

RESEARCH ARTICLE

# With Reference to Reference Genes: A Systematic Review of Endogenous Controls in Gene Expression Studies

Joanne R. Chapman\*, Jonas Waldenström

Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, Kalmar, Sweden

\* [joanne.chapman@lnu.se](mailto:joanne.chapman@lnu.se)

## Abstract

The choice of reference genes that are stably expressed amongst treatment groups is a crucial step in real-time quantitative PCR gene expression studies. Recent guidelines have specified that a minimum of two validated reference genes should be used for normalisation. However, a quantitative review of the literature showed that the average number of reference genes used across all studies was 1.2. Thus, the vast majority of studies continue to use a single gene, with  $\beta$ -actin (*ACTB*) and/or glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) being commonly selected in studies of vertebrate gene expression. Few studies (15%) tested a panel of potential reference genes for stability of expression before using them to normalise data. Amongst studies specifically testing reference gene stability, few found *ACTB* or *GAPDH* to be optimal, whereby these genes were significantly less likely to be chosen when larger panels of potential reference genes were screened. Fewer reference genes were tested for stability in non-model organisms, presumably owing to a dearth of available primers in less well characterised species. Furthermore, the experimental conditions under which real-time quantitative PCR analyses were conducted had a large influence on the choice of reference genes, whereby different studies of rat brain tissue showed different reference genes to be the most stable. These results highlight the importance of validating the choice of normalising reference genes before conducting gene expression studies.



## OPEN ACCESS

**Citation:** Chapman JR, Waldenström J (2015) With Reference to Reference Genes: A Systematic Review of Endogenous Controls in Gene Expression Studies. PLoS ONE 10(11): e0141853. doi:10.1371/journal.pone.0141853

**Editor:** Ramona Natacha PENA i SUBIRÀ, University of Lleida, SPAIN

**Received:** August 26, 2015

**Accepted:** October 13, 2015

**Published:** November 10, 2015

**Copyright:** © 2015 Chapman, Waldenström. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors received no specific funding for this work.

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

## Introduction

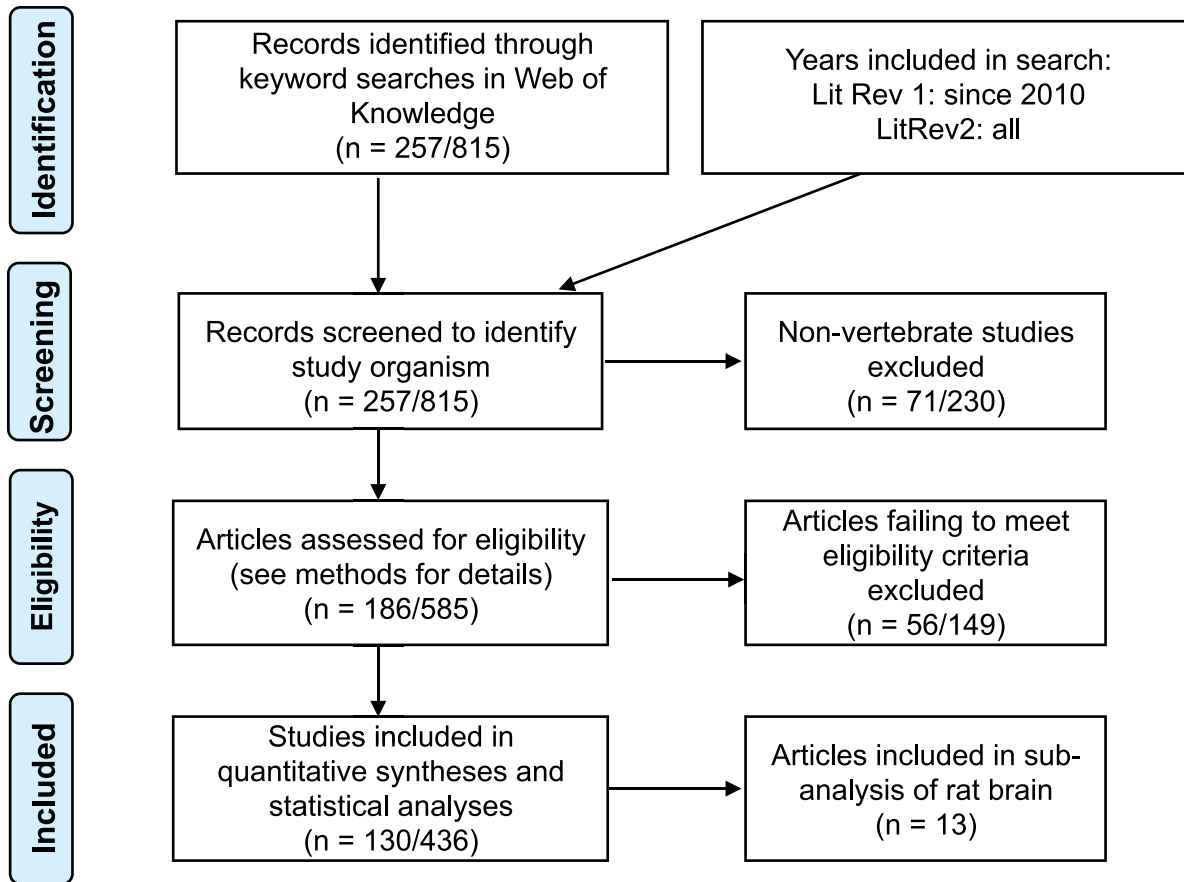
The use of real-time quantitative PCR (RT-qPCR, hereafter abbreviated to qPCR) to quantify mRNA transcription has revolutionised our understanding of cellular responses triggered by the developmental progression or experimental manipulation of individuals [1]. As qPCR is an extremely sensitive method that allows detection of small dynamic changes in gene expression between samples, it is imperative that due care is used in every step of sample preparation and processing. When properly employed, data normalisation helps to correct for inevitable experimental variations, the most common of which are disparities in the amount of starting material [2] and/or sample loading [3].

While several methods of data normalisation are available (reviewed in [4–6]), the most commonly employed is the use of stably expressed reference genes (RGs) as endogenous (*i.e.* internal) controls. This method is essential when results are expressed in terms of relative quantification (fold-change) using the popular  $2^{-\Delta\Delta CT}$  method [7] or modified versions thereof (*e.g.* [8–11]). Thus, an essential step in development of qPCR assays is identification of appropriate RGs for normalisation of the data. Crucially, the RG must be unaffected by the experimental treatment(s) and/or developmental stages under investigation. Thus, identifiable changes in expression of the genes(s) of interest (GOI(s)) must not be biased by changes in expression of the RGs between experimental groups. As correct normalisation is such an integral component of gene expression analysis, it has been strongly advocated that at least two RGs be employed and the geometric mean of RG expression levels be used for normalisation (*e.g.* [2, 4, 11]).

The most frequently chosen RGs are so called ‘housekeeping genes’—being those required for basic cellular processes. To constitute an ideal RG, a gene must display constitutive, stable expression in all cell types/tissues and treatment regimens under consideration. It should also be expressed at similar levels (as indicated by comparable C<sub>q</sub> values) as the GOI. The C<sub>q</sub>, or quantification cycle, (also known as the C<sub>T</sub>, or threshold cycle) is the number of cycles of amplification required to reach a pre-selected threshold of fluorescence. In vertebrate studies, glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*),  $\beta$ -actin (*ACTB*) and 18S rRNA ribosomal RNA (*18S rRNA*) have historically been popular choices [12]. However, it is becoming increasingly apparent that the utility of these genes is limited in many cases due to differential expression across species, tissue types, cell lines, developmental stages and/or in response to experimental treatments (*e.g.* [11–14]). If gene expression analyses are normalised against RGs that fluctuate under the experimental conditions used, then small but potentially biologically relevant changes in expression of the GOI(s) of interest can easily be missed or overemphasised [15]. For example, using unstable RGs to normalise can result in overestimating (*e.g.* [16, 17]) or underestimating (*e.g.* [18, 19]) GOI expression levels. Several methods are available for testing RG stability. Currently, the most popular [5, 20] are the BestKeeper [21], geNorm [11] and NormFinder [22] programmes. While the algorithms employed by each of these methods differs somewhat (reviewed in [23]), all are based on estimating the variance in C<sub>q</sub> values of each RG tested across treatment groups.

In February 2009, the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines were published [24]. These guidelines lay the foundation for the optimum methodology in gene expression studies to ensure reliable and repeatable results, including a comprehensive checklist of procedures that should be followed encompassing sample preparation and handling, quality control, cDNA creation, qPCR and data analysis. With respect to RG selection, the key recommendations are that two or more RGs should be employed and the choice of RGs should be validated to ensure stable expression across treatment groups for the given experimental setting and sample set. Despite the widespread dissemination and advocacy of these guidelines, it is currently unclear to what extent these recommendations have been implemented in practical terms.

Here, we undertake two systematic quantitative reviews of the gene expression literature to assess current trends in the selection and use of RGs in: (1) studies quantifying expression of GOI(s); and (2) studies screening a panel of RGs to determine those with highest stability. We posited that over time, as the MIQE guidelines [24] have become more widely disseminated, fewer studies would be published without first providing evidence that RG stability had been quantified, increasing numbers of RGs would be tested for stability, and the use of *GAPDH* and *ACTB* as normalisers would become less widespread.



**Fig 1. PRISMA flowchart.** The number (n) of studies included and excluded at various stages of the quantitative review process, where numbers are specified as LitRev1/LitRev2.

doi:10.1371/journal.pone.0141853.g001

## Materials and Methods

Two quantitative literature reviews (LitRev) were conducted to assess the current methodology with respect to RG selection and use in qPCR studies. We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [25] for reporting methodology (Fig 1, S1 Table). The full tables of papers used and data extracted are provided as Supplementary Information (S2–S4 Tables).

