Session 2

Pete Vallone

Quantitation Using PCR

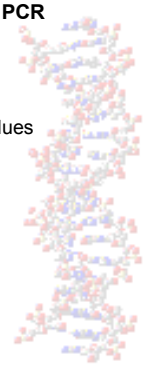
Exponential Phase

Setting Baseline and Threshold values

C_T – Cycle Threshold

Standard Curve parameters
- Slope (m)
- R^2

Importance of the Calibrant



Quantitation Using PCR

Visually inspect qPCR curves

Set Baseline and Threshold values

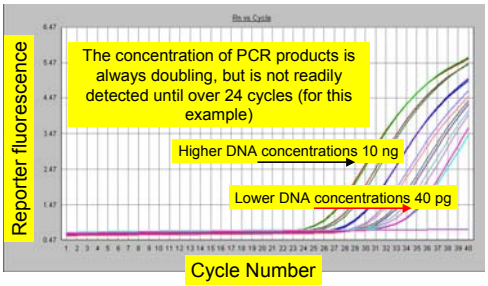
Construct and evaluate a Calibrant Curve

Review estimated DNA concentrations

This can be done rapidly in the instrument software package

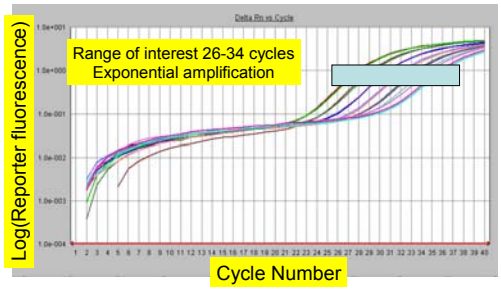
Estimated DNA concentrations can be easily manipulated in Excel

Fluorescence vs Cycle Number



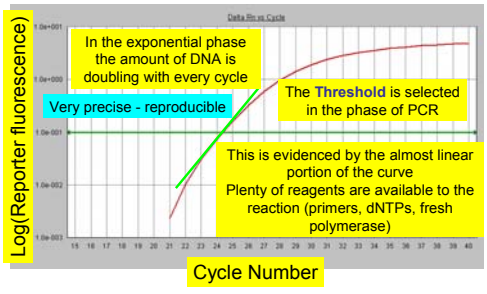
Quantifiler data

Log View of Data



Quantifiler data

Data Measured in the Exponential Phase



Quantifiler data

Setting the Baseline

A low and high value are set

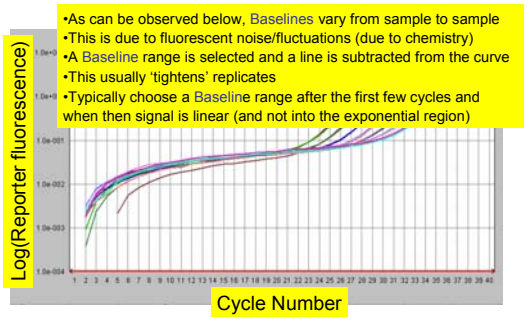
The Baseline is set to eliminate the background signal found in the early cycles of amplification

The Baseline should not interfere with the exponential phase of the amplification

The Baseline is set to allow for accurate C_T determination

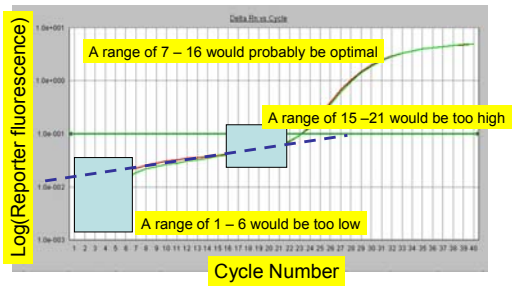
Many qPCR methods have a prescribed Baseline

