



Estimation of the Relative Sensitivity of qPCR Analysis Using Pooled Samples

Ana Muniesa¹, Chelo Ferreira², Héctor Fuertes¹, Nabil Halaihel¹, Ignacio de Blas^{1*}

¹ Department of Animal Pathology, Facultad de Veterinaria, Universidad de Zaragoza, Zaragoza, Spain, ² Department of Applied Mathematics, Facultad de Veterinaria, Universidad de Zaragoza, Zaragoza, Spain

Abstract

The high sensitivity of qPCR makes it a desirable diagnostic method in epidemiological surveillance programs. However, due to high costs, the use of pooling has been suggested. In this paper, an algorithm based on the Montecarlo method has been designed and implemented. The algorithm had been tested in many different situations, and finally it was validated with a real dataset. Moreover, based on the results obtained and depending on pooling conditions, a drastic decrease of sensitivity is observed.

Citation: Muniesa A, Ferreira C, Fuertes H, Halaihel N, de Blas I (2014) Estimation of the Relative Sensitivity of qPCR Analysis Using Pooled Samples. PLoS ONE 9(4): e93491. doi:10.1371/journal.pone.0093491

Editor: Joao Inacio, University of Brighton, United Kingdom

Received: December 17, 2013; **Accepted:** March 5, 2014; **Published:** April 10, 2014

Copyright: © 2014 Muniesa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Ministerio de Educación of Spain is funding a FPU/MEC doctorate fellowship for Ana Muniesa. Also the Departamento de Industria e Innovación del Gobierno de Aragón, Fondo Social Europeo and Dirección General de Ciencia y Tecnología (REF. MTM2010-21037) provides financial support to Research Groups A35 and T53 for their basic running costs. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: deblas@unizar.es

Introduction

The Polymerase Chain Reaction (PCR) is one of the most powerful technologies in molecular biology. Using PCR, specific sequences within a DNA or cDNA template can be copied, or “amplified”, many thousand- to a million- fold. PCR is a technique requiring a specific fragment of DNA, and it is useful for different applications: to identify anomalies in the sequence of nucleotides that point to possible genetic diseases [1], to identify an individual or to determine their relationships with others or to detect the presence of DNA of microorganisms useful in the diagnosis of disease or for testing the effectiveness of a treatment [2].

In traditional (endpoint) PCR, detection of the amplified sequence are performed at the end of the reaction after the last PCR cycle, and involve post-PCR analysis such as gel electrophoresis and image analysis.

In real-time quantitative PCR (qPCR), the amount of PCR product is measured at each cycle by the use of fluorescent markers that are incorporated into the PCR product [3], [2]. The increase in fluorescent signal is directly proportional to the number of PCR product molecules (amplicons) generated in the exponential phase of the reaction. Fluorescent reporters used include double-stranded DNA (dsDNA)-binding dyes, or dye molecules attached to PCR primers or probes that are incorporated into the product during amplification. The change in fluorescence over the course of the reaction is measured by an equipment that combines thermal cycling with scanning capability. By plotting fluorescence against the cycle number, the qPCR equipment generates an amplification plot that represents the accumulation of product over the duration of the entire PCR reaction (Figure 1).

The cycle threshold (C_t) or cycle quantification (C_q) records the cycle when the sample fluorescence exceeds a chosen threshold

above background fluorescence. This value is correlated with the number of copies of the target sequence originally present in the reaction mixture [4]. The samples with a high number of initial copies of target nucleic acid, are detected sooner and therefore they will have low C_t values (usually around 20–25). However, those samples with very low numbers of copies are later detected, and the C_t values are above 30–35 [5]. The sample is defined as positive when the C_t analyzed by the qPCR technique is less than the established C_t (threshold value). In other cases it is considered as negative.

Using the Standard Curve Method based on known quantities, it is possible to extrapolate a value of a sample. The target DNA gene copies in the pathogen are to be considered to determine the absolute number of the agent in the processed sample, so qPCR provides us the number of copies of a particular pathogen obtained from a sample of an infected individual. The slope of the linear regression curve determines the efficiency of amplification, which is 100% if a dilution of 1:2 results in a C_t difference of 1 [6].

