

Quantifying gene expression with RT-PCR

- ◆ you are quantifying the cDNA, not the mRNA (RT means reverse transcription, not real time)
- ◆ START WITH GOOD RNA; batch for simultaneous RT

Detection in PCR products in gels

- ◆ run 30-40 cycles and stain gel for presence of amplicon of expected size
not quantitative, ok for “there or not there” experiments
- ◆ stop PCR while amplification still is exponential and measure band intensity
trial and error to determine where to stop
optimal # of cycles varies across genes
- ◆ Add competitive internal standard and measure relative band intensities
competitor uses same primers as target cDNA
competitor amplicon longer or shorter than target amplicon

Real time detection

- ◆ Has become the “gold standard,” but you still have to be careful and understand the pitfalls

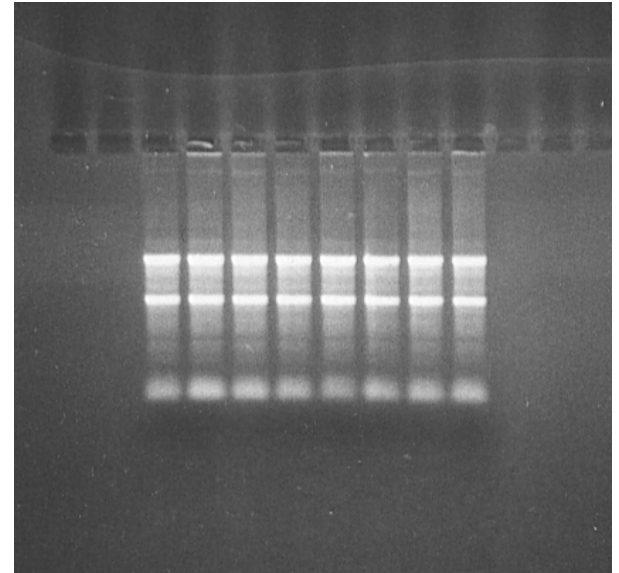
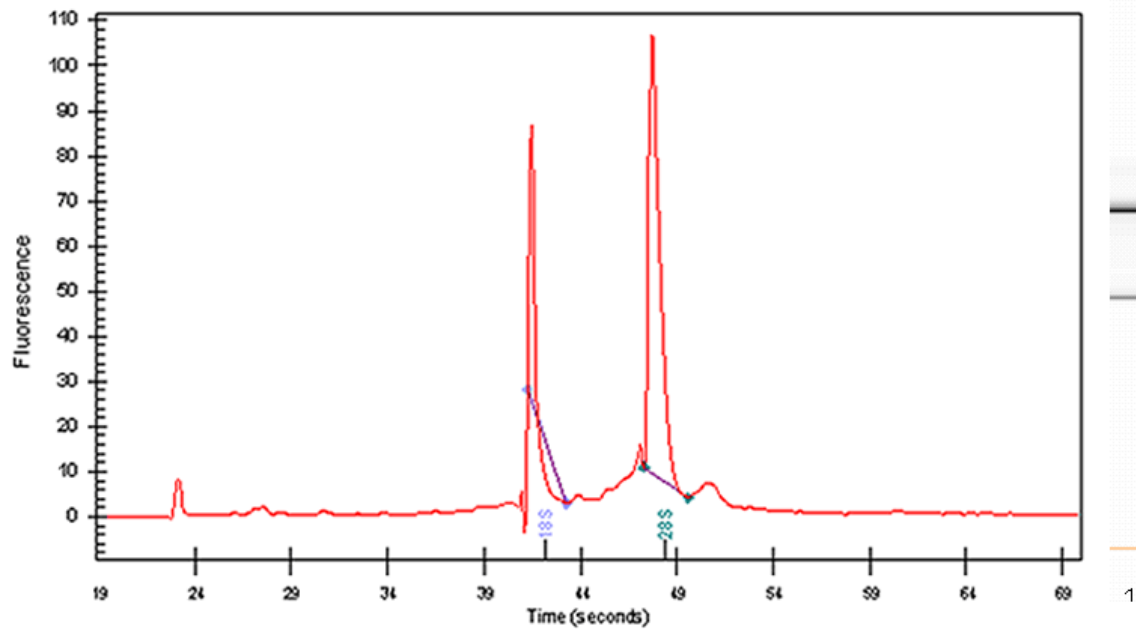
Good RNA

- ribosomal bands stand out in gel (or Bioanalyzer)
- $A_{260}/A_{280} > 1.8$ (lower does not necessarily mean bad RNA, but quantification by A_{260} is not accurate if ratio low)
- partially degraded RNA might be usable [primers must be close to 3' end of mRNA with oligo(dT) priming, or use random priming]
- free of genomic DNA (important for accurate quantification of transcripts expressed at low levels); DNase it or use RT-negative controls; or design PCR primers (or probes) so genomic DNA is not amplified

Consistent cDNA synthesis

- Prepare both experimental and control cDNA samples at the same time with the same lot of RT with the same amount of RNA template using the same kind of primer
- Oligo(dT) and random primers can give you different Ct and Δ Ct values, but $\Delta\Delta$ Ct should be similar if samples are batched as indicated above
- without consistent reverse transcription efficiency, accuracy of method becomes much more dependent on selection of a good reference gene

Universal Mouse Reference Total RNA



Training_Total-RNA-Nano_2002-05-21

Universal Mouse Reference (Stratagene)

Fragment	Name	Start_Time(secs)	End_Time(secs)	Area	% of total Area
1	18S	41.25	43.30	40.39	9.90
2	28S	47.30	49.60	131.85	32.31

RNA Area 408.10

RNA Concentration(ng/ul) 245.37

rRNA Ratio [28S / 18S] 3.26

Ideally, absorbance ratios should be > 1.8 for both 260 nm/280 nm and 260 nm/230 nm. This allows more accurate quantification from A_{260} and makes it more likely that there are no inhibitors of reverse transcription.