Log View of Data



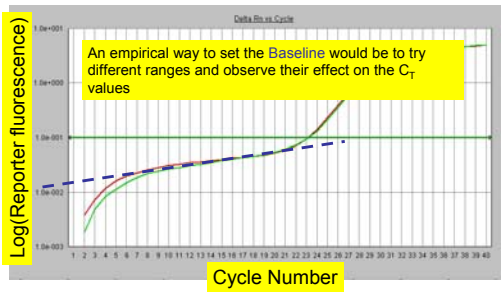
Quantifiler data

Setting the Baseline



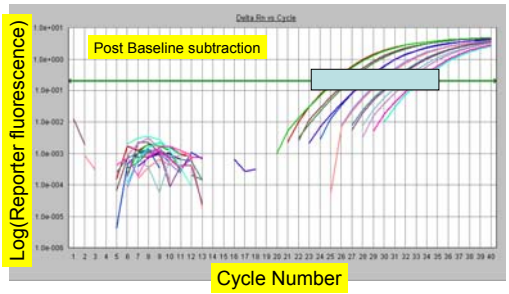
Quantifiler data

Setting the Baseline



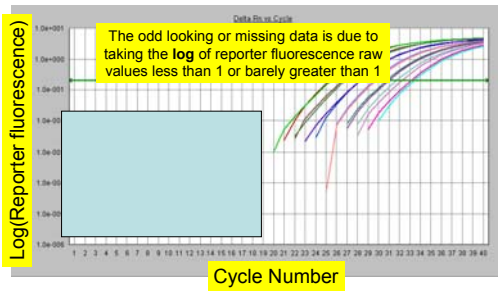
Quantifiler data

Log View of Data



Quantifiler data

What is with the Confetti?



Quantifiler data

The C_T Value

C_T is simply the cycle number selected at a specific threshold value

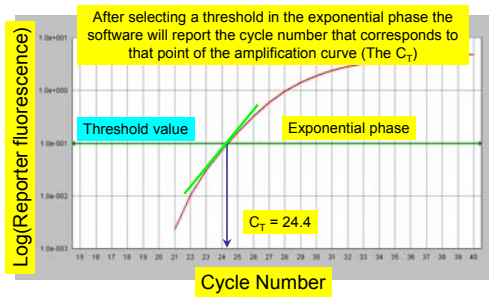
The threshold value is selected where all the data is undergoing exponential amplification

The threshold value can be selected manually or by the software

The threshold value for different methods may vary

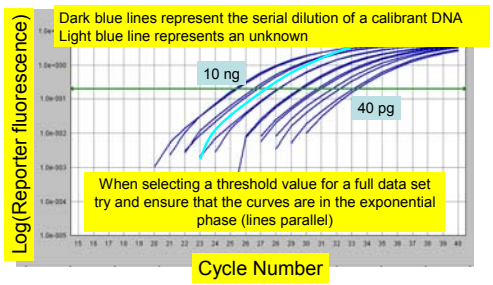
Selected in the log(signal) plot view

Selecting the Threshold Value



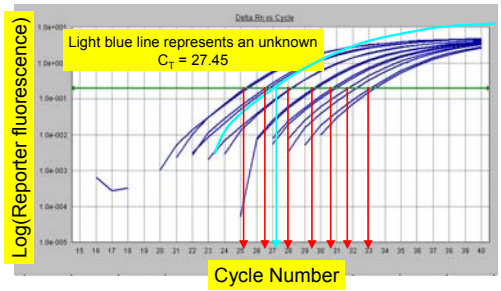
Quantifiler data

Selecting the Threshold Value



Quantifiler data

Selecting the Threshold Value



Quantifiler data

C_T Value and the Standard Curve

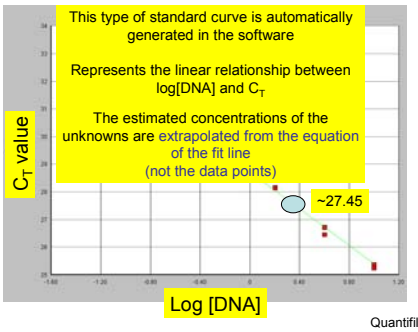
After a suitable threshold has been selected the data is analyzed and the C_T values are determined