Currently PCR is the best-known and most successfully implemented diagnostic molecular technology. PCR, specifically qPCR, can detect slow-growing or difficult-to-culture microorganisms and can be used in situations in which clinical microbiology diagnostic procedures are inadequate, time-consuming, difficult, expensive, or hazardous to laboratory staff [2]. The analytical specificity and sensitivity of qPCR assay is considered as perfect for diagnostic of clinical cases (i.e. identification of bovine mastitis pathogens [7]). A general review over the use of qPCR in clinical microbiology testing showed an increased specificity and sensitivity over standard serological tests or culturing methods [8], and for these reasons qPCR is considered as “gold standard” for direct diagnosis in most of pathogens.

The high sensitivity of qPCR makes it as a desirable diagnostic method to use in epidemiological surveillance programs in animal

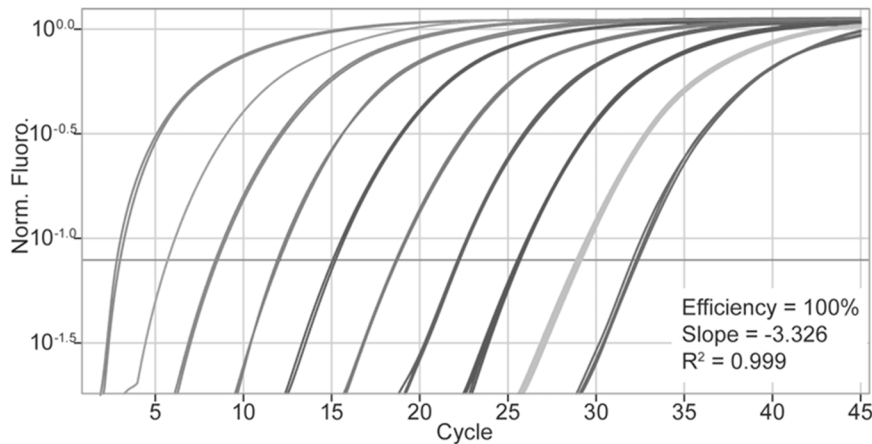


Figure 1. Amplification plots represent the accumulation of product over the duration of the real-time PCR experiment.
doi:10.1371/journal.pone.0093491.g001

health [9], [10], but qPCR is a relatively expensive technique that limits its generalized application. In order to minimize this problem, the use of pooled samples has been suggested [10]. Thus, it can result in major savings (consumables and labor), and reduced costs [8], [11]. So pooling is now routinely used for health status monitoring purposes.

The theoretical probability of including at least one infected individual in a pool (p_{infected}) is increased when the pool size (n_{pool}) is bigger and the prevalence (P) is higher. This probability is calculated as [12], [13]:

$$p_{\text{infected} \geq 1} = 1 - (1 - P)^{n_{\text{pool}}} \quad (1)$$

However, we carefully note the significant decrease in pooled sensitivity due to the dilution effect. Two factors should be taken into account: the low proportion of infected samples in the pool (i.e. pools of big size from a population with low prevalence) and the low number of DNA copies of the infected individuals (i.e. low pathogen loads in individuals from asymptomatic populations [8]). Unfortunately, this is a common scenario of most of the epidemiological surveillance programmes: low prevalence of asymptomatic infected animals in the investigated population.

In most of cases qPCR is considered as gold standard (it means that sensitivity and specificity are perfect). However, it is not completely true and the accuracy may be unknown, so it would be possible to estimate the relative sensitivity of pooled qPCR, assuming individual qPCR as gold standard. An algebraic solution of this problem is not possible and a simulation procedure is suggested.

The objective of this paper is to estimate the relative sensitivity (rS) of a qPCR analysis of a pooled sample.

Materials and Methods

We have designed the following stochastic algorithm using the Montecarlo method. For the later convenience, we define the following variables:

- n_{pool} : pool size
- iCt : cycle threshold (defined for diagnosis of an individual sample)
- mCt : mean of Ct in an infected and asymptomatic population.
- $difCt = iCt - mCt$

- sCt : standard deviation of Ct in an infected and asymptomatic population
- P : prevalence of infection in an asymptomatic population

Firstly, we define a pooled sample as a mix of n_{pool} individual samples.