### 1) LitRev 1: Current patterns of RG use in gene expression studies

To determine whether the MIQE guidelines are being adhered to, we conducted a restricted literature review on 10 July 2015 of gene expression studies in the years following their publication. The search was restricted to papers published since 2010, to allow one year for dissemination of the guidelines and implementation by researchers. As a vast number of gene expression papers are published every year, the search (conducted in Web of Science (WoS), Thomson Reuters) was restrictive, as follows: TITLE: "gene expression", TOPIC: "real-time quantitative PCR" and limited to TYPE: "article". While this restricted search meant that many gene expression studies were excluded from analysis, we felt that it provided an unbiased overview of current gene expression studies.

The returned papers ( $n = 257$ ) were manually searched and papers retained if they met the following three additional criteria: (1) the study was conducted on a vertebrate species; (2) we had access to the full-text article; and (3) the paper reported the results of a gene expression analysis and was not solely a RG selection paper (such papers were analysed in LitRev 2, described below). Of the papers retained ( $n = 128$ ), the methodology and results were searched to determine: (1) the number of RGs used; (2) the identity of the RG(s); (3) the number of GOIs subsequently analysed; and (4) whether there was evidence that the authors had tested a panel of RGs to select those with the most stable expression, either in the current work or in previous work conducted in the same lab with the same/similar samples.

## 2) LitRev 2: Testing panels of RGs for expression stability

Our second literature review focused on papers specifically testing panels of RGs to determine their stability. Papers that subsequently included analysis of GOI expression were included, but this was not a pre-requisite for inclusion in LitRev2 (in contrast to LitRev1). The aim was to determine whether methodology is changing over time or differs between fields (such as model versus non-model species) and the consequences this has on the number and identity of RGs tested and selected. We conducted a literature search on 15 July 2015 in WoS using the following search string: TOPIC: "Gene expression" and TOPIC: "bestkeeper OR genorm OR norm-finder" and limited to TYPE: "article". All years were included.

The returned results ( $n = 815$ ) were manually searched and studies retained if they fit the following criteria: (1) they were conducted on a vertebrate species; (2) they did not solely test micro-RNA (miRNA) RGs; and (3) they used qPCR to test a panel of at least three RGs (thus, papers using one of the software programs solely on micro-array data were eliminated). Many of the retained papers ( $n = 436$ ) included analyses on multiple tissues or treatments. These studies were listed once in our table of results to avoid pseudo-replication. However, some papers ( $n = 17$ ) provided results for multiple species. The results for each different species were recorded separately as there is no *a priori* reason to expect the same RGs to be optimal in different species. Thus, our final dataset comprised 457 results from 436 studies. Of the papers retained, the methodology was searched to determine: (1) the number of RGs used; (2) the study species; (3) whether *GAPDH*, *ACTB* or *18S rRNA* were included in the panel tested; and (4) if so, whether these genes were subsequently selected within the most stable (two to four) genes required for accurate normalisation. For point (1), three studies used different numbers of RGs for different tissues or treatments. In such cases, we took the largest number of RGs tested in a single analysis as our number of RGs tested. For point (4), we used the authors' recommendations of which RGs provided the best normalisation for the given tissue and/or experimental treatment or the number of genes required to obtain a stability  $V$  value below 0.15 in GeNorm as suggested by Vandesompele *et al.* [11]. In some cases, authors simply listed stability values for each gene without providing a recommendation. In such cases, we searched the results to determine which two to four genes had the highest stability for the given tissue and/or experimental treatment. By default, we selected the two most stable unless the authors showed that three or four RGs were required to reach a threshold stability value. In cases where the authors showed that five or more RGs would be required to provide reliable normalisation, we considered that no RGs were optimal—given the methodological and financial limitations of normalising with such a high number of RGs.

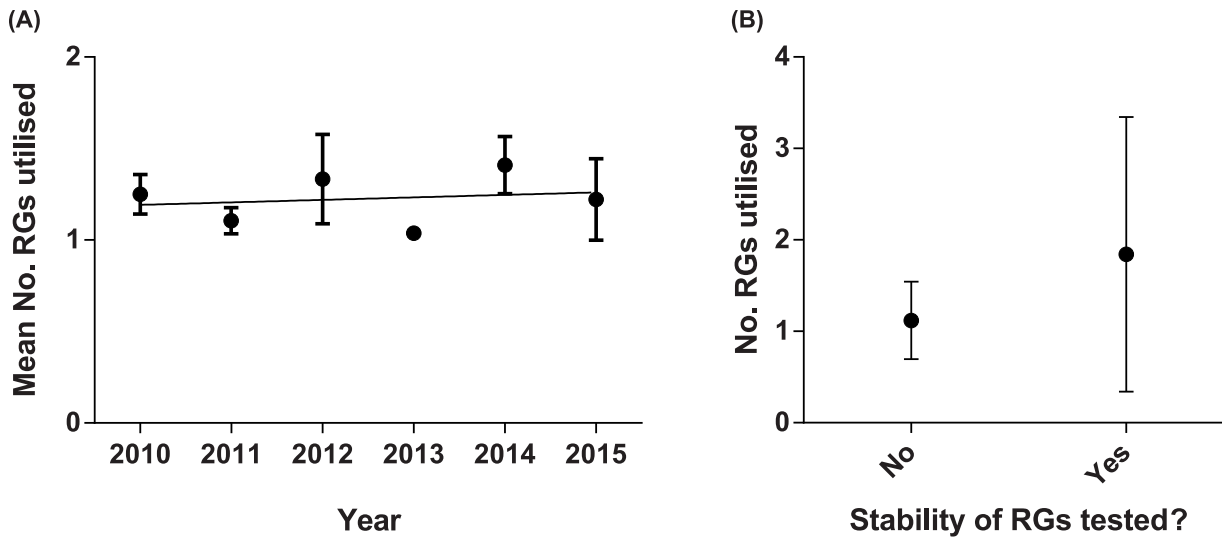
In order to provide a more comprehensive examination of the RGs tested and selected within a single tissue type from one species, we determined that the most common non-human tissue examined in LitRev 2 was rat brain. We chose to exclude human cell-line studies from this analysis because tissue samples were the more common source of RNA across all studies,

and likely represent a more homogeneous source material than cell-line extracts. We examined these studies further to determine if different studies, with different treatment regimens and/or different laboratory methods, resulted in the same RGs being chosen. This should be the case if RGs are universally stable for a given tissue, independent of treatment effects and differences between laboratories. We classified studies by the structure(s) of the brain examined and treatment regime(s) and determined the RGs tested and selected. Some studies analysed their data with more than one stability testing algorithm. In such cases, we preferentially recorded the results from the GeNorm analysis in the main analysis (11/13 studies) and noted when other algorithms provided differing results as Supplementary Information (S4 Table). When provided, we listed all genes required for stability based on V values falling below the standard threshold of 0.15 [11]. When this analysis was not provided we listed the two most stable genes. We excluded studies that performed analyses on rat cell lines ( $n = 3$ ) and those that included both mRNA and microRNA RGs in the same analysis ( $n = 1$ ). This resulted in 25 results from 13 studies being retained in the analysis.

## Statistical analyses

**LitRev 1: Current patterns of RG use in gene expression studies.** We assessed the influence of time on the number of RGs utilised via a generalised linear model (GLM) with Poisson errors since the response variable was count data, and deviance goodness of fit to assess the fit of the model. We used a Mann-Whitney U test to determine whether studies that test a panel of RGs for stability subsequently use more RGs when testing expression of the GOI than those that do not. The use of a non-parametric test was required as the response variable (number of RGs utilised) did not fit normality in this dataset. We investigated the influence of the number of GOIs investigated on the amount of effort expended in testing and using reference genes via GLMs with the following three response variables and associated error distributions: 1) whether or not RG stability was tested (binomial error); 2) the number of RGs used to normalise data (Poisson error); and 3) the number of RGs included in the panel amongst the subset of studies that tested RG stability (Poisson error).

**LitRev 2: Testing panels of RGs for expression stability.** We investigated the influence of time on the number of RGs included in panels for stability testing using a linear model after log transforming the response variable (number of RGs tested). This was owing to the fact that we could not use Poisson regression due to overdispersion of the data. In this case, we used a Kolmogorov-Smirnov test [26] to assess the Gaussian error distribution. We used GLMs with binomial errors to determine whether the selection of *GAPDH*, *ACTB* or *18S rRNA* as a stable RG was influenced by the number of RGs included in the panel and deviance goodness of fit to assess the fit of the model. For these tests, we first recorded whether *GAPDH*, *ACTB* or *18S rRNA* was included amongst the panel of RGs in each paper as a categorical variable with four possible outcomes: *NA* (not tested), *yes* (tested and selected), *no* (tested and not selected) or *sometimes* (tested, selected in a subset of the treatments and/or tissue types analysed). We then only considered the variables *yes* and *no* (binomial response variable) in the models. We assessed whether the number of RGs tested differs with respect to the type of species the study was conducted on in two ways. Firstly, we used a Kruskal-Wallis test followed by Dunn's post-hoc test for multiple unequal group comparisons to determine whether broad species groupings differed. To facilitate this, we first produced a seven-level categorical variable for the type of species the study was conducted on, as follows: *human*; *non-human primate*; *rodent model* (rat, mouse, guinea pig); *livestock mammal* (cow, pig, sheep, goat, horse, buffalo, yak, deer); *other mammal* (dog, cat, rabbit, wild rodent, dolphin); *bird*; *fish*. Secondly, we used a Mann-Whitney U test to determine whether the number of RGs used differs between model and non-



**Fig 2. Publication of studies testing expression of GOI(s).** Papers included in LitRev1 showing (A) The mean number (no.) of RGs utilised in gene expression studies per year. (B) Mean no. of RGs utilised, according to whether or not a panel of RGs was first tested for stability. In both cases, data is represented as mean (dot)  $\pm$  standard deviation (bars).

