

Notes & Tips

The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization

K. Dheda^{a,b}, J.F. Huggett^{a,*}, J.S. Chang^a, L.U. Kim^a, S.A. Bustin^c, M.A. Johnson^b,
G.A.W. Rook^a, A. Zumla^a

^a Centre for Infectious Diseases and International Health, Royal Free and University College Medical School, London W1T 4JF, UK

^b Department of Thoracic and HIV Medicine, Royal Free Hospital, London NW3 2QG, UK

^c Centre for Academic Surgery, Barts and the London, Queen Mary's School of Medicine and Dentistry, London E1 2AD, UK

Received 3 March 2005

Available online 8 June 2005

Real-time reverse transcription PCR (RT-PCR)¹ has become a routine technique for gene expression analysis, allowing accurate high-throughput RNA quantification over a wide dynamic range at a relatively low cost. Although the technique has become well established over the past few years, data normalization remains problematic and is subject to frequent criticism. Normalization is essential to control for experimental error between samples that can be introduced at a number of stages throughout the procedure. Several different strategies can be applied to the normalization of the expression of an RNA of interest (for a review, see [1]). The most common strategy is to use a reference gene as an internal standard that is assumed to remain constant between experimental groups. It is now clear, however, that in most experimental situations the use of conventional reference genes, such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β -actin, is inappropriate due to their variability [1,2]. If the reference gene chosen has a large expression error, the noise of the assay is increased and detection of small changes becomes unfeasible. Far worse, if expression of the reference gene is altered by the experimental condition or the disease under study, the results obtained may be entirely incorrect. Therefore, it is essential to validate potential reference genes to establish whether they are appropriate

for a specific experimental purpose. We have shown previously that human acidic ribosomal protein (HuPO) was the least variable reference gene investigated in our model of pulmonary tuberculosis (PTB) [3].

Despite the fact that many validation exercises have been undertaken confirming the variability of conventionally used reference genes [1], it has been argued that overall study conclusions remain unchanged because this variability is often similar in the control and intervention groups [4]. Furthermore, reports that address the net effect of reference gene variability on experimental conclusions are sparse and do not provide the final proof that reference gene variability may have a variable influence on study conclusions. To address the impact of a validation exercise on intergroup comparisons, we normalized different sets of data with differing reference genes, including conventionally chosen genes. This report demonstrates that experimental results are highly dependent on the reference gene chosen. Moreover, whether statistically significant differences between study groups were present or absent depended on which reference gene was used.

In the study, 21 patients with proven PTB and 21 matched healthy volunteers were recruited. Effector T-cell IFN- γ ELISPOT responses to TB-specific antigens were used to confirm that control subjects were not latently infected with *Mycobacterium tuberculosis* [5]. The relevant ethical review committees approved the study.

Whole blood (2.5 ml) was immediately transferred into PAXgene Blood RNA tubes (PreAnalytix, Qiagen), and total RNA was extracted (Qiagen). Levels of mRNAs encoding IL-4 were measured in whole blood from all patients at baseline (before or within the first 2 weeks of

* Corresponding author. Fax: +44 207 6368175.

E-mail address: j.huggett@ucl.ac.uk (J.F. Huggett).

¹ Abbreviations used: RT-PCR, reverse transcription PCR; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HuPO, human acidic ribosomal protein; PTB, pulmonary tuberculosis.