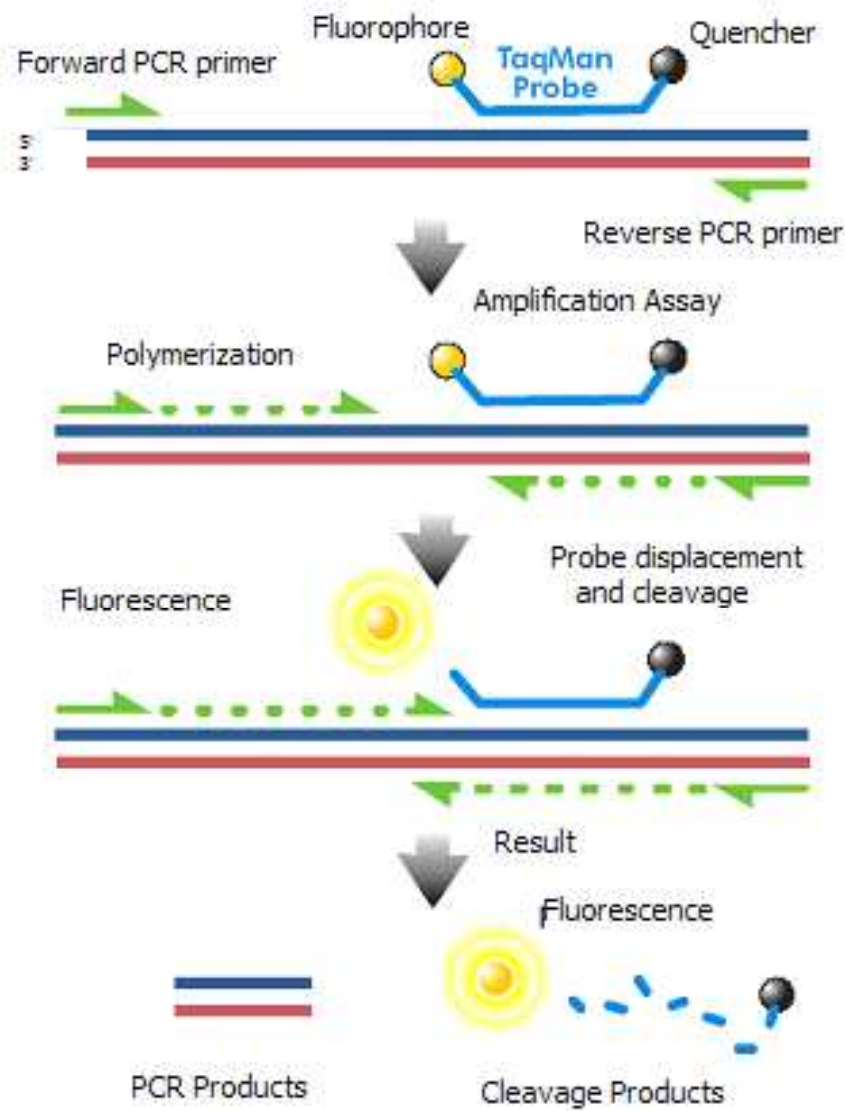
Two basic approaches to real-time PCR

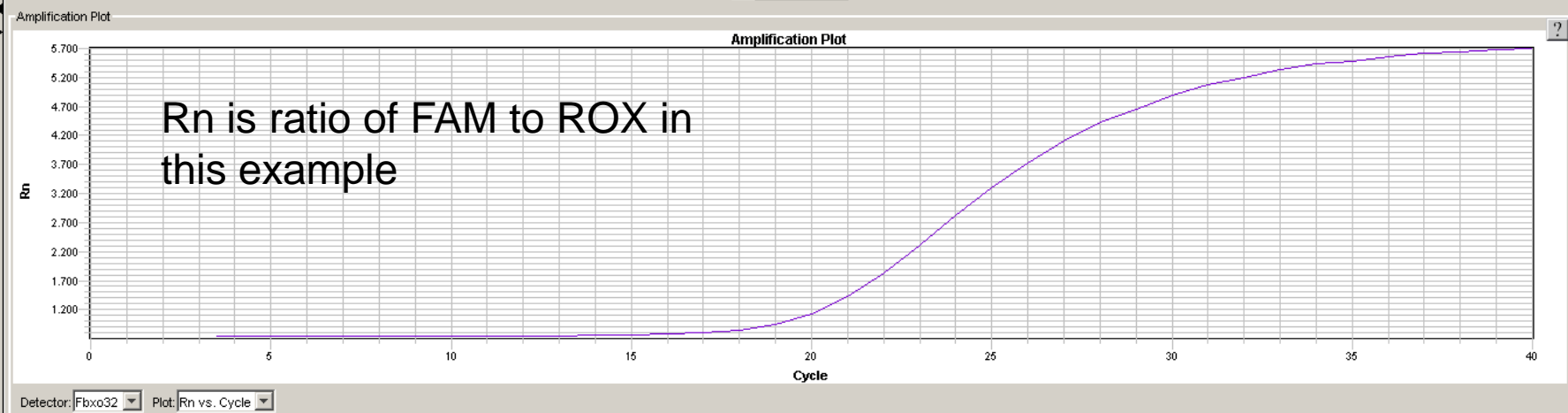
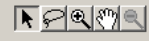
1. sequence-specific fluorescent probes

- adds specificity - nonspecific products should be invisible
- multiplexing and allelic discrimination possible
- probes are expensive
- single reporter molecule per new amplicon molecule produced

2. SYBR Green dye

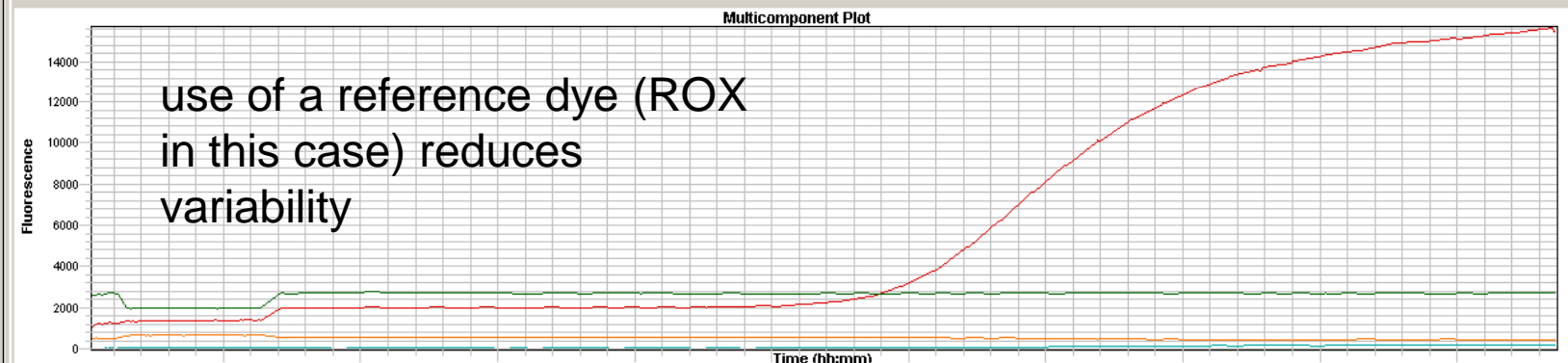
- less expensive than using specific probes
- all nonspecific DNA products detected (melting curve or gel analysis should be done)
- cannot multiplex
- multiple SYBR Green molecules per amplicon molecule



