

# Validation of adequate endogenous reference genes for reverse transcription-qPCR studies in human post-mortem brain tissue of SIDS cases

Noha El-Kashef<sup>1,2</sup> · Iva Gomes<sup>1</sup> · Katja Mercer-Chalmers-Bender<sup>1</sup> · Peter M. Schneider<sup>1</sup> · Markus A. Rothschild<sup>1</sup> · Martin Juebner<sup>1</sup>

Accepted: 14 September 2015 / Published online: 5 October 2015  
© Springer Science+Business Media New York 2015

**Abstract** Sudden infant death syndrome (SIDS) is the main cause of post-neonatal infant death in most developed countries. It is still of ambiguous etiology. Gene expression studies of relevant target genes using reverse transcription quantitative real-time PCR (RT-qPCR) in SIDS cases, and comparing them with age-matched controls, could help in understanding the pathogenesis of SIDS. However, selecting inadequate reference genes used for normalization of the RT-qPCR gene expression data can give misleading results. The aim of the present study was to identify reference genes with the most stable expression in post-mortem brainstem samples of SIDS and control cases. Among the five candidate reference genes (*GAPDH*, *GUSB*, *HMBS*, *SDHA*, *UBXN6*) studied in both groups, *SDHA* and *UBXN6* were identified as the most stable. To further demonstrate the importance of using validated genes for RT-qPCR data normalization, the expression of a potential gene of interest in SIDS, the *RPS27A* gene, was evaluated using validated versus non-validated reference genes for normalization. This gene encodes the ubiquitin protein that has been shown in other pathological studies to be induced in SIDS. Using the identified most stable genes for normalization of *RPS27A* gene expression data revealed, as expected, a statistically significant up-regulation in SIDS as compared to the controls. However, using a single unstable reference gene for normalization resulted in no significant differences in transcript abundance of

*RPS27A* between SIDS and the controls. This emphasizes the need for validation of the suitability of reference genes used in a given tissue type under certain experimental conditions.

**Keywords** Reference genes · RT-qPCR · qbase<sup>PLUS</sup> · Brainstem · SIDS

## Introduction

Gene expression analysis using reverse transcription quantitative real-time PCR (RT-qPCR) is increasingly applied nowadays due to the high sensitivity and accuracy of the technique as well as its ease of performance and rapid quantification. Either absolute or relative quantification methods can be chosen for the qPCR step. Absolute quantification usually requires a calibration curve that is generated using known quantities of a serially diluted sample to extrapolate the quantity of an unknown sample. The relative quantification method is commonly used to detect the relative physiological changes in gene expression. It relies on the normalization of a gene of interest to one or several endogenous reference gene(s), then comparing the expression of the normalized gene of interest in a given sample (e.g., a treated sample), to another reference sample (e.g., an untreated control sample) [1, 2].

The normalization to the endogenous control gene or reference gene, commonly known as housekeeping gene, is important to correct for non-specific experimental variability, such as the amount and quality of the starting material, and PCR enzymatic efficiencies. Ideally, a housekeeping gene should show stable expression in different tissue types and individuals, in different disease states, or under different experimental conditions [3–7].

✉ Martin Juebner  
martin.juebner@uk-koeln.de

<sup>1</sup> Medical Faculty, Institute of Legal Medicine, University of Cologne, Cologne, Germany

<sup>2</sup> Department of Forensic Medicine and Clinical Toxicology, Faculty of Medicine, Assiut University, Assiut, Egypt

However, this is not always the case. Several studies have shown variability in the expression of the commonly used housekeeping genes [8–13]. These findings emphasize the importance of validation of the used reference genes.

Since the success of gene expression analysis in post-mortem human brain tissue has been previously demonstrated [14, 15], RT-qPCR can be used to study the expression of different genes that are thought to be enrolled in the mechanism of diseases with vague etiology, like sudden infant death syndrome (SIDS). SIDS is defined as the sudden unexpected death of an infant less than 1 year of age, with onset of the fatal episode apparently occurring during sleep, that remains unexplained after a thorough investigation, including performance of a complete autopsy and review of the circumstances of death and the clinical history [16]. It is one of the leading causes of death in most developed countries after the first month to the end of the first year of life [17]. Though some risk factors have been identified [18–20] and several hypotheses about its etiology have been proposed [21–25], its pathophysiology remains unclear.

To our knowledge, this is the first validation study of endogenous control genes for qPCR data normalization in human post-mortem brain tissue from SIDS cases. Casale et al. [26] have performed a RT-qPCR based study for the analysis of the *5-HTT* gene expression in post-mortem brain specimens of SIDS cases, however, only a single reference gene was used for normalization of the expression data, *GAPDH*, without previous validation of its stability in their sample set of SIDS. As discussed above, the use of a single reference gene, without evaluation of gene expression stability, can lead to inaccurate results and data interpretation [3]. This motivated us to perform this study to investigate the expression stability of several frequently used endogenous reference genes in human post-mortem brainstem medulla oblongata tissues of SIDS cases and age-matched controls. A total of five potential reference genes were selected, namely, *GAPDH*, *GUSB*, *HMBS*, *SDHA*, and *UBXN6*, and their gene expression stability was evaluated. We aimed to determine the optimal endogenous reference genes and the minimum required number of genes for accurate qPCR data normalization.

In addition, to show the effects of choosing valid versus non-valid normalizing reference gene(s), different normalization strategies were tested to analyze the expression levels of the *RPS27A* (*ribosomal protein S27a*) target gene. This gene encodes a fusion protein consisting of ubiquitin and ribosomal protein S27a. Ubiquitin is a highly conserved protein found in every eukaryotic cell. It plays an important role in degradation of proteins. It is induced by cell stress such as hypoxia to help in protecting cells from damage. Quan et al. [27] showed that intranuclear ubiquitin immunoreactivity of the midbrain neurons was induced by

a fatal severe stress on the central nervous system in asphyxiation and drowning. Moreover, immunohistochemical positivity was found in six SIDS infants with antibodies against ubiquitin in brainstem neurons [28]. In addition, significant correlations were found between pathological data referring to ubiquitin and physiological data in SIDS cases [29]. Based on these evidences, up-regulation of ubiquitin gene expression can be expected in cases of SIDS when compared to the controls.

## Materials and methods

### Tissue samples

This study included post-mortem brainstem medulla oblongata samples obtained from 72 infants that died under the age of 1 year. Fifty cases of the 72 were classified as SIDS cases according to the criteria of the San Diego definition [16]. The remaining twenty-two were age-matched controls with determined causes of death other than SIDS. Post-mortem interval (PMI), age, and gender groups of the study subjects as well as causes of death of the controls are summarized in Table 1. Samples were collected by cutting a section of the *medulla oblongata* in the region of the crossing of the pyramids (*decussatio pyramidum*) during forensic autopsies performed at the Institute of Legal Medicine, Medical Faculty, University of Cologne, Cologne, Germany, and were immediately stored at  $-80^{\circ}\text{C}$  until RNA extraction. This work was approved by the local ethical committee (no. 072/33, University of Cologne).