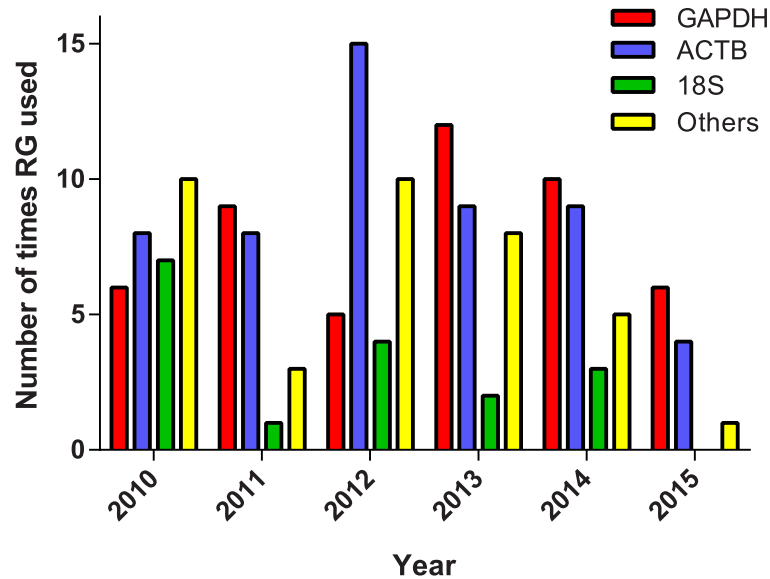
doi:10.1371/journal.pone.0141853.g002

model organisms. The use of non-parametric tests (Kruskal-Wallis and Mann-Whitney) were required because the response variable (number of reference genes) is count data and thus does not follow a Gaussian distribution. This classification was based on whether there is a full genome sequence available for the study species in the Ensembl database (<http://www.ensembl.org/info/about/species.html>) on 30 July 2015. Species with a full genome sequence available were classified as *model* species and the remainder were classified as *non-model* species. While undoubtedly some of the studies on model species were published before the availability of their genome sequence, we felt that this classification provided a proxy for the amount of molecular research conducted on the species and therefore, the availability of genetic resources such as primers. With the exception of non-parametric tests, all statistical analyses were performed in R v 3.0.2 [27]. Non-parametric tests and plots were conducted in GraphPad Prism v.6.03 (GraphPad Software Inc, San Diego, CA, Figs 2–6) or R (S1 Fig).

## Results

### LitRev 1: Current patterns of RG use in gene expression studies

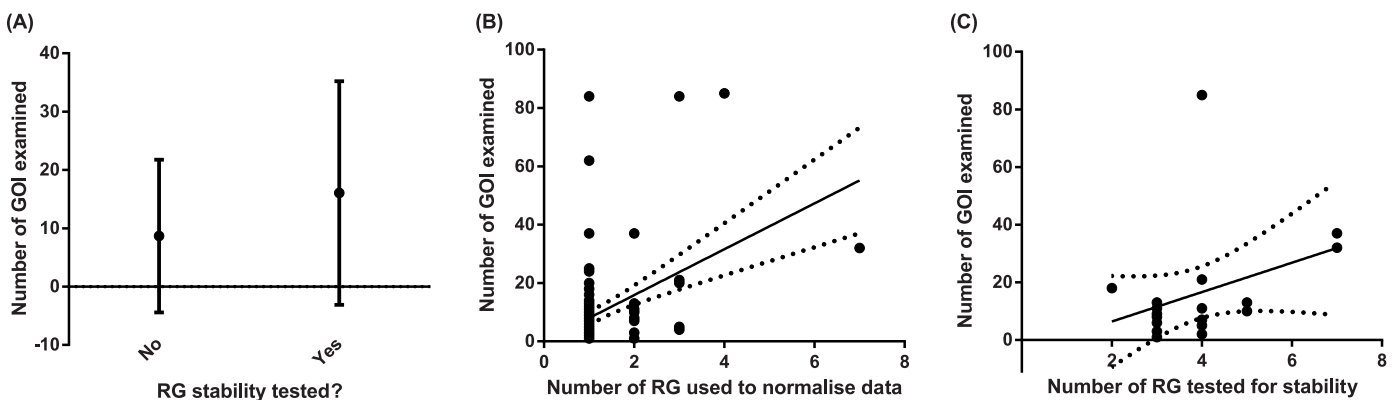
The average number of RGs used was 1.23 ( $n = 128$ ; range 1–7) with only 13% (17/128) of studies using more than a single RG for normalisation (Fig 2A). Furthermore, there was no evidence that the number of RGs used for normalisation increased over time (test for the slope:  $z = 0.206$ ,  $P = 0.837$ ). Only 19 studies (15%) provided evidence that a panel of RGs had been tested to find those with stable expression. Amongst these, the mean number of RGs tested was 3.9 (range 2–7). Studies that tested a panel of RGs for stability subsequently used more RGs in their analyses of gene expression than those that did not (two-tailed Mann-Whitney  $U = 683.5$ ,  $P = 0.0003$ ,  $n = 128$ , Fig 2B). However, such studies continued to use very few RGs for normalisation of qPCR data (mean = 1.84, range 1–7 versus studies that did not test a panel: mean = 1.10, range = 1–3). Amongst all included studies, the most common choice of RGs were *ACTB* (used in 38% of the studies) and *GAPDH* (37%) with *18S rRNA* (12%) being somewhat less utilised (Fig 3).



**Fig 3. Utilisation of *GAPDH* and *ACTB* and *18S rRNA* as RGs in gene expression studies since the publication of the MIQE guidelines.**

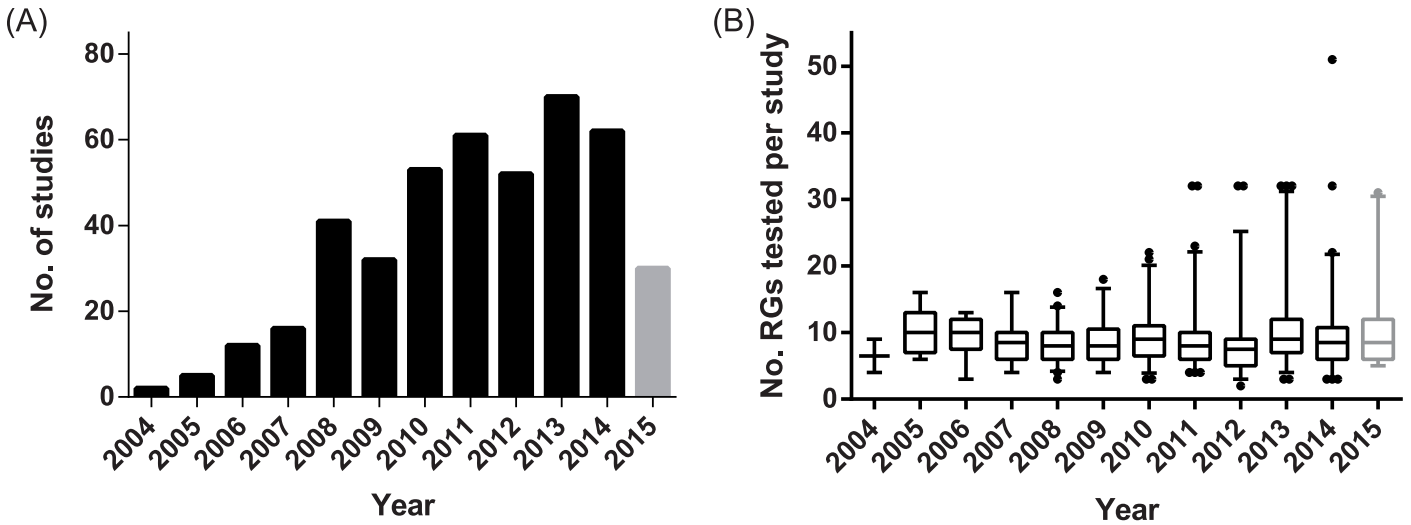
doi:10.1371/journal.pone.0141853.g003

We further tested whether the number of GOIs investigated influenced the research effort expended finding suitable RGs. We found that the number of GOIs analysed was marginally unrelated to whether or not stability of RGs was tested (test for the slope:  $z = 1.887$ ,  $P = 0.059$ ,  $n = 128$ , Fig 4A). However, there was a significant relationship between the number of RGs used to normalise data and the number of GOIs subsequently tested (test for the slope:  $z = 2.960$ ,  $P = 0.003$ ,  $n = 128$ , Fig 4B), and amongst the 19 studies that did test for RG stability, the number tested in the panel was influenced by the number of GOIs under investigation (test for the slope:  $z = 2.740$ ,  $P = 0.006$ ,  $n = 19$ , Fig 4C). In both cases, the number of RGs used/assessed increased in studies subsequently analysing more GOIs.



