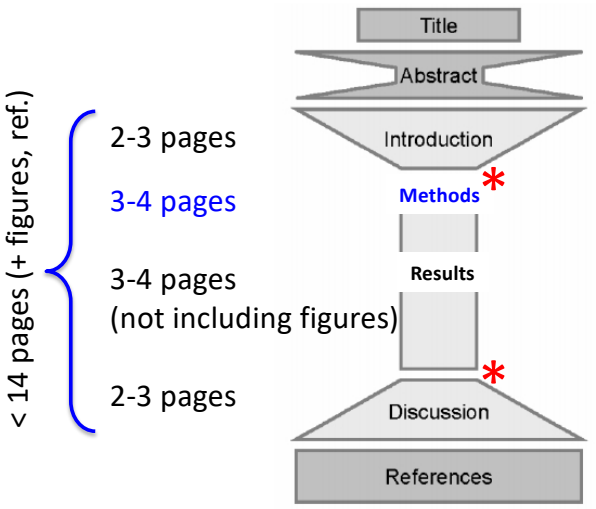
M2D7: Review qPCR experiment and complete statistical analysis

1. Email distribute Quiz, due on Stellar at 10pm
2. Prelab discussion
3. Review qPCR experiment
4. Statistical analysis exercise
5. Continue working on R.studio.cloud Ex3

Mod2 major assignments

- Research Article (20%)
  - individual, submit on Stellar
  - due Monday April 20th at 10pm
  - format: word document, figures can be submitted separately
- Journal Club Presentation (17.5%)
  - presentation **slides** due on Stellar April 11th 10pm
  - format: powerpoint or pdf
  - Presentation **video** due to Dropbox April 11th 10pm (details on wiki)
- Lab quizzes M2D7, M2D9
- Homework and Notebook (10%)
- Blog (5%), 3 posts for full credit
  - 4/6 at 10 pm, 4/13 at 10 pm, 4/21 at 10 pm, 5/12 at 10pm

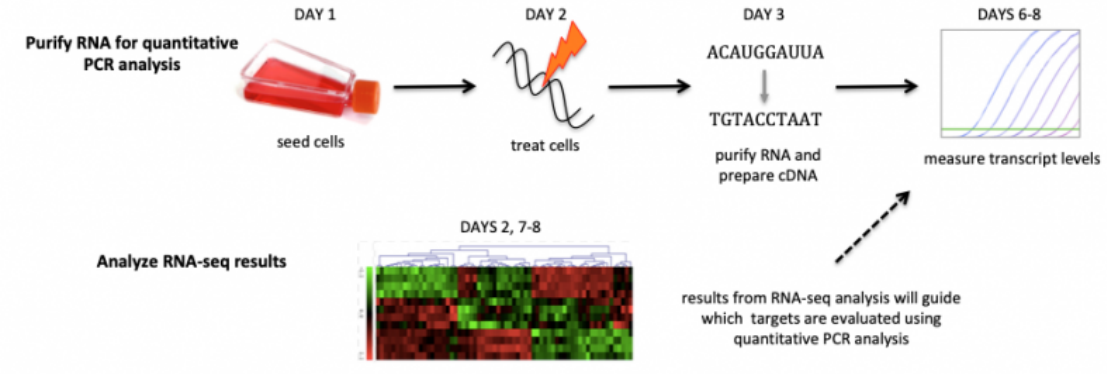
Mod2 Research Report (20% of final grade)



- Title, Abstract (10%)
- Introduction (10%)
- Methods (20%)
- Results, Figures and captions (50%)
- Discussion (10%)
- References

* Last paragraph of introduction and first paragraph of discussion reiterate similar message

Mod2: Experimental overview

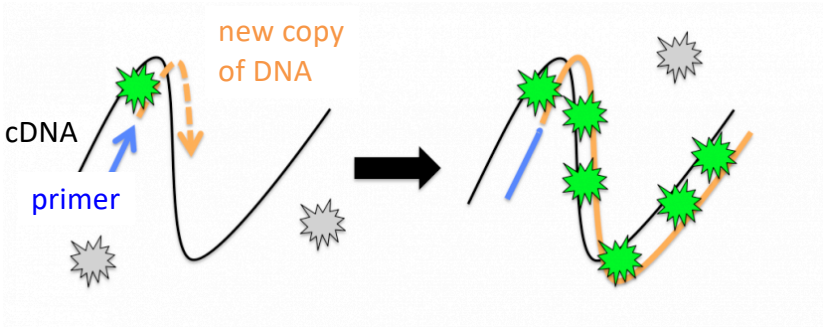


What is our overall question in this module?

