Method to Calculate Biological Group Variation for Real-Time PCR Gene Expression Experiments

Version 1.0 2-7-2020

Introduction

Ct is the method of choice for real-time PCR quantification. Quantification experiments involving biological samples typically consist of two types of assays. Target assays are designed to amplify and detect gene sequences of scientific interest. Control assays are designed to amplify a gene, RNA or DNA, whose measure is used to normalize the data for sample mass.

In real-time PCR, gene expression is an application where changes in RNA transcript quantities are measured. Real-time PCR data reported as molecules of DNA or RNA target qualifies as absolute quantification. Relative quantification occurs when the data is uncalibrated or when the calibration unit is arbitrary. Most gene expression experiments are designed to produce relative quantities, so the term RQ will be used in this document, but the same principles described below apply to absolute quantities (AQ) as well.

A common practice in relative quantitative real-time PCR experiments is to designate one sample, called the “calibrator sample” or “reference sample,” to calibrate the data set. If the molecular concentration of the target is unknown in the calibrator sample, it is a relative calibrator. The calibrator sample value is normalized to itself, resulting in an RQ of 1, and to all other sample values in the experiment, producing RQ values relative to 1. Calibrating to a value of 1 makes RQ comparisons easier.

Biological variation for gene expression can be defined as the natural variation in target expression across different samples belonging to the same group, referred to as a “biogroup.” For example, target expression will vary in liver samples from multiple mice, even if they are from the same in-bred strain and treated the same way. Gene expression experiments often involve multiple biogroups that are compared to each other, e.g. control group and treated group. The average RQ of each biogroup can be compared to obtain a fold-change.

Average biogroup RQ values can be calibrated to one group to make biogroup RQ comparisons easier. In that case, one biogroup is selected as the calibrator group, often the control group. The average calibrator group RQ is divided by itself, producing an RQ of 1, and the average RQ values from other biogroups. Sample calibration is irrelevant when performing biogroup calibration, because biogroup calibration overrides any sample calibration.

If a fold change in target expression is observed between biogroups, was the treatment responsible for that change? A statistical test is needed to assess whether the observed fold change could reasonably be accounted for by biological variation. If the answer is no, in a well-designed experiment the treatment is thought to be responsible for the fold change. A measure of biological variation of each biogroup is required to perform such a test.

One approach to measure biological variation in gene expression experiments is to calculate standard deviation (SD) RQ. Use of SD RQ assumes biogroup variation is normally distributed. In real-time PCR experiments, it is unlikely that biogroup distribution shapes could be determined with high confidence, due to the large number of samples required to provide that confidence.

Coefficient of variation (CV) is a measure of the variation of a normally distributed population. Experimental results provide only an estimate of biogroup variation due to the impact of random variation from the processes involved in generating the results. An important use of CV values is to monitor for abnormally large or small biogroup variations, as either case will cause problems for statistical testing. Abnormal biogroup variation data should not be used for statistical testing.

The method to calculate SD RQ for biogroups shown below is based on the analysis described in example 4 of Schmittgen and Livak[1], in which SD RQ is calculated from quantities instead of Ct or ΔCt values.

1. Average ΔCt is calculated for each sample.

   Singleplex (one assay per well)
   Avg Ct = Avg Ct_Target – Avg Ct_Control

   Multiplex (Target and Control assays in the same well)
   Ct_well = Ct_Target – Ct_Control from the same well.
   Avg Ct = Avg Ct_well of all wells (replicates) from the same sample.

2. Uncalibrated RQ values are calculated for each sample.
   Sample RQ = $2^{-\text{Avg } \Delta Ct}$

3. RQ standard deviation is calculated for each biological group.
   For example, Excel provides a STDEV function using the formula:

$$\sqrt{\frac{\sum (x-\bar{x})^2}{(n-1)}}$$

   For biogroup SD RQ:

   x is the sample RQ
   $\bar{x}$ is group average RQ
   n is the number of samples in the biogroup

4. All average biogroup RQ values and biogroup SD RQ values are normalized to the group calibrator RQ.

5. CV = SD RQ/Average Biogroup RQ

An example is provided in the table below to illustrate the method.

Control Group (Calibrator Group)			Treated Group		
	Average ΔCt	RQ ($2^{-\Delta Ct}$)		Average ΔCt	RQ ($2^{-\Delta Ct}$)
Sample C1	2.0	0.25	Sample T1	0.5	0.71
Sample C2	2.5	0.18	Sample T2	0.6	0.66
Sample C3	2.2	0.22	Sample T3	0.3	0.81
Sample C4	1.8	0.29	Sample T4	0.4	0.76
Sample C5	1.9	0.28	Sample T5	0.8	0.57
RQ Mean		0.244	RQ Mean		0.7
RQ SD*		0.045	RQ SD		0.093
Calibrated RQ		1			2.87
Calibrated RQ SD		0.18			0.38
CV without calibration		0.18			0.13
CV with calibration		0.18			0.13

References

1. Schmittgen TD and Livak KJ (2008): Analyzing real-time PCR data by the comparative Ct method. Nature Protocols 3(6), 1101-1108

2. User Bulletin: Applied Biosystems Real-Time PCR Systems - Relative Quantification (RQ) Algorithms in Applied Biosystems Real-Time PCR Systems Software. July 2007, Part Number 4378622.

3. Livak KJ and Schmittgen TD (2001): Analysis of Relative Gene Data Using Real-Time Quantitative PCR and the 2-ddCt Method. Methods 25, 402-408