BIOINFORMATICS APPLICATIONS NOTE Vol. 19 no. 11 2003, pages 1439–1440
DOI: 10.1093/bioinformatics/btg157

Q-Gene: processing quantitative real-time RT–PCR data

Perikles Simon

Section for Neurobiology of the Eye, University Eye Hospital Tuebingen, Calwerstr. 7/1, 72076 Tuebingen, Germany

Received on December 6, 2002; revised and accepted on January 30, 2003

ABSTRACT

Summary: *Q-Gene* is an application for the processing of quantitative real-time RT–PCR data. It offers the user the possibility to freely choose between two principally different procedures to calculate normalized gene expressions as either means of Normalized Expressions or Mean Normalized Expressions. In this contribution it will be shown that the calculation of Mean Normalized Expressions has to be used for processing simplex PCR data, while multiplex PCR data should preferably be processed by calculating Normalized Expressions. The two procedures, which are currently in widespread use and regarded as more or less equivalent alternatives, should therefore specifically be applied according to the quantification procedure used.

Availability: Web access to this program is provided at http://www.biotechniques.com/softlib/qgene.html

Contact: perikles@uni-tuebingen.de

Quantification of RNA levels using real-time RT–PCR has considerable potential for a variety of biomedical applications (Walker, 2002). In principle this method allows for online monitoring of the increase of amplicons generated by PCR and offers a possibility to determine the relative starting concentrations of different RNAs. Typically, the expression level of a target gene is normalized to the expression level of a reference or, for absolute quantification, to a standard of known copy number (Pfaffl *et al.*, 2002; Vu *et al.*, 2000). This process of normalization usually takes the different efficiencies of PCR amplification for the target (E_{target}) and the reference ($E_{reference}$) into consideration and transforms the logarithmic scaled raw data unit Cycle Threshold (CT) into the linear unit of normalized expressions.

Since marginal differences in the procedural course of a PCR lead to noticeable changes in results, repetitions of the same probes are used to determine the variability of real-time PCR procedures. One of the major hurdles of the new technology is the evaluation and the mathematical and statistical analysis of data, especially since this issue is still a matter of debate (Bustin, 2002).

Currently two different data processing procedures are used, which transform means of raw data CT values and the related standard errors (SEs) into means of normalized expression levels and their respective SEs. Both of these procedures are provided by the software package *Q-Gene* and regarded as equivalent alternatives (Muller *et al.*, 2002). In the current analysis major differences between both procedures will be addressed.

The first procedure calculates in a series with *n* repetitions 'Normalized Expressions (NEs)' according to Equation (1) by *n* pairwise comparisons of a $CT_{reference}$ with a CT_{target} of the same series. Afterwards a mean and an SE for a repetition series are calculated using the NEs of this series.

$$NE = \frac{(E_{reference})^{CT_{reference}}}{(E_{target})^{CT_{target}}} \quad (1)$$

The second procedure calculates means and SE for the references (CTreference, mean and $SE_{CTreference,mean}$) as well as for the targets (CTtarget,mean and $SE_{CTtarget,mean}$) and determines a Mean Normalized Expression (MNE) according to Equation (2) subsequently.

$$MNE = \frac{(E_{reference})^{CTreference,mean}}{(E_{target})^{CTtarget,mean}} \quad (2)$$

By applying the differential equation of Gauss for error propagation the SE for this MNE (SE_{MNE}) is calculated according to Equation 3.

$$SE_{MNE} = MNE \cdot \left((\ln(E_{target}) \cdot SE_{CTtarget,mean})^2 + (\ln(E_{reference}) \cdot SE_{CTreference,mean})^2\right)^{\frac{1}{2}} \quad (3)$$

Table 1 shows a fictitious example of a triplicate of a reference compared to a triplicate of a target gene, in which the data were processed by *Q-Gene* calculating NEs and SE_{NE}, or MNEs and SE_{MNE}. It is first of all assumed that raw data were obtained by performing a simplex real-time RT–PCR, where the CTs of three repetitions for the reference and three for the target were determined in six independent wells. For Case 1 and 2 E_{target} and $E_{reference}$ were 2.0 and 1.9 and for Case 3 they were 1.9 both.