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

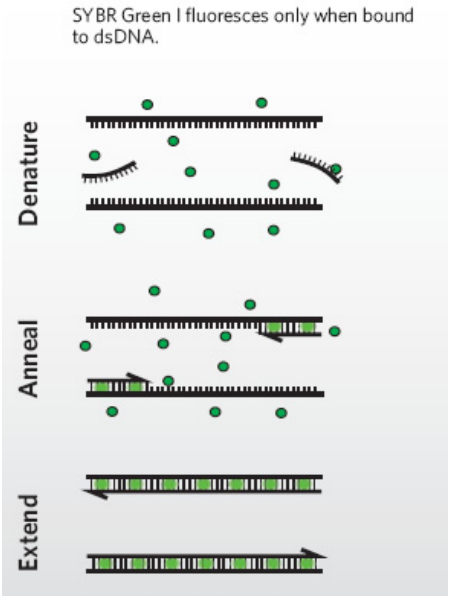
qPCR: quantitative polymerase chain reaction

- Monitor PCR product with fluorescence
  - using dye (SYBR green) that is fluorescent (green below) when bound to double stranded DNA; little or no fluorescence when not bound to dsDNA (grey below)
  - signal proportional to initial amount of cDNA (-> mRNA -> gene expression)



SYBR Green qPCR reagents & cycling conditions

PCR ingredients
SYBR Green
cDNA mix (template)
buffer and water
sequence-specific primers
iTaq DNA polymerase
dNTPs

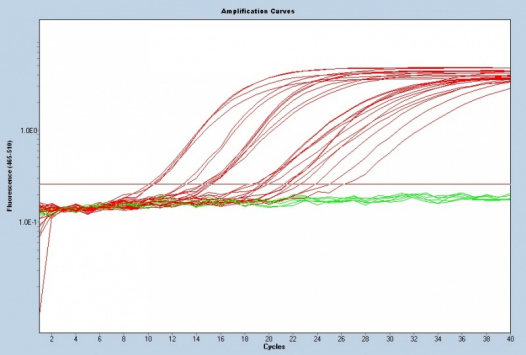


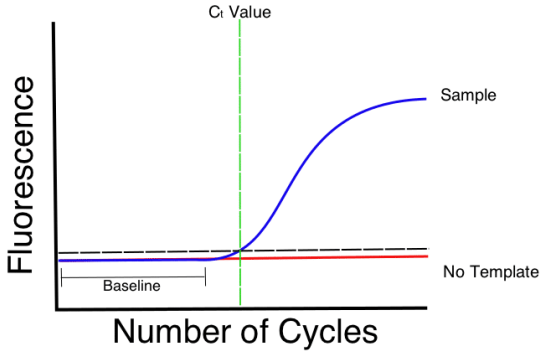
95 °C
(15 sec)

60 °C
(30 sec)

qPCR data output is the threshold cycle (C_T)

- Plotted as Fluorescence vs. cycle number
- C_T threshold cycle
  - fluorescent signal significantly above the background fluorescence
  - relative measure of the initial number of copies of cDNA

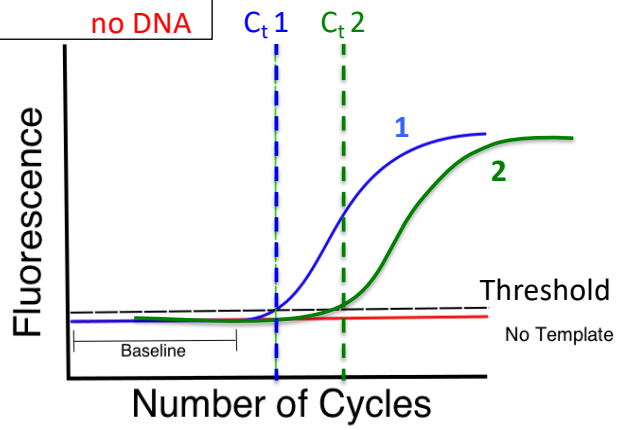




https://bitesizebio.com/24581/what-is-a-ct-value/

C_T related to amount of template present at the start of the amplification reaction

Blue: qPCR rxn 1
Green: qPCR rxn 2
No template: qPCR rxn, no DNA



- Related to baseline fluorescence in the run
- C_T is calculated from qPCR after all cycles complete

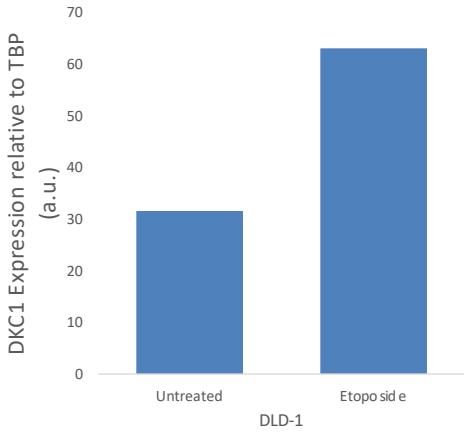
If C_t 1=18 and C_t 2=22 which reaction has a higher signal? --What does a "higher signal" mean?