### RNA extraction

Total RNA was extracted from the frozen brainstem medulla tissue samples using the EZ1 RNA universal tissue kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Briefly, about 60 mg of each frozen brainstem tissue sample were quickly cut and weighed, and then homogenized in 750  $\mu\text{L}$  QIAzol lysis reagent using an Ultra-Turrax<sup>®</sup> homogenizer (Ika, Staufen, Germany). After addition of chloroform, the homogenate was separated into aqueous and organic phases by centrifugation at  $12,000\times g$  for 15 min at  $4^{\circ}\text{C}$ . The upper aqueous phase containing the RNA was manually removed, and the RNA was purified using the automated, magnetic-particle based procedure of the BioRobot EZ1 system from QIAGEN in a 50  $\mu\text{L}$  elution volume. The DNase I digestion step was included using the RNase-free DNase Set (QIAGEN) following the recommended protocol of the manufacturer. To minimize the risk of RNA degradation, recommended laboratory practices were used such as thorough RNase/

**Table 1** Summary of study subjects included in the present study

Parameter	SIDS group ( <i>n</i> = 50)	Control group ( <i>n</i> = 22)
Cause of death	SIDS	Congenital malformations/cardiac causes (6) Pulmonary causes (5) Asphyxia (6) Traffic accident trauma (2) Others (3) <sup>a</sup>
Age at death (months)		
≤1	6	6
2–4	26	2
5–6	11	2
≥7	7	12
Gender		
Male	33	10
Female	17	12
PMI (h)		
0–24	14	4
>24–48	21	8
>48	15	10
Mean (range)	44.4 (10–144)	48.8 (12.5–125)

<sup>a</sup> Including one case of each of gastrointestinal infection, hyperthermia, and hypoxic brain damage

DNase decontamination of the working surfaces and equipment as well as using nuclease-free water (QIAGEN) and reagents. The concentration of RNA was determined by measuring the absorbance at 260 nm (A<sub>260</sub>) using the NanoDrop 2000 spectrophotometer (Thermoscientific Inc., Wilmington, DE, USA) and its purity was assessed by calculating the 260/280 absorbance ratio.

### cDNA synthesis

An equal amount of total RNA from each sample was used to generate complementary DNA (cDNA) using the iScript cDNA Synthesis Kit (Bio-Rad Inc., Hercules, CA, USA) with a blend of oligo (dT) and random hexamer primers, according to the manufacturer's protocol. In brief, 20 µL of total RNA containing ca. 580–585 ng were used as a starting material for the reverse transcription reaction, to which we added 6 µL of 5× iScript reaction mix, 1.5 µL of iScript reverse transcriptase and 2.5 µL of nuclease-free H<sub>2</sub>O. After mixing, the reactions were incubated in a thermal cycler at 25 °C for 5 min, at 42 °C for 30 min, and at 85 °C for 5 min.

### Selection of genes

Five potential reference genes were analyzed in this study (Table 2): *glyceraldehyde-3-phosphate dehydrogenase (GAPDH)*; *glucuronidase, beta (GUSB)*; *hydroxy methyl bilane synthase (HMBS)*; *succinate dehydrogenase*

*complex, subunit A, flavoprotein (SDHA)*; and *UBX domain protein 6 (UBXN6)*. The genes, except for *UBXN6* which was a novel reference gene, were commonly used endogenous control genes chosen from the relevant literature and have been previously validated in human post-mortem brain tissues without obvious neuropathological changes [5, 30], or in human post-mortem brain tissues of neurodegenerative diseases [31, 32]. The selected reference genes belong to different molecular pathways to minimize the risk of co-regulation between genes. In addition, *RPS27A* gene, a potential gene of interest in SIDS, was used to test different normalization strategies using different reference genes. For each selected gene, the “best coverage” TaqMan real-time PCR assay (Life Technologies, Darmstadt, Germany) was chosen (i.e., detecting the maximum number of gene transcripts and not detecting off-target sequences). When not possible, assays at least with probes or amplicons spanning an exon junction were then selected to increase specificity and ensure amplification of the target cDNA only. Assays selected had small amplicon sizes of 112 bps or less. Additional detailed information on the assays is given in Table 2.

### Quantitative real-time PCR (qPCR)

PCR reaction mix was prepared using 7.5 µL of 2× TaqMan<sup>®</sup> Gene Expression Master Mix, 0.75 µL of 20× TaqMan<sup>®</sup> Gene Expression Assay Mix (Life Technologies, Darmstadt, Germany), 5.25 µL nuclease-free water, and

**Table 2** Different reference and target genes investigated in this study

Gene symbol	Gene name	Assay ID (ABI) <sup>a</sup>	Amplicon size (bp)	NCBI location chromosome	Molecular function	Context sequence
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Hs02758991_g1	93	Chr.12: 6643657–6647536	Dehydrogenase in glycolysis	GACTCATGACCACAGTCCATGCCAT
GUSB	Glucuronidase, beta	Hs99999908_m1	81	Chr.7: 65425671–65447301	Galactosidase in glycosaminoglycan metabolism	TGAACAGTCA CCGACGAGAGTGCTG
HMBS	Hydroxymethylbilane synthase	Hs00609293_g1	62	Chr.11: 118955587–118964259	Deaminase in porphyrin metabolism	ATGCTGTGGCCAGGGGGCCTTGGG
SDHA	Succinate dehydrogenase complex, subunit A, flavoprotein	Hs00188166_m1	70	Chr.5: 218356–256815	Dehydrogenase in the TCA cycle and respiratory chain electron transport	TGTTGTTGCCACAGGAGGCTACGGG
UBXN6	UBX domain protein 6	Hs00228853_m1	95	Chr.19: 4445003–4457791	Unclassified, involved in protein binding	CCATTCTTTCACACTTCTCCACCGA
RPS27A	Ribosomal protein S27a	Hs01923841_uH	112	Chr.2: 55459039–55462989	Involved in protein degradation and nucleic acid binding	GTTGAGACTTCGTGGTGGTGCTAAG

<sup>a</sup> Purchased from Life Technologies (Darmstadt, Germany)

1.5 µL cDNA in a final 15 µL reaction volume. Life Technologies 7500 Fast Real-Time PCR System ABI with SDS software v 2.0 was used to run standard thermal cycling conditions: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s, and 60 °C for 1 min in a MicroAmp<sup>®</sup> Fast Optical 96-Well Reaction Plate sealed with MicroAmp<sup>®</sup> Optical Adhesive Film (Life Technologies). All the qPCR reactions were performed in duplicate with no-template controls (NTCs) included in each assay run. Due to the large number of samples investigated for each gene, samples of each gene were spread over two plates using a “gene maximization set-up” with the inclusion of three inter-run calibrators (IRCs) in each plate to remove any possible inter-run technical variation [33]. No-reverse transcriptase (no-RT) control reactions were also performed for all assays using RNA that was not reverse transcribed in order to ensure that no genomic DNA was amplified during the PCR reactions.