**Fig 4. Influence of the number of GOI investigated on the rigor with which RGs are selected and used.** (A) Number of GOIs analysed does not affect whether or not the stability of RGs is tested. (B) Number of GOIs analysed does not affect the number of RGs used to normalise data. (C) Amongst those studies that do test a panel of RGs for stability ( $n = 19$ ), more RGs are tested in studies subsequently analysing more GOI. For panels (B) and (C), a linear regression (solid line) and 95% confidence intervals (dotted lines) were fitted solely to show the direction of the relationship.

doi:10.1371/journal.pone.0141853.g004

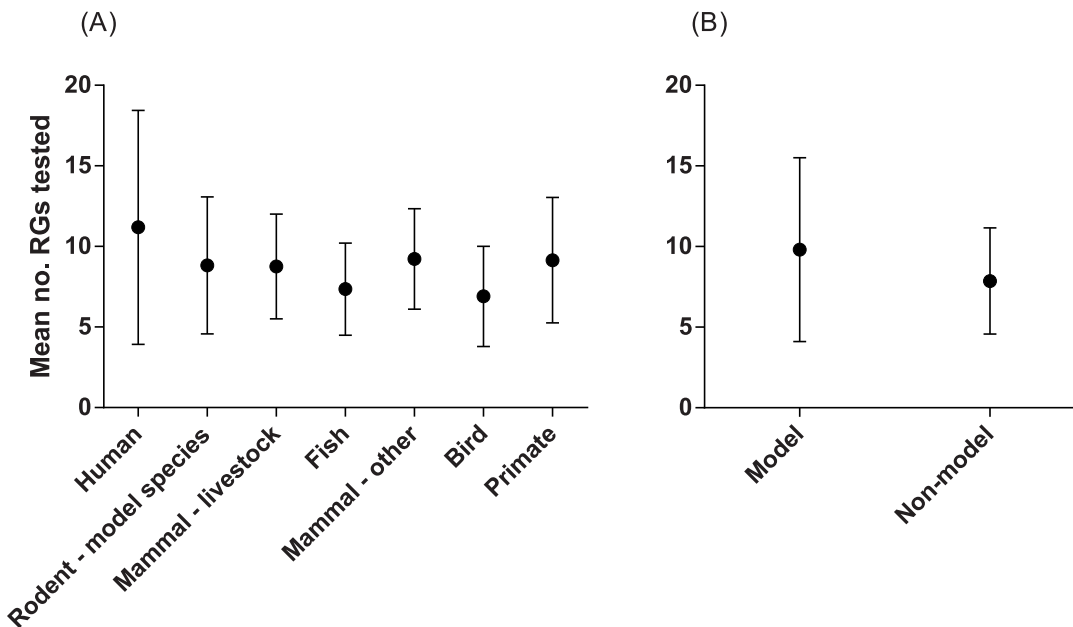


**Fig 5. Publication of studies testing a panel of RGs for expression stability.** Papers included in LitRev2 showing (A) Temporal trend, and (B) Median and interquartile range of number (no.) of RGs included in stability panels per year, where whiskers span the 5–95 percentiles. Data for 2015 are shown in grey as this comprises the months January to June 2015 only.

doi:10.1371/journal.pone.0141853.g005

### LitRev2: Testing panels of RGs for expression stability

The average number of RGs tested for stability was 9.53 ( $n = 456$ ; range 2–51). An increasing number of studies testing panels of RGs for expression stability are being published (Fig 5A), but we found no evidence that amongst these studies an increasing number of RGs are included in the panel over time (test for the slope:  $t_{456} = 1.412$ ,  $P = 0.159$ , Fig 5B). *GAPDH* and *ACTB* were commonly included in the panel of genes tested (*GAPDH* = 89%, *ACTB* = 86%); in contrast, *18S rRNA* was included in slightly less than half (48%) of the studies. Importantly, we



**Fig 6. Number of RGs tested for stability based on taxonomic categorisation of study organism.** (A) Number of RGs used per species grouping. (B) Number of RGs used in model and non-model species. In both cases, data is represented as mean (dot)  $\pm$  standard deviation (bars).

doi:10.1371/journal.pone.0141853.g006



**Table 1. Number (no.) of studies and reference genes (RGs) analysed according to: (A) species groupings (rodent = rodent model species, livestock = livestock mammal) and (B) model versus non-model species.**

	(A)							(B)	
	Human	Rodent	Livestock	Fish	Mammal—other	Bird	Primate	Model	Non-model
<b>No. studies</b>	174	101	91	51	22	10	7	390	66
<b>Mean no. RGs tested</b>	11.18	8.83	8.76	7.35	9.23	6.90	9.14	9.81	7.86
<b>Std. Dev of mean</b>	7.25	4.25	3.26	2.86	3.12	3.11	3.89	5.70	3.30

doi:10.1371/journal.pone.0141853.t001

found that there was a significant relationship between the number of RGs tested and the likelihood that *GAPDH* ( $z = -5.113, P < 0.001, n = 307$ ), *ACTB* ( $z = -3.506, P < 0.001, n = 328$ ) or *18S rRNA* ( $z = -4.311, P < 0.001, n = 177$ ) were chosen as the most stable RGs for normalisation, whereby these three genes are significantly less likely to be chosen when more RGs were screened in the panel (S1 Fig).

Most studies were conducted on *humans* (38% of studies), followed by *rodent model* species (22%) and *livestock* (19%) (Table 1A). On average, *human* studies tested the highest number of potential RGs (11.2) and *bird* studies the fewest (6.9) (Table 1A). Across all groupings, significantly different numbers of RGs were tested (Kruskal-Wallis statistic = 25.08,  $P = 0.0003$ , Fig 6A). Posthoc testing via Dunn’s multiple comparisons test revealed that the only significant difference between groupings was between *humans* and *fish*; whereby more RGs were tested in human studies (S5 Table). For other comparisons, some tests were underpowered due to few studies in some of the categories (Table 1A). Similarly, we found that significantly more RGs were tested for stability in studies of *model* than *non-model* species (Mann-Whitney  $U = 9989, P = 0.003$ , Table 1B, Fig 6B).

Given that a suitable RG should display stable expression independent of experimental treatment, one should expect that the same RGs will be chosen within a given species and tissue, independent of study-level effects. We tested this for a subset of LitRev2 studies focussing on rat brain (the most common non-human tissue investigated) [28–40] and found that different researchers selected different RGs even within the same cerebral sub-structures (Fig 7). Five genes were tested in more than half of the studies (*18S rRNA*, *Actb*, *β2m*, *Gapdh* and *Hprt1*). However, none of these popular RGs were selected as stable more than 50% of time. In fact, amongst RGs tested more than once, the most commonly selected was *Cypa* (although it should be noted that while this gene was tested in seven separate analyses, this comprised just two individual studies; Fig 7).

## Discussion

Accurate normalisation of transcripts is a fundamental requirement for gene expression studies. As such, we aimed to investigate the current state of research practice with respect to data normalisation in qPCR studies of gene expression. Taken together, these two quantitative reviews reveal there is a fundamental disconnect between studies testing panels of RGs for stability and the practical implementation of RGs in gene expression studies. While there is increasing awareness of the need to test RG stability before choosing which genes to normalise qPCR data with (LitRev 2; Fig 5A), few studies are currently converting this awareness into a more robust methodological approach when analysing the expression of GOIs (LitRev 1; Fig 2). Given that a clear set of guidelines are now available to assist researchers in performing gene expression analyses [24], the failure of the vast majority of LitRev1 studies to thoroughly investigate RG stability and use adequate numbers of RGs for data normalisation calls is surprising and concerning. Indeed, only two studies [41, 42] in LitRev1 noted that they followed

Structure	Substructure	Reference																	
		18S	Actb	$\beta$ 2m	Cypa	Cypb	Gapdh	Gusb	Hmbs	Hprt1	Oaz1	Pkg1	Ppia	Rpl13a	Rps6	Sdha	Tbp	Ubc	Yhwaz
Hippocampus	1																		[28]
	1																		[29]
	1																		[30]
	1																		[31]
	1																		[31]
	1																		[31]
	1																		[32]
	1																		[33]
	2																		[34]
Cortex	3																		[35]
	4																		[33]
	5																		[30]
	6																		[30]
	1																		[36]
	1																		[32]
	7																		[37]
Basal ganglia	8																		[38]
	8																		[33]
Hypothalamus	1																		[30]
	1																		[39]
Mesencephalon	9																		[38]
	9																		[38]
OB	1																		[33]
Whole	1																		[40]
	1																		[40]

**Fig 7. Testing and selection of RGs in different structures of the rat brain after various treatments.** Shading in grey indicates the gene was tested and black indicates it was selected as within the group of most stable genes required for accurate normalisation. OB is the olfactory bulb with cerebral substructure codes provided as a footnote. Only genes tested more than once included. See S4 Table for further details of each study, including treatment regime and full list of all RGs tested and selected. <sup>a</sup>Substructures: 1 = Whole, 2 = Dentate gyrus, 3 = Auditory, 4 = Prefrontal, 5 = Frontal, 6 = Temporal, 7 = Whole + Basal ganglia, 8 = Striatum, 9 = Nigra.

doi:10.1371/journal.pone.0141853.g007

the MIQE guidelines during experimental procedures. Furthermore, though papers testing panels of potential RGs for stability are increasing in number (LitRev2), the number of individual RGs tested has not significantly increased. However, there does appear to be a trend for increasing numbers of RGs to be tested in the last few years (Fig 5B). Interestingly, more effort appears to be expended on RG selection and use in those studies subsequently analysing more GOIs (Fig 4B and 4C). While this is encouraging, analysis of one or a few GOI(s) does not negate the requirement that data is first normalised with RGs that have been shown to be stable within the experimental context.

Of the 23 studies in LitRev2 that included twenty or more RGs, all have been conducted since 2010. All but one of these were conducted on model species: 18 (78%) were conducted on humans, three (13%) on model rodent species (rats and mice), and one (4%) on zebrafish. The remaining study was conducted on a less well characterised, but nevertheless domestic, species: sika deer (*Cervus nippon*). This bias towards larger numbers of RGs being tested when the study species is a model organism can likely be attributed to many more PCR primers being available in such species. Furthermore, two of the most promising avenues for robust identification of stable RGs are more tractable in model species. The first is the use of microarray data, whereby a comparative analysis of microarray results from the species (and tissue) of interest,

across treatment and experimental regimes and originating from different laboratories, can identify potentially stable genes for further testing via qPCR. This approach has been increasingly employed in the last few years (e.g. [43–45]), and often results in the discovery of novel RGs for data normalisation. The RefGenes tool in Genevestigator (<http://www.genevestigator.com>) has been developed to help identify stably expressed RGs from publically available microarray data [20]. The second is the use of commercially available plates preloaded with primers for a range of potential RGs, such as the TaqMan<sup>®</sup> Express Endogenous Control Plates and SYBR<sup>®</sup> OpenArray<sup>®</sup> Real-Time PCR Plates (Applied Biosystems) or the PrimePCR Assay system (BioRad). For certain model species, particularly human, rat and mouse, the use of these plates allows efficient and rapid screening of a large number of common potential RGs for expression stability within the context of a particular set of samples and experimental conditions. This approach is increasingly popular (e.g. [46–48]) and is likely to spread as such plates become available for a variety of species.