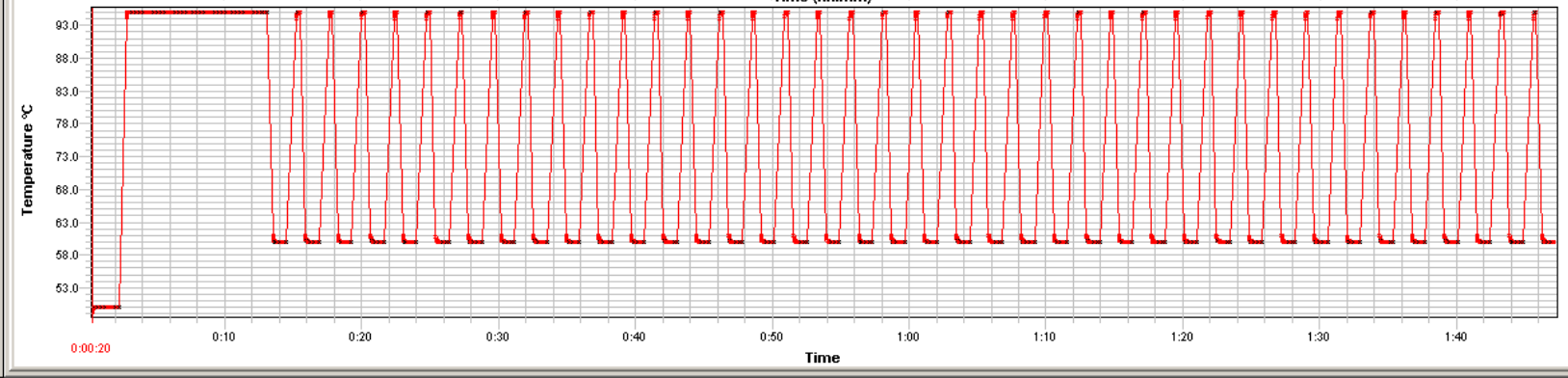


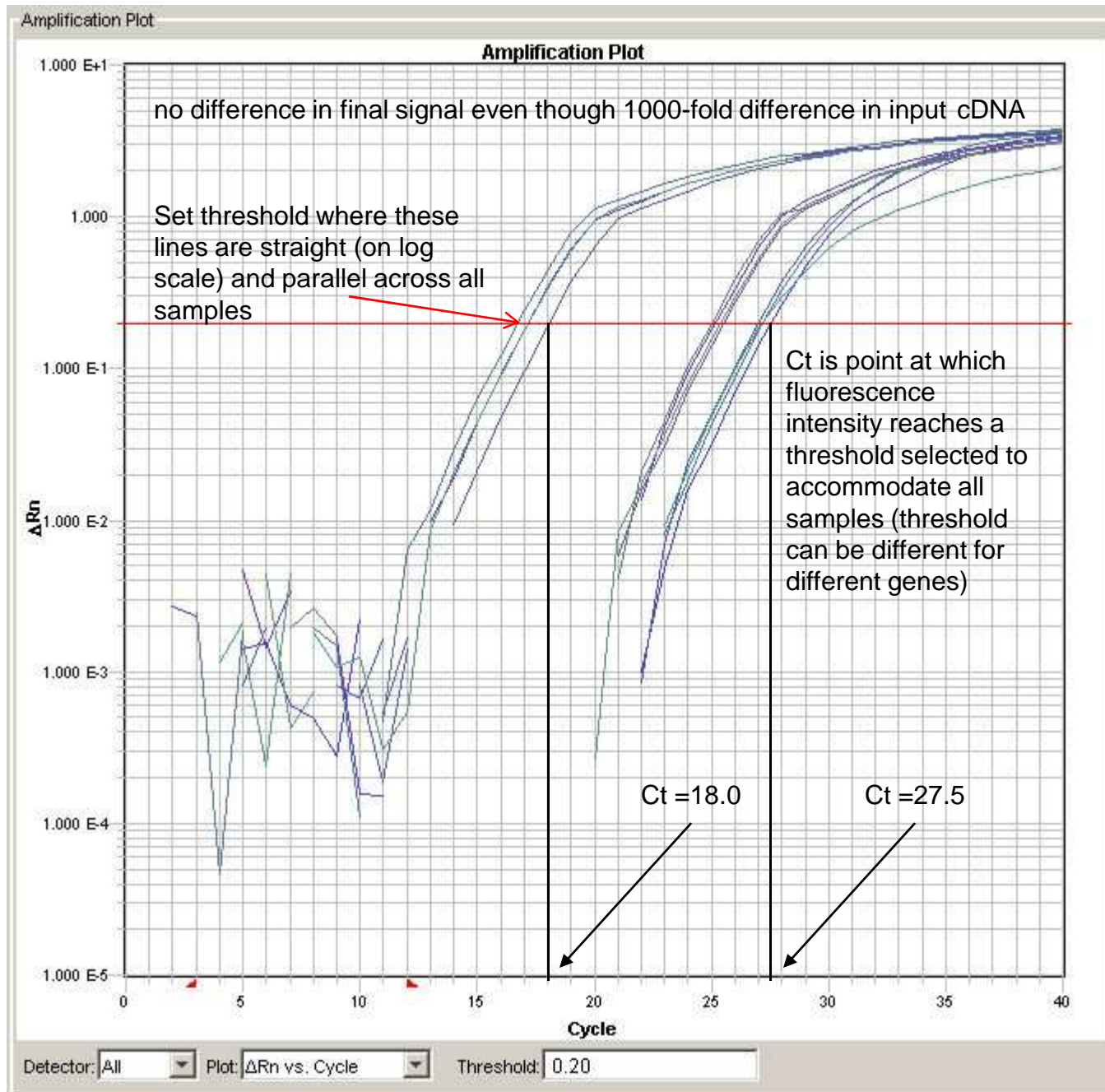
Multicomponent

vWell: 11



Name
FAM
ROX
Background
mse





SYBR Green assay

3 targets

4 samples/target

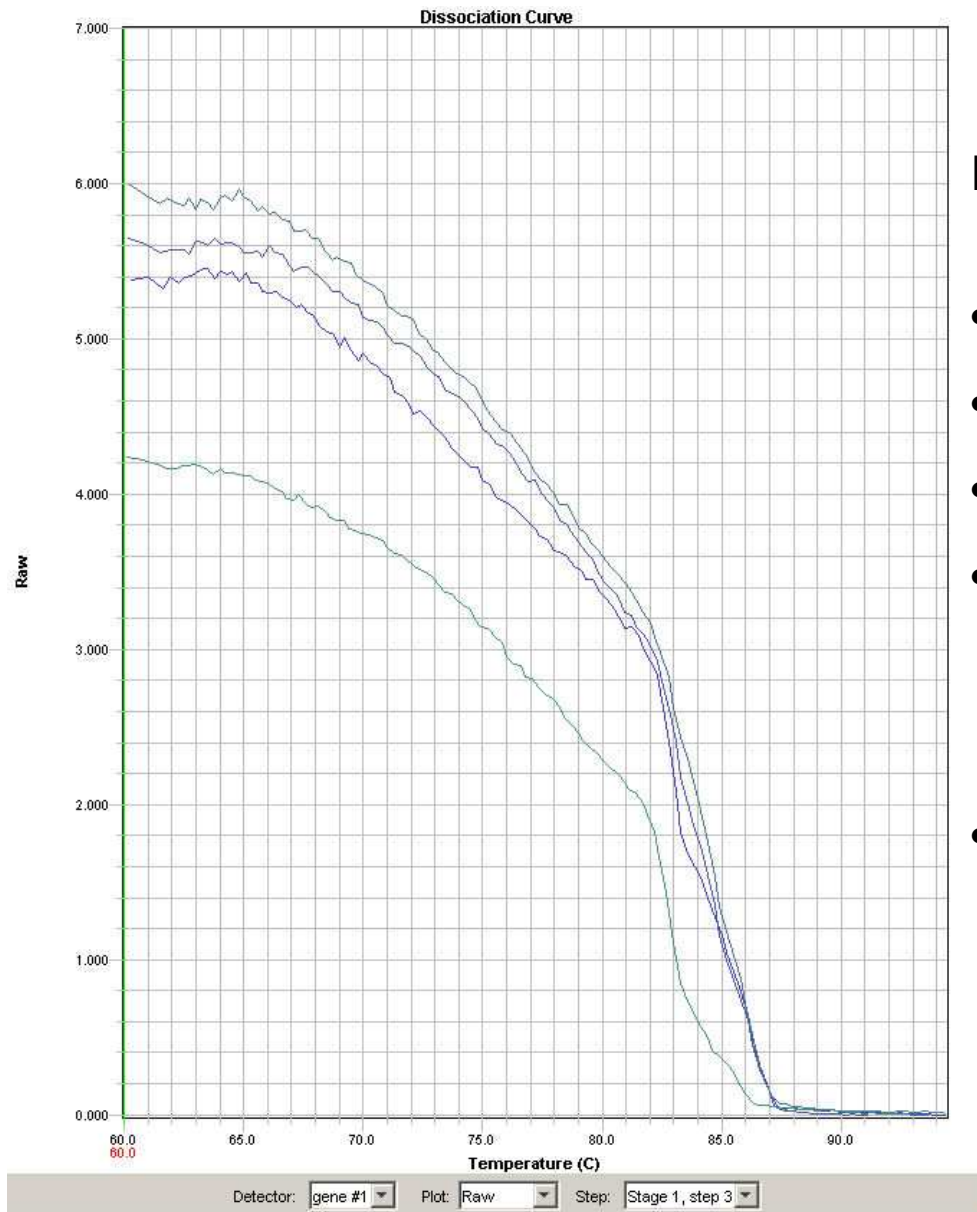
all curves look ok

$$\Delta Ct = 27.5 - 18 = 9.5$$

$$2^{9.5} = 724\text{-fold}$$

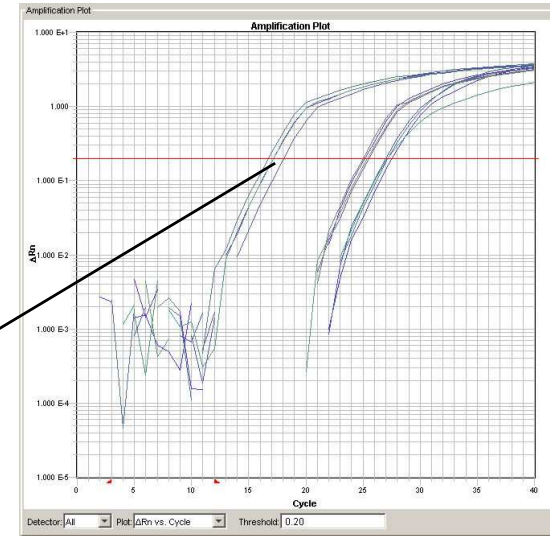