### Data analysis

The quantification cycle ( $C_q$ ) values as well as the baseline settings of each amplification plate were automatically calculated by the qPCR instrument software (SDS software v. 2.0, Life Technologies). However, manual inspection was performed for both parameters after each run to ensure correct automated assignment by the software. Quality control of the replicates variability was performed and 0.5 cycles were set as a limit of the accepted  $C_q$  difference between replicates. The range of the  $C_q$  values, expressed as the average of the replicates, for each candidate reference gene in both studied groups was represented in the form of boxplot using the SPSS v. 16.0 software (SPSS Inc., Chicago, USA).

The qbase<sup>PLUS</sup> software v. 2.2 (Biogazelle, Zwijnaarde, Belgium) was used for further analysis of the  $C_q$  values. Firstly geNorm module, integrated in the software, was used to determine the most stable reference genes in our set of samples (using the calculated geNorm stability  $M$  values), and the optimal number of reference genes required for reliable normalization (using the pairwise variation). Secondly two normalization strategies were implemented to demonstrate the effects of using inadequately validated reference genes in the interpretation of the results. The first strategy was the use of the geometric mean of the identified most stable reference genes for normalization of the gene expression data of the *RPS27A* target gene, giving the normalized relative quantities (NRQ). The final calibrated normalized relative quantities (CNRQ) were calculated by calibrating the NRQ in both runs using the integrated IRCs. The second normalization strategy was carried out by normalizing the same target gene, *RPS27A*, against a

single unstable reference gene using the  $2^{-\Delta\Delta C_q}$  method of Livak, Schmittgen [1].

Statistical analysis of the relative quantification data was conducted using Mann–Whitney  $U$  test. One-way ANOVA was used to assess the effect of four different parameters on gene expression data of the rather unstable reference genes as well as *RPS27A* target gene after their normalization by the geometric mean of the stable genes. The different parameters were cause of death in controls as well as PMI, age at death, and gender in both study groups. The difference was regarded as statistically significant when the  $p$  value was less than 0.05.

## Results

### RNA yield

Total RNA was successfully extracted from all the samples. The RNA concentrations in the different samples ranged from 48.4 to 263 ng/ $\mu$ L. RNA purity, determined using 260/280 absorbance ratios, ranged from 2.07 to 2.15.

### Expression levels of the candidate genes

RT-qPCR reactions were performed successfully for all candidate genes in all samples with good reproducibility of the replicates; 444 out of 468 sample-target combinations passed the  $C_q$  difference limit of  $<0.5$  cycles between replicates (94.9 %). The expression levels of the candidate reference genes, expressed as the average  $C_q$  values of the replicates, ranged from 20.60 to 32.70. The most abundant transcript was *GAPDH* with the lowest mean  $C_q$  value (SD) of 22.76 (1.36). In contrast, *HMBS* showed the lowest expression with the highest mean  $C_q$  value (SD) of 30.39 (1.40). According to Kolmogorov–Smirnov normality test, the  $C_q$  values of all genes in both SIDS and control cases showed normal distribution except for *GUSB* and *HMBS* in SIDS cases. The range of distribution of the  $C_q$  values of all of the reference genes was larger in SIDS cases than in controls with outliers only observed in SIDS cases especially for *GUSB* and *HMBS* (Fig. 1). This denotes the higher variability in gene expression levels in SIDS even for the so called endogenous reference genes.

Regarding the NTCs included for each gene to exclude cDNA contamination of the qPCR reactions, no amplification was detected in any case. In addition, qPCR reactions using no-RT controls were performed for each gene to evaluate the possibility of genomic amplification. No amplification of non-specific products was detected in all genes, except for some samples of the *GAPDH* gene. This gene is known to have pseudogenes in its genome. The

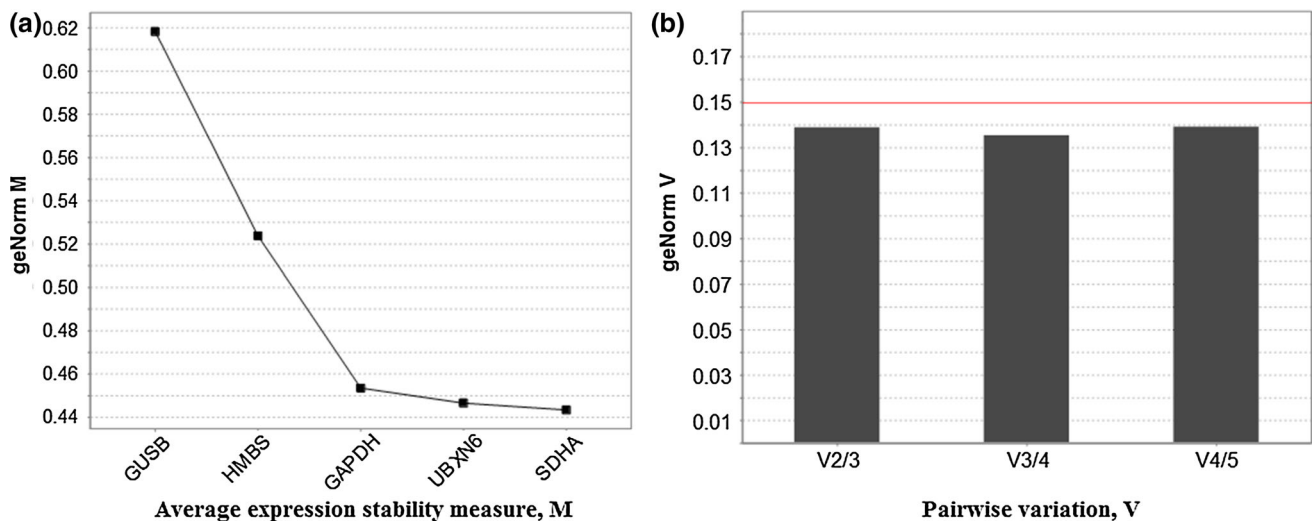
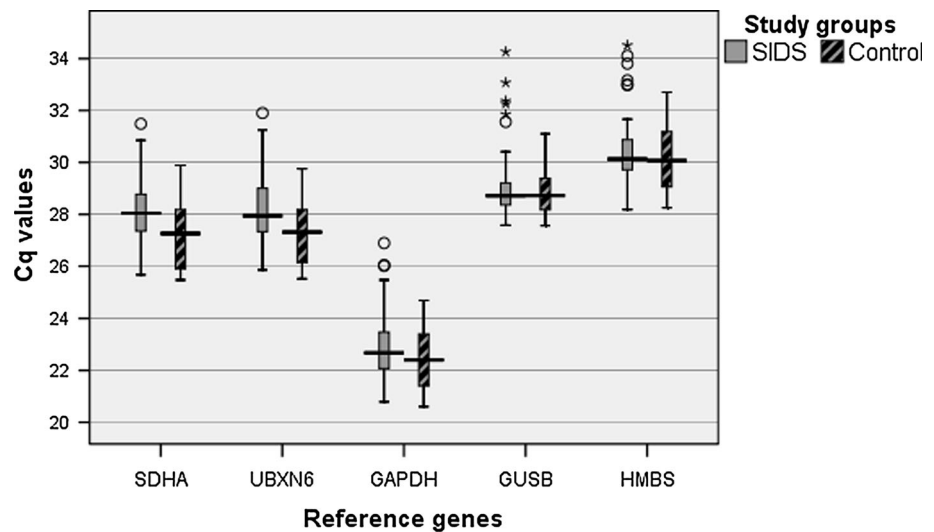
non-specific co-amplification in *GAPDH* was detected even with the inclusion of procedures to minimize the possibility of gDNA contamination, including DNase digestion before the reverse transcription step and choosing exon–exon spanning probes. However, when comparing the  $C_q$  values of the *GAPDH* gene between the RT samples (containing the reversely transcribed cDNA) and the no-RT control samples, we observed that 10.5 cycles were the lowest difference ( $\Delta C_q$ ) between the  $C_q$  values of the positive amplifications in the RT samples (ranging from 20.6 to 24.7) and the  $C_q$  values of the contaminant genomic amplifications in the no-RT controls (ranging from 35 to undetermined). Thus, we can conclude that the amount of gDNA in the no-RT controls was about 1000-fold less than that of cDNA in RT samples, assuming 100 % efficiency using the  $E^{\Delta C_q}$  equation (10.5 cycles; each is a twofold doubling of the initial template amount, where  $E$  is the PCR efficiency). This indicates that 0.1 % (1:1000) of the amplification for the *GAPDH* gene in the RT sample was attributable to the genomic DNA template which is very low.