For each of the individual samples ($k; k = 1, \dots, n_{\text{pool}}$), the infection status (infected/not infected) is randomly determined as a function of the prevalence (P).

For non-infected individual samples, the pathogen load is assumed as zero ($L_k = 0$). For infected samples we calculate a random Ct assuming a gaussian distribution with a mean (mCt) and standard deviation (sCt) given

$$Ct_k = \text{random}(\mathcal{N}(mCt, sCt)). \quad (2)$$

Next, the load of an infected individual sample (L_k) is estimated by ΔCt -method [14].

$$L_k = 2^{(iCt - Ct_k)}. \quad (3)$$

The load of the pooled sample (L_{pool}) is estimated as the average of the individual samples loads (L_k)

$$L_{\text{pool}} = \frac{1}{n_{\text{pool}}} \sum_{k=1}^{n_{\text{pool}}} L_k. \quad (4)$$

Then, the Ct_{pool} is given as

$$Ct_{\text{pool}} = \log_2 L_{\text{pool}}. \quad (5)$$

As we have explained in the introduction section, when the Ct_{pool} is lower than the iCt the pool is considered as positive; in other case, as negative.

During the simulation, the algorithm is iterated until a desired quantity of infected pools (I_{pool}) is reached. So, the number of simulations depends on the required precision for the relative sensitivity. We consider as infected pool any pool that included at