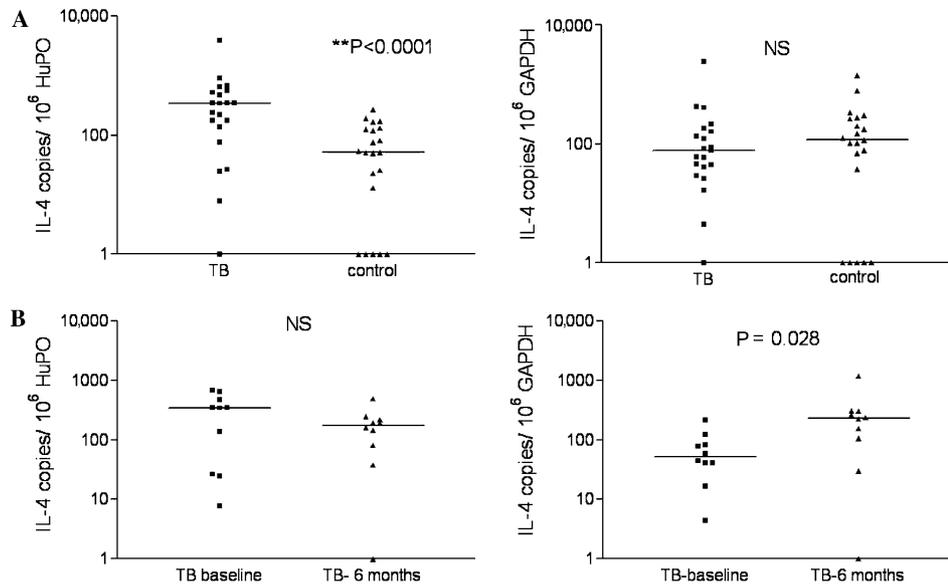


Fig. 1. (A) Differences in IL-4 mRNA expression in patients with PTB and in healthy controls. Intergroup gene expression differences were statistically significant when data were normalized to HuPO (left) but not when data were normalized to GAPDH (right). (B) Differences in IL-4 mRNA expression in untreated patients with PTB and in patients treated for 6 months. Intergroup gene expression differences were statistically significant when data were normalized to GAPDH (right) but not when data were normalized to HuPO (left).

anti-TB treatment) and in 10 patients at approximately 6 months after initiating anti-TB treatment.

The human myelomonocytic cell line, THP1, was cultured in RPMI 1640 medium (supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 10% fetal bovine serum) at 37°C and 5% CO₂. Differentiated macrophages (2 × 10⁶/ml) were treated with 30 µg/ml sonicated *M. tuberculosis* (n = 3) or phosphate-buffered saline (n = 3). Cells were harvested at 6 h, and RNA was extracted using the RNeasy Mini Kit (Qiagen).

Extracted RNA, after on-column DNase digestion (Qiagen), was reverse transcribed in parallel for each experiment (20-µl reactions, Omniscript, Qiagen). The RNA template was qualitatively assessed and quantified using an Agilent Technologies 2100 bioanalyzer. For each cDNA reaction, 150 and 50 ng of input total RNA was used from whole blood and THP1 cells, respectively.

The PCR reactions were performed using the Rotor-gene (Corbett Research). In each reaction, ~15 and ~2.5 ng of reverse transcribed RNA (based on the initial RNA concentration) were used for reactions containing whole blood and THP1 cells, respectively. Reaction parameters and primer sequences can be found as [supplementary material](#). All reactions were run in duplicate and nontemplate controls were used as recommended [6].

Cytokine mRNA values from TB samples were normalized to a conventional reference gene (GAPDH) or a validated reference gene (HuPO). The latter was validated by assessing the variability of 13 reference genes in samples (blood and mononuclear cell cultures) from patients with tuberculosis and from healthy volunteers [3]. Validation of an optimal reference gene for stimulated THP1 cells (n = 6) was achieved by assessing the variability of

HuPO, GAPDH, and β-actin [3]. For the purposes of data analysis, raw data replicates that were nonreproducible and below the detection limit of the assay (10 copies) were given an arbitrary value of 1 copy in the final analysis.

Results were analyzed using the Mann–Whitney *U* test, the Wilcoxon matched-pairs test, and the Student's *t* test.

We (and others) have demonstrated previously that many conventionally used reference genes vary widely in mRNA expression [1,3]. We present here the effect of using a validated reference versus an unvalidated reference on the target gene expression. Validation refers to a reference gene that has been assessed for the degree of fluctuation within an experimental system as well as for “directional shift” (i.e., whether the experimental condition itself up- or down-regulates reference gene expression). Our approach to reference gene validation has been discussed in detail elsewhere [1].

In our first experiment, we showed that an increase in expression of IL-4 in tuberculosis patients can be masked when real-time RT–PCR is used with an inappropriate reference gene (Fig. 1A). Data normalized to HuPO RNA expression resulted in an approximately 7-fold highly significant ($P < 0.0001$) increase in IL-4 mRNA expression in TB that was abrogated when the data were normalized using GAPDH RNA expression. This may partly explain previous conflicting results when looking at IL-4 mRNA expression in tuberculosis [7,8]; the reference genes used in those earlier studies were not validated.

In our second experiment, normalization of IL-4 gene expression with HuPO suggested that anti-TB treatment results in a nonsignificant decrease in IL-4 mRNA expression. When GAPDH was used, an approximately 5-fold significant increase was observed (Fig. 1B). This is biologi-

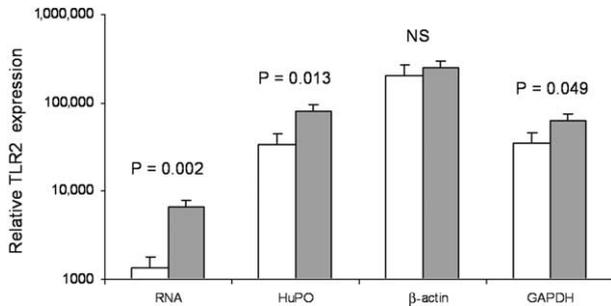


Fig. 2. Differences in TLR2 gene expression with three different reference genes and total RNA. Intergroup statistical differences of TLR2 expression in stimulated THP1 cells (shaded bars) and unstimulated THP1 cells (unshaded bars) differ when normalized to HuPO, GAPDH, total RNA, or β -actin. *P* value indicates statistical confidence.

cally less likely because it suggests persistent disease despite clinical, biochemical, and radiological improvement.