### Stability of the reference gene expression: geNorm analysis

Based on the raw, non-normalized expression levels of the different candidate reference genes, the geNorm module calculates their stability measure  $M$  which is the average pairwise variation of the expression level of one particular reference gene compared to all the other genes tested. The gene stability increases with lowering of the  $M$  value so that the most stable reference gene is the one with the lowest  $M$  value [3]. In relatively homogeneous sample panels, average  $M$  values  $\leq 0.5$  are typically observed for stably expressed reference genes [33]. In our sample panel, geNorm analysis revealed that *SDHA* (with the lowest  $M$  value) followed by *UBXN6* and *GAPDH* were found to be the most stable reference genes with average  $M$  values below 0.5, whereas *GUSB* and *HMBS* were the least stable genes. The average expression  $M$  values of the five tested reference genes are depicted in Fig. 2a.

Figure 2b shows that the geNorm pairwise variation between the two normalization factors based on the 2 or 3 most stable reference genes in our sample panel ( $V_{2/3}$ ) is below the cut-off of 0.15 recommended by Vandesompele et al. [3]; this means that there is no need to include more than two reference genes for post-mortem brainstem medulla tissue samples of SIDS and control groups, and therefore, the optimal reference normalization factor can be calculated as the geometric mean of the two most stable genes (with the lowest  $M$  value), namely *SDHA* and *UBXN6*.

**Fig. 1** The distribution of the expression levels of the candidate reference genes. For each reference gene, the range of quantification cycles ( $C_q$ ) values, expressed as the average of the replicates, in SIDS cases (gray box,  $n = 50$ ) and controls (shaded box,  $n = 22$ ) is represented by boxplots. The box indicates the interquartile range (25th–75th percentile). The median is depicted by the line across the box. Whiskers represent the maximum and minimum values. Outlier values are indicated by closed circles and extremes by asterisks



**Fig. 2** Stability values of candidate reference genes for RT-qPCR gene expression studies in post-mortem brainstem medulla samples of SIDS and control groups using geNorm analysis. **a** Expression stability plot showing the average expression stability measure  $M$  values of 5 reference genes after stepwise exclusion of the least stable genes. The plotted gene with the highest  $M$  value is the least stable gene (*far left*), whereas the most stable gene is the one with the

lowest  $M$  value (*far right*). **b** Pairwise variation analysis to determine the optimal number of reference genes required for reliable normalization. It indicates that a normalization factor based on the geometric mean of the two most stable reference genes is sufficient for accurate normalization (V2/3, or pairwise variation between the two normalization factors based on the 2 or 3 most stable reference genes, is below the cut-off of 0.15)

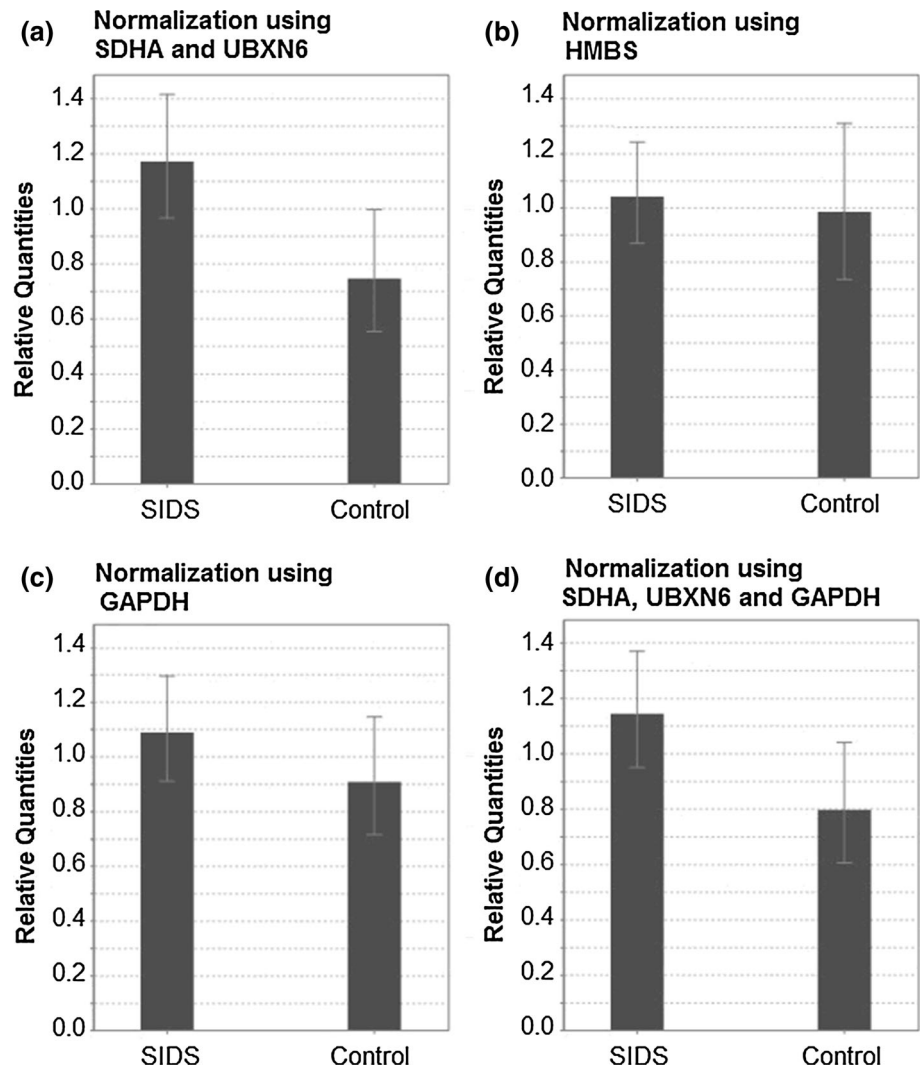
### Effects of different normalization strategies on target gene expression

Different strategies for normalization of the *RPS27A* target gene expression were compared in this study in order to show the effects of selecting inadequately validated reference genes in qPCR experimental results. The first normalization strategy was the use of the geometric mean of the set of the most stable reference genes that was validated by geNorm (*SDHA* and *UBXN6*) in our sample panel. By applying this strategy, the *RPS27A* transcript abundance

was found to be 1.57 times higher in SIDS cases than that in controls with a  $p$  value of 0.016 (Fig. 3a).

