



OPEN

Validation of reference genes for quantitative real-time PCR in chemical exposed and at different age's brackish water flea *Diaphanosoma celebensis*

Young-Mi Lee¹✉, Hayoung Cho¹, Ryeo-Ok Kim^{1,4}, Soyeon In¹, Se-Joo Kim² & Eun-Ji Won³✉

Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR), a primary approach for evaluating gene expression, requires an appropriate normalization strategy to confirm relative gene expression levels by comparison, and rule out variations that might occur in analytical procedures. The best option is to use a reference gene whose expression level is stable across various experimental conditions to compare the mRNA levels of a target gene. However, there is limited information on how the reference gene is differentially expressed at different ages (growth) in small invertebrates with notable changes such as molting. In this study, expression profiles of nine candidate reference genes from the brackish water flea, *Diaphanosoma celebensis*, were evaluated under diverse exposure to toxicants and according to growth. As a result, four different algorithms showed similar stabilities of genes for chemical exposures in the case of limited conditions using the same developmental stage (*H2A* was stable, whereas *Act* was fairly unstable in adults), while the results according to age showed a significantly different pattern in suite of candidate reference genes. This affected the results of genes *EcRA* and *GST*, which are involved in development and detoxification mechanisms, respectively. Our finding is the first step towards establishing a standardized real-time qRT-PCR analysis of this environmentally important invertebrate that has potential for aquatic ecotoxicology, particularly in estuarine environments.

Environmental pollutants, such as heavy metals and persistent organic pollutants, have been of great concern because of their persistence in aquatic environments and harmful effects on aquatic organisms^{1,2}. In order to manage pollutants in aquatic environments, the diagnosis of biological responses is considered to be a significant factor in identifying the pollution or for further actions such as setting the guidelines³. Gene expression in the cell is the first step in the biological response to environmental stressors, resulting in biochemical, morphological, and functional changes. Reverse transcription—polymerase chain reaction (RT-PCR) is a highly sensitive and reproducible method for investigating gene expression⁴. In recent decades, RT-PCR has become a leading technique with the advantages of high sensitivity, specificity, and repeatability in several studies conducted in the field, including aquatic environments⁵. However, since variations in the initial RNA amount and efficiency of cDNA synthesis can lead to errors in RT-PCR results, a normalization step using stable and constantly expressed genes is crucial⁶. Additionally, the range of impacts of environmental pollutants depends on several factors, including concentration and duration of exposure to chemicals, as well as the stage of development, especially in species with molting systems⁷.

¹Department of Biotechnology, College of Convergence Engineering, Sangmyung University, Seoul 03016, Republic of Korea. ²Genome Editing Research Center, Korea Research Institute Bioscience and Biotechnology, Daejeon 34141, Republic of Korea. ³Department of Marine Science and Convergent Technology, Hanyang University, Ansan 15588, Republic of Korea. ⁴Present address: Division of Chemical Research, National Institute of Environmental Research, Hwangyeong-ro 42, Seo-gu, Incheon 22689, Korea. ✉email: ymlee70@smu.ac.kr; ejwon@hanyang.ac.kr

In general, the expression value (Ct) of the target gene was normalized to that of the reference genes, called the housekeeping genes, to compensate for differences among RNA samples^{8,9}. Commonly used reference genes include the *18S rRNA*, β -actin, and *GAPDH* genes^{10–12} in crustaceans (mainly large crustaceans). However, few studies have validated reference genes in small crustaceans^{13,14}, although many concerns arise in selection of appropriate genes. For example, *GAPDH*, *EF- α* , and *18S rRNA*, the most popular genes still preferred in aquatic invertebrate for normalization, remain unvalidated¹⁵, and some genes suggested as reference genes can be modulated by different environmental stresses¹⁶, indicating that even housekeeping genes may be involved in pathways other than cellular homeostasis¹⁷. Thus no gene is truly stable in all cells under all experimental conditions. Thus, reliable reference genes should be selected prior to experiments, and it is recommended that at least two reference genes be employed to avoid incorrect results^{18,19}.

In aquatic ecosystems, coastal areas are vulnerable to land pollution due to substances derived from industries, wastewater, farming, and tourism. Many studies have shown that aquatic organisms, whether in situ environments or cultured in the laboratory, can be used to evaluate pollution and to test ecotoxicological responses^{20,21}. In particular, brackish water species have a wide range of adaptations to environmental physico-chemical conditions, such as salinity and temperature, and are considered suitable for monitoring and risk assessment of various pollutants originating from land and marine environments. The brackish water flea *Diaphanosoma celebensis*, that belongs to Crustacea, Branchiopoda, and Cladocera, is one of the species of interest to many scientists as potential non-model species for marine pollution^{22–24}. They are widely distributed in tropical Asia at a wide range of salinity and temperature. In particular, *D. celebensis* has several advantages such as a short life cycle (about 4–5 days from hatching to reproductive maturity, and 2–3 weeks lifespan), filter-feeders that might be exposed to several pollutants both in water or particles, small body length, and easy culture under laboratory conditions. Most importantly, single-breeding makes genetically identified results in genomic information. This is a great advantage that can be used for transcriptomic research with high efficiency. In fact, in recent studies, the genome and transcriptome of *D. celebensis* were published^{25,26}, leading to their potential use at the molecular level in marine ecotoxicology. However, studies on the validation of reference genes are not explored yet in this species.