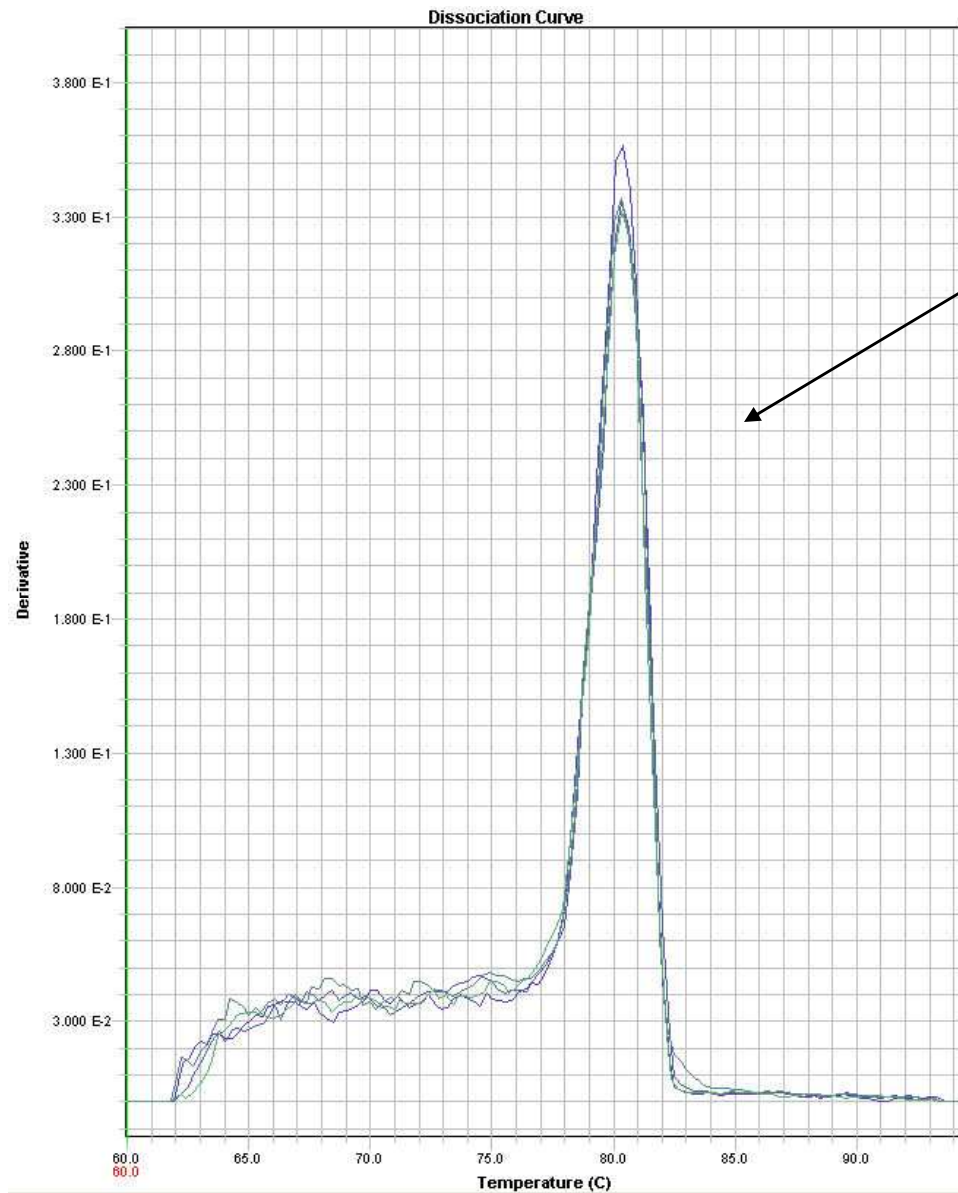
difference if amplicons same size and PCR efficiency is 100% for all targets

As long as threshold is chosen within exponential phase of amplification, ΔCt should be the same regardless of absolute Ct. If comparing data across different plates, use same threshold for all plates to avoid increasing variability.