The second normalization strategy was the use of a single reference gene for normalization of the same target gene (*RPS27A*). We tested the use of only one reference gene for normalization that was found to be quite unstable in our sample panel, the *HMBS* gene. This revealed a non significant difference in transcript abundance of *RPS27A* between SIDS and the controls (Fig. 3b). Interestingly by using *GAPDH* alone for normalization, which is one of the most commonly used reference genes for

**Fig. 3** The relative expression of *RPS27A* gene in SIDS and control groups using different normalization strategies. Normalization using the geometric mean of the most stable reference genes revealed a statistically significant difference in *RPS27A* gene expression between SIDS ( $n = 50$ ) and control ( $n = 22$ ) groups;  $p$  value  $<0.05$  (a). In contrast, normalization using a single reference gene whether it is unstable (*HMBS*) or apparently stable (*GAPDH*) resulted in non-significant difference in *RPS27A* gene expression (b, c). Significance was observed again using *GAPDH* combined with the other two most stable reference genes (d). Scale was set to average



qPCR data normalization, and which was found to have a low  $M$  value in our study and thus apparently could be regarded to be stable in SIDS cases, the difference in transcript abundance between the study groups was not significant (Fig. 3c). In contrast, when using *GAPDH* in combination with the other two most stable reference genes, *SDHA* and *UBXLN6* (applying the first normalization approach), a significant up-regulation of the *RPS27A* transcript abundance in SIDS cases as compared to the controls by a factor of 1.44 ( $p$  value of 0.03) was observed again (Fig. 3d).

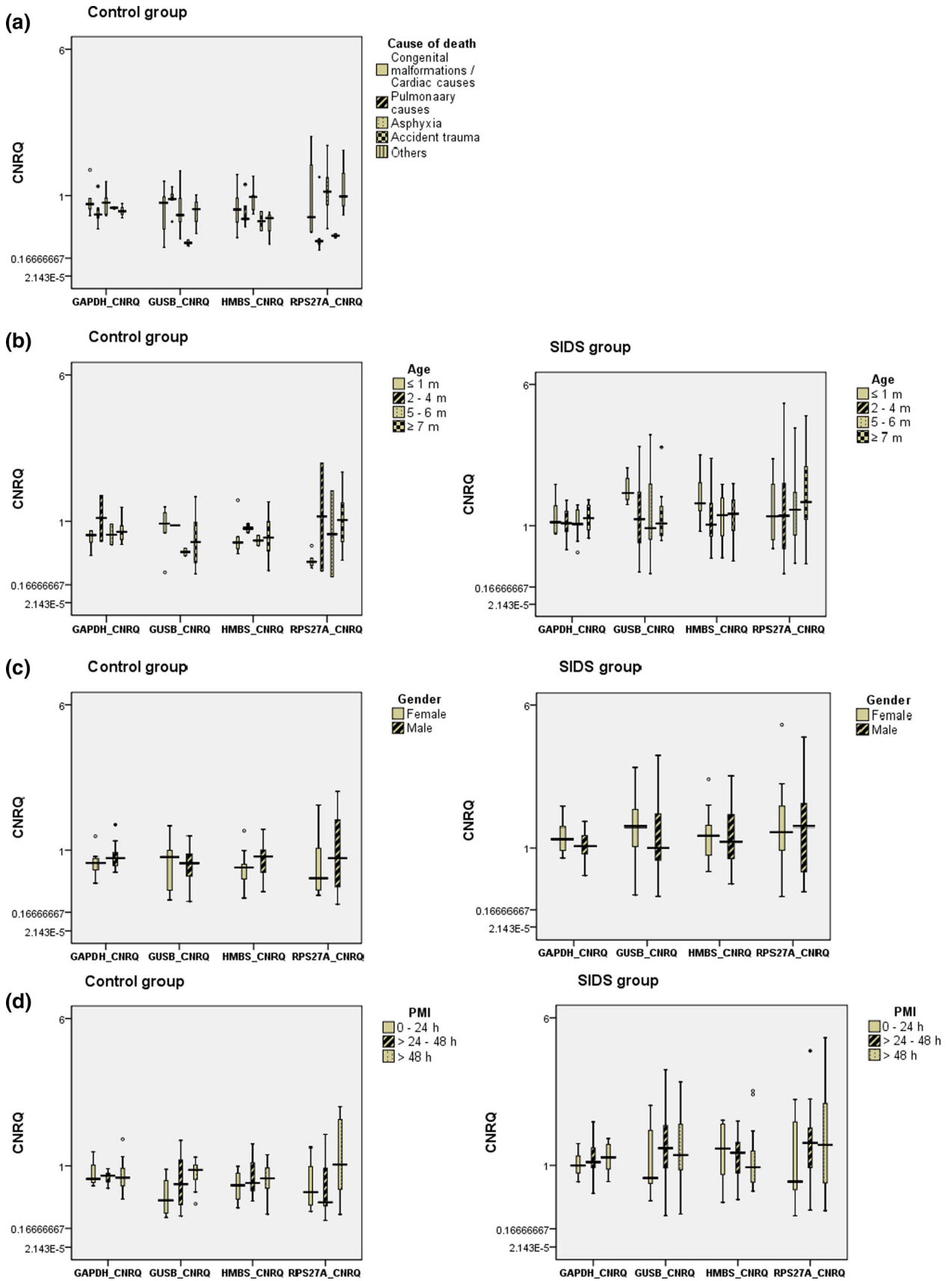
#### Effects of different tissue sample parameters on gene expression

Four different parameters of the tissue samples that might have an influence on gene expression were investigated. They include cause of death in controls, PMI, age at death, and gender. The expression data of the reference genes that

were shown to be relatively unstable in the study sample panel (*GAPDH*, *GUSB*, and *HMBS*) as well as *RPS27A* target gene data were normalized using the geometric mean of the stable genes (*SDHA* and *UBXLN6*). The analyses revealed no significant differences between the different sample groups of each parameter as shown in Fig. 4.

#### Discussion

To our knowledge, this is the first study evaluating the stability of gene expression of multiple reference genes for qPCR data normalization in human post-mortem brain tissue from SIDS cases. From a set of five candidate reference genes, *SDHA* and *UBXLN6* were identified as the most stable reference genes in our sample panel of SIDS cases and controls. We also tested different normalization strategies to analyze the expression levels of the *RPS27A* target gene and demonstrate the importance of using





◀ **Fig. 4** Analysis of gene expression level depending on different tissue sample parameters (cause of death in controls, age at death, gender, and PMI in both study groups). Genes were normalized using the most stable reference genes. No significant differences could be observed

validated reference genes. At first, the geometric mean of the validated *SDHA* and *UBXN6* genes in SIDS and control sample panels was used for normalization of *RPS27A* gene expression data. By using this normalization strategy, a significant up-regulation of *RPS27A* expression in SIDS as compared to the controls was found. This up-regulation in SIDS was expected and was concordant with the results of the previous pathological studies about ubiquitin described in the introduction.

