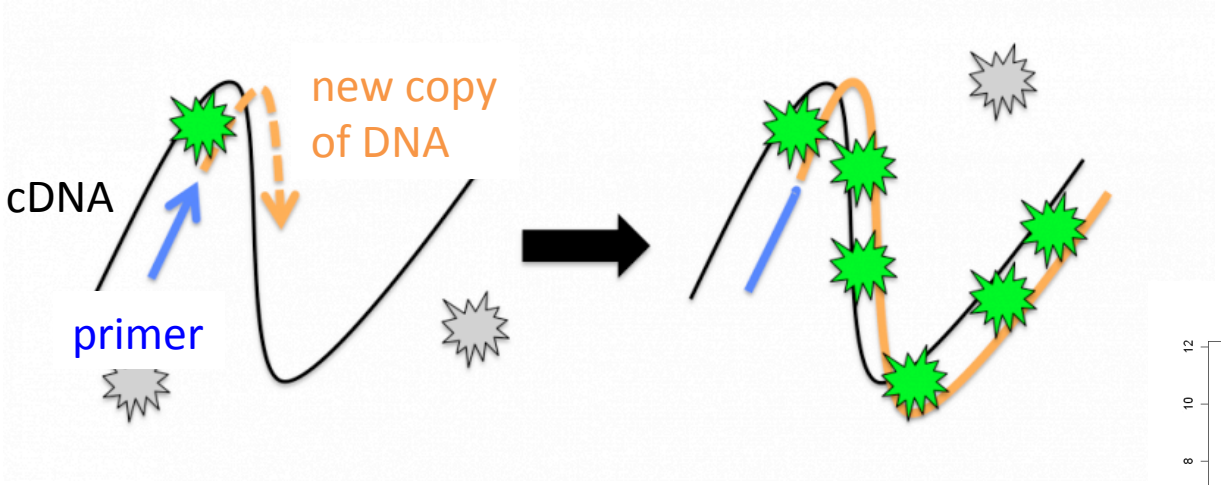
✧ The Student's t-test only applies to two data sets.
Only compare two conditions at a time.

- $p < 0.05$
  - typically considered "statistically significant"
  - the two data sets are different, have a different mean

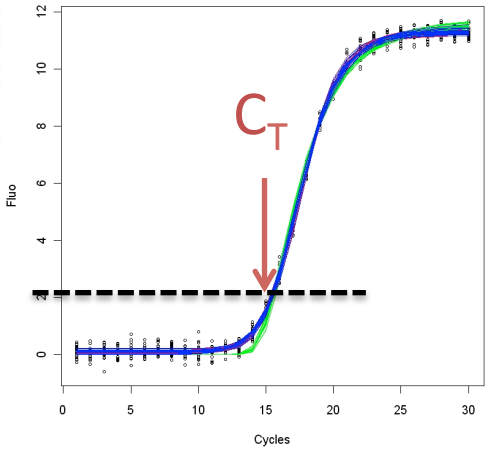
Now that you've practiced…

Analyze quantitative PCR data

- Monitor PCR as it occurs
  - using dye that is fluorescent when DNA is double-stranded
  - signal proportional to initial amount of cDNA (= original RNA)