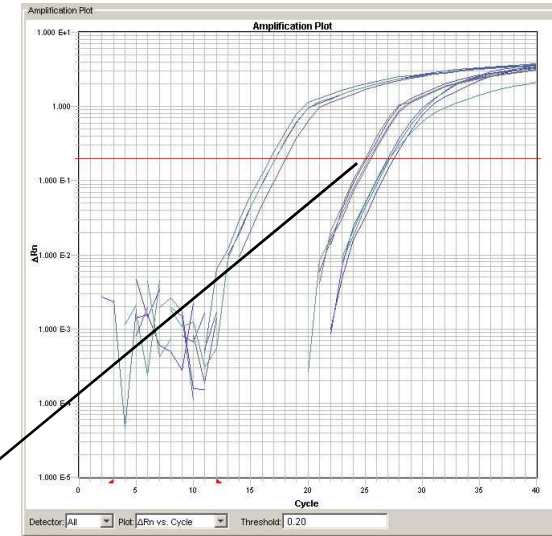
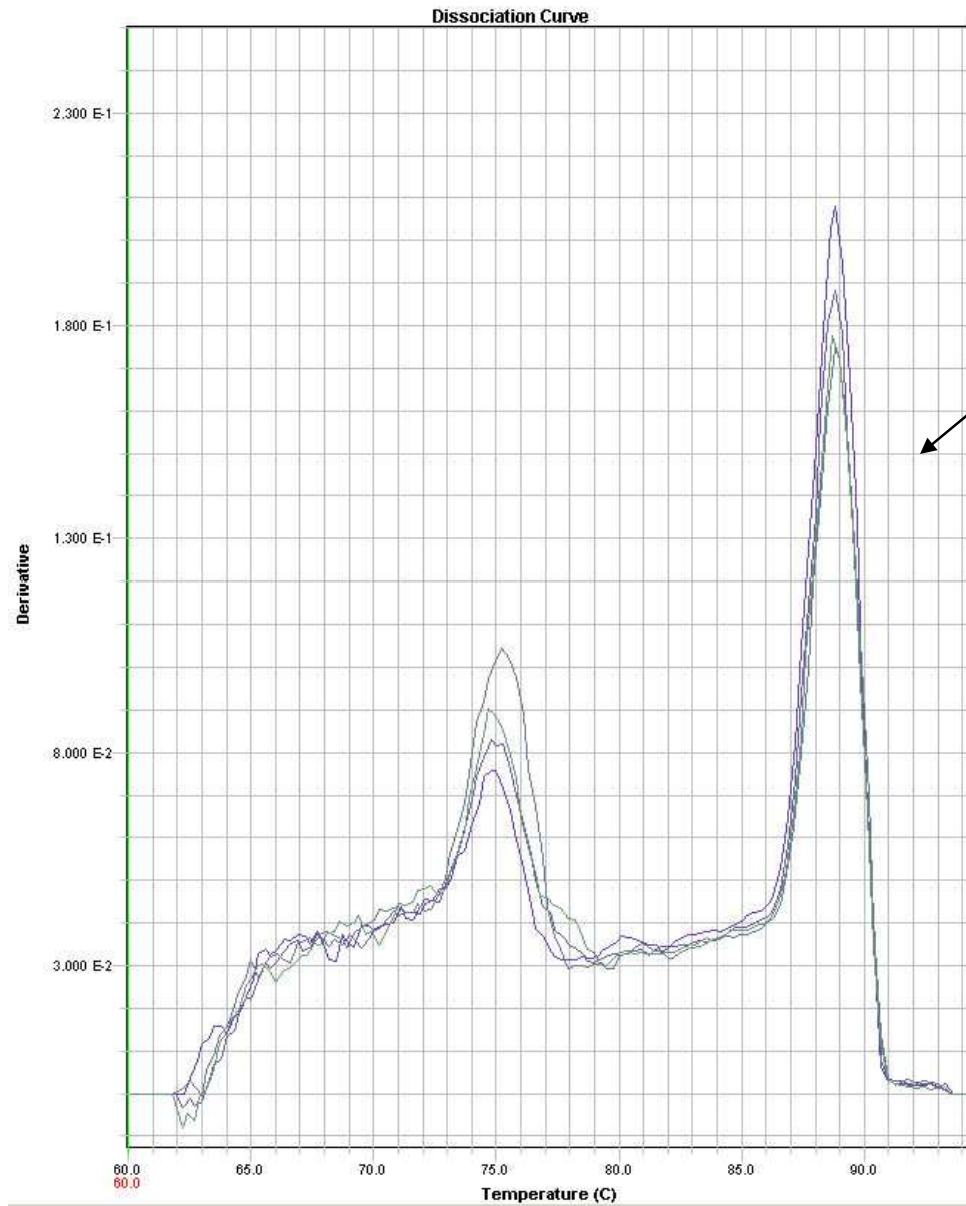


Melting curve – raw fluorescence data

- % dsDNA drops as T increases
- hard to interpret curves in this form
- look at derivative instead
- derivative is $-\Delta y/\Delta x$ over very small intervals of x, in other words rate of drop in fluorescence as T increases
- rate is fastest at T_m

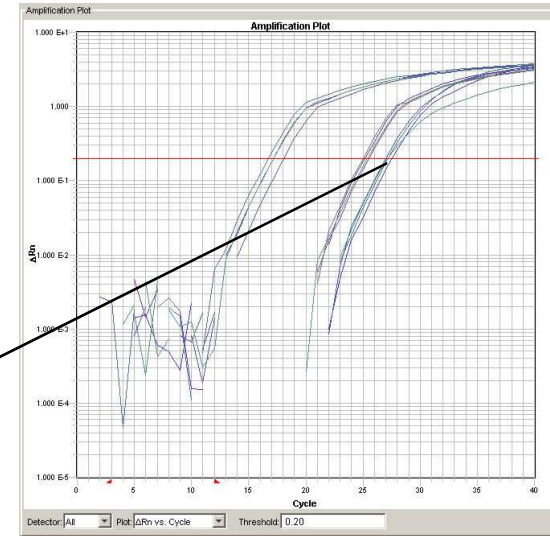
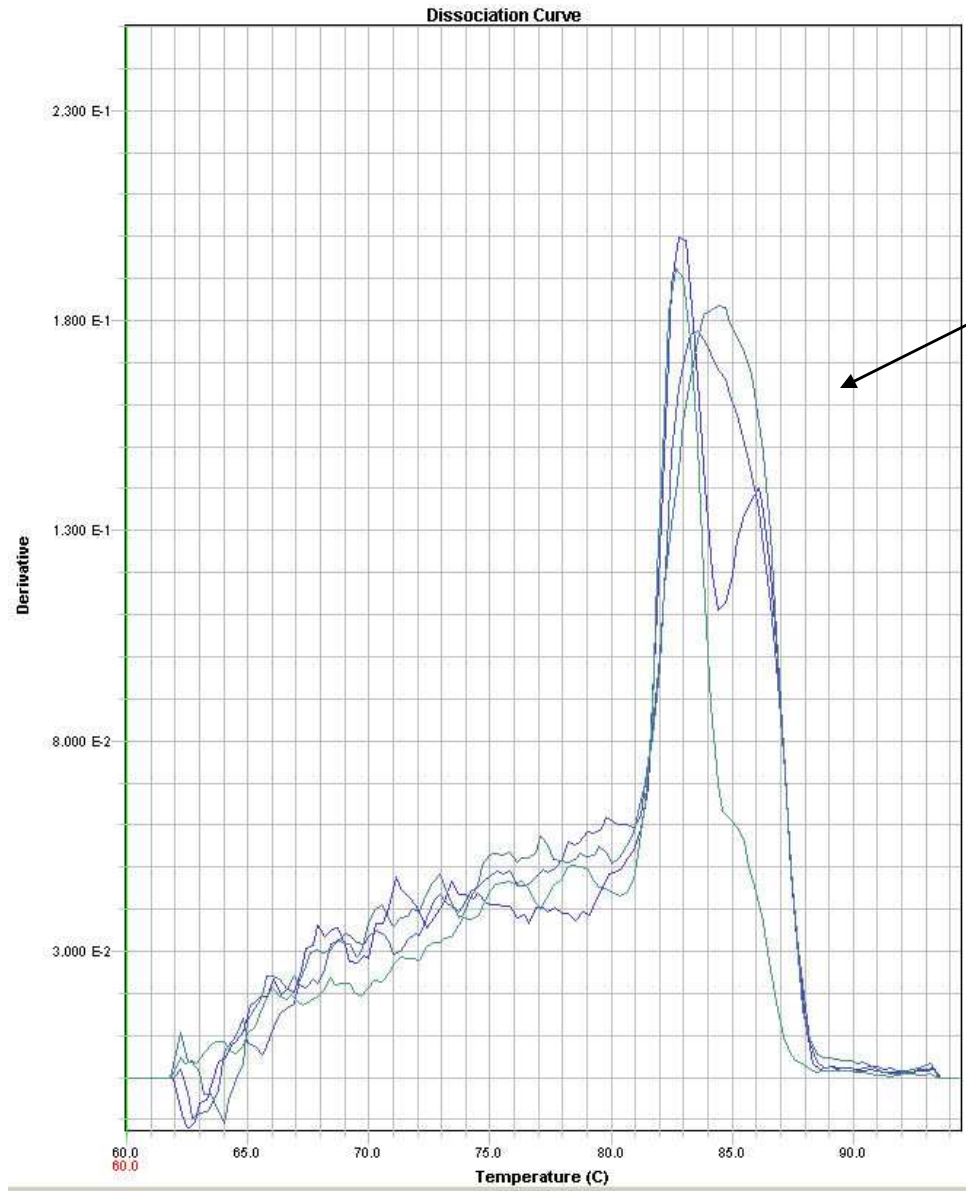


This is what you want to see



two products, one small (primer dimer?)