### The idea of a universal reference gene is a fallacy

Our first literature review revealed that *GAPDH* and *ACTB*, and to a lesser extent *18S rRNA*, continue to be common RG choices (Fig 3). In over 72% of the studies included (93/128), *GAPDH*, *ACTB* or *18S rRNA* were used as the single normalising gene. We show in the second literature review that, as increasing numbers of potential RGs are screened within a study, there is a decreasing chance these three genes are selected as having the most stable expression. If any of these genes were ‘universal’ endogenous controls, as they are clearly still viewed by many researchers, then they would be preferentially selected regardless of the number of other potential RGs screened. *GAPDH* is known to be stimulated by diverse biological factors ([12] and references therein), as is *ACTB* ([49] and references therein). Moreover, both *GAPDH* and *ACTB* have numerous pseudogenes in many species [50–52], adding a further potential source of error to qPCR studies, if primers have not been thoroughly tested to ensure they do not co-amplify pseudogenes in those species where they occur. The use of *18S rRNA* as a RG has also been questioned. First, ribosomal RNA (rRNA) comprises approximately 80% of total RNA [53]. Thus, rRNA transcripts (e.g. *18S rRNA* and 28S) tend to be much more abundant than mRNA transcripts (e.g. other RGs and GOIs) in cDNA generated by random oligonucleotides, and consequently have lower Cq values [54]. As recommended by Vandesompele *et al.* [11], transcript levels (and therefore Cq values) between GOIs and RGs should be similar as this ensures all transcripts are subject to the same kinetic interactions during qPCR. Second, rRNA is transcribed by RNA polymerase I while mRNA is transcribed by RNA polymerase II. Thus, mRNA and rRNA synthesis are regulated via different pathways, and changes in one may not affect the other [55]. Third, rRNA RGs can only be used when the cDNA template is reverse transcribed from total RNA using random hexamers. This is because rRNA lacks a polyA tail, and thus synthesis of cDNA with oligo d(T) primers will only amplify mRNA transcripts. It has been suggested that random hexamer primed cDNA can greatly over-estimate mRNA copy numbers [56] and that oligo dT-primed cDNA will provide a more accurate estimation of the mRNA pool [57]. Fourth, a common approach to circumvent potential genomic DNA contamination of RNA extracts is to design primers spanning introns, thus nucleic acid containing exonic regions (*i.e.* DNA) will fail to amplify. *18S rRNA* is devoid of introns, and as such this approach cannot be used with this marker. Fifth, as with *GAPDH* and *ACTB*, *18S rRNA* has been shown to be regulated in response to various biological stimuli (e.g. [30, 34, 46, 58]). We found no evidence for the presence of a universal RG amongst the studies included in LitRev2. The fact that *GAPDH* and *ACTB* remain popular choices is because they were popular choices as standards during the rise of Northern blotting in the 1980s and early 1990s [59]. Thus,

many research laboratories were familiar with the use of these genes and primer sequences were widely available as qPCR became popular. While the methodology moved on, it appears many researchers have been slow to update their protocols and move beyond using *GAPDH* and/or *ACTB* to a thorough investigation of the expression stability of multiple potential RGs.

## Methodological and analytical pitfalls

We noted that many studies in LitRev 2 added a layer of unnecessary experimental variability by performing a single RG stability analysis across multiple tissue types (e.g. [60–64]) or cell lines (e.g. [65–67]); we even found one paper which combined multiple tissues from two different species in a single analysis [68]. This is an unsound approach, as there is no *a priori* reason to expect that the same RG will be stable across different tissues, or even cell lines from the same tissue type. Studies that perform independent analyses on different tissues, for the same species and treatment regime, nearly universally identify different RGs in different tissues/cell types as the most stable (e.g. [16, 69–71]). Indeed, it has been shown that tissue/cell type can have a larger influence on RG stability than treatment regime [72]. Furthermore, within a single tissue, developmental stage can influence RG selection (e.g. [73–75]). Studies that perform analyses on both pooled tissues and single tissue extracts tend to reveal much higher GeNorm *V* values in the tissue mixture analyses—in other words they find that more RGs are required for accurate normalisation when disparate tissues are pooled (e.g. [76–78]). As such, studies comprising multiple tissue types or cell lines should either test RG stability in every tissue independently, or use RG independent normalisation methods (e.g. [5, 79, 80]). In our analysis of studies testing RGs in rat brains, we found little consensus over the choice of the most stable RGs, even within cerebral substructures (Fig 7). While *Actb*, *Gapdh* and *18S rRNA* were commonly tested in these studies, they were each selected less than half the time, showing that they cannot be considered universal RGs. Similarly, studies on different cell lines from the same tissue source tend to find different optimal RGs (e.g. liver cell lines [81]), as do studies on the same cell line performed by different researchers (e.g. HepG2 cell line [81, 82]). Unless a primary study of gene expression specifically aims to assess the differences in gene expression across multiple tissues/cell lines or treatments in a single analysis, we would strongly advocate that RG selection is performed separately for each different tissue/cell line and treatment regime under investigation. Indeed, RG selection should be performed amongst the exact same pool of samples that will be compared in the analysis of GOI expression. Thus, one can be confident that the correct normalisation has been applied for their given set of experimental conditions and samples.

## Perspectives

The incorrect use of normalising genes can have a profound effect on the conclusions drawn from studies of gene expression [15, 57, 83]. Studies that test the use of stable versus unstable RGs in the estimation of GOI expression show that the use of unstable RGs nearly universally results in spurious estimation of GOI expression (e.g. [18, 19, 84]). Yet our first literature review clearly demonstrates that many researchers continue to use a single, unvalidated reference gene to normalise data (Fig 2). One of the justifications often appears to be that the search for, and use of, multiple reference genes is cost prohibitive. We argue that this cost is insignificant compared to the cost of publishing erroneous results, given that this can lead to future research being predicated on incorrect or biased data. Given that many such studies are performed in the field of medical research, the consequences could also have negative impacts on drug development and human health. Furthermore, it is clear that validation of a RG as stable given the particular conditions of one laboratory may not mean that this RG can automatically

be considered to be stable for the given species and tissue/cell type when used in other laboratories (e.g. Fig 7). We advocate that researchers should routinely be testing the stability of a large panel of potential RGs (>10, preferably >20) for use as endogenous controls before commencing any study of gene expression. Additionally, newly available methods and analytical software for normalising qPCR data (e.g. [5, 79, 85–87]) should be investigated where appropriate. Our second literature review revealed that studies specifically testing RG stability are increasing over time (Fig 5). Given our previous statement (that RGs should be tested within the context of one's own samples); one might question the utility of publishing such studies. However, we would advocate that reference gene stability studies continue to be published, as they provide a useful starting point to identify which RGs might be appropriate to include in a test panel for their species and experimental set-up as well as primers and methodological details that can be invaluable to others. This is especially the case for non-model species, which are currently lagging behind well characterised species with respect to the number of RGs tested for stability (Table 1b; Fig 6b). Furthermore, the publication of such studies allows identification of patterns when research is synthesised quantitatively, such as the fact that selection of *GAPDH*, *ACTB* and *18S rRNA* is highly dependent on the number of potential RGs screened. Given the numerous potential pitfalls that exist when undertaking qPCR studies [59], it can seem an overwhelming task to obtain reliable results. We would argue that so long as care and attention is taken when choosing which RGs with which to normalise data, one of the major hurdles can be overcome.

## Supporting Information

**S1 Fig. Likelihood that three common RGs are selected as amongst the most stability expressed.** (A) *GAPDH*, (B) *ACTB* and (C) *18S rRNA*. All three genes are less likely to be selected as stable when more RGs are tested.

(EPS)

**S1 Table. PRISMA summary.** Details and locations within the text of specific components of the quantitative reviews conducted in this paper, based on the 2009 PRISMA group guidelines.

(DOC)

**S2 Table. LitRev1 studies.** Full details of data extracted and citations for all studies included in LitRev1.

(XLSX)

**S3 Table. LitRev2 studies.** Full details of data extracted and citations for all studies included in LitRev2.

(XLSX)

**S4 Table. Studies of gene expression in rat brains.** Full details of the rat brain studies summarised in Fig 7, including the strain and sex of rats used, the substructure of the brain analysed, the experimental treatment applied/investigated, the reference genes (RGs) tested and selected, the number (no.) required to reach a threshold stability value of below 0.15 in GeNorm (NA = this analysis not presented) and selected notes. Results summarised preferentially based on GeNorm results.

(DOCX)

**S5 Table. Dunn's test for posthoc multiple unequal group comparisons.** Differences in the number of RGs used in each of the seven categories of study species (rodent = rodent model species and livestock = livestock mammal).

(DOCX)

## Acknowledgments

We thank Alexis Avril for help with statistical analyses and comments on the manuscript, Michelle Wille for help with formatting figures and helpful discussions about the experimental procedures and Mark Dopson plus two anonymous reviewers for comments on the manuscript.

## Author Contributions

Conceived and designed the experiments: JRC JW. Performed the experiments: JRC. Analyzed the data: JRC. Contributed reagents/materials/analysis tools: JRC JW. Wrote the paper: JRC JW.

