Statistical analysis of relative expression results in real-time PCR: Development of a Relative Expression Software Tool (REST-2007©)

Ross Saad[1], David Chiew[1], Matthew Herrmann[1], Valin Reja[1], Michael W. Pfaffl[2]
(1) Corbett Life Science, (2) Technical University of Munich

Relative expression is not enough

When comparing treated and untreated samples, relative quantitation using qRT-PCR allows the estimation of gene expression (1). This technique provides relative expression; a single value that indicates the relative quantities of untreated and treated sample groups.

$$Relative\ Expression\ = \frac{E_{target}^{\Delta C_{T\ target\ (control - sample)}}}{E_{target}^{\Delta C_{T\ ref\ (control - sample)}}}$$

Figure 1: Calculation of relative expression.

While useful, it does not provide statistical information on the variation in relative expression, nor does it provide a means of understanding the significance of the result.

Bootstrapping is key

Because simple ratios are used in relative expression, it becomes very complex to perform traditional statistical analysis, e.g. ratio distributions do not have a standard deviation. By using bootstrapping techniques(2) we were able to overcome many assumptions of normality made with traditional statistical tests.

Confidence in relative expression

By using bootstrapping techniques we were able to generate a rich distribution curve of relative expression. With this distribution we were then able to calculate standard error and confidence intervals for the original relative expression value.
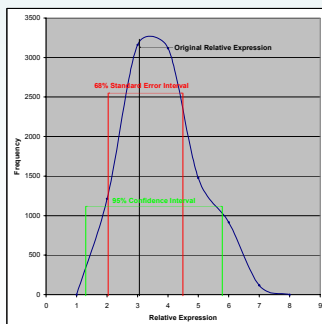

Figure 2: Example frequency distribution chart of relative expression.

We use the following method to generate our distribution curve:

- Given the C_Ts and efficiencies of treated and untreated samples, pick a random pair of C_Ts from our given sample set.
- Feed this random C_T pair into the relative expression formula (Figure 1), this represents a single instance in the frequency distribution chart in Figure 2.
- This process of picking random C_T pairs is repeated many times.
- Once a large data set of relative expressions has been generated, we then numerically sort the data set.
- Having a sorted data set allows us to see where 95% of the population lies, and determine the limits where this occurs – the 95% confidence interval.

How significant are the results?

Our primary goal was to determine if there is a significant difference between treated and untreated samples. We set up the null hypothesis (H_0): the difference between treated and untreated groups are only due to chance (3). The aim was to disprove H_0 which will suggest that the result is significant.

If we assume H_0 is true, then we should be able to randomly swap values between the two groups and not see a greater difference than what we see between the labelled groups.

We again used bootstrapping techniques to attempt to disprove H_0 as follows:

- Randomly reclassify the entire C_T sample set as a treated or untreated sample.
- Check to see if the relative expression of this random set is larger than the original relative expression.
- Repeat this process many times, and keep track of how often the above condition is met.
- Once the process is complete we can generate $P(H_0)$ the probability that H_0 is true. This is calculated using the number of occurrences of the above condition.

The closer that $P(H_0)$ is to zero, the more confident we can be that the difference is not only due to chance, hence our result being significant.

Development of REST 2007

As a result of this work, the REST© (relative expression software tool) was developed, a tool that would easily allow the researcher to determine significance and confidence measurements in their studies. REST 2007 is the latest installment of the REST series, building on its predecessor REST 2005 (4) with significant improvements to randomization algorithms. The new revision introduces alternative data inputs such as single sample efficiency and amplification take-off points, alleviating the need to set amplification plot thresholds.

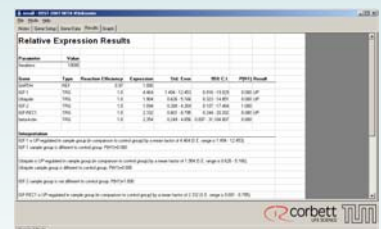Figure 3: Box and whisker plots provided by REST 2007.

Figure 4: Results and interpretation provided by REST 2007.

References

1. 6-8; http://docs.appliedbiosystems.com/pebiodocs/04303859.pdf
2. [Davidson] A.C. Davidson & D.V. Hinkley: Bootstrap Methods and their Application (ISBN 0-521-57391-2, Cambridge University Press 2002)
3. [Pfaffl] M. W. Pfaffl, G. W. Horgan & Leo Dempfile: "Relative Expression Software Tool (REST) for group-wise comparison and statistical analysis of relative expression results in Real-Time PCR" (Nucleic Acids Research 2002 May 1; 30 (9): E36)
4. http://rest.gene-quantification.info/