Detector: Plot: Step:



- Two products close in size
- abundance of larger varies a lot
- in two samples almost looks like a single peak

Detector: gene #1 Plot: Derivative Step: Stage 1, step 3

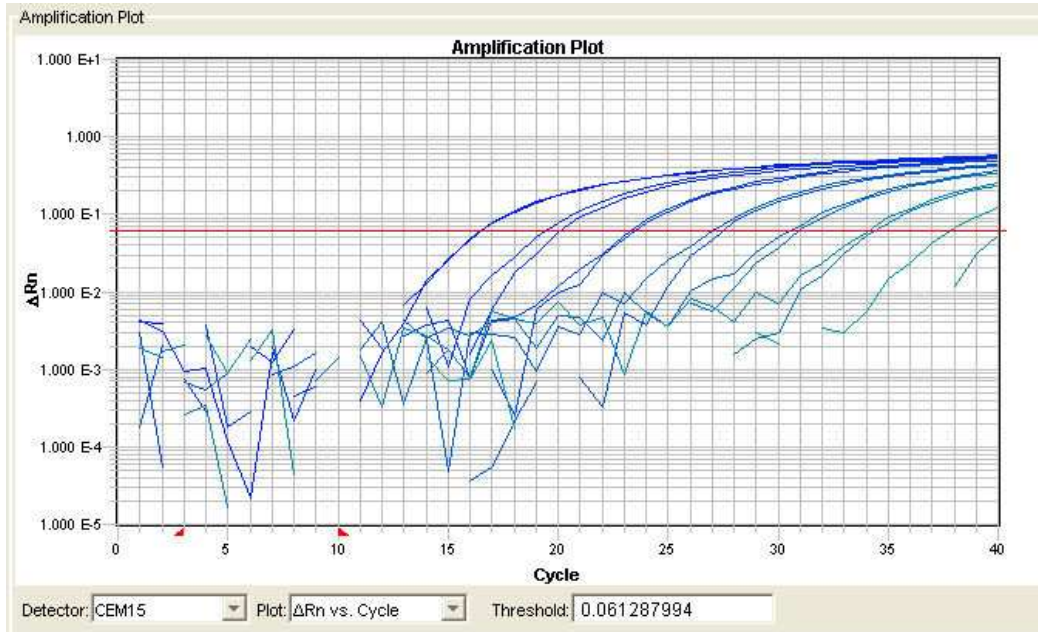
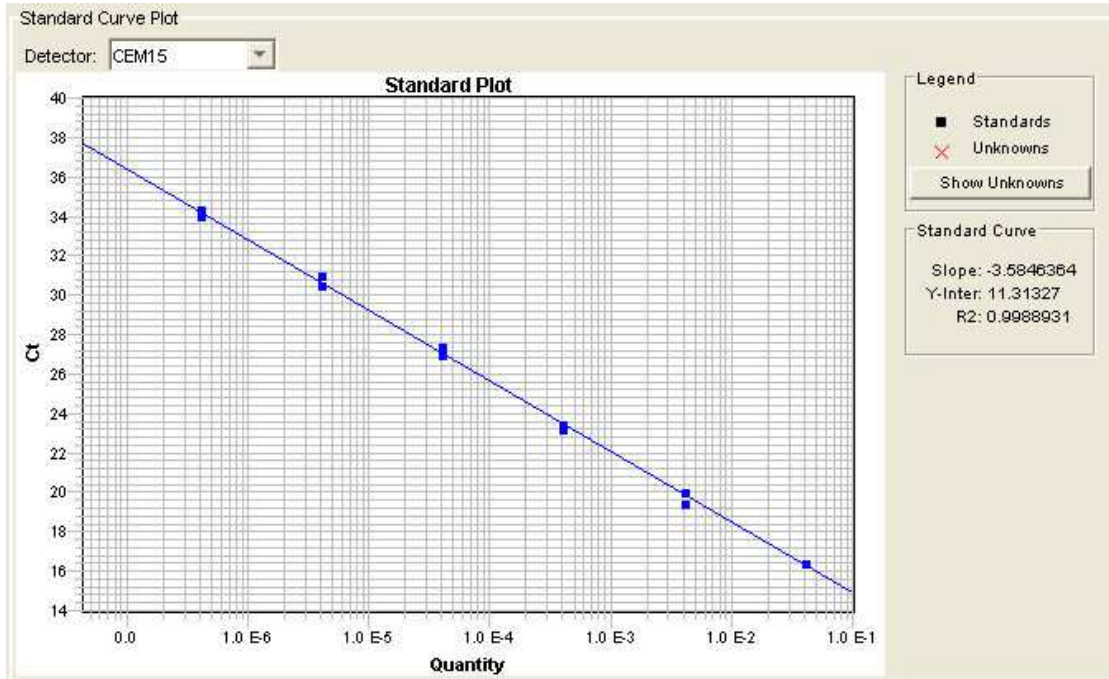
From Ct to gene expression

absolute quantification

- # molecules mRNA (really cDNA) per ? (cell, mg protein, ng total RNA, or # molecules reference RNA)
- only as accurate as your denominator even if quantification of target cDNA is very accurate
- make RNA standard by IVT if you need to quantify mRNA rather than cDNA, but even this might not be 100% accurate

relative quantification

- ΔCt : ratio of target cDNA to reference (“housekeeping”) cDNA
- $\Delta\Delta Ct$: target/reference ratio in experimental sample relative to target/reference ratio in “calibrator” sample
- can check amplification efficiencies of target and reference by dilution curve
- reference (housekeeping) genes are not necessarily invariant



Standard curve for absolute quantification

serial 10-fold dilutions of plasmid

amount (log 10) in unknown derived from slope (a) and intercept (b) of regression line:

$$Y = aX + b$$

$$X = (Y - b)/a$$

“ideal” slope is -3.32, i.e efficiency* is 1 because:

$$2^{3.32} = 10$$

here, slope is -3.58, meaning it takes 3.58 cycles to get 10-fold amplification, so

$$X^{3.58} = 10$$

$$X = 10^{1/3.58} = 1.9$$

$$\text{Efficiency} = X - 1 = .9$$

*efficiency is number of amplicon molecules produced per cycle per number of template molecules. If one molecule is produced per template (maximum possible) then number of amplicons doubles

Issues in absolute quantification

- quantification of DNA standard not always accurate
- must make RNA standard by in vitro transcription if you want to quantify RNA rather than cDNA
- quantification of standard RNA can be imprecise (fluorescence-based methods more accurate than UV absorbance)
- Even with RNA standard, accuracy of quantification is affected by RNA instability (freeze small aliquots) and possibly differences in RT efficiency between RNA standard and target mRNA

Relative quantification

$$X_C = X_0 \times (1+E)^C$$

where X_C is amount of amplicon after C cycles, X_0 is number of template molecules before amplification, and E is amplification efficiency.

We want to know X_0 ; $X_0 = X_C \div (1+E)^C$

Because relation between X_C and detected fluorescence is not necessarily constant for all values of C, we get relative quantity by picking a fluorescence level, well above the background fluorescence, that all amplification curves cross while still in exponential phase of amplification (linear part of amplification plot). X_C at C_t is a constant (all samples have the same fluorescence level), which for simplicity is set at an arbitrary value of 1. Thus,

$$\text{Relative } X_0 \text{ (arbitrary units)} = 1 \div (1+E)^{C_t} = (1+E)^{-C_t}$$

If we are sure that every sample in the experiment has an identical amount of total cDNA of the same quality, this information might be enough to compare across samples. Usually, we do not want to make this assumption but prefer to express X_0 in relation to a reference gene to control for variations in amount of RNA in RT reaction, RNA quality, RT efficiency, etc. In other words, we want to know $X_0/X_{0\text{ref}}$.

