Published in final edited form as:

Exp Dermatol. 2012 August ; 21(8): 625–629. doi:10.1111/j.1600-0625.2012.01537.x.

Identification and validation of reference genes for expression studies in human keratinocyte cell lines treated with and without interferon- γ – a method for qRT-PCR reference gene determination

Angelika B. Riemer^{1,2,3}, Derin B. Keskin², and Ellis L. Reinherz^{1,2}

¹Cancer Vaccine Center, Department of Medical Oncology, Dana-Farber Cancer Institute, and Department of Medicine, Harvard Medical School, Boston, MA, USA;

²Laboratory of Immunobiology, Department of Medical Oncology, Dana-Farber Cancer Institute, and Department of Medicine, Harvard Medical School, Boston, MA, USA;

³Immunotherapy and -prevention, German Cancer Research Center (DKFZ), Heidelberg, Germany

Abstract

Based on the exquisite sensitivity, reproducibility, and wide dynamic range of quantitative reverse-transcription real-time polymerase-chain-reaction (qRT-PCR), it is currently the gold standard for gene expression studies. Target gene expression is calculated relative to a stably expressed reference gene. An ideal reference should be uniformly expressed during all experimental conditions within the given experimental system. However, no commonly applicable “best” reference gene has been identified. Thus, endogenous controls must be determined for every experimental system. As no appropriate reference genes have been reported for immunological studies in keratinocytes, we aimed at identifying and validating a set of endogenous controls for these settings.

An extensive validation of sixteen possible endogenous controls in a panel of 8 normal and transformed keratinocyte cell lines in experimental conditions with and without interferon- γ was performed. RNA and cDNA quality was stringently controlled. Candidate reference genes were assessed by TaqMan® qRT-PCR. Two different statistical algorithms were used to determine the most stably and reproducibly expressed housekeeping genes.

mRNA abundance was compared and reference genes with widely different ranges of expression than possible target genes were excluded. Subsequent geNorm and NormFinder analyses identified GAPDH, PGK1, IPO8, and PPIA as the most stably expressed genes in the keratinocyte panel under the given experimental conditions.

We conclude that the geometric means of expression values of these four genes represents a robust normalization factor for qRT-PCR analyses in interferon- γ -dependent gene expression studies in keratinocytes. The methodology and results herein may help other researchers by facilitating their choice of reference genes.

Corresponding author: Angelika B. Riemer, Immunotherapy and -prevention, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, Tel: +49-6221-42-3820, Fax: +49-6221-42-4852, a.riemer@dkfz.de.

Conflict of Interest

The authors state no conflict of interest.

Keywords

RT-PCR; endogenous control; normalization, keratinocyte; immunology; IFN-gamma; inflammatory milieu

Introduction

Assessing gene expression by quantification of mRNA levels is a standard technique in the analysis of cellular responses. Quantitative reverse transcription real-time polymerase chain reaction (qRT-PCR) is the most widely used approach in the detection and quantification of mRNA. Due to its high sensitivity, reproducibility, and wide dynamic range, qRT-PCR is the gold standard for gene expression studies, and is commonly used for the validation of results from microarray experiments.

To obtain gene expression data, it is critical that qRT-PCR results are normalized to a reference gene (1, 2). This internal reference gene is exposed to the same preparation steps as the genes of interest and controls for non-biological variation introduced during sample preparation (3). The endogenous control is usually a housekeeping gene, which, ideally, is uniformly expressed during all environmental or experimental conditions in the given experimental system. The range of expression should be similar to the target gene analysed (4).

The most commonly used reference genes are β -actin, glyceraldehyde-3-phosphatedehydrogenase (GAPDH), hypoxanthine-guanine-phosphoribosyl-transferase (HPRT), and 18S-rRNA. However, the use of these housekeeping genes is largely a historical holdover - they were adequate references in non- or semi-quantitative techniques (northern blots, RNase protection assays, conventional RT-PCR). Although they are acceptable for these techniques, there is increasing evidence that many reference genes are not as resistant to internal and external environmental changes as previously thought (4). Consequently, their qualification as appropriate reference genes for the highly sensitive qRT-PCR has been reconsidered. Various recent studies investigate appropriate reference gene selection in human tissues (5, 6), vertebrate models (7-9), *C. elegans* (10), plants (11, 12), and even cyanobacteria (13).