## References

1. Ginzinger DG. Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Exp Hematol*. 2002; 30(6):503–12. doi: [10.1016/S0301-472x\(02\)00806-8](https://doi.org/10.1016/S0301-472x(02)00806-8) PMID: [12063017](https://pubmed.ncbi.nlm.nih.gov/12063017/)
2. Nestorov J, Matic G, Elakovic I, Tanic N. Gene expression studies: How to obtain accurate and reliable data by quantitative real-time RT-PCR. *J Med Biochem*. 2013; 32(4):325–38. doi: [10.2478/jomb-2014-0001](https://doi.org/10.2478/jomb-2014-0001)
3. VanGuilder HD, Vrana KE, Freeman WM. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques*. 2008; 44(5):619–26. doi: [10.2144/000112776](https://doi.org/10.2144/000112776) PMID: [18474036](https://pubmed.ncbi.nlm.nih.gov/18474036/)
4. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun*. 2005; 6(4):279–84. doi: [10.1038/sj.gene.6364190](https://doi.org/10.1038/sj.gene.6364190) PMID: [15815687](https://pubmed.ncbi.nlm.nih.gov/15815687/)
5. Feuer R, Vlaic S, Arlt J, Sawodny O, Dahmen U, Zanger UM, et al. LEMming: A linear error model to normalize parallel quantitative real-time PCR (qPCR) data as an alternative to reference gene based methods. *PLoS One*. 2015; 10(9):e0135852. doi: [10.1371/journal.pone.0135852](https://doi.org/10.1371/journal.pone.0135852) PMID: [26325269](https://pubmed.ncbi.nlm.nih.gov/26325269/)
6. Pabinger S, Rödiger S, Kriegner A, Vierlinger K, Weinhäusel A. A survey of tools for the analysis of quantitative PCR (qPCR) data. *Biomol Detect Quantif*. 2014; 1(1):23–33. doi: [10.1016/j.bdq.2014.08.002](https://doi.org/10.1016/j.bdq.2014.08.002)
7. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods*. 2001; 25(4):402–8. doi: [10.1006/meth.2001.1262](https://doi.org/10.1006/meth.2001.1262) PMID: [11846609](https://pubmed.ncbi.nlm.nih.gov/11846609/)
8. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*. 2007; 8(2). doi: [10.1186/Gb-2007-8-2-R19](https://doi.org/10.1186/Gb-2007-8-2-R19)
9. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001; 29(9):e45. doi: [10.1093/nar/29.9.e45](https://doi.org/10.1093/nar/29.9.e45) PMID: [11328886](https://pubmed.ncbi.nlm.nih.gov/11328886/)
10. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc*. 2008; 3(6):1101–8. doi: [10.1038/nprot.2008.73](https://doi.org/10.1038/nprot.2008.73) PMID: [18546601](https://pubmed.ncbi.nlm.nih.gov/18546601/)
11. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002; 3(7):0034.1. doi: [10.1186/Gb-2002-3-7-Research0034](https://doi.org/10.1186/Gb-2002-3-7-Research0034)
12. Suzuki T, Higgins PJ, Crawford DR. Control selection for RNA quantitation. *Biotechniques*. 2000; 29(2):332–7. PMID: [10948434](https://pubmed.ncbi.nlm.nih.gov/10948434/)
13. Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics*. 2005; 21(3):389–95. doi: [10.1152/physiolgenomics.00025.2005](https://doi.org/10.1152/physiolgenomics.00025.2005) PMID: [15769908](https://pubmed.ncbi.nlm.nih.gov/15769908/)
14. Selvey S, Thompson EW, Matthaei K, Lea RA, Irving MG, Griffiths LR.  $\beta$ -actin—an unsuitable internal control for RT-PCR. *Mol Cell Probe*. 2001; 15(5):307–11. doi: [10.1006/mcpr.2001.0376](https://doi.org/10.1006/mcpr.2001.0376)
15. Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques*. 2004; 37(1):112–9. PMID: [15283208](https://pubmed.ncbi.nlm.nih.gov/15283208/)
16. Kosir R, Acimovic J, Golcink M, Perse M, Majdic G, Fink M, et al. Determination of reference genes for circadian studies in different tissues and mouse strains. *BMC Mol Biol*. 2010; 11. doi: [10.1186/1471-2199-11-60](https://doi.org/10.1186/1471-2199-11-60)
17. Liu MC, Yao BJ, Zhang H, Guo HY, Hu DY, Wang Q, et al. Identification of novel reference genes using sika deer antler transcriptome expression data and their validation for quantitative gene expression analysis. *Genes Genom*. 2014; 36(5):573–82. doi: [10.1007/s13258-014-0193-x](https://doi.org/10.1007/s13258-014-0193-x)

18. Anderson KC, Elizur A. Hepatic reference gene selection in adult and juvenile female Atlantic salmon at normal and elevated temperatures. *BMC Res Notes*. 2012; 5. doi: [10.1186/1756-0500-5-21](https://doi.org/10.1186/1756-0500-5-21)
19. Xiang-Hong J, Yan-Hong Y, Han-Jin X, Li-Long A, Ying-Mei X, Pei-Rong J, et al. Selection of reference genes for gene expression studies in PBMC from Bama miniature pig under heat stress. *Vet Immunol Immunop*. 2011; 144(1–2):160–6. doi: [10.1016/j.vetimm.2011.07.004](https://doi.org/10.1016/j.vetimm.2011.07.004)
20. Hruz T, Wyss M, Docquier M, Pfaffl MW, Masanetz S, Borghi L, et al. RefGenes: Identification of reliable and condition specific reference genes for RT-qPCR data normalization. *BMC Genomics*. 2011; 12. doi: [10.1186/1471-2164-12-156](https://doi.org/10.1186/1471-2164-12-156)
21. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnol Lett*. 2004; 26(6):509–15. doi: [10.1023/B:BILE.0000019559.84305.47](https://doi.org/10.1023/B:BILE.0000019559.84305.47) PMID: [15127793](https://pubmed.ncbi.nlm.nih.gov/15127793/)
22. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*. 2004; 64(15):5245–50. doi: [10.1158/0008-5472.CCR-04-0496](https://doi.org/10.1158/0008-5472.CCR-04-0496) PMID: [15289330](https://pubmed.ncbi.nlm.nih.gov/15289330/)
23. Robledo D, Hernandez-Urcera J, Cal RM, Pardo BG, Sanchez L, Martinez P, et al. Analysis of qPCR reference gene stability determination methods and a practical approach for efficiency calculation on a turbot (*Scophthalmus maximus*) gonad dataset. *BMC Genomics*. 2014; 15. doi: [10.1186/1471-2164-15-648](https://doi.org/10.1186/1471-2164-15-648)
24. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: Minimum Information for publication of Quantitative real-time PCR Experiments. *Clin Chem*. 2009; 55(4):611–22. doi: [10.1373/clinchem.2008.112797](https://doi.org/10.1373/clinchem.2008.112797) PMID: [19246619](https://pubmed.ncbi.nlm.nih.gov/19246619/)
25. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009; 6(7):e1000097. doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097) PMID: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)
26. Massey FJ. The Kolmogorov-Smirnov Test for goodness of fit. *J Am Stat Assoc*. 1951; 46(253):68–78. doi: [10.2307/2280095](https://doi.org/10.2307/2280095)
27. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
28. Marques TEBS, de Mendonca LR, Pereira MG, de Andrade TG, Garcia-Cairasco N, Paco-Larson ML, et al. Validation of suitable reference genes for expression studies in different pilocarpine-induced models of mesial temporal lobe epilepsy. *PLoS One*. 2013; 8(8):e71892. doi: [10.1371/journal.pone.0071892](https://doi.org/10.1371/journal.pone.0071892) PMID: [24009668](https://pubmed.ncbi.nlm.nih.gov/24009668/)
29. Bonefeld BE, Elfving B, Wegener G. Reference genes for normalization: a study of rat brain tissue. *Synapse*. 2008; 62(4):302–9. doi: [10.1002/syn.20496](https://doi.org/10.1002/syn.20496) PMID: [18241047](https://pubmed.ncbi.nlm.nih.gov/18241047/)
30. Julian GS, de Oliveira RW, Perry JC, Tufik S, Chagas JR. Validation of housekeeping genes in the brains of rats submitted to chronic intermittent hypoxia, a sleep apnea model. *PLoS One*. 2014; 9(10):e109902. doi: [10.1371/journal.pone.0109902](https://doi.org/10.1371/journal.pone.0109902) PMID: [25289636](https://pubmed.ncbi.nlm.nih.gov/25289636/)
31. Langnaese K, John R, Schweizer H, Ebmeyer U, Keilhoff G. Selection of reference genes for quantitative real-time PCR in a rat asphyxial cardiac arrest model. *BMC Mol Biol*. 2008; 9. doi: [10.1186/1471-2199-9-53](https://doi.org/10.1186/1471-2199-9-53)
32. Tanic N, Perovic M, Mladenovic A, Ruzdijic S, Kanazir S. Effects of aging, dietary restriction and glucocorticoid treatment on housekeeping gene expression in rat cortex and hippocampus-evaluation by real time RT-PCR. *J Mol Neurosci*. 2007; 32(1):38–46. doi: [10.1007/s12031-007-0006-7](https://doi.org/10.1007/s12031-007-0006-7) PMID: [17873286](https://pubmed.ncbi.nlm.nih.gov/17873286/)
33. Moura AC, Lazzari VM, Agnes G, Almeida S, Giovenardi M, Veiga AB. Transcriptional expression study in the central nervous system of rats: what gene should be used as internal control? *Einstein (São Paulo)*. 2014; 12(3):336–41. doi: [10.1590/S1679-45082014AO3042](https://doi.org/10.1590/S1679-45082014AO3042)
34. Swijsen A, Nelissen K, Janssen D, Rigo JM, Hoogland G. Validation of reference genes for quantitative real-time PCR studies in the dentate gyrus after experimental febrile seizures. *BMC Res Notes*. 2012; 5. doi: [10.1186/1756-0500-5-685](https://doi.org/10.1186/1756-0500-5-685)
35. Chen J, Ruan R. Identifying stable reference genes for evaluation of antioxidative enzyme gene expression in auditory cortex and cochlea of young and old Fischer 344 rats. *Acta Otolaryngol*. 2009; 129(6):644–50. doi: [10.1080/00016480802311015](https://doi.org/10.1080/00016480802311015) PMID: [18720057](https://pubmed.ncbi.nlm.nih.gov/18720057/)
36. Gubern C, Hurtado O, Rodriguez R, Morales JR, Romera VG, Moro MA, et al. Validation of housekeeping genes for quantitative real-time PCR in in-vivo and in-vitro models of cerebral ischaemia. *BMC Mol Biol*. 2009; 10. doi: [10.1186/1471-2199-10-57](https://doi.org/10.1186/1471-2199-10-57)

