**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

---

**Visualize 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

---

**Log View of Data**

- Range of interest 26-34 cycles
- Exponential amplification

---

**Data Measured in the Exponential Phase**

- In the exponential phase, the amount of DNA is doubling with every cycle
- Very precise, reproducible
- The **Threshold** is selected in the phase of PCR
- This is evidenced by the almost linear portion of the curve
- Plenty of reagents are available to the reaction (primers, dNTPs, fresh polymerase)

---

**Fluorescence vs Cycle Number**

- The concentration of PCR products is always doubling, but is not readily detected until over 24 cycles (for this example)
- Higher DNA concentrations 10 ng
- Lower DNA concentrations 40 pg

---

**Quantification Using PCR**

- **Quantifiler data**
- **Cycle Number**
- **Log** View of Data
- **Data Measured in the Exponential Phase**
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 Cₜ determination. Many qPCR methods have a prescribed Baseline.

Log View of Data

- As can be observed below, Baselines vary from sample to sample.
- This is due to fluorescent noise/fluctuations (due to chemistry).
- A Baseline range is selected and a line is subtracted from the curve.
- Typically choose a Baseline range after the first few cycles and when the signal is linear (and not into the exponential region).

Log View of Data

- Setting the Baseline with different ranges will observe their effect on the Cₜ values.

Log View of Data

- The Log View of Data shows the odd looking or missing data is due to the log of reporter fluorescence raw values less than 1 or barely greater than 1.

What is with the Confetti?

- The odd looking or missing data is due to the log of reporter fluorescence raw values less than 1 or barely greater than 1.

Quantifier data
qPCR Workshop held at NFSTC
Peter Vallone (NIST) and Cristián Orrego (CA DOJ)

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

CT is simply the cycle number selected at a specific threshold value.

After selecting a threshold in the exponential phase the software will report the cycle number that corresponds to that point of the amplification curve ($C_T$).

CT = 24.4

Exponential phase

Threshold value

Quantifiler data

Selecting the Threshold Value

When selecting a threshold value for a full data set try and ensure that the curves are in the exponential phase (lines parallel).

Dark blue lines represent the serial dilution of a calibrant DNA

Light blue line represents an unknown

10 ng

40 pg

Quantifiler data

Selecting the Threshold Value

CT = 27.45

Quantifiler data

Selecting the Threshold Value

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.

CT Value and the Standard Curve

CT value

Log [DNA]

This type of standard curve is automatically generated in the software.

Represents the linear relationship between log(DNA) and $C_T$.

The estimated concentrations of the unknowns are extrapolated from the equation of the fit line (not the data points).

Log [DNA] versus $C_T$

Quantifiler data
The equation $Y = mX + b$ defines a straight line.

$m$ is the slope:
- $\frac{(y_2 - y_1)}{(x_2 - x_1)}$
- 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 best fit line associated with the $n$ points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

$y = mx + b$

Where

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

$$b = \frac{\sum y - m\sum x}{n}$$

Here, $\sum$ means "the sum of". Thus

$$\sum xy = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

$$\sum x = x_1 + x_2 + \ldots + x_n$$

$$\sum y = y_1 + y_2 + \ldots + y_n$$

$$\sum x^2 = x_1^2 + x_2^2 + \ldots + x_n^2$$

$\sum y^2$ can be calculated similarly.

**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.

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

$$E = \left[10^{(-1/-3.317296)}\right] - 1$$

$$E = 2.0019 - 1 = 1.019$$

Just over 100% efficient.
qPCR Workshop held at NFSTC
Peter Vallone (NIST) and Cristián Orrego (CA DOJ)

President’s DNA Initiative Training Materials

Page 5

R² (R-squared)

Coefficient of determination

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

\[
R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}
\]

A perfectly fitting model yields an R² of 1 (all points fall directly on the line)

R² (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 R² 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

\[\begin{align*}
C_T &\quad \text{CT value} \\
\log[DNA] &\quad \text{Log [DNA]}
\end{align*}\]

Removing an Outlier

\[\begin{align*}
C_T &\quad \text{CT value} \\
\log[DNA] &\quad \text{Log [DNA]}
\end{align*}\]

Solving for an Unknown

From the data

\[Y = mX + b\]

\[C_T = m\log[DNA] + b\]

Solving for [DNA]

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

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

Solving for an Unknown

From the data

Solving for [DNA]

\(\begin{align*}
\text{DNA} &\quad 10^{-3.3172} \\
27.45 - 28.71 &\quad \text{[DNA]} = 10^{-3.3172}
\end{align*}\)

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…
President’s DNA Initiative Training Materials

qPCR Workshop held at NFSTC
Peter Vallone (NIST) and Cristián Orrego (CA DOJ)

July 26-27, 2006

Data Report

<table>
<thead>
<tr>
<th>Sample</th>
<th>Threshold</th>
<th>Ct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>2</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>3</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>4</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>5</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>6</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>7</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>8</td>
<td>0.004</td>
<td>23.51</td>
</tr>
<tr>
<td>9</td>
<td>0.004</td>
<td>23.51</td>
</tr>
</tbody>
</table>

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 Ct values will change
– But it will be consistent for that data set

You don’t want to compare Ct 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

x

Log(Reporter fluorescence)

Try different Threshold values

Quantifier data

Varying the Threshold Value

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

<table>
<thead>
<tr>
<th>Threshold</th>
<th>A</th>
<th>B</th>
<th>Avg</th>
<th>Stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>23.51</td>
<td>24.48</td>
<td>24.00</td>
<td>0.69</td>
</tr>
<tr>
<td>Low</td>
<td>23.18</td>
<td>21.12</td>
<td>22.15</td>
<td>1.46</td>
</tr>
<tr>
<td>Below Opt</td>
<td>18.83</td>
<td>18.10</td>
<td>18.47</td>
<td>0.52</td>
</tr>
<tr>
<td>Optimal</td>
<td>17.13</td>
<td>18.13</td>
<td>17.63</td>
<td>0.71</td>
</tr>
<tr>
<td>High</td>
<td>17.58</td>
<td>16.68</td>
<td>17.13</td>
<td>0.64</td>
</tr>
<tr>
<td>Above Opt</td>
<td>17.5</td>
<td>16.83</td>
<td>17.17</td>
<td>0.47</td>
</tr>
</tbody>
</table>

~6.8 ng difference (max)

Varying the Threshold Value

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

x

Rxn efficiency

Varying the Threshold Value

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

<table>
<thead>
<tr>
<th>Threshold</th>
<th>R2</th>
<th>slope</th>
<th>E</th>
<th>E-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.989</td>
<td>-3.474</td>
<td>1.94</td>
<td>0.94</td>
</tr>
<tr>
<td>Low</td>
<td>0.991</td>
<td>-3.336</td>
<td>1.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Below Opt</td>
<td>0.994</td>
<td>-3.289</td>
<td>2.01</td>
<td>1.01</td>
</tr>
<tr>
<td>Optimal</td>
<td>0.995</td>
<td>-3.157</td>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>High</td>
<td>0.993</td>
<td>-3.421</td>
<td>1.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Above Opt</td>
<td>0.995</td>
<td>-3.322</td>
<td>2.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

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
Important 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 CT Value

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

Error in the CT Value

Differences Between Calibrants

Summary

Data is collected in the exponential range

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

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

R² > 0.990

Experiment with baselines and CT values

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