Current consensus is that there is no single best reference gene for every condition (1, 14). Appropriate endogenous controls should be determined for every qRT-PCR experiment. Sometimes, validated endogenous controls for the desired experimental conditions can be derived from the literature. In most cases, however, systematic validation of reference genes must be the first step of the experiment.

As no published data were available on appropriate reference genes for immunological studies in keratinocytes, we performed an extensive validation of possible endogenous controls in a panel of 8 normal and transformed keratinocyte cell lines with and without interferon (IFN)- γ . The housekeeping genes determined herein to be most stable may be used as references by other investigators working with keratinocytes, saving them the effort of repeating the endogenous control determination and validation procedures. Alternatively, the described procedure can be employed for straightforward reference gene validation for other experimental setups.

Materials and methods

Cell lines and culture conditions

Two normal and six human papilloma virus (HPV)-immortalized keratinocyte cell lines were used in this study (Supplementary Table 1). CaSki (ATCC CRL-1550), SiHa (ATCC HTB-35), and 93VU147T (15) (kindly supplied by Prof. Steenberg, Dept. of Pathology, VUMC, Amsterdam, Netherlands) were grown in DMEM (Sigma) supplemented with 10% FCS, 1% L-glutamine, and 1% penicillin/streptomycin. C66-7 (p46) were kindly supplied by Prof. Lee (Dept. of Otolaryngology, University of Iowa, Iowa City, IA, USA) and grown in E media consisting of 3:1 DMEM:Ham's F-12 supplemented as described in (16). SCC090 was kindly supplied by Prof. Gollin (Dept. of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA) and grown in MEM supplemented as described in (17). HPV18HTE (p33), pHFK (p4), and NOK (p18) were kindly supplied by Prof. Munger (Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA) and grown in K-SFM (Invitrogen). All cells were subjected to authentication, and regularly tested for mycoplasma, acholeplasma and SMRV contamination. IFN- γ treatment was performed at a concentration of 10ng/ml for 48 hours.

RNA extraction, quantification and purity

Total RNA was isolated using the QIAgen RNeasy Mini kit. RNA concentration and purity (260/280 and 260/230 ratios) were analysed using an ND-1000 Spectrophotometer (NanoDrop Technologies), on gel, and with the Agilent Bioanalyzer 2100.

Quantitative reverse transcription real-time polymerase chain reaction (qRT-PCR)

Total RNA (1 μ g) was reverse-transcribed using Applied Biosystems' cDNA Reverse Transcription kit. All samples were run in the same reaction, and resulting cDNA controlled for purity and transcription efficiency on the NanoDrop spectrophotometer. qRT-PCR for 16 candidate housekeeping genes was performed using Applied Biosystems' TaqMan[®] Human Endogenous Control Array. This array is a 384-well microfluidic card containing 16 human gene expression assays used as endogenous controls, and was run on an Applied Biosystems 7900HT Fast Real-Time PCR System. The control genes tested were 18S, ACTB, B2M, GAPDH, GUSB, HMBS, HPRT, IPO8, PGK1, POLR2A, PPIA, RPLP, TBP, TFRC, UBC, and YWHAZ. Full gene names and functions are listed in Table 1. Assay IDs, reference gene number, assay location, and amplicon length are given in Supplementary Table 2. Reproducibility of results with the SYBR-Green detection method was tested with three randomly selected genes (B2M, ACTB, HPRT) in three cell lines (CaSki, SiHa, 93VU147T). All experiments were performed in triplicate.

Statistical analysis

Differential expression levels of candidate housekeeping genes were analysed by directly comparing Ct values. The stability of each gene was assessed using geNorm version 3.5 (18). The geNorm algorithm defines the internal control gene-stability measure M as the average pairwise variation of a particular gene with all other candidate genes. Genes with the lowest M-values exhibit the most stable expression. Genes with the highest M-value are excluded from the analysis in a stepwise approach, allowing for ranking of the candidate genes from least to most stable. geNorm results were confirmed with the NormFinder algorithm, version 0.953 (19).