However, when using a single reference gene for normalization of the expression of the same target gene, *RPS27A*, the obtained results were contradictory with the previous approach. Although there is an increasing consensus regarding the fact that normalization using a single reference gene can lead to misleading results [3, 5, 6, 34, 35], this approach is still widely implemented in different RT-qPCR experiments [26, 36], as it is less expensive and data analysis can be easily done using the  $2^{-\Delta\Delta Cq}$  algorithm. The selected reference gene is usually a commonly used reference gene described before in the literature.

Normalization using *HMBS* gene, which was found to be variable and unstably expressed in our sample panel, masked the true variation in the expression of the *RPS27A* target gene between SIDS and control groups. Data suggested in such case a non significant difference in transcript abundance of *RPS27A* between the study groups. Non significant differences in gene expression were also observed between study groups when using the *GAPDH* gene as a reference gene. Although the *M* value of *GAPDH* was found to be  $<0.5$  and thus apparently could be regarded to be stable in SIDS cases, data analysis showed that it could not be used alone for normalization in SIDS cases. This conclusion is in accordance with the findings by Huth et al. [11] which showed the very low expression stability of *GAPDH* under hypoxic conditions. These hypoxic conditions were found to precede SIDS [37–39]. Using *GAPDH* after geometric averaging with *SDHA* and *UBXN6* to calculate the normalization factor returned the difference in transcript abundance among the study groups to be shown as significantly up-regulated in SIDS. These results confirm that any variation in the normalizing reference gene(s) can obscure real differences in the expression of the genes of interest. Therefore, we do not recommend the use of *GAPDH* as a reference gene in cases where hypoxia could have played an important role in the gene expression pattern of brain tissue such as SIDS.

Although four of the reference genes that have been investigated in our study were chosen from those previously validated in human post-mortem brain tissues [5, 30–32], our results have shown that some of these reference genes were not sufficiently stable in human post-mortem brainstem tissue of SIDS cases, e.g., *HMBS* and *GUSB*. Moreover, we showed that some of the reference genes were only stable when used within a group of stable reference genes and not alone, e.g., *GAPDH*. This demonstrates that the selection of suitable reference genes is highly dependent on the experimental set-up, in particular, on the type of biological samples being investigated. In addition, the results corroborate the fact that no single gene expression has been shown to be truly stable so far [3, 4, 34]. Validated reference genes should therefore be used for normalization of gene expression data and results should be interpreted rationally. When this is not possible, at least three reference genes should be used for normalization as recommended by Vandesompele et al. [3].

RNA degradation can occur due to various reasons such as natural fragmentation, inadequate sample handling, suboptimal storage conditions and others, especially when dealing with post-mortem samples [40]. The most common procedure nowadays to assess RNA integrity is the estimation of the RIN (RNA Integrity Number). The RIN algorithm is calculated using electrophoretic RNA measurements through specific platforms that are not always available in all laboratory settings due to the high costs of instrumentation and consumables. On the other hand, traditional methods, such as denaturing agarose gels for RNA, require extremely high amounts of material, are not always possible especially when valuable and rare samples are being studied as is the case of the SIDS samples. Due to the limitations of these methods, RNA integrity could not be estimated in this study.

Several studies have suggested the use of proxy indices of RIN value such as post-mortem brain tissue pH as indicators of the RNA quality [41–43]. Brain pH has been shown to be significantly correlated with RIN values; however there is no specific pH value below which the RNA will be definitely degraded; similarly, a high pH value does not guarantee intact RNA. In addition, ante-mortem metabolic abnormalities may contribute to the low pH encountered in some cases [44].

Nevertheless, it is important to consider the implications of potential RNA degradation in the present work. In a review on RNA degradation and its effect on RT-qPCR performance by Fleige and Pfaffl [45] results demonstrated that the importance of RNA integrity raises considerably with the increasing length of the amplified product, i.e., that the amplicon length of the target transcript is strongly dependent on the degradation level. The authors claim that amplification of long products (over 400 bps) is strongly

dependent on a good RNA quality, and on the other hand, shorter “RT-qPCR products of 70–250 bp are more or less ‘independent’ of the RNA quality” and can still generate reliable qPCR data. It is concluded that “...while all efforts should be made to obtain high-quality RNA samples that reflect the natural state most reliably, moderately degraded samples with a degradation signature may still lead to a reasonable qRT-PCR expression profile” [45]. Similarly, in another study, also on gene expression analysis on degraded RNA by RT-qPCR [46], genes with different primer sets generating different size targets from short (54 bp) to long amplicons (147 bp) were tested in a set of controlled, artificially degraded and non-degraded RNA. In this study, similarly to Fleige and Pfaffl [45], Antonov and co-authors [46] also demonstrated that short amplicons are more robust against RNA fragmentation than longer ones. We support the argument that the impact of degradation on small amplicons is quite low compared to longer fragments and that the crucial point in downstream RT-qPCR regarding RNA integrity is the extent of degradation. We are aware that there is a correlation between RNA fragmentation and gene expression values, however based on clear results from the latter experiments [45, 46], RNA degradation has higher impact on amplification of long size targets and so primer sets for small amplicons should be used. Therefore, in the present study, it was of major importance to select assays for the candidate reference genes following this recommendation. The chosen candidates all had small amplicon sizes, between 62 and 95 bps.

The non-normalized  $C_q$  values obtained for the reference genes in the present work further support the argument of RNA integrity and fragment length. As known, degradation of RNA leads to a loss of intact amplicons and consequently in cDNA, and thus, an increase in  $C_q$  values is observed. This means the increase of  $C_q$  values reflects the loss of amplicons due to RNA fragmentation [46]. If extreme degradation was present in our sample set (high fragmentation would mean in this study amplicons below the transcript size of the longest reference gene <95 bp) a significant effect of degradation would become apparent in our  $C_q$  results. This means a correlation of fragment size and raw  $C_q$  values would be expected: higher  $C_q$  values would be observed in longer fragments and vice versa. Actually, the opposite was detected: *HMBS* showed higher range of  $C_q$  values but is the smallest target (62 bp) and on the other hand, *GAPDH* had the lowest  $C_q$  values and has the longest size amplicon (93 bp) among the reference genes. Most likely we are dealing with true biological variation of transcript abundance instead of degradation effects.