In this study, we selected nine common housekeeping genes as candidate reference genes from previous studies^{27–30} and compared their stability in expressions to identify appropriate reference genes under conditions such as chemical exposure and age. For this, three representative chemicals, benzo[a]pyrene (B[a]P), bisphenol A (BPA), and mercury (Hg), were selected. In particular, we focused on their growth ages, as this experimental species has a molting process with different sensitivities to growth^{31,32}. Therefore, we investigated the stability of the nine candidate reference genes under different ages (24-h, 4-day, 7-day, and 10-day) and measured the relative expression of target genes based on the conditions of reference genes.

Results

PCR amplification of nine candidate reference genes. Nine common reference genes in *D. celebensis*, alpha-tubulin (*Atb*), β -actin (*Act*), 18S ribosomal RNA (*18S*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), elongation factor 1-beta (*EF-1b*), ubiquitin conjugating enzyme (*UBC*), histone H2A (*H2A*), TATA-box binding protein (*TBP*), and succinate dehydrogenase (*SDH*) were selected as candidate references for the analysis of gene expression under stress conditions made by three representative chemicals (e.g., BPA, B[a]P, Hg), and different ages (24-h, 4-day, 7-day, and 10-day). The information of all nine candidates, such as their gene IDs, gene, size of the PCR products, and Tm values, is shown in Table 1. In general, the sizes of the qPCR products of all the tested genes ranged from 81 to 104 bp. The applicability of the primers designed for the qPCR amplification of these nine candidate genes was verified by RT-PCR and sequencing of the amplified products (Supplementary Information, Text S1). The unimodal melting curves indicate that the primers used in this study have high specificity (Text S1). Furthermore, the efficiency values ranged from 90.8 to 106.7% (Table 1). Therefore, our qRT-PCR results were confirmed to be valid and reliable.

Expression level of nine reference genes. As shown in the results and box and whisker plots (Tables 2, 3, and Fig. S1), the Ct value distribution of nine reference genes under different conditions varied from 8.7 to 28.3. The *18S* gene encoded the most abundant transcripts in *D. celebensis*, reaching the threshold fluorescence peak after approximately 10 cycles with the lowest median Ct value (Ct < 10, Tables 2, 3, and Fig. S1) compared to those of other candidate reference genes during qPCR amplification when the same amounts of total RNA (500 ng) were used as templates for reverse transcription reactions. The least abundant transcripts were *Atb*, which had a Ct value of 21 or higher (21.93 ± 0.38), for chemical treatment. In a test using different ages, *UBC* had the highest Ct mean values (25.80 ± 0.54). However, with regard to variations in Ct values, *Atb* genes showed significant differences between all ages of development (ANOVA, $p < 0.05$).

The stability of candidate reference genes under chemical exposures. The stability of the reference genes showed different results according to the chemicals (Fig. 1, Fig. S2–S5). Based on the M values evaluated by geNorm, all genes and chemicals showed M values of less than 1, an acceptable value considering suitable reference genes for RT-qPCR normalization in all chemical exposure conditions (Fig. 1, Fig. S2). *Act* and *GAPDH* were the best reference genes for samples exposed to B[a]P (Fig. S2A) and BPA (Fig. S2B). Upon exposure to Hg, *UBC* and *TBP* were the most suitable reference genes (Fig. S2C). When considering all samples exposed to three different chemicals, the M values of all genes were stable in the following order: *H2A*, *EF-1b*, *UBC*, *TBP*, *Act*, *18S*, *SDH*, *GAPDH*, and *Atb* (Fig. 1).

NormFinder also showed different results depending on the sample group exposed to the chemical (Fig. S3). *Act*, *Atb*, and *TBP* were the most stable genes for exposure to B[a]P, BPA, and Hg, respectively (Fig. S3A–S3C).