Results

To determine valid reference genes for keratinocytes under different immunological conditions, eight keratinocyte cell lines were analysed with and without IFN- γ treatment.

The cell line panel (Supplementary Table 1) included primary foreskin keratinocytes (pHFK), oral mucosal keratinocytes (NOK), and various HPV-immortalized lines from different anatomical sites (cervix [CaSki, SiHa, C66-7], oropharynx [SCC090, 93VU147T]). We also included cells that have been propagated in cell culture for a very long time (e.g. SiHa, CaSki), and lines that have been established more recently (e.g. C66-7, SCC090). An HPV18-immortalized line (HPV18HTE) was included to assess possible differences caused by HPV type. The expression of 16 candidate housekeeping genes was assessed (Table 1). The panel represents housekeeping genes involved in a variety of cellular maintenance functions, and includes commonly used endogenous controls.

Comparisons of mRNA expression levels

Fluorescence detection curves of the qRT-PCR reactions were compared visually and immediately suggested differences in expression patterns (uniform vs. changing) of the candidate control genes. Examples of a gene stably expressed across all samples and experimental conditions (PGK1) and a gene with unstable expression (B2M) are shown in Supplementary Figure 1. To directly compare different RNA transcription levels before normalization, crossing threshold (Ct)-values were evaluated. Ct is defined as the number of cycles required for fluorescence to reach a specific threshold level of detection. Ct is inversely correlated with the amount of template RNA. As can be seen from Figure 1a, 18S-rRNA is present at much higher copy numbers than the other candidate control genes (and also future target genes). Figure 1b magnifies the Ct range of interest, showing two very stably expressed genes (GAPDH and PGK1), and two genes with differing expression. TFRC shows an irregular expression pattern, whereas B2M expression is influenced by the experimental conditions, being consistently upregulated by IFN- γ .

To assess if these results are also applicable for the SYBR-Green detection method, three randomly selected genes (B2M, ACTB, HPRT) were tested in three cell lines (CaSki, SiHa, 93VU147T) with and without IFN- γ treatment. All the tested genes yielded essentially the same Ct values with both detection methods (Supplementary Figure 2). Therefore, our results are valid for laboratories working with the SYBR-Green system. (For the complete set of fluorescence detection curves and Ct-values see Supplementary Figures 3 and 4, respectively.)

Determination of the most suitable reference genes

To reliably identify the genes that are most stably expressed across all cell lines and experimental conditions, two statistical programs specifically developed for this purpose were used. The geNorm algorithm uses pairwise comparisons, and provides a list of ranked gene stability, resulting in the definition of the two most stable genes (which cannot be compared anymore). The slope of the graph between two data points is indicative of the difference in expression stability between one gene and the remaining more stable genes (18). As can be seen from Figure 1c, GAPDH and PGK1 were the most stable genes in our panel, followed by 18S, IPO8 and PPIA. As the slope of the graph is relatively flat until HMBS, it can be assumed that these genes (RPLP0, HPRT1, and HMBS) could also be used as reference genes if a given experiment precludes the inclusion of the most stable genes. The instability of B2M, TFRC and UBC in the present experimental set-up can again be seen from their position in Figure 1c and the slope of the graph. These results were confirmed with the NormFinder algorithm (19), also returning 18S, IPO8, PGK1, GAPDH and PPIA as the most stably expressed genes in our panel (Supplementary Table 3).

Paradigm normalizations

To exemplify the effects of choosing an inappropriate reference gene on the calculation of relative mRNA expression, one of the two most stable genes in our panel (GAPDH) was

normalized on the other most stable gene (PGK1), on a gene with irregular unstable expression (TFRC), and on a gene which is itself regulated by the experimental conditions (B2M). Figure 2 (top, Normalization on PGK1) shows the expected stable GAPDH expression across all cell lines and conditions. However, when GAPDH is normalized on TFRC (center), there is apparent overexpression in HPV18-transformed tonsillar epithelial cells. The bottom panel demonstrates the effects of normalizing on a gene that is co-regulated. β -2-microglobulin, the small subunit of MHC-I molecules, is upregulated upon IFN- γ exposure (cf Figure 1). When used as reference gene for relative expression calculation, all other genes appear to be downregulated by IFN- γ .