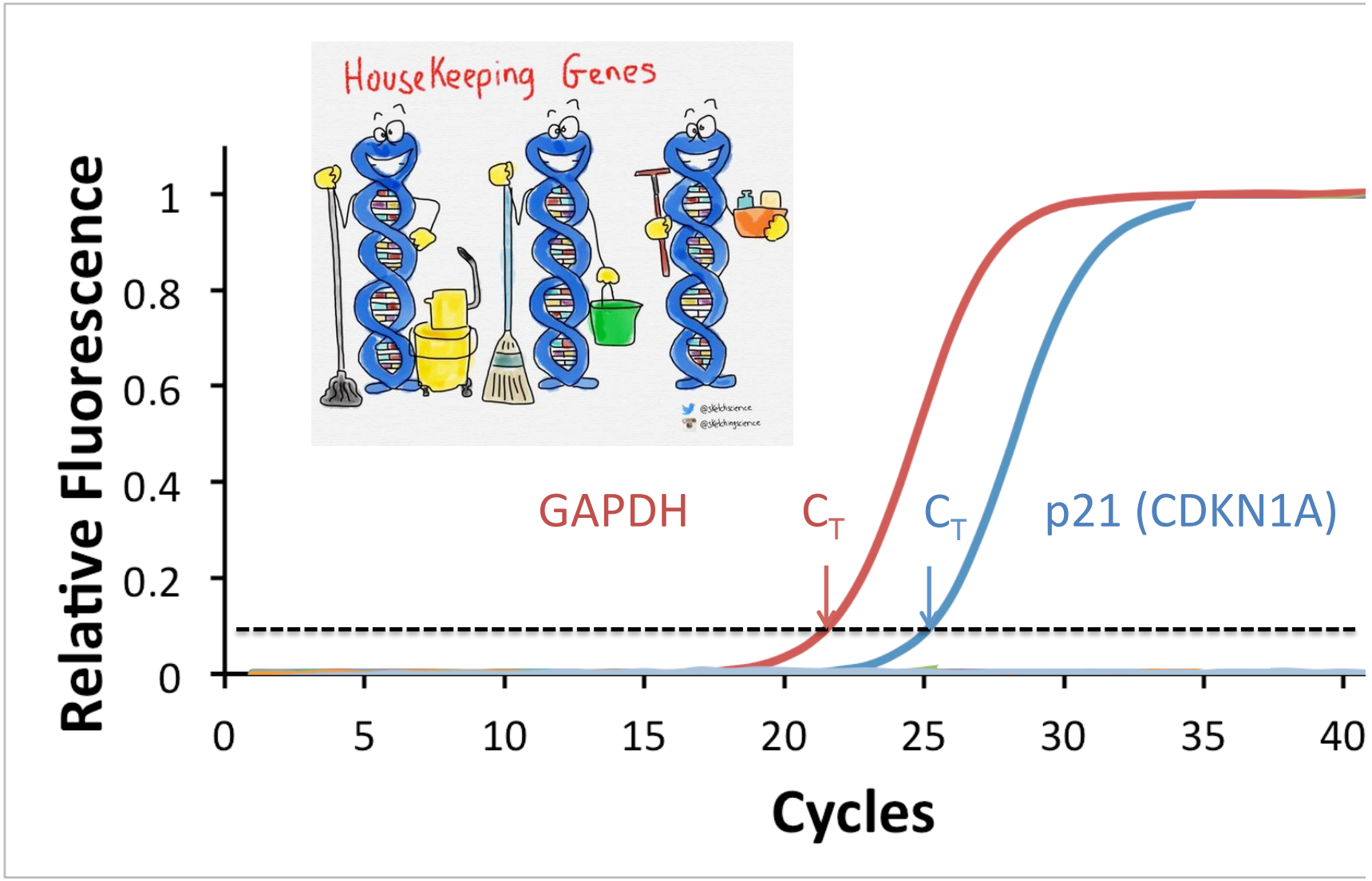


- Plot fluorescence *vs*. cycle number
- Extract C_T threshold cycle



0 40 cycles

C_T value ~ amount of cDNA template present at the start of the amplification reaction



The students' Ct values were ~ 40 (end of qPCR!)
Most likely the RNA had been degraded (by RNases),
the cDNA was of poor quality.

What do we calculate from these C_T values?

$\Delta C_T = C_T$ (p21) – C_T (GAPDH)

	DLD-1	DLD-1 + etoposide	BRCA2-/-	BRCA2-/- + etoposide
1	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$
2	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$
3	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$	$2^{-\Delta Ct}$
average				

- Compare p21 mRNA levels in 4 conditions

Analyze cell viability data

- etoposide:

creates double-stranded breaks

- compound 401:

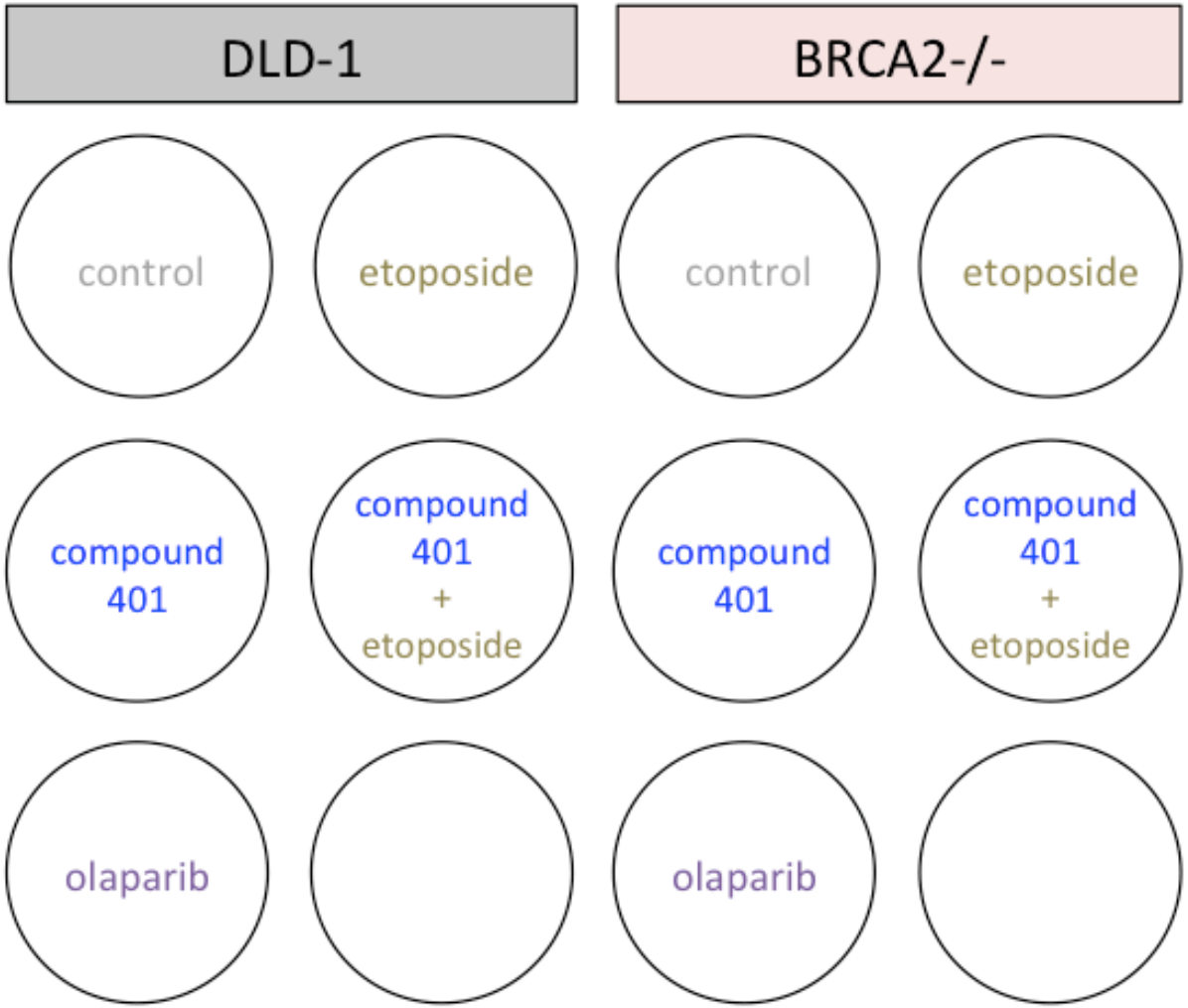
inhibitor of DNA-PK

NHEJ

- olaparib:

inhibitor of PARP

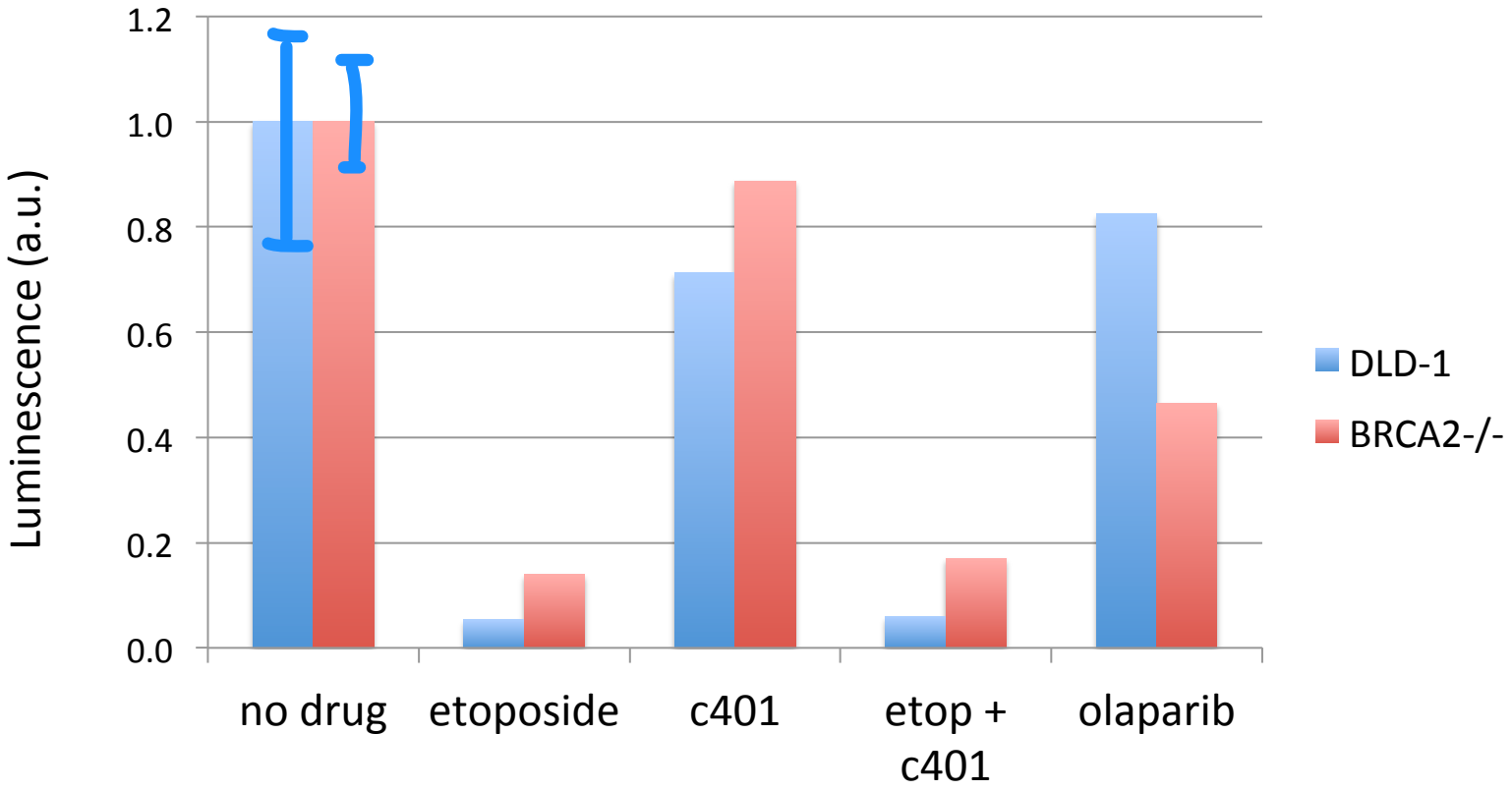
BER



Analyze cell viability data, with error bars

using the class's 13 replicates

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



Today in lab

- Practice statistics
- Analyze qPCR data
- Analyze CellTiter Glo cell viability data

❖ Review Prof. Samson's M2L8 lecture for ideas

- In Discussion / Future Directions, propose 2 years of research
  - to further elucidate DLD-1 / and BRCA2-/- genotype
  - and/or to continue answer "How does DNA repair affect the ability of chemotherapeutic drugs to kill cancer cells?" and "How does chemotherapy affect gene expression?"

M2: Experimental overview