37. Cook NL, Kleinig TJ, van den Heuvel C, Vink R. Reference genes for normalising gene expression data in collagenase-induced rat intracerebral haemorrhage. *BMC Mol Biol.* 2010; 11. doi: [10.1186/1471-2199-11-7](https://doi.org/10.1186/1471-2199-11-7)
38. He Y, Yu S, Bae E, Shen H, Wang Y. Methamphetamine alters reference gene expression in nigra and striatum of adult rat brain. *Neurotoxicology.* 2013; 39:138–45. doi: [10.1016/j.neuro.2013.08.009](https://doi.org/10.1016/j.neuro.2013.08.009) PMID: [24042092](https://pubmed.ncbi.nlm.nih.gov/24042092/)
39. Li B, Matter EK, Hoppert HT, Grayson BE, Seeley RJ, Sandoval DA. Identification of optimal reference genes for RT-qPCR in the rat hypothalamus and intestine for the study of obesity. *Int J Obes.* 2014; 38(2):192–7. doi: [10.1038/ijo.2013.86](https://doi.org/10.1038/ijo.2013.86)
40. Lee KS, Alvarenga TA, Guindalini C, Andersen ML, Castro RMRPS, Tufik S. Validation of commonly used reference genes for sleep-related gene expression studies. *BMC Mol Biol.* 2009; 10. doi: [10.1186/1471-2199-10-45](https://doi.org/10.1186/1471-2199-10-45)
41. Gonzalez-Herrera L, Valenzuela A, Marchal JA, Lorente JA, Villanueva E. Studies on RNA integrity and gene expression in human myocardial tissue, pericardial fluid and blood, and its postmortem stability. *Forensic Sci Int.* 2013; 232(1–3):218–28. doi: [10.1016/j.forsciint.2013.08.001](https://doi.org/10.1016/j.forsciint.2013.08.001) PMID: [24053884](https://pubmed.ncbi.nlm.nih.gov/24053884/)
42. Tucker JD, Grever WE, Joiner MC, Konski AA, Thomas RA, Smolinski JM, et al. Gene expression-based detection of radiation exposure in mice after treatment with granulocyte colony-stimulating factor and lipopolysaccharide. *Radiat Res.* 2012; 177(2):209–19. doi: [10.1667/RR2749.1](https://doi.org/10.1667/RR2749.1) PMID: [22128785](https://pubmed.ncbi.nlm.nih.gov/22128785/)
43. Chang Z, Ling CY, Yamashita M, Welham NV. Microarray-driven validation of reference genes for quantitative real-time polymerase chain reaction in a rat vocal fold model of mucosal injury. *Anal Biochem.* 2010; 406(2):214–21. doi: [10.1016/J.Ab.2010.07.021](https://doi.org/10.1016/J.Ab.2010.07.021) PMID: [20670610](https://pubmed.ncbi.nlm.nih.gov/20670610/)
44. Guo Y, Chen JX, Yang S, Fu XP, Zhang Z, Chen KH, et al. Selection of reliable reference genes for gene expression study in nasopharyngeal carcinoma. *Acta Pharmacol Sin.* 2010; 31(11):1487–94. doi: [10.1038/Aps.2010.115](https://doi.org/10.1038/Aps.2010.115) PMID: [21052085](https://pubmed.ncbi.nlm.nih.gov/21052085/)
45. Noriega NC, Kohama SG, Urbanski HF. Microarray analysis of relative gene expression stability for selection of internal reference genes in the rhesus macaque brain. *BMC Mol Biol.* 2010; 11. doi: [10.1186/1471-2199-11-47](https://doi.org/10.1186/1471-2199-11-47)
46. Das RK, Banerjee S, Shapiro BH. Extensive sex- and/or hormone-dependent expression of rat housekeeping genes. *Endocr Res.* 2013; 38(2):105–11. doi: [10.3109/07435800.2012.723294](https://doi.org/10.3109/07435800.2012.723294) PMID: [22992156](https://pubmed.ncbi.nlm.nih.gov/22992156/)
47. Ribeiro MA, dos Reis MB, de Moraes LN, Briton-Jones C, Rainho CA, Scarano WR. Defining suitable reference genes for RT-qPCR analysis on human sertoli cells after 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) exposure. *Mol Biol Rep.* 2014; 41(11):7063–6. doi: [10.1007/s11033-014-3637-0](https://doi.org/10.1007/s11033-014-3637-0) PMID: [25078986](https://pubmed.ncbi.nlm.nih.gov/25078986/)
48. Schildberg T, Rauh J, Bretschneider H, Stiehler M. Identification of suitable reference genes in bone marrow stromal cells from osteoarthritic donors. *Stem Cell Res.* 2013; 11(3):1288–98. doi: [10.1016/j.scr.2013.08.015](https://doi.org/10.1016/j.scr.2013.08.015) PMID: [24080205](https://pubmed.ncbi.nlm.nih.gov/24080205/)
49. Ruan WJ, Lai MD. Actin, a reliable marker of internal control? *Clin Chim Acta.* 2007; 385(1–2):1–5. doi: [10.1016/j.cca.2007.07.003](https://doi.org/10.1016/j.cca.2007.07.003) PMID: [17698053](https://pubmed.ncbi.nlm.nih.gov/17698053/)
50. Sun Y, Li Y, Luo DZ, Liao DJ. Pseudogenes as weaknesses of ACTB (Actb) and GAPDH (Gapdh) used as reference genes in reverse transcription and polymerase chain reactions. *PLoS One.* 2012; 7(8): e41659. doi: [10.1371/journal.pone.0041659](https://doi.org/10.1371/journal.pone.0041659) PMID: [22927912](https://pubmed.ncbi.nlm.nih.gov/22927912/)
51. Liu YJ, Zheng D, Balasubramanian S, Carriero N, Khurana E, Robilotto R, et al. Comprehensive analysis of the pseudogenes of glycolytic enzymes in vertebrates: the anomalously high number of GAPDH pseudogenes highlights a recent burst of retrotrans-positional activity. *BMC Genomics.* 2009; 10. doi: [10.1186/1471-2164-10-480](https://doi.org/10.1186/1471-2164-10-480)
52. Ng SY, Gunning P, Eddy R, Ponte P, Leavitt J, Shows T, et al. Evolution of the functional human beta-actin gene and its multi-pseudogene family—conservation of noncoding regions and chromosomal dispersion of pseudogenes. *Mol Cell Biol.* 1985; 5(10):2720–32. doi: [10.1128/MCB.5.10.2720](https://doi.org/10.1128/MCB.5.10.2720) PMID: [3837182](https://pubmed.ncbi.nlm.nih.gov/3837182/)
53. Goidin D, Mamessier A, Staquet MJ, Schmitt D, Berthier-Vergnes O. Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and  $\beta$ -actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations. *Anal Biochem.* 2001; 295(1):17–21. doi: [10.1006/abio.2001.5171](https://doi.org/10.1006/abio.2001.5171) PMID: [11476540](https://pubmed.ncbi.nlm.nih.gov/11476540/)
54. Solanas M, Moral R, Escrich E. Unsuitability of using ribosomal RNA as loading control for Northern blot analyses related to the imbalance between messenger and ribosomal RNA content in rat mammary tumors. *Anal Biochem.* 2001; 288(1):99–102. doi: [10.1006/abio.2000.4889](https://doi.org/10.1006/abio.2000.4889) PMID: [11141312](https://pubmed.ncbi.nlm.nih.gov/11141312/)
55. Radonić A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A. Guideline to reference gene selection for quantitative real-time PCR. *Biochem Bioph Res Co.* 2004; 313(4):856–62. doi: [10.1016/j.bbrc.2003.11.177](https://doi.org/10.1016/j.bbrc.2003.11.177)