Discussion

This study underscores the need to properly validate endogenous reference genes before data normalization in qRT-PCR experiments. Systematic validation of reference genes is essential for producing accurate, reliable data in qRT-PCR analyses, and should be an integral component of such experiments (20, 21). To be used as a reference gene, several mandatory criteria are required (22, 23): A gene should be constitutively non-regulated, and the range of expression values should be similar to those of the target gene to be analysed. The detection of the reference gene should be RNA specific. Pseudogene- and DNA-free amplification should be ensured by stringent primer design.

Several strategies for the selection of suitable control genes have been devised. The applet geNorm was developed by Vandesompele et al. (18). It consecutively eliminates less stable genes from a selection of putative housekeeping genes, resulting in suitable reference genes for a given experimental system. Alternatively, NormFinder (19) or Bestkeeper (24) are options for validating reference genes. In many studies, normalization is based on co-amplification of a single reference gene. However, recent studies suggest to measure at least three housekeeping genes (18, 25), followed by calculating their geometric means as a reference for normalization (18). This adds another criterion to the selection of reference genes: housekeeping genes with independent functions in cellular maintenance should be selected, as this significantly reduces the chance that the genes chosen might be co-regulated (23).

Validation studies are often circumvented by choosing reference genes from literature describing similar experiments. In keratinocytes, however, the literature on housekeeping genes contains an abundance of inconsistent findings (4). β -actin and GAPDH have been described as highexpression reference genes. Yet, their expression was found to vary across tissues and cell types and during cell proliferation and differentiation (26). Particular attention must be paid to reference genes in diseased skin (4, 27). Dedicated studies on reference gene selection for qRT-PCR analyses have been performed comparing adult and juvenile normal human epidermal keratinocytes with and without short hairpin (sh)RNA (2), neonatal human epidermal keratinocytes with and without UVB irradiation (28), immortalized normal keratinocytes (HaCaT) and two head and neck squamous cell carcinoma lines with and without EGF and TGF- 1β (23), primary keratinocytes from different donors at various culture densities (29, 30), and UVA-exposed dermal fibroblasts in normoxic and hypoxic conditions (31). No studies on keratinocytes in different immunological conditions were available, prompting the current set of experiments.

Our validation procedure emphasised the following criteria: i) stringent quality and quantity control of isolated RNA and reverse transcribed cDNA (32), ii) the simultaneous investigation of 16 candidate reference genes, and iii) the use of two different software algorithms for data analysis. The requirement for stringent primer design was met by

utilizing a set of validated primers provided by Applied Biosystems. Additionally, the panel included genes from different functional classes.

We found considerable differences in mRNA expression across the cell lines and experimental conditions (Figure 1). Both geNorm (Figure 1c) and NormFinder (Supplementary Table 3) results found GAPDH, PGK1, and 18S rRNA to be the most stably expressed genes. We explicitly chose not to use 18S rRNA as a reference gene because rRNA is present in much higher abundance when compared to target mRNA transcripts, and is not an ideal reference because of the imbalance between rRNA and mRNA fractions and its absence from purified mRNA samples (18). Because PGK1 and GAPDH both are involved in glycolysis (Table 1), we recommend using four instead of three endogenous control genes for immunological studies in keratinocytes, namely GAPDH, PGK1, IPO8 and PPIA. In accordance with the recommendation of Vandesompele et al, the geometric means of these genes will be used as a robust normalization factor in subsequent qRT-PCR analyses.

To illustrate typical errors of data interpretation in consequence of normalization to an inadequate reference gene, the relative mRNA expression of GAPDH (being one of the two most stable genes in our panel), was determined relative to different candidate endogenous controls (Figure 2). As can be seen from Figure 2a, a stable gene normalized on a stable gene leads to the expected result, stable expression across all cell lines and conditions. However, when normalizing on a gene with unstable expression, the calculation of the relative amount of target gene expression leads to erroneous conclusions. In the example shown in Figure 2b, one assumes that the amount of GAPDH is markedly higher in HPV18-transformed tonsillar epithelial cells than in HPV16-transformed or normal keratinocytes. In fact, only the expression of the transferrin receptor (TFRC) is lower. The worst scenario that can occur if reference genes are not properly validated is normalization on a gene that is itself regulated by the experimental conditions. This is exemplified in Figure 2c, where GAPDH is normalized on B2M. The choice of a co-regulated gene as a reference leads to vast misinterpretations and must be avoided.