Since X_{0ref} is calculated the same way as X_0

$$X_0/X_{0ref} = (1+E)^{-Ct} \div (1+E_{ref})^{-Ct_{ref}}$$

If $E = E_{ref}$, then you can just subtract exponents

$$X_0/X_{0ref} = (1+E)^{-Ct - (-Ct_{ref})} = (1+E)^{Ct_{ref}-Ct}$$

$Ct_{ref} - Ct$ is known as ΔCt , and often it is assumed that $E = 1$, so

$$X_0/X_{0ref} = 2^{\Delta Ct}$$

[although note that ΔCt often refers to $Ct - Ct_{ref}$, so $2^{-\Delta Ct}$ is used]

Use of $\Delta\Delta Ct$

This method is used to normalize all X_0/X_{0ref} ratios relative to a single sample (often a “control” sample, but not necessarily). The X_0/X_{0ref} ratio in the calibrator sample is set arbitrarily at a value of 1.0.

Example: calibrator ΔCt (ref – gene of interest) is -5, experimental ΔCt is -10, so gene is downregulated by experimental condition relative to calibrator by:

$$2^{\Delta\Delta Ct} = 2^{[-5 - (-10)]} = 2^5 \text{ (32-fold)}$$

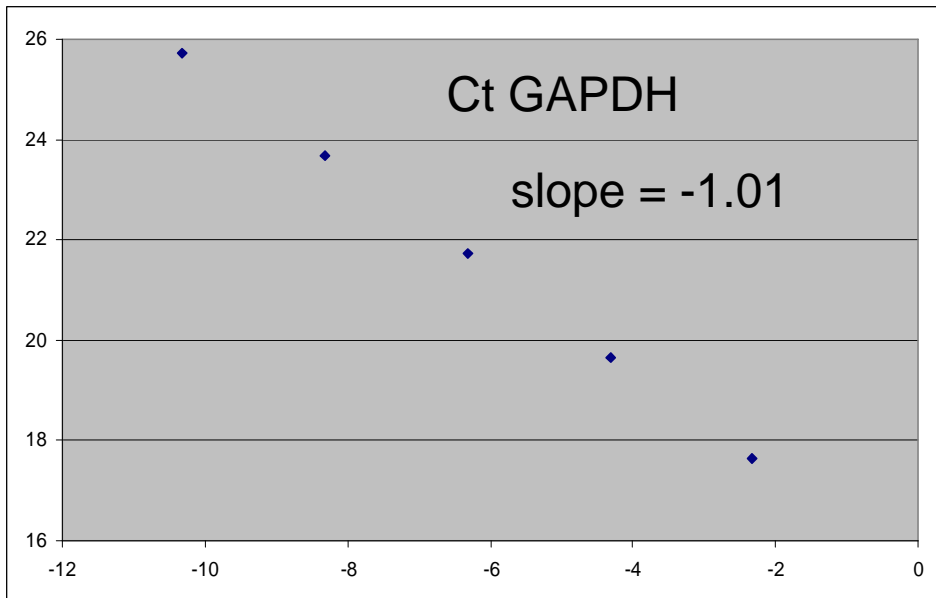
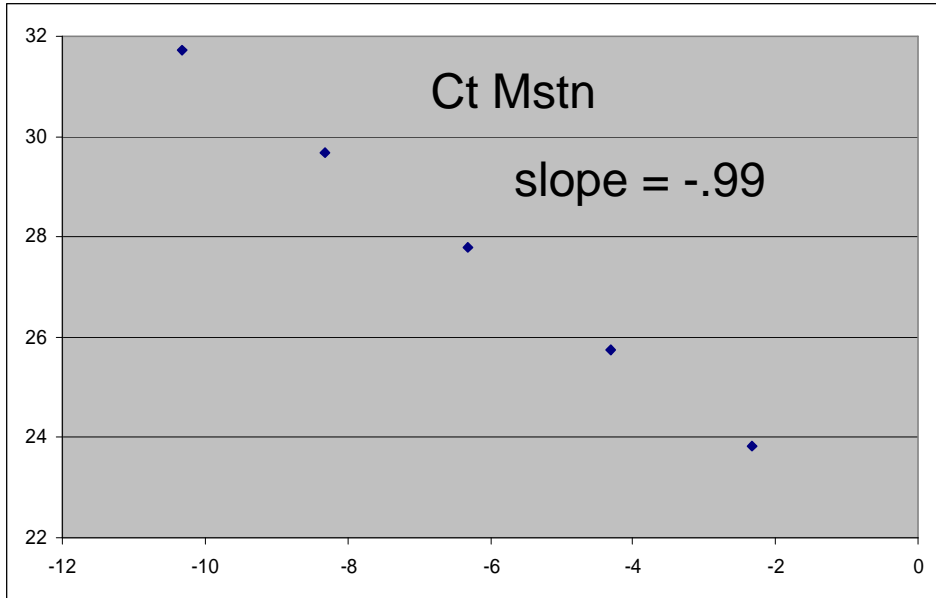
This calculation should be used only if $E = 1$ for both reference gene and gene of interest (use $X_0/X_{0ref} = (1+E)^{-Ct} \div (1+E_{ref})^{-Ct_{ref}}$ if E and E_{ref} are not same).

Alternatively (my preference), you can just compare average X_0/X_{0ref} in experimental samples with X_0/X_{0ref} in a group of control samples (convert ΔCt to X_0/X_{0ref} before averaging). This gives you the same relative quantification except absolute units are different (i.e., no sample has to have a value of 1.0).

Magnitude of errors when E is different for gene of interest and reference gene, but you use $2^{\Delta\Delta Ct}$ anyway:

Suppose actual $E_{ref} = 1$ and actual $E_{goi} = 0.9$, sample 1 is calibrator

ref Ct 1	ref Ct 2	goi Ct 1	goi Ct 2	$2^{\Delta\Delta Ct}$	correct ratio	%error
15	15	20	22	0.25	0.2770	10.8
15	18	21	26	0.25	0.3231	29.2
15	20	22	29	0.25	0.3580	43.2
15	22	23	32	0.25	0.3967	58.7
15	25	20	32	0.25	0.4627	85.1
12	22	17	29	0.25	0.4627	85.1
12	12	17	19	0.25	0.2770	10.8
12	12	17	21	0.0625	0.0767	22.8
12	12	17	23	0.01563	0.0213	36.0



Log₂ total cDNA

Validating Δ Ct method

dilute cDNA to determine efficiency for both gene of interest and control gene (use the “strongest” sample for the dilution series)