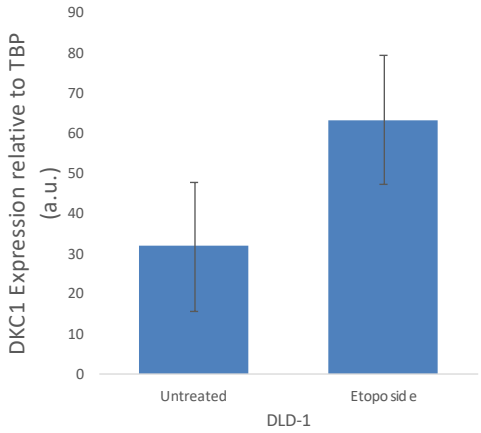
https://bitesizebio.com/24581/what-is-a-ct-value/

Practically-- Calculating Δ C_T for a gene

- The analysis output of a qPCR reaction is a C_T value
  - You can not directly compare C_T values of samples due to variation between qPCR reactions and experiments
- Excel data sheet has 3 C_T values for each gene
  - Represents triplicate wells in the experiment
  - The three tabs in the excel represent 3 experiments carried out on different days
- Must normalize your C_T to a gene you know should not change between your samples and in response to treatment
  - Housekeeping gene
  - Ours is TBP (TATA Binding Protein)
  - Use the TBP C_T values in the experiment you chose to calculate
    - (on the same tab of the excel)
- Finally exponentially transform each normalized value to the ΔCT expression

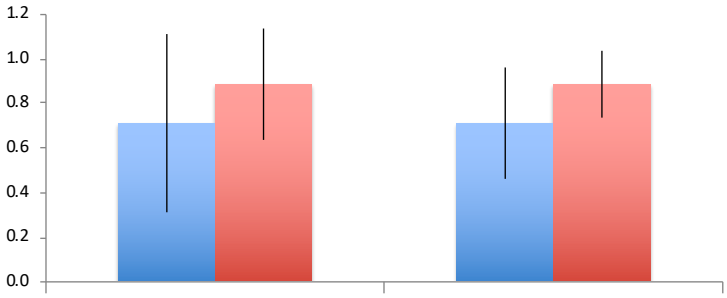
Use statistics to analyze qPCR data





Confidence intervals show the variance in the data set

- Assumes data follow a t-distribution
- At 95% confidence interval, there is a 95% chance that the true mean is within the defined range



Error bars depict: 95% CI Standard Deviation

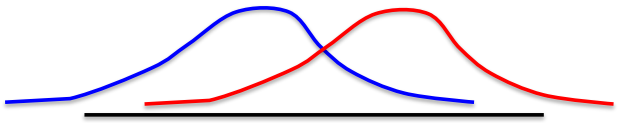
Calculating Confidence interval in excel

= CONFIDENCE(confidence level, standard dev., size)

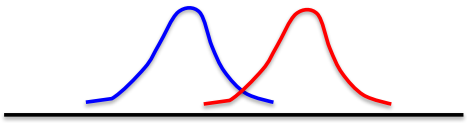
Once you have calculated the confidence interval you will enter this value as your “custom” error bar in excel

Student's *t*-test used to determine if populations are significantly different

- Assume data follows *t*-distribution
- At $p < 0.05$, there is less than a 5% chance that populations are the same (95% chance that populations are different)
- Examines signal (means):noise (variance) ratio



high noise



low noise

Calculating Student's *t* in excel

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

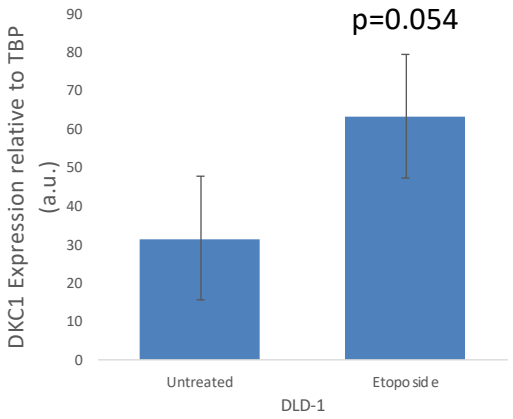
Use the fewest assumptions: two-tailed

unequal variance

Can only compare two data sets at a time

*Make sure it is clear on your plots/writing which conditions are being compared

How will you use statistics in your data analysis?



What do you write if the data are not statistically significant, but almost?

Today in "lab"…

1. Watch qPCR video at the top of the Protocol section and read through Part 1.
2. Calculate the ΔC_T values for the 2-3 genes you would like to investigate further
3. Calculate confidence intervals and p-values with the Student's t-test using the ΔC_T values of the genes you will use for your research article
4. Continue working on R studio cloud Ex3.
   - We will send out code to assist with the refresher exercise today

You are turning in Methods by 10pm tonight– make sure to look over Noreen's comments on your last methods so you don't repeat mistakes!

M2D8 Homework

Peer review methods

- We will email you another student's methods
  - You will be blind to the identify of the other student

- You should comment on the methods similarly to instructor's feedback – Write comments on the document, or write them on a separate page with numbers corresponding to comment

- There are overview questions in the homework prompt you should address at the bottom of your specific comments