These data highlight the critical importance of characterizing housekeeping gene expression profiles in each specific cell type and for every experimental condition prior to choosing a set of housekeeping genes for expression studies. The reference gene validation procedure described represents a straightforward approach to do so. The present results can expedite experiments by other researchers who are interested in gene expression studies in keratinocytes under different immunological conditions. According to our investigation, the most suitable reference genes in keratinocytes for IFN- γ -dependent gene expression studies are GAPDH, PGK1, IPO8 and PPIA.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

A.B. Riemer (ABR) performed the research and analysed the data, ABR, D.B. Keskin and E.L. Reinherz (ELR) designed the research study and wrote the paper. ABR was supported by an Erwin Schrödinger fellowship of the Austrian Science Fund. This work was supported by DoD & NIH grants to ELR. The authors want to thank Mel Hernandez for excellent technical and scientific advice, and Linda K. Clayton for expert scientific advice and proofreading of the manuscript.

References

1. Huggett J, Dheda K, Bustin S, Zumla A. *Genes Immun.* 2005; 6:279–284. [PubMed: 15815687]
2. Allen D, Winters E, Kenna PF, Humphries P, Farrar GJ. *J Dermatol Sci.* 2008; 49:217–225. [PubMed: 18061409]
3. de Kok JB, Roelofs RW, Giesendorf BA, et al. *Lab Invest.* 2005; 85:154–159. [PubMed: 15543203]
4. Bär M, Bär D, Lehmann B. *J Invest Dermatol.* 2009; 129:535–537. [PubMed: 19209154]
5. Chechi K, Gelinas Y, Mathieu P, Deshaies Y, Richard D. *PLoS One.* 2012; 7:e32265. [PubMed: 22511915]
6. Mohelnikova-Duchonova B, Oliverius M, Honsova E, Soucek P. *Dis Markers.* 2012; 32:203–210. [PubMed: 22377737]
7. Veazey KJ, Golding MC. *PLoS One.* 2011; 6:e27592. [PubMed: 22102912]
8. Cabiati M, Raucci S, Caselli C, et al. *J Mol Endocrinol.* 2012
9. Cinar MU, Islam MA, Uddin MJ, et al. *BMC Res Notes.* 2012; 5:107. [PubMed: 22340302]
10. Zhang Y, Chen D, Smith MA, Zhang B, Pan X. *PLoS One.* 2012; 7:e31849. [PubMed: 22438870]
11. de Oliveira LA, Breton MC, Bastolla FM, et al. *Plant Cell Physiol.* 2012; 53:405–422. [PubMed: 22197885]
12. Mafra V, Kubo KS, Alves-Ferreira M, et al. *PLoS One.* 2012; 7:e31263. [PubMed: 22347455]
13. Pinto F, Pacheco CC, Ferreira D, Moradas-Ferreira P, Tamagnini P. *PLoS One.* 2012; 7:e34983. [PubMed: 22496882]
14. Derveaux S, Vandesompele J, Hellemans J. *Methods.* 2010; 50:227–230. [PubMed: 19969088]
15. Steenbergen RD, Hermesen MA, Walboomers JM, et al. *Cancer Res.* 1995; 55:5465–5471. [PubMed: 7585617]
16. Lee JH, Yi SM, Anderson ME, et al. *Proc Natl Acad Sci U S A.* 2004; 101:2094–2099. [PubMed: 14769917]
17. Ferris RL, Xi L, Raja S, et al. *Cancer Res.* 2005; 65:2147–2156. [PubMed: 15781625]
18. Vandesompele J, De Preter K, Pattyn F, et al. *Genome Biol.* 2002; 3:RESEARCH0034. [PubMed: 12184808]
19. Andersen CL, Jensen JL, Orntoft TF. *Cancer Res.* 2004; 64:5245–5250. [PubMed: 15289330]
20. Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. *J Exp Bot.* 2009; 60:487–493. [PubMed: 19264760]
21. Khan-Malek R, Wang Y. *Methods Mol Biol.* 2011; 691:227–241. [PubMed: 20972756]
22. Bustin SA, Benes V, Nolan T, Pfaffl MW. *J Mol Endocrinol.* 2005; 34:597–601. [PubMed: 15956331]
23. Campos MS, Rodini CO, Pinto-Junior DS, Nunes FD. *Oral Oncol.* 2009; 45:121–126. [PubMed: 18621570]
24. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. *Biotechnol Lett.* 2004; 26:509–515. [PubMed: 15127793]
25. Wagner M, Hintner H, Bauer JW, Onder K. *Exp Dermatol.* 2012; 21:111–117. [PubMed: 22168818]
26. Zhong H, Simons JW. *Biochem Biophys Res Commun.* 1999; 259:523–526. [PubMed: 10364451]
27. Turabelidze A, Guo S, DiPietro LA. *Wound Repair Regen.* 2010; 18:460–466. [PubMed: 20731795]
28. Balogh A, Paragh G Jr, Juhasz A, et al. *J Photochem Photobiol B.* 2008; 93:133–139. [PubMed: 18789713]
29. Minner F, Poumay Y. *J Invest Dermatol.* 2009; 129:770–773. [PubMed: 18719611]
30. Paragh G, Ugocsai P, Vogt T, et al. *Exp Dermatol.* 2010; 19:297–301. [PubMed: 19961536]
31. Bruguè F, Venditti E, Tiano L, Littarru GP, Damiani E. *J Biotechnol.* 2011; 156:153–162. [PubMed: 21963587]
32. Vermeulen J, De Preter K, Lefever S, et al. *Nucleic Acids Res.* 2011; 39:e63. [PubMed: 21317187]