The C_T values of the serial dilution are plotted versus the log[DNA] – your serial dilution of a calibrant DNA

The line is visually inspected and the parameters are reviewed

If the standard curve is linear and the line parameters are acceptable, the unknown concentrations can then be estimated

Log [DNA] versus C_T



Quantifiler data

Equation of a Straight Line

The equation Y = mX + b defines a straight line

m is the slope
- $(y_1-y_2)/(x_1-x_2)$
- The "steepness" of the line
- Relates to the efficiency of the PCR

b is the Y-intercept (where the line crosses the Y-axis)

X is your log[DNA] concentration (serial dilutions)

Y is the C_T value

Linear Least Squares Regression

The most widely used modeling method

"regression," "linear regression," or "least squares"

Many processes in science and engineering are well-described by linear models

Good results can be obtained with relatively small data sets

Main disadvantages: limitations in the shapes that linear models can assume over long ranges, possibly poor extrapolation properties, and sensitivity to outliers

Linear Least Squares Regression

Carried out by the instrument software

Can also be easily performed in Excel, Sigma Plot etc

Briefly, the method solves for m and b from the data points (remember X and Y are constants)

Finds numerical values for the parameters that minimize the sum of the squared deviations between the observed responses (your data!) and the functional portion of the model (the line!)

Linear Least Squares Regression

The best fit line associated with the n points (x_1,y_1), (x_2,y_2),…,(x_n,y_n)

y = mx + b

Where

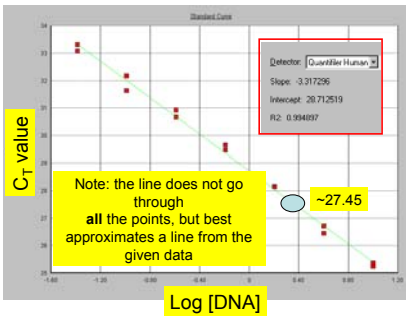
$$m = \frac{n(\sum xy)-(\sum x)(\sum y)}{n(\sum x^2)-(\sum x)^2} \quad b = \frac{\sum y - m(\sum x)}{n}$$

Here, Σ means "the sum of". Thus

$\sum xy = x_1y_1 + x_2y_2 \ldots + x_ny_n$ $\quad \sum x = x_1 + x_2 \ldots + x_n$

$\sum x^2 = x_1^2 + x_2^2 \ldots + x_n^2$ $\quad \sum y = y_1 + y_2 \ldots + y_n$

Log [DNA] versus C_T



Quantifiler data

Calculating PCR Efficiency

Taking the relationship between log(copies) and cycles of PCR one can rearrange the equation $X_N = X_0 (1 + E)^N$ in order to determine efficiency

$$\text{Rxn Efficiency} = [10^{(-1/m)}] - 1$$

slope(m) = -3.317296

$$= [10^{(-1/-3.317296)}] - 1$$

E = 2.0019-1

Just over 100% efficient

E = (2.0019 – 1) = 1.019

R^2 (R-squared)

Coefficient of determination

A statistic for a predictive model's lack of fit using the data from which the model was derived

$$R^2\text{-squared} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y}_i)^2}$$

A perfectly fitting model yields an R^2 of 1 (all points fall directly on the line)

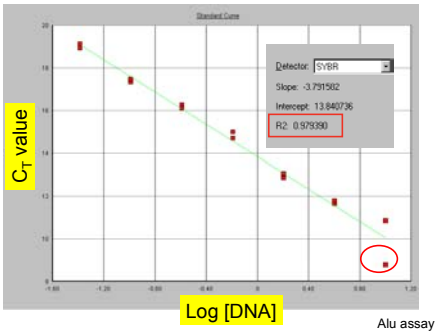
R^2 (R-squared)