Gene symbol	Gene description	GenBank accession no	Primer sequences (5'-3')	Amplicon size (bp)	T _m (°C)	Efficiency (%)
<i>Atb</i>	<i>Alpha-tubulin</i>	MH636293	GCATGATTTCCA ACACGAC TACCAGTGAACG AAGGCT	97	55.2	93.92
					53.9	
<i>Act</i>	<i>Beta-actin</i>	MH636289	CAAGATTGTCCG TCCTCCTG CATCTGCTGGAA GGTGG	84	60.5	90.79
					59.5	
<i>18 s</i>	<i>18 s ribosomal RNA</i>	AF144210.1	TGGAAGGATTGA CAGATTGA CTTAGTTGGTGG AGCGATTT	81	54.3	101.89
					56.4	
<i>GAPDH</i>	<i>Glyceraldehyde-3-phosphate dehydrogenase</i>	MH636290	AACTGTCGCCGC TGTTGA ATGGAATGTCT TGGGGTCG	86	56.1	96.31
					58.4	
<i>EF-1b</i>	<i>Elongation factor 1-beta</i>	MH636295	CGGCTGTGTCGT TGAAGA GGCAATGTCCAC ACTCTG	94	56.1	106.67
					56.1	
<i>UBC</i>	<i>Ubiquitin Conjugating enzyme</i>	MH636294	CCTTTTGACGGA CCCATAT TAGTCCAACAGC GAGCAATA	104	55.2	103.89
					56.4	
<i>H2A</i>	<i>Histone H2A</i>	MH636296	GAATACCTGGCT GCTGAAG CAATTGGAGATG ACGGGG	90	56.1	101.35
					57.3	
<i>TBP</i>	<i>TATA-box binding protein</i>	MH636292	TTTCCTGGCTTA ATTTACCGTA ATCTCTGTCTG ACTTTGGC	101	54.3	96.37
					56.4	
<i>SDH</i>	<i>Succinate dehydrogenase</i>	MH636291	CATCGAGTCTCA ACAGAAGA GAGCTCAACCTT TCCAGTT	91	56.4	100.16
					55.2	

Table 1. Summary of nine reference genes.

Gene	Ct value statistics						
	Min	Max	Max-Min	Mean	SD	SE	CV
<i>Atb</i>	21.18	22.48	1.30	21.93	0.38	0.06	0.02
<i>Act</i>	16.06	16.52	0.46	16.26	0.09	0.01	0.01
<i>18S</i>	8.70	9.61	0.91	9.19	0.21	0.03	0.02
<i>GAPDH</i>	17.00	18.27	1.27	17.32	0.23	0.03	0.01
<i>EF-1b</i>	17.34	18.39	1.05	17.80	0.20	0.03	0.01
<i>UBC</i>	21.40	22.23	0.83	21.82	0.20	0.03	0.01
<i>H2A</i>	18.26	18.98	0.72	18.49	0.17	0.02	0.01
<i>TBP</i>	20.85	21.65	0.8	21.21	0.20	0.03	0.01
<i>SDH</i>	18.60	19.54	0.94	19.00	0.19	0.03	0.01

Table 2. Cycle threshold (Ct) values of nine putative reference genes in 12 mRNA at different conditions (three chemicals and four concentrations) from a monoculture of *Diaphanosoma celebensis* exposed to chemicals. Min., minimum; Max., maximum; SD, standard deviation; SE, standard error; CV, coefficient of variation.

The results calculated by NormFinder were somewhat similar to those from geNorm in some representative genes in stability order for each chemical and whole samples (Fig. 1B).

For all cases, including each chemical and all samples, *Act* was the most stable reference gene in BestKeeper (Fig. 1C and Fig. S4). This method showed similar ranking results to other methods, but the order was observed to be different for Hg exposure (Fig. S4C). Upon exposure to Hg, other methods (geNorm and NormFinder) showed that *TBP* and *UBC* were suitable, but only the BestKeeper showed *Act* and *H2A* as suitable candidates. This result is also the same in the case where all samples were gathered.

ReffFinder, a web-based analysis tool, integrates three major computational programs, geNorm, NormFinder, and BestKeeper, and the comparative Δ Ct method to comprehensively rank the tested candidate genes. This

Gene	Ct value statistics						
	Min	Max	Max-Min	Mean	SD	SE	CV
<i>Atb</i>	18.85	28.27	9.42	23.82	3.41	0.57	0.14
<i>Act</i>	18.97	21.29	2.32	19.75	0.57	0.10	0.03
<i>18S</i>	9.59	13.43	3.84	11.18	1.14	0.19	0.10
<i>GAPDH</i>	19.41	21.79	2.38	20.31	0.63	0.11	0.03
<i>EF-1b</i>	21.51	24.23	2.72	22.20	0.69	0.12	0.03
<i>UBC</i>	24.99	26.80	1.81	25.80	0.54	0.09	0.02
<i>H2A</i>	22.22	23.86	1.64	22.88	0.45	0.08	0.02
<i>TBP</i>	24.21	26.85	2.64	25.39	0.90	0.15	0.04
<i>SDH</i>	22.00	24.12	2.12	22.90	0.46	0.08	0.02

Table 3. Cycle threshold (Ct) values of nine putative reference genes in 12 mRNA samples from a monoculture of *Diaphanosoma celebensis* at different ages (Min., minimum; Max., maximum; SD, standard deviation; SE, standard error; CV, coefficient of variation).