Figure 1.

Comparisons of mRNA expression levels and geNorm analysis. (a) Ct values of three stably expressed genes (GAPDH, 18S, PGK1) and two unstable genes (B2M, TFRC), highlighting the much higher expression level (i.e. lower Ct values) of 18S ribosomal RNA in comparison to the other genes. (b) Detail of the Ct results (range 18-29) for GAPDH, PGK1, B2M and TFRC, showing irregular expression of TFRC and regulation of B2M expression by the experimental conditions. (c) Ranked expression stability of candidate endogenous control genes (geNorm analysis). Average expression stability values (M) of the remaining control genes during stepwise exclusion of the least stable candidate genes in the panel of eight keratinocyte lines with and without IFN- γ treatment. The slope of the graph between

two data points is indicative of the difference in expression stability between one gene and the remaining more stable genes.

Figure 2.

Examples of gene expression calculations. Relative mRNA expression calculated by normalization of a stably expressed gene (GAPDH) on a stable gene (PGK1, top panel), an unstable gene (TFRC, center panel), and a co-regulated gene (B2M, bottom panel).

Table 1

Endogenous control genes and their functions

Abbreviation	Name	Function
18S	18S ribosomal RNA	Small ribosomal subunit, translation
ACTB	β -actin	Cytoskeleton
B2M	β -2-microglobulin	Subunit of MHC class I, antigen presentation
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Glycolysis
GUSB	β -glucuronidase	Breakdown of mucopolysaccharides
HMBS	Hydroxymethylbilane synthase (Porphobilinogen deaminase)	Porphyrin metabolism
HPRT1	Hypoxanthine guanine phosphoribosyl transferase	Generation of purine nucleotides
IPO8	Importin-8	Nuclear protein import
PGK1	Phosphoglycerate kinase-1	Glycolysis
POLR2A	RNA polymerase II, subunit A	Transcription
PPIA	Peptidylproline isomerase A (Cyclophilin A)	Protein folding
RPLP0	Large ribosomal protein P0	Translation
TBP	TATA box binding protein	Transcription initiation
TFRC	Transferrin receptor	Endocytosis of iron
UBC	Ubiquitin C	Protein degradation
YWHAZ	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (Phospholipase A2)	Signal transduction