For most log[DNA] versus C_T standard curves R^2 should be greater than 0.990

Sometimes outliers can be removed to improve the R^2 values

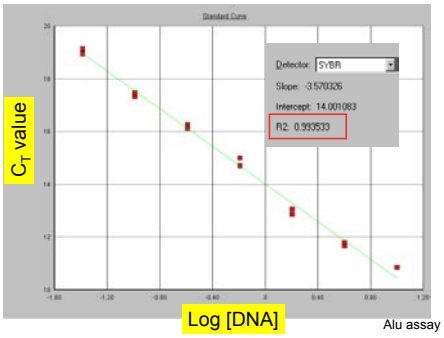
Outliers can be at low/high concentrations or outside the performance range of the qPCR assay (or just a bad point – pipet error, dirty well etc)

Removing An Outlier



Alu assay data

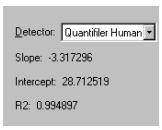
Removing An Outlier



Alu assay data

Solving for an Unknown

From the data



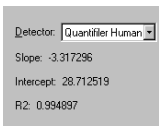
Y = mX + b
C_T = m*log[DNA] + b
Solving for [DNA]

$$[DNA] = 10^{\frac{CT-b}{m}}$$

The equation above is used to estimate the [DNA] of the unknowns

Solving for an Unknown

From the data



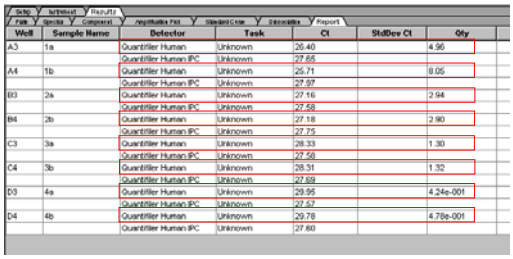
Solving for [DNA]

$$[DNA] = 10^{\frac{27.45-28.71}{-3.3172}}$$

After solving for the equation when C_T = 27.45 this corresponds to a [DNA] of 2.39 ng

The software will do this for you…

Data Report



Well	Sample Name	Detector	Task	Ct	StdDev Ct	Qty
A3	1a	Quantifiler Human	Unknown	26.40		4.96
		Quantifiler Human IPC	Unknown	27.65		
A4	1b	Quantifiler Human	Unknown	25.71		9.05
		Quantifiler Human IPC	Unknown	27.97		
B3	2a	Quantifiler Human	Unknown	27.16		2.94
		Quantifiler Human IPC	Unknown	27.58		
B4	2b	Quantifiler Human	Unknown	27.18		2.90
		Quantifiler Human IPC	Unknown	27.75		
C3	3a	Quantifiler Human	Unknown	28.33		1.30
		Quantifiler Human IPC	Unknown	27.58		
C4	3b	Quantifiler Human	Unknown	28.31		1.32
		Quantifiler Human IPC	Unknown	27.69		
D3	4a	Quantifiler Human	Unknown	29.95		4.24e-001
		Quantifiler Human IPC	Unknown	27.57		
D4	4b	Quantifiler Human	Unknown	29.78		4.78e-001
		Quantifiler Human IPC	Unknown	27.60		

An example of a data report from the 7500 collection software
Report can be exported and manipulated in a spreadsheet

Varying the Threshold Value

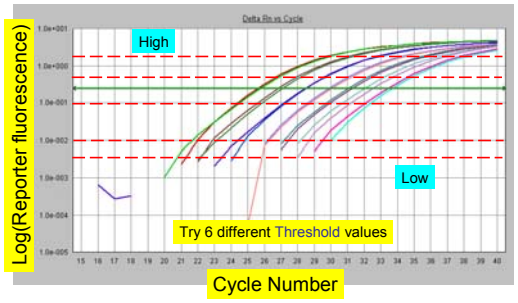
What happens when we change the Threshold value?

Of course the absolute C_T values will change
- But it will be consistent for that data set

You don't want to compare C_T values from different methods or even runs

What is the effect of varying Threshold on the standard curve and the estimated values for the unknowns?