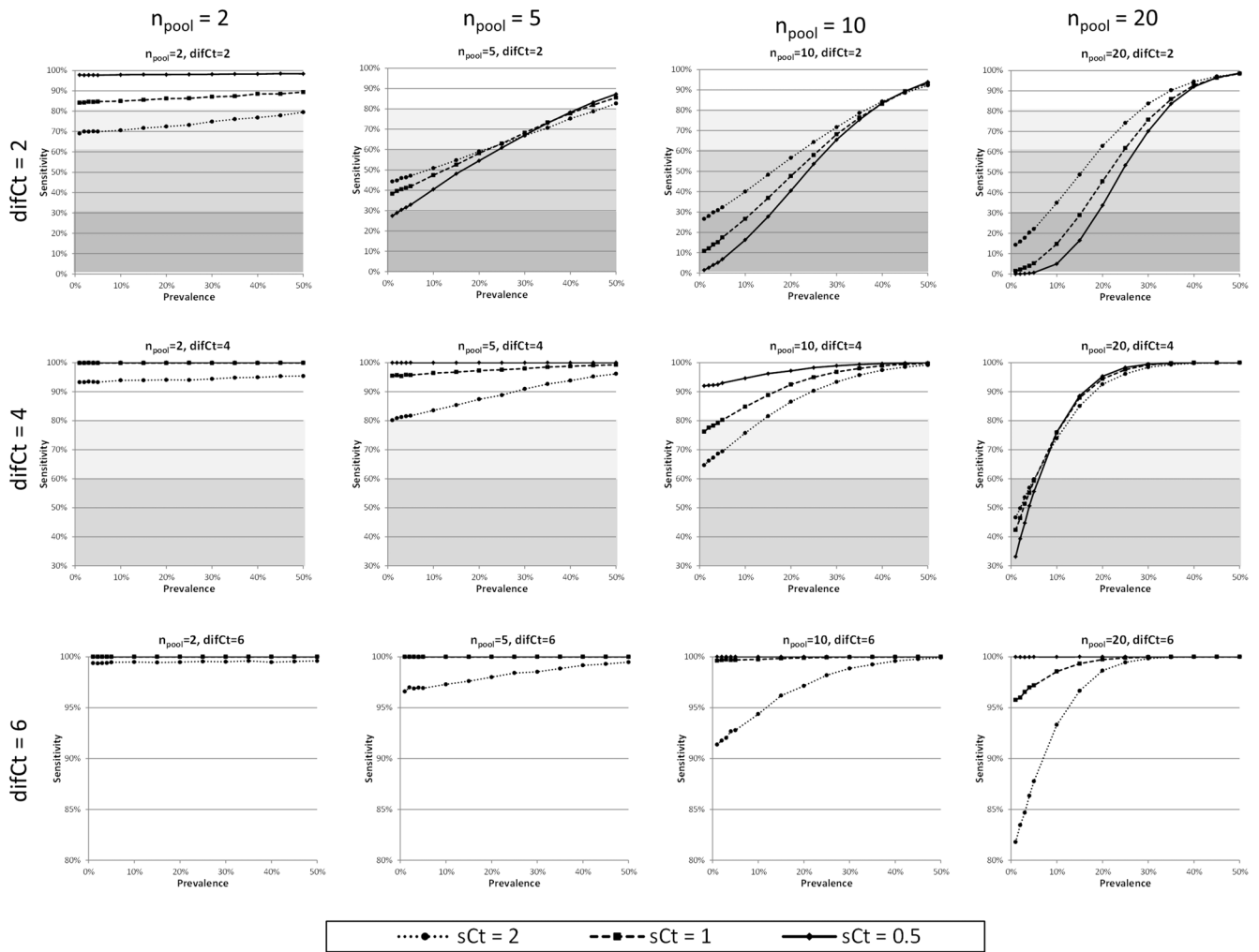


Figure 2. Relative sensitivities corresponding to the simulation results for different scenarios.
 doi:10.1371/journal.pone.0093491.g002

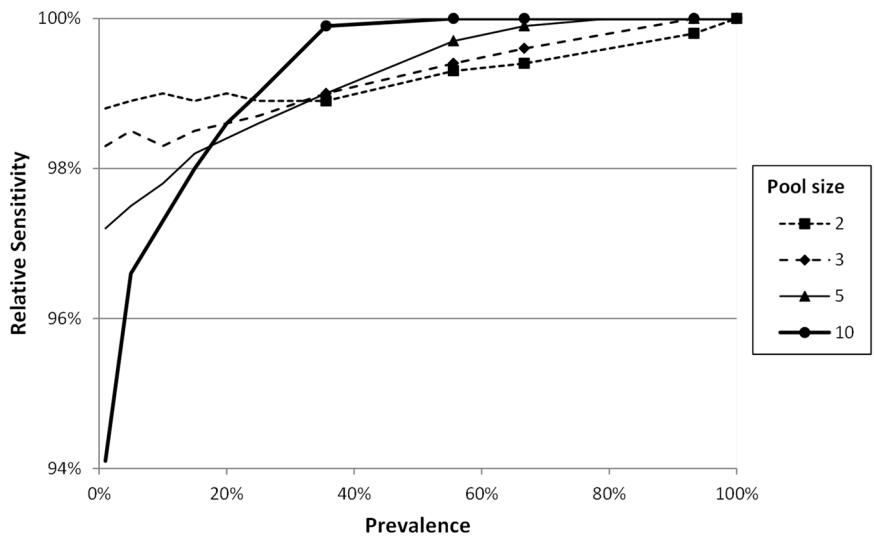


Figure 3. Relative sensitivities corresponding to a pooled diagnostic for PRRSV using global samples of Ct ($mCt = 27.75$, $sCt = 4.625$)
 (data from Gerber et al, 2013)[8].
 doi:10.1371/journal.pone.0093491.g003