Table 1. Comparison of the calculation of normalized expressions with mean normalized gene expressions for three cases

	CT target gene	CT reference gene	Normalized expression (NE)	Mean normalized expression (MNE)
Case 1 ($E_{target} \neq E_{reference}$)				
Repetition 1	24.60	17.10	$1.95E{-}02_1$	
Repetition 2	24.50	17.40	2.56E−02	
Repetition 3	24.20	17.50	3.33E−02	
Mean	24.43	17.33	2.61E−02	$2.55E{-}02_2$
SE	0.12	0.12	3.98E−03	$2.90E{-}03_4$
SE%	**0.49**	**0.69**	**15.24**	**11.35**
Case 2 ($E_{target} \neq E_{reference}$)				
Repetition 1	24.60	17.50	2.57E−02	
Repetition 2	24.50	17.40	2.56E−02	
Repetition 3	24.20	17.10	2.52E−02	
Mean	24.43	17.33	2.55E−02	2.55E−02
SE	0.12	0.12	1.57E−04	2.90E−03
SE%	**0.49**	**0.69**	**0.62**	**11.35**
Case 3 ($E_{target} = E_{reference}$)				
Repetition 1	24.60	17.50	1.05E−02	
Repetition 2	24.50	17.40	1.05E−02	
Repetition 3	24.20	17.10	1.05E−02	
Mean	24.43	17.33	1.05E−02	1.05E−02
SE	0.12	0.12	0	1.14E−03
SE%	**0.49**	**0.69**	**0**	**10.91**

[1,2 and 3]Values were calculated according to the respective Equations (1), (2) and (3).

Case 1 illustrates, that due to the transformation of the raw data the relative SE (SE%) increases from a level below 1% for reference and target to 15.24% for SE_{NE} and 11.35% for SE_{MNE}. Additionally the mean of NEs (2.61E−02) and MNEs (2.55E−02) differ by 2.3%. In Case 2 only the order of two reference values was switched. It becomes evident that a simple rearrangement of the raw data decreases the SE_{NE}, while the SE_{MNE} does not change. Moreover, if both efficiencies are set to 1.9 as in the third case, the SE_{NE} reaches zero. Again the SE_{MNE} changes only slightly, due to the manipulation of $E_{control}$.

The presented analysis shows that the value of SE_{NE} scatters rather according to the selection of target and reference values, which are compared to each other, than reflecting the scattering of raw data. Therefore it has to be concluded, that the calculation of NE and SE_{NE} is an inappropriate way of processing simplex quantitative real-time RT–PCR data. The calculation of MNE and SE_{MNE}, is therefore proposed for simplex PCR data processing.

In multiplex PCR, however, the analysis calculating NE and SE_{NE} is useful, since CTs of target and internal standard are measured in the same well directly. Therefore, the links of data in pair-wise comparison of internal standard and target are well defined by the starting conditions of the experiment. It remains to be studied whether this difference contributes to a higher sensitivity of the multiplex approach. It is hoped that these findings will further alleviate data processing for quantitative real-time PCR users and will be considered for determining the upper limits of sensitivity of different procedures.

ACKNOWLEDGEMENTS

I would like to thank Frank Schaeffel for critically reading the manuscript.

REFERENCES

Bustin,S.A. (2002) Quantification of mRNA using real-time reverse transcription PCR (RT–PCR): trends and problems. *J. Mol. Endocrinol.*, **29**, 23–39.

Muller,P.Y., Janovjak,H., Miserez,A.R. and Dobbie,Z. (2002) Processing of gene expression data generated by quantitative real-time RT–PCR. *Biotechniques*, **32**, 1372–1379.

Pfaffl,M.W., Horgan,G.W. and Dempfle,L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.*, **30**, e36.

Vu,H.L., Troubetzkoy,S., Nguyen,H.H., Russell,M.W. and Mestecky,J. (2000) A method for quantification of absolute amounts of nucleic acids by (RT)-PCR and a new mathematical model for data analysis. *Nucleic Acids Res.*, **28**, E18.

Walker,N.J. (2002) A technique whose time has come. *Science*, **296**, 557–559.