Still regarding the topic on RNA integrity, the power of gene expression normalization in RNA degradation has also been demonstrated [45, 46]. In an experiment by

Antonov et al. [46] four reference genes were selected from a pool of 15 candidates by using the same software, GeNorm, and procedure described by Vandesompele et al. [3], as was done in the present work. The mean of the four reference genes was then used to calculate the delta  $C_q$  value between these and a target gene on several degraded RNA samples. Even after advanced fragmentation, authors observed that the delta  $C_q$  values remained quite consistent and concluded that the used normalization method is useful for an efficient normalization of expression levels for low and highly degraded RNA [46]. Consistently, this argument is also presented by Fleige and Pfaffl [45], where the authors highlight that normalization of RT-qPCR applying the relative quantification method using a stably expressed endogenous control (or controls) are similar in fragmented RNA when compared to high quality samples. Only the non-normalized  $C_q$  data show a correlation with RNA integrity and fragment length. Therefore, when analyzing the expression levels of the target gene, *RPS27A*, potential RNA degradation has a smaller or inconsequential role due to the observations that normalized data and RIN values show no or little correlation [45, 46]. The final conclusions reached for *RPS27A* in this work are a result of normalized data against two stably expressed reference genes (*SDHA* and *HMBS*) as these presented low  $M$  values ( $M < 0.5$ ) across our sample set of controls and SIDS.

Besides the selection of short amplicons for the gene expression assays, we tried to maintain RNA integrity by implementing adequate tissue handling and storage procedures to decrease the unavoidable effects of post-mortem changes on the RNA. The same automated purification procedure was used in extracting all samples as it is known that RNA quality can be different between two extraction methods [45]. Spectrophotometer readings estimated that RNA quality and purity of all samples were acceptable and more importantly they did not differ between the SIDS and the control groups. In the reverse transcription step, as expression differences obtained from degraded samples using random-primed cDNA can correlate with high quality samples [40], a mix of random hexamer primers with oligo (dT) primers was used to insure priming at random sites throughout the RNA strand in addition to the poly (A) tail priming and thus decrease the effect of any possible RNA degradation on cDNA synthesis by capturing all RNA fragments.

Schoor et al. [40] and van der Linden et al. [47] stated that while all efforts should be made to obtain high quality RNA samples that reflect the natural state most reliably, partially degraded samples may still lead to a reasonable expression profile. Therefore, due to all of these aspects, we believe that RNA integrity did not have a high impact in the expression profiling and analysis of the present work. With all the precautions that we have taken, we can argue

that there is a minor influence of RNA integrity on our short RT-qPCR products.

In addition, we investigated several sample parameters that have been previously shown to affect the transcript abundance of some genes. PMI or the interval between the time of death and the time the tissue has been placed in the storage freezer has varying effects on individual gene expression with a general tendency to decreased transcription levels with increasing PMI [48]. However, others have reported that there was no significant effect of PMI on RNA integrity [42, 49, 50]. Age at death and gender are other factors that can affect gene expression [50, 51]. As it is known that SIDS has a higher incidence in the first 6 months of age especially from the second to the fourth month and that there is a higher mortality in male infants (60 %) than in female infants [21, 52], we divided our cohort accordingly. Moreover, we tested the influence of the different causes of death in the control group, although we have done a validation of reference genes to be used for normalization in the given experimental setup of cases. We found no significant differences between the transcript levels of the different sample groups for each parameter, concluding that the stability results were not affected by the demographic characteristics of cases or by the possible tissue RNA degradation resulting from prolonged PMI or agonal events.

## Conclusion

This study supports the need to adequately validate reference genes in a given tissue type and under the determined experimental conditions before their use for normalization of gene expression, even if they are commonly used reference genes. In addition, the demonstrated significant up-regulation of *RPS27A* gene in SIDS cases compared to the controls suggested an independent evidence of the potential role of ubiquitin protein in the pathophysiology of SIDS at the mRNA level established initially by previous immunohistochemical findings. Nevertheless, it should be highlighted that additional studies investigating SIDS at the proteomic level are needed to further clarify this role due to the complexity of gene expression patterns behind SIDS.

## Key points

1. To our knowledge, this is the first study investigating the stability of different reference genes used for data normalization in gene expression studies in human post-mortem brainstem samples of SIDS cases.
2. Among the tested five candidate reference genes, it was shown that *SDHA* and *UBXN6* (a novel reference gene) could be used confidently in such studies.

3. Normalization of the *RPS27A* target gene expression data by the validated *SDHA* and *UBXN6* reference genes revealed its significant up-regulation in SIDS cases compared to the controls.
4. The results of the tested different normalization strategies demonstrated the importance of using validated reference genes for a given experimental set-up, and confirmed the possibility of obscuring real differences in the expression of genes of interest by any variation in the normalizing reference gene(s).

**Acknowledgments** We gratefully thank Ruth Volland for assistance in the SPSS software.

## References

1. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*. 2001;25(4):402–8.
2. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29(9):e45.
3. 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):RESEARCH0034.
4. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun*. 2005;6(4):279–84.
5. Koppelkamm A, Vennemann B, Fracasso T, Lutz-Bonengel S, Schmidt U, Heinrich M. Validation of adequate endogenous reference genes for the normalisation of qPCR gene expression data in human post mortem tissue. *Int J Legal Med*. 2010;124(5):371–80.
6. Hoogewijs D, Houthoofd K, Matthijssens F, Vandesompele J, Vanfleteren JR. Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in *C. elegans*. *BMC Mol Biol*. 2008;9:9.
7. Kozera B, Rapacz M. Reference genes in real-time PCR. *J Appl Genet*. 2013;54(4):391–406.
8. Bemeur C, Ste-Marie L, Desjardins P, Hazell AS, Vachon L, Butterworth R, et al. Decreased beta-actin mRNA expression in hyperglycemic focal cerebral ischemia in the rat. *Neurosci Lett*. 2004;357(3):211–4.
9. Bas A, Forsberg G, Hammarstrom S, Hammarstrom ML. Utility of the housekeeping genes 18S rRNA, beta-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand J Immunol*. 2004;59(6):566–73.
10. 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):112–4, 6, 8–9.
11. Huth A, Vennemann B, Fracasso T, Lutz-Bonengel S, Vennemann M. Apparent versus true gene expression changes of three hypoxia-related genes in autopsy derived tissue and the importance of normalisation. *Int J Legal Med*. 2013;127(2):335–44.
12. Suzuki T, Higgins PJ, Crawford DR. Control selection for RNA quantitation. *Biotechniques*. 2000;29(2):332–7.
13. Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, et al. Housekeeping genes as internal standards: use and limits. *J Biotechnol*. 1999;75(2–3):291–5.