We demonstrated that GAPDH, as compared with HuPO, is considerably more variable in this experimental system [3]. These first two experiments outlined the dangers of normalizing to this commonly used conventional reference gene in our system. If we had simply used GAPDH as our reference, as with many previously published studies, we would have generated study conclusions that were very different biologically. Moreover, these conclusions, which are diametrically different from those generated by HuPO, would have influenced how we conducted further experiments and modified our working hypothesis. Without the initial validation exercise [3], we would have had no means of cautioning against the use of GAPDH in our system.

Our third experiment investigated the expression of Toll-like receptors in response to *M. tuberculosis* antigens. To establish which of three potential reference genes was most appropriate, we assessed their variability in antigen-driven cultures. β -Actin had the greatest variation at approximately 7.5-fold, followed by GAPDH (~3.5-fold) and HuPO (~3-fold), suggesting that HuPO is the most favorable reference gene. The TLR2 expression differed, depending on which normalization method was used (Fig. 2), because the intergroup TLR2 differences were small (<5-fold). However, it is questionable whether a 3-fold change between study groups, although statistically significant, is meaningful when the variability of the reference gene itself is 3-fold. An alternative approach using the geometric mean of multiple reference genes can address this problem [9]. This method uses the error-induced trends of multiple reference genes for validation. Furthermore, by measuring multiple reference genes with each experiment, this strategy allows the validation to be ongoing, controlling for variables that might not have been accounted for in the initial validation exercise. However, this strategy is also susceptible to reference gene fluctuation, or “noise,” dictating final assay resolution. Furthermore, this method is not suitable in limited resource

settings (e.g., small laboratories, developing countries) and may require the measurement of at least three reference genes; hence, it is reagent and labor intensive.

In conclusion, this study has demonstrated that when a carefully validated reference is used to normalize data, the results can be significantly different from those obtained when an unvalidated reference gene is used. Consequently, experimental results may be erroneous. The findings of this study refute the suggestion that variability in reference gene expression does not statistically affect study conclusions, strongly supporting the argument for validation of reference genes prior to their use.

Acknowledgments

We thank Paul Kellam of the Department of Bioinformatics for advice and use of equipment. This work was funded by a grant from the British Lung Foundation. S.A. Bustin is grateful to Bowel and Cancer Research for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ab.2005.05.022](https://doi.org/10.1016/j.ab.2005.05.022).

References

- [1] J. Huggett, K. Dheda, S. Bustin, A. Zumla, Real-time RT-PCR normalisation: strategies and considerations, *Genes Immunol.* 6 (2005) 279–284.
- [2] S.A. Bustin, T. Nolan, Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction, *J. Biomol. Tech.* 15 (2004) 155–166.
- [3] K. Dheda, J.F. Huggett, S.A. Bustin, M.A. Johnson, G. Rook, A. Zumla, Validation of housekeeping genes for normalizing RNA expression in real-time PCR, *BioTechniques* 37 (2004) 112–119.
- [4] K. Dheda, J.F. Huggett, L.U. Kim, A. Zumla, Type2 cytokines in respiratory syncytial virus bronchiolitis, *Am. J. Respir. Crit. Care Med.* 169 (2004) 1167–1168.
- [5] A. Alvani, Counting antigen-specific T cells: a new approach for monitoring response to tuberculosis treatment?, *Clin. Infect. Dis.* 38 (2004) 757–759.
- [6] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, *J. Mol. Endocrinol.* 29 (2002) 23–39.
- [7] Y. Lin, M. Zhang, F.M. Hofman, J. Gong, P.F. Barnes, Absence of a prominent Th2 cytokine response in human tuberculosis, *Infect. Immunol.* 64 (1996) 1351–1356.
- [8] G.T. Seah, G.M. Scott, G.A. Rook, Type2 cytokine gene activation and its relationship to extent of disease in patients with tuberculosis, *J. Infect. Dis.* 181 (2000) 385–389.
- [9] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, *Genome Biol.* 3 (2002) 1–12.