Varying the Threshold Value



Quantifiler data

Varying the Threshold Value

Selecting 6 Threshold values then estimating [DNA] for a sample run in duplicate

	Threshold	A	B	Avg	Stdev
Low	0.004	23.51	24.48	24.00	0.69
Low	0.01	23.18	21.12	22.15	1.46
Below Opt	0.1	18.83	18.1	18.47	0.52
Optimal	0.2	17.13	18.13	17.63	0.71
High	1.7	17.58	16.68	17.13	0.64
Above Opt	0.25	17.5	16.83	17.17	0.47

~6.8 ng difference (max)

Varying the Threshold Value

Selecting 6 Threshold values then estimating [DNA] for a sample run in duplicate

Rxn efficiency

	Threshold	R2	slope	E	E -1
Low	0.004	0.989	-3.474	1.94	0.94
Low	0.01	0.991	-3.336	1.99	0.99
Below Opt	0.1	0.994	-3.289	2.01	1.01
Optimal	0.2	0.994	-3.317	2.00	1.00
High	1.7	0.993	-3.421	1.96	0.96
Above Opt	0.25	0.995	-3.322	2.00	1.00

Amp efficiency

Importance of the Calibrant!

All qPCR results are relative to the standard curve

Serial dilutions of the Calibrant DNA comprise the standard curve

Any errors involving the Calibrant DNA directly effect the estimates of your unknown DNA concentrations
- Pipetting errors
- Miscalculation of concentrations
- New lots or vendors of Calibrant DNA
- Contamination of Calibrant
- Evaporation of Calibrant DNA

Importance of the Calibrant!

Things to keep in mind about Calibrants

The Calibrant is usually a pristine well-characterized DNA sample

- Not extracted the same as the unknown
- Not subjected to the same environment as your unknown(s)
- Will not contain inhibitors, Heme, Ca^{++} etc
- May be from a cell line or mixed source sample
- May exhibit lot-to-lot variation (monitor this)

Error in the C_T Value

Be aware that relatively small changes in C_T result in large variations in estimated concentration

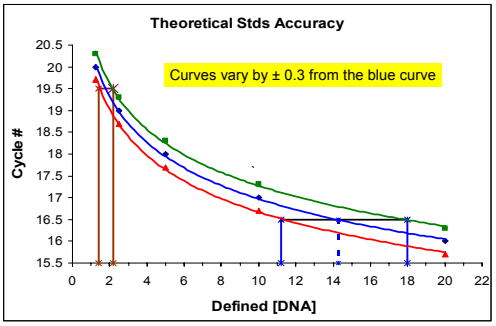
m	b	CT	[DNA]	%	delta
-3.3219	26	25.1	1.87	6.70	0.13
-3.3219	26	25	2.00		
-3.3219	26	24.9	2.14	6.70	-0.14
-3.3219	26	20.1	59.72	6.70	4.29
-3.3219	26	20	64.00		
-3.3219	26	19.9	68.60	6.70	-4.59

± 0.1 C_T

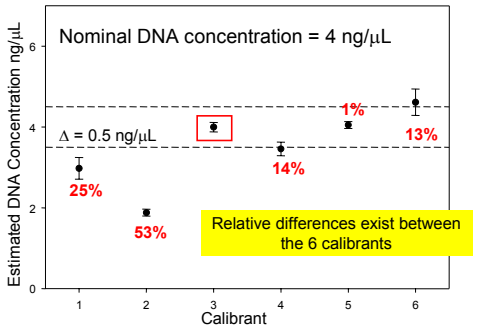
m	b	CT	[DNA]	%	delta
-3.3219	26	25.3	1.62	18.77	0.38
-3.3219	26	25	2.00		
-3.3219	26	24.7	2.46	18.77	-0.46
-3.3219	26	20.3	51.99	18.77	12.02
-3.3219	26	20	64.00		
-3.3219	26	19.7	78.80	18.77	-14.79

± 0.3 C_T

Error in the C_T Value



Differences Between Calibrants



Summary

Data is collected in the exponential range

After threshold selection, amplification curve data is reduced down to C_T values

The log[DNA] vs C_T standard curve is the backbone of data interpretation

$R^2 > 0.990$

Experiment with baselines and C_T values

Errors/variations in the DNA Calibrant concentration are directly translated into the estimates for the unknowns