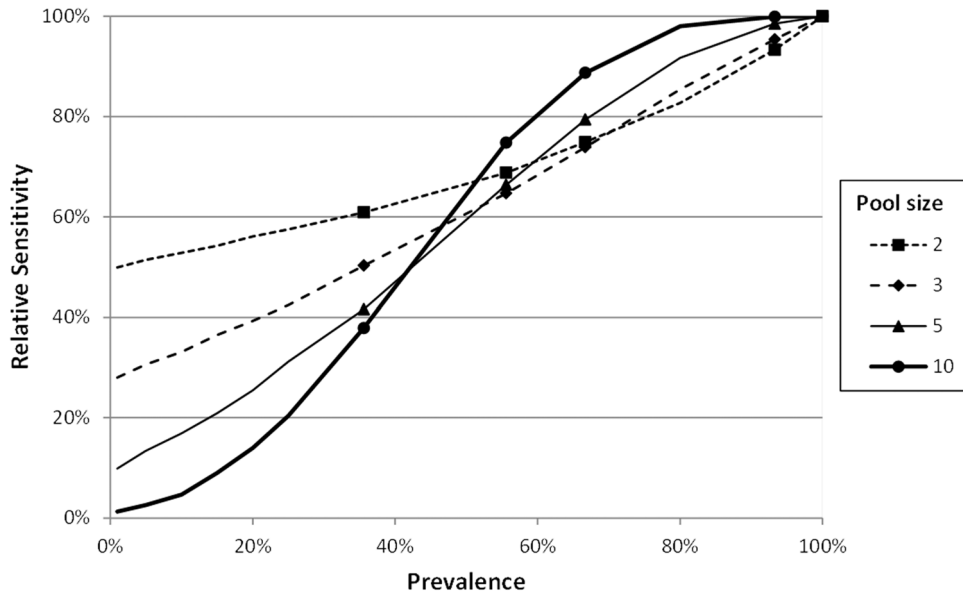


Figure 4. Relative sensitivities corresponding to a pooled diagnostic for PRRSv using low load samples of Ct ($mCt=36.0$, $sCt=1$) (data from Gerber et al, 2013)[8].

doi:10.1371/journal.pone.0093491.g004

least one infected individual sample. Moreover, following the previous criteria, the number of positive infected pools (P_{pool}) is determined. Therefore, the relative sensitivity is estimated as [12]

$$rS = \frac{P_{pool}}{I_{pool}} \quad (6)$$

The algorithm is implemented with *php* language (it is possible to obtain the code just asking for the authors). It has also been implemented in a web page (<http://www.winepi.net/f302.php?ID=2>) in order to make it available to the scientific community and biomedical practitioners. Accuracy of the results depends on the number of iterations.

In order to validate the algorithm, we have used published data about prevalence of PRRS and QPCR results [8].

Results and Discussion

The numerical experiments have been given for the all combinations of the next values of the variables n_{pool} (2, 5, 10, 20), $difCt$ (2, 4, 6), sCt (0.5, 1, 2), and P (1, 2, 3, 4, 5, 10, 15, ..., 50). The results showed in the graphics are the average of three simulations of 10000 iterations each one.

By direct observation of the Figure 2, we can extract the following statements:

- Both, low n_{pool} as well as high $difCt$, provide high sensitivity.
- With high values of $difCt$ (>5), the influence of n_{pool} is low, except for high n_{pool} with low prevalence.
- The decreasing values of $difCt$, the increasing influence of n_{pool} . Therefore, higher values of n_{pool} , and low prevalence, lower values of the sensitivity.
- The extreme cases are situated at the right upper zone of the Figure 2 ($n_{pool} \geq 5$ and $difCt \leq 4$). The sensitivity values are only

acceptable for very high prevalence (highly endemic diseases and epidemic outbreaks).

- Finally, the lower prevalence, the higher effect of the standard deviation effect.

In order to assess the consequences of pooling, we used real data from the experimental work of Gerber et al[8], about the qPCR diagnosis of PRRSv with pooled samples. Based on the individual diagnostic results from serum, the prevalences of infection, for days 1, 3, 5, 7, 14 and 21 post-infection (p.i.), were calculated. And the variabilities of Ct , in all samples globally and in a group defined as low load, were estimated.

Firstly, the observed prevalences varied from 66.7% (day 1 p.i.) to 93.3–100% during the acute phase (days 3, 5 & 7 p.i.). Then, they decreased progressively to 55.6% (day 14 p.i.) and 35.6% (day 21 p.i.). When we estimated the global Ct ($mCt=27.75$, $sCt=4.625$), the high variability of the pathogen load is observed. Next, we applied our algorithm, for prevalences from 0% to 100%, to estimate the relative sensitivity in these conditions (Figure 3). In that figure, the specific values corresponding to the prevalences in 1, 3, 5, 7, 14, 21 days p.i. were marked.