For log₂ data, slope should be close to -1, for log₁₀ data, close to -3.32

for slope from log₂ data

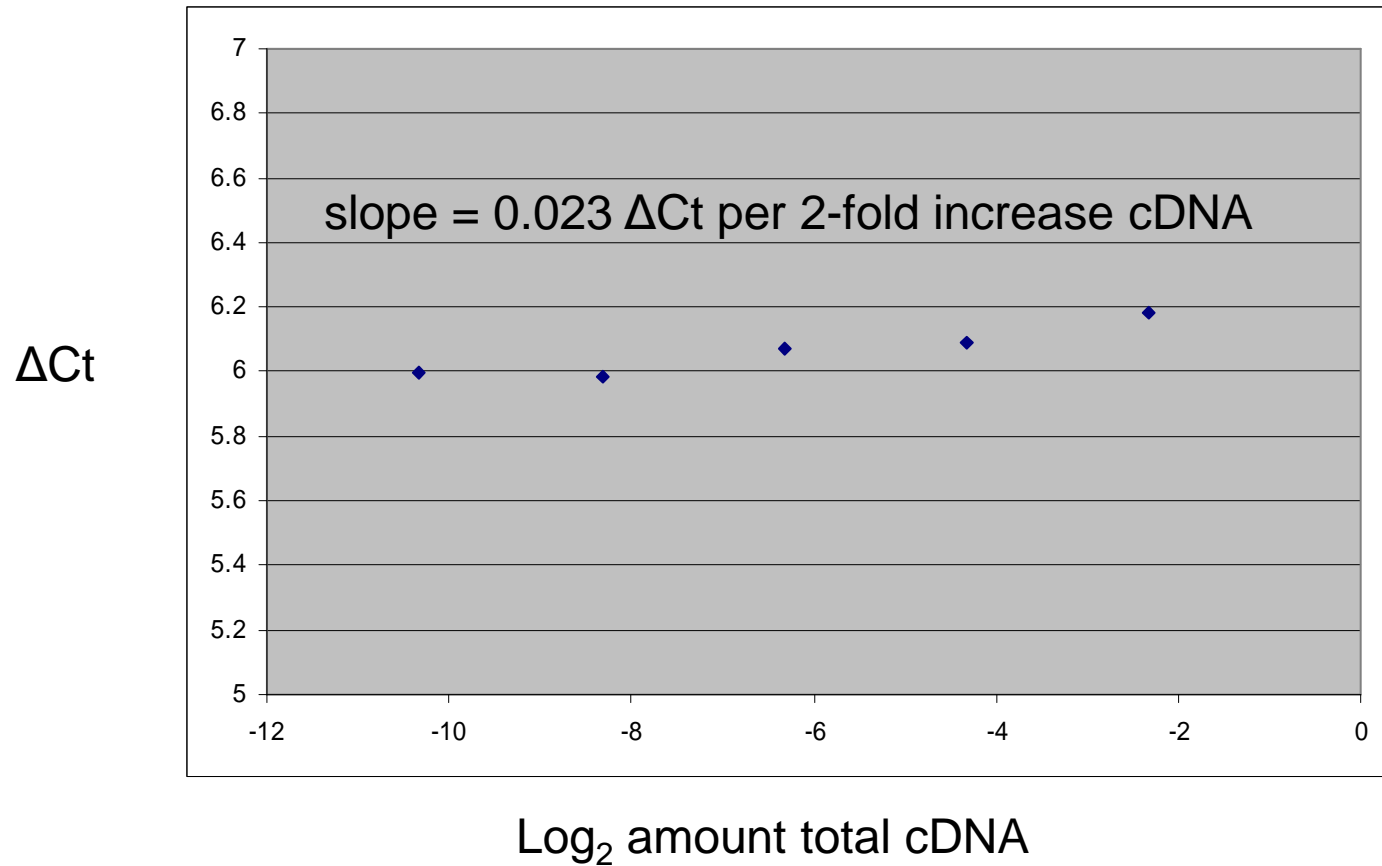
$$E = 2^{(-1/\text{slope})} - 1$$

= 1.01 for Mstn, 0.99 for GAPDH

for slope from log₁₀ data

$$E = 10^{(-1/\text{slope})} - 1$$

ΔCt for Mstn Ct – GAPDH Ct



If amplification efficiencies are similar enough to use the simple $(1+E)^{\Delta Ct}$ calculation method, then slope must be close to zero

Selecting a reference gene

- consistent expression in all samples both within and across experimental conditions
- easier to choose if validating microarray data, otherwise usually you are not certain that reference gene is totally unaffected by experimental conditions
- can try several reference genes, or see if $C_{t_{ref}}$ changes with experimental conditions when you carefully standardize cDNA input
- more critical if hoping to find relatively small effects
- An alternative is to measure total cDNA concentration, use same amount of total cDNA in every well, and calculate relative X_0 ($=[1+E]^{-C_t}$), for every well without worrying about reference genes

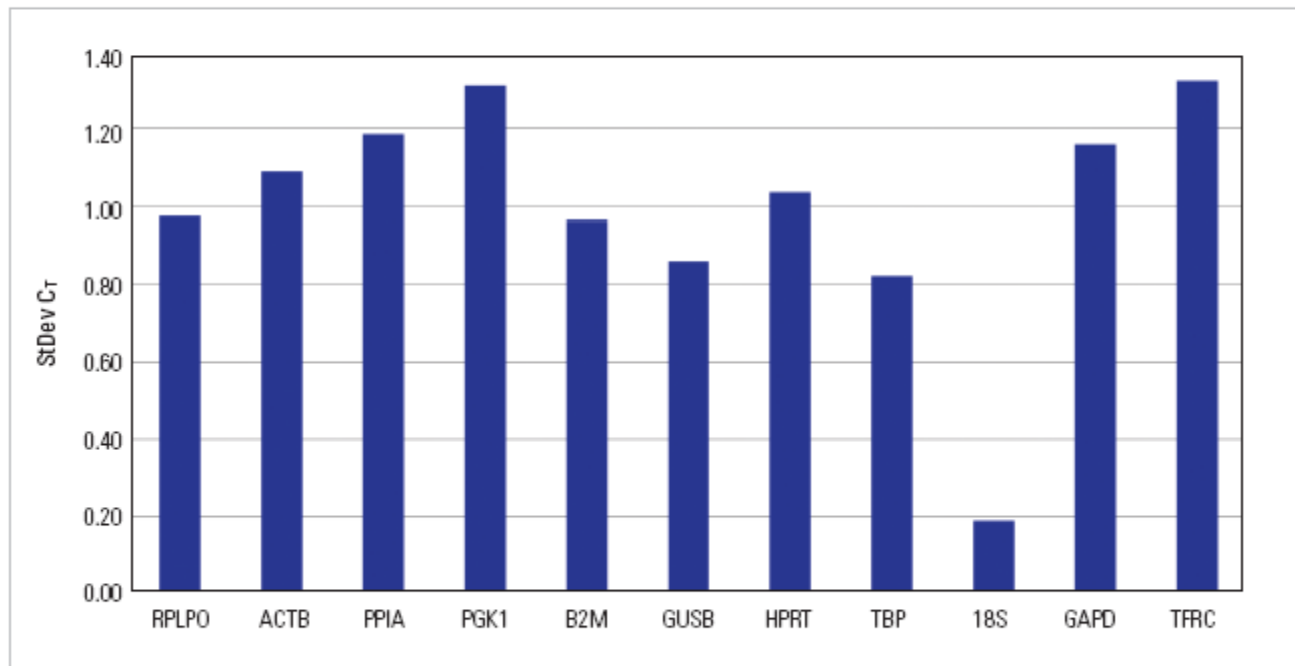


Figure 4. Human Endogenous Control. Variation of 11 Human Endogenous Controls across 11 samples as measured by Standard Deviation of C_t .

Replicate, Replicate, Replicate

- Always run replicate wells (we prefer 3) of each assay for each sample. Precision (SD) with careful pipetting can be less than 0.1 Ct units (on our 7900HT instrument). Thus, it is possible to detect small differences between two samples with very careful plate set up.
- Do as many biological replicates as you can if you want to detect relatively small effects of a treatment. Selecting a good reference transcript (stable expression across conditions) or accurate determination of total cDNA concentrations will enhance the ability to detect small effects. With 8-10 subjects per group we have been able to detect effects on gene expression as small as 25% at $P < 0.05$.