14. Heinrich M, Lutz-Bonengel S, Matt K, Schmidt U. Real-time PCR detection of five different “endogenous control gene” transcripts in forensic autopsy material. *Forensic Sci Int Genet.* 2007;1(2):163–9.
15. Preece P, Virley DJ, Costandi M, Coombes R, Moss SJ, Mudge AW, et al. An optimistic view for quantifying mRNA in post-mortem human brain. *Brain Res Mol Brain Res.* 2003;116(1–2):7–16.
16. Krous HF, Beckwith JB, Byard RW, Rognum TO, Bajanowski T, Corey T, et al. Sudden infant death syndrome and unclassified sudden infant deaths: a definitional and diagnostic approach. *Pediatrics.* 2004;114(1):234–8.
17. Kinney HC, Thach BT. The sudden infant death syndrome. *N Engl J Med.* 2009;361(8):795–805.
18. Carpenter RG, Irgens LM, Blair PS, England PD, Fleming P, Huber J, et al. Sudden unexplained infant death in 20 regions in Europe: case control study. *Lancet.* 2004;363(9404):185–91.
19. Vennemann MM, Findeisen M, Butterfass-Bahloul T, Jorch G, Brinkmann B, Kopcke W, et al. Modifiable risk factors for SIDS in Germany: results of GeSID. *Acta Paediatr.* 2005;94(6):655–60.
20. Hunt CE, Darnall RA, McEntire BL, Hyma BA. Assigning cause for sudden unexpected infant death. *Forensic Sci Med Pathol.* 2015. doi:10.1007/s12024-014-9650-8.
21. Courts C, Madea B. Genetics of the sudden infant death syndrome. *Forensic Sci Int.* 2010;203(1–3):25–33.
22. Klintschar M, Heimbold C. Association between a functional polymorphism in the MAOA gene and sudden infant death syndrome. *Pediatrics.* 2012;129(3):e756–61.
23. Millat G, Kugener B, Chevalier P, Chahine M, Huang H, Malicier D, et al. Contribution of long-QT syndrome genetic variants in sudden infant death syndrome. *Pediatr Cardiol.* 2009;30(4):502–9.
24. Paterson DS, Trachtenberg FL, Thompson EG, Belliveau RA, Beggs AH, Darnall R, et al. Multiple serotonergic brainstem abnormalities in sudden infant death syndrome. *JAMA.* 2006;296(17):2124–32.
25. Opdal SH, Rognum TO. Gene variants predisposing to SIDS: current knowledge. *Forensic Sci Med Pathol.* 2011;7(1):26–36.
26. Casale V, Oneda R, Maturri L, Lavezzi AM. Investigation of 5-HTT expression using quantitative real-time PCR in the human brain in SIDS Italian cases. *Exp Mol Pathol.* 2013;94(1):239–42.
27. Quan L, Zhu BL, Ishida K, Oritani S, Taniguchi M, Fujita MQ, et al. Intranuclear ubiquitin immunoreactivity of the pigmented neurons of the substantia nigra in fatal acute mechanical asphyxiation and drowning. *Int J Legal Med.* 2001;115(1):6–11.
28. Toupalik P, Bouska I. Immunohistochemical findings in the central nervous system in sudden infant death. *Soud Lek.* 1999;44(2):17–20.
29. Sawaguchi T, Patricia F, Kadhim H, Groswasser J, Sottiaux M, Nishida H, et al. The correlation between ubiquitin in the brainstem and sleep apnea in SIDS victims. *Early Hum Dev.* 2003;75(Suppl):S75–86.
30. Wang Q, Ishikawa T, Michiue T, Zhu BL, Guan DW, Maeda H. Stability of endogenous reference genes in postmortem human brains for normalization of quantitative real-time PCR data: comprehensive evaluation using geNorm, NormFinder, and BestKeeper. *Int J Legal Med.* 2012;126(6):943–52.
31. Barrachina M, Castano E, Ferrer I. TaqMan PCR assay in the control of RNA normalization in human post-mortem brain tissue. *Neurochem Int.* 2006;49(3):276–84.
32. Coulson DT, Brockbank S, Quinn JG, Murphy S, Ravid R, Irvine GB, et al. Identification of valid reference genes for the normalization of RT qPCR gene expression data in human brain tissue. *BMC Mol Biol.* 2008;9:46.
33. Hellemans J, Mortier G, De Paep 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):R19.
34. 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.
35. Derveaux S, Vandesompele J, Hellemans J. How to do successful gene expression analysis using real-time PCR. *Methods.* 2010;50(4):227–30.
36. Demendi C, Borzsonyi B, Nagy ZB, Rigo J Jr, Pajor A, Joo JG. Gene expression patterns of insulin-like growth factor 1, 2 (IGF-1, IGF-2) and insulin-like growth factor binding protein 3 (IGFBP-3) in human placenta from preterm deliveries: influence of additional factors. *Eur J Obstet Gynecol Reprod Biol.* 2012;160(1):40–4.
37. Neary MT, Breckenridge RA. Hypoxia at the heart of sudden infant death syndrome? *Pediatr Res.* 2013;74(4):375–9.
38. Rognum TO, Saugstad OD. Hypoxanthine levels in vitreous humor: evidence of hypoxia in most infants who died of sudden infant death syndrome. *Pediatrics.* 1991;87(3):306–10.
39. Jones KL, Krous HF, Nadeau J, Blackbourne B, Zielke HR, Gozal D. Vascular endothelial growth factor in the cerebrospinal fluid of infants who died of sudden infant death syndrome: evidence for antecedent hypoxia. *Pediatrics.* 2003;111(2):358–63.
40. Schoor O, Weinschenk T, Hennenlotter J, Corvin S, Stenzl A, Rammensee HG, et al. Moderate degradation does not preclude microarray analysis of small amounts of RNA. *Biotechniques.* 2003;35(6):1192–6, 8–201.
41. Stan AD, Ghose S, Gao XM, Roberts RC, Lewis-Amezcuca K, Hatanpaa KJ, et al. Human postmortem tissue: what quality markers matter? *Brain Res.* 2006;1123(1):1–11.
42. Durrenberger PF, Fernando S, Kashefi SN, Ferrer I, Hauw JJ, Seilhean D, et al. Effects of antemortem and postmortem variables on human brain mRNA quality: a BrainNet Europe study. *J Neuropathol Exp Neurol.* 2010;69(1):70–81.
43. Gomez-Nicola D, Boche D. Post-mortem analysis of neuroinflammatory changes in human Alzheimer’s disease. *Alzheimers Res Ther.* 2015;7(1):42.
44. Webster MJ. Tissue preparation and banking. *Prog Brain Res.* 2006;158:3–14.
45. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Asp Med.* 2006;27(2–3):126–39.
46. Antonov J, Goldstein DR, Oberli A, Baltzer A, Pirotta M, Fleischmann A, et al. Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. *Lab Invest.* 2005;85(8):1040–50.
47. van der Linden A, Blokker BM, Kap M, Weustink AC, Riegman PH, Oosterhuis JW. Post-mortem tissue biopsies obtained at minimally invasive autopsy: an RNA-quality analysis. *PLoS ONE.* 2014;9(12):e115675.
48. Birdsill AC, Walker DG, Lue L, Sue LI, Beach TG. Postmortem interval effect on RNA and gene expression in human brain tissue. *Cell Tissue Bank.* 2011;12(4):311–8.
49. Tomita H, Vawter MP, Walsh DM, Evans SJ, Choudary PV, Li J, et al. Effect of agonal and postmortem factors on gene expression profile: quality control in microarray analyses of postmortem human brain. *Biol Psychiatry.* 2004;55(4):346–52.
50. Preece P, Cairns NJ. Quantifying mRNA in postmortem human brain: influence of gender, age at death, postmortem interval, brain pH, agonal state and inter-lobe mRNA variance. *Brain Res Mol Brain Res.* 2003;118(1–2):60–71.

51. Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, et al. Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biol Psychiatry*. 2005;57(5):549–58.
52. Moon RY. SIDS and other sleep-related infant deaths: expansion of recommendations for a safe infant sleeping environment. *Pediatrics*. 2011;128(5):1030–9.