56. Zhang JL, Byrne CD. Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. *Biochem J*. 1999; 337:231–41. doi: [10.1042/0264-6021:3370231](https://doi.org/10.1042/0264-6021:3370231) PMID: [9882620](https://pubmed.ncbi.nlm.nih.gov/9882620/)
57. Bustin SA, Benes V, Nolan T, Pfaffl MW. Quantitative real-time RT-PCR—a perspective. *J Mol Endocrinol*. 2005; 34(3):597–601. doi: [10.1677/Jme.1.01755](https://doi.org/10.1677/Jme.1.01755) PMID: [15956331](https://pubmed.ncbi.nlm.nih.gov/15956331/)
58. Schmittgen TD, Zakrajsek BA. Effect of experimental treatment on housekeeping gene expression: Validation by real-time, quantitative RT-PCR. *J Biochem Bioph Meth*. 2000; 46(1–2):69–81. doi: [10.1016/S0165-022x\(00\)00129-9](https://doi.org/10.1016/S0165-022x(00)00129-9)
59. Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech*. 2004; 15(3):155–66. PMID: [15331581](https://pubmed.ncbi.nlm.nih.gov/15331581/)
60. Mitter K, Kotoulas G, Magoulas A, Mulero V, Sepulcre P, Figueras A, et al. Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus labrax*). *Comp Biochem Physiol B Biochem Mol Biol*. 2009; 153(4):340–7. doi: [10.1016/j.cbpb.2009.04.009](https://doi.org/10.1016/j.cbpb.2009.04.009) PMID: [19398033](https://pubmed.ncbi.nlm.nih.gov/19398033/)
61. Nygard AB, Jorgensen CB, Cirera S, Fredholm M. Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Mol Biol*. 2007; 8. doi: [10.1186/1471-2199-8-67](https://doi.org/10.1186/1471-2199-8-67)
62. Park SJ, Kwon SG, Hwang JH, Park DH, Kim TW, Kim CW. Selection of appropriate reference genes for RT-qPCR analysis in Berkshire, Duroc, Landrace, and Yorkshire pigs. *Gene*. 2015; 558(1):152–8. doi: [10.1016/j.gene.2014.12.052](https://doi.org/10.1016/j.gene.2014.12.052) PMID: [25550045](https://pubmed.ncbi.nlm.nih.gov/25550045/)
63. Shimamoto Y, Kitamura H, Niimi K, Yoshikawa Y, Hoshi F, Ishizuka M, et al. Selection of suitable reference genes for mRNA quantification studies using common marmoset tissues. *Mol Biol Rep*. 2013; 40(12):6747–55. doi: [10.1007/s11033-013-2791-0](https://doi.org/10.1007/s11033-013-2791-0) PMID: [24068436](https://pubmed.ncbi.nlm.nih.gov/24068436/)
64. Altmann S, Rebl A, Kuhn C, Goldammer T. Identification and de novo sequencing of housekeeping genes appropriate for gene expression analyses in farmed maraena whitefish (*Coregonus maraena*) during crowding stress. *Fish Physiol Biochem*. 2015; 41(2):397–412. doi: [10.1007/s10695-014-9991-y](https://doi.org/10.1007/s10695-014-9991-y) PMID: [25249196](https://pubmed.ncbi.nlm.nih.gov/25249196/)
65. Cicinnati VR, Shen QL, Sotiropoulos GC, Radtke A, Gerken G, Beckebaum S. Validation of putative reference genes for gene expression studies in human hepatocellular carcinoma using real-time quantitative RT-PCR. *BMC Cancer*. 2008; 8. doi: [10.1186/1471-2407-8-350](https://doi.org/10.1186/1471-2407-8-350)
66. Ersahin T, Carkacioglu L, Can T, Konu O, Atalay V, Cetin-Atalay R. Identification of novel reference genes based on MeSH categories. *PLoS One*. 2014; 9(3):e93341. doi: [10.1371/journal.pone.0093341](https://doi.org/10.1371/journal.pone.0093341) PMID: [24682035](https://pubmed.ncbi.nlm.nih.gov/24682035/)
67. Liu LL, Zhao H, Ma TF, Ge F, Chen CS, Zhang YP. Identification of valid reference genes for the normalization of RT-qPCR expression studies in human breast cancer cell lines treated with and without transient transfection. *PLoS One*. 2015; 10(1):e0117058. doi: [10.1371/journal.pone.0117058](https://doi.org/10.1371/journal.pone.0117058) PMID: [25617865](https://pubmed.ncbi.nlm.nih.gov/25617865/)
68. Bonnet M, Bernard L, Bes S, Leroux C. Selection of reference genes for quantitative real-time PCR normalisation in adipose tissue, muscle, liver and mammary gland from ruminants. *Animal*. 2013; 7(8):1344–53. doi: [10.1017/S1751731113000475](https://doi.org/10.1017/S1751731113000475) PMID: [23552195](https://pubmed.ncbi.nlm.nih.gov/23552195/)
69. Cabiati M, Raucci S, Caselli C, Guzzardi MA, D'Amico A, Prescimone T, et al. Tissue-specific selection of stable reference genes for real-time PCR normalization in an obese rat model. *J Mol Endocrinol*. 2012; 48(3):251–60. doi: [10.1530/Jme-12-0024](https://doi.org/10.1530/Jme-12-0024) PMID: [22493144](https://pubmed.ncbi.nlm.nih.gov/22493144/)
70. Stephens AS, Stephens SR, Morrison NA. Internal control genes for quantitative RT-PCR expression analysis in mouse osteoblasts, osteoclasts and macrophages. *BMC Res Notes*. 2011; 4. doi: [10.1186/1756-0500-4-410](https://doi.org/10.1186/1756-0500-4-410)
71. Ji H, Wang JF, Liu JX, Guo JR, Wang ZW, Zhang X, et al. Selection of reliable reference genes for real-time qRT-PCR analysis of Zi geese (*Anser anser domestica*) gene expression. *Asian Austral J Anim*. 2013; 26(3):423–32. doi: [10.5713/ajas.2012.12417](https://doi.org/10.5713/ajas.2012.12417)
72. Jacobsen AV, Yemaneab BT, Jass J, Scherbak N. Reference gene selection for qPCR is dependent on cell type rather than treatment in colonic and vaginal human epithelial cell lines. *PLoS One*. 2014; 9(12):e115592. doi: [10.1371/journal.pone.0115592](https://doi.org/10.1371/journal.pone.0115592) PMID: [25526394](https://pubmed.ncbi.nlm.nih.gov/25526394/)
73. Rocha-Martins M, Njaine B, Silveira MS. Avoiding pitfalls of internal controls: Validation of reference genes for analysis by qRT-PCR and Western blot throughout rat retinal development. *PLoS One*. 2012; 7(8):e43028. doi: [10.1371/journal.pone.0043028](https://doi.org/10.1371/journal.pone.0043028) PMID: [22916200](https://pubmed.ncbi.nlm.nih.gov/22916200/)
74. Uddin M, Cinar M, Tesfaye D, Looft C, Tholen E, Schellander K. Age-related changes in relative expression stability of commonly used housekeeping genes in selected porcine tissues. *BMC Res Notes*. 2011; 4. doi: [10.1186/1756-0500-4-441](https://doi.org/10.1186/1756-0500-4-441)

75. Fernandes JMO, Mommens M, Hagen O, Babiak I, Solberg C. Selection of suitable reference genes for real-time PCR studies of Atlantic halibut development. *Comp Biochem Phys B*. 2008; 150(1):23–32. doi: [10.1016/j.cbpb.2008.01.003](https://doi.org/10.1016/j.cbpb.2008.01.003)
76. Taki FA, Abdel-Rahman AA, Zhang BH. A comprehensive approach to identify reliable reference gene candidates to investigate the link between alcoholism and endocrinology in Sprague-Dawley rats. *PLoS One*. 2014; 9(5):e94311. doi: [10.1371/journal.pone.0094311](https://doi.org/10.1371/journal.pone.0094311) PMID: [24824616](https://pubmed.ncbi.nlm.nih.gov/24824616/)
77. Olias P, Adam I, Meyer A, Scharff C, Gruber AD. Reference genes for quantitative gene expression studies in multiple avian species. *PLoS One*. 2014; 9(6):e99678. doi: [10.1371/journal.pone.0099678](https://doi.org/10.1371/journal.pone.0099678) PMID: [24926893](https://pubmed.ncbi.nlm.nih.gov/24926893/)
78. Saremi B, Sauerwein H, Danicke S, Mielenz M. Technical note: Identification of reference genes for gene expression studies in different bovine tissues focusing on different fat depots. *J Dairy Sci*. 2012; 95(6):3131–8. doi: [10.3168/jds.2011-4803](https://doi.org/10.3168/jds.2011-4803) PMID: [22612949](https://pubmed.ncbi.nlm.nih.gov/22612949/)
79. Matz MV, Wright RM, Scott JG. No control genes required: Bayesian analysis of qRT-PCR data. *PLoS One*. 2013; 8(8):e71448. doi: [10.1371/journal.pone.0071448](https://doi.org/10.1371/journal.pone.0071448) PMID: [23977043](https://pubmed.ncbi.nlm.nih.gov/23977043/)
80. Steibel JP, Poletto R, Coussens PM, Rosa GJ. A powerful and flexible linear mixed model framework for the analysis of relative quantification RT-PCR data. *Genomics*. 2009; 94(2):146–52. doi: [10.1016/j.ygeno.2009.04.008](https://doi.org/10.1016/j.ygeno.2009.04.008) PMID: [19422910](https://pubmed.ncbi.nlm.nih.gov/19422910/)
81. Hashemi A, Roohvand F, Ghahremani MH. Selection of valid reference genes for expression studies of hepatic cell lines under IFN-alpha treatment. *Biochem Bioph Res Co*. 2012; 426(4):649–53. doi: [10.1016/j.bbrc.2012.09.009](https://doi.org/10.1016/j.bbrc.2012.09.009)
82. Fox BC, Devonshire AS, Schutte ME, Foy CA, Minguez J, Przyborski S, et al. Validation of reference gene stability for APAP hepatotoxicity studies in different in vitro systems and identification of novel potential toxicity biomarkers. *Toxicol in Vitro*. 2010; 24(7):1962–70. doi: [10.1016/j.tiv.2010.08.007](https://doi.org/10.1016/j.tiv.2010.08.007) PMID: [20732408](https://pubmed.ncbi.nlm.nih.gov/20732408/)
83. Bustin SA. Why the need for qPCR publication guidelines?—The case for MIQE. *Methods*. 2010; 50(4):217–26. doi: [10.1016/j.ymeth.2009.12.006](https://doi.org/10.1016/j.ymeth.2009.12.006) PMID: [20025972](https://pubmed.ncbi.nlm.nih.gov/20025972/)
84. Kadegowda AKG, Bionaz M, Thering B, Piperova LS, Erdman RA, Loor JJ. Identification of internal control genes for quantitative polymerase chain reaction in mammary tissue of lactating cows receiving lipid supplements. *J Dairy Sci*. 2009; 92(5):2007–19. doi: [10.3168/jds.2008-1655](https://doi.org/10.3168/jds.2008-1655) PMID: [19389958](https://pubmed.ncbi.nlm.nih.gov/19389958/)
85. Ballester M, Cordon R, Folch JM. DAG expression: High-throughput gene expression analysis of real-time PCR data using standard curves for relative quantification. *PLoS One*. 2013; 8(11):e80385. doi: [10.1371/journal.pone.0080385](https://doi.org/10.1371/journal.pone.0080385) PMID: [24260380](https://pubmed.ncbi.nlm.nih.gov/24260380/)
86. Ling D. SASqPCR: robust and rapid analysis of RT-qPCR data in SAS. *PLoS One*. 2012; 7(1):e29788. doi: [10.1371/journal.pone.0029788](https://doi.org/10.1371/journal.pone.0029788) PMID: [22238653](https://pubmed.ncbi.nlm.nih.gov/22238653/)
87. Mar JC, Kimura Y, Schroder K, Irvine KM, Hayashizaki Y, Suzuki H, et al. Data-driven normalization strategies for high-throughput quantitative RT-PCR. *BMC Bioinformatics*. 2009; 10. doi: [10.1186/1471-2105-10-110](https://doi.org/10.1186/1471-2105-10-110)