The relative sensitivities calculated with our method were over to 98%. Therefore, it is consistent with the results of [8].

However, these authors described a group of samples with low pathogen load ($mCt=36$, $sCt=1$). In the acute phase the relative sensitivity was greater than 90% but the marks corresponding to the 1, 14 and 21 days p.i. (early infection and recovery) where it was from 40 to 80% (Figure 4). And also this is consistent with the results of [8].

The use of pooled samples could be a good strategy in order to reduce analytical cost in surveillance programmes, but loss of sensitivity could be a critical issue due to existence of false negative results.

By way of conclusion, the effect of n_{pool} on the relative sensitivity depends on such as the values of the prevalence as the quantity of pathogen load.

Author Contributions

Conceived and designed the experiments: AM CF HF NH IdB. Performed the experiments: AM CF IdB. Analyzed the data: AM CF IdB.

Contributed reagents/materials/analysis tools: AM CF HF NH IdB. Wrote the paper: AM CF IdB.

References

1. Reiss J, Cooper DN (1990) Application of the polymerase chain reaction to the diagnosis of human genetic disease *Hum Genet* 85(1): 1–3.
2. Yang S, Rothman RE (2004) PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. *Lancet Infect Dis* 4(6): 337–348.
3. Abbott MA, Poiesz BJ, Byrne BC, Kwok S, Sninsky JJ, et al. (1988) Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in vitro. *J Infect Dis* 158: 1158–1169.
4. Bustin SA, Nolan T (2004) Pitfalls of quantitative reverse transcription polymerase chain reaction. *J Biomol Tech* 15: 155–166.
5. Mehta T, McGrath E, Bheemreddy S, Salimnia H, Abdel-Haq N, et al. (2010) Pranatharthi Chandrasekar and George J. Alangaden. Detection of Oseltamivir Resistance during Treatment of 2009 H1N1 Influenza Virus Infection in Immunocompromised Patients: Utility of Cycle Threshold Values of Qualitative Real-Time Reverse Transcriptase PCR. *J Clin Microbiol* 48(11): 4326–4328.
6. Carr AC, Moore SD (2012) Robust Quantification of Polymerase Chain Reactions Using Global Fitting. *PLoS ONE* 7(5): e37640.
7. Koskinen MT, Holopainen J, Pyörälä S, Bredbacka P, Pitkälä A, et al. (2009) Analytical specificity and sensitivity of a real-time polymerase chain reaction assay for identification of bovine mastitis pathogens. *J Dairy Sci* 92(3): 952–959.
8. Gerber PF, O'Neill K, Owolodun O, Wang C, Harmon K, et al. (2013) Comparison of Commercial Real-Time Reverse Transcription-PCR Assays for Reliable, Early, and Rapid Detection of Heterologous Strains of Porcine Reproductive and Respiratory Syndrome Virus in Experimentally Infected or Noninfected Boars by Use of Different Sample Types. *J Clin Microbiol* 51(2): 547.
9. OIE (2013) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals Paris: World Organisation for Animal Health. Available: <http://www.oie.int/en/international-standard-setting/terrestrial-manual/access-online/>. Accessed 25 February 2014.
10. OIE (2013) Manual of Diagnostic Tests for Aquatic Animals. Paris: World Organisation for Animal Health. Available: <http://www.oie.int/en/international-standard-setting/aquatic-code/access-online/>. Accessed 25 January 2014.
11. Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, et al. (2006) Real-Time PCR in Clinical Microbiology: Applications for Routine Laboratory Testing. *Clin Microbiol Rev* 19(1): 165–256.
12. Thrusfield M (2010) *Veterinary Epidemiology* (3rd edition). Oxford: Blackwell Science Ltd. 610 p.
13. Martin SW, Shoukri M, Thorburn MA (1992) Evaluating the health status of herds based on tests applied to individuals. *Prev Vet Med* 14: 33–43.
14. Scheffe JH, Lehmann KE, Buschmann IR, Unger T, Funke-Kaiser H (2006) Quantitative real-time RT-PCR data analysis: current concepts and the novel gene expressions CT difference formula. *J Mol Med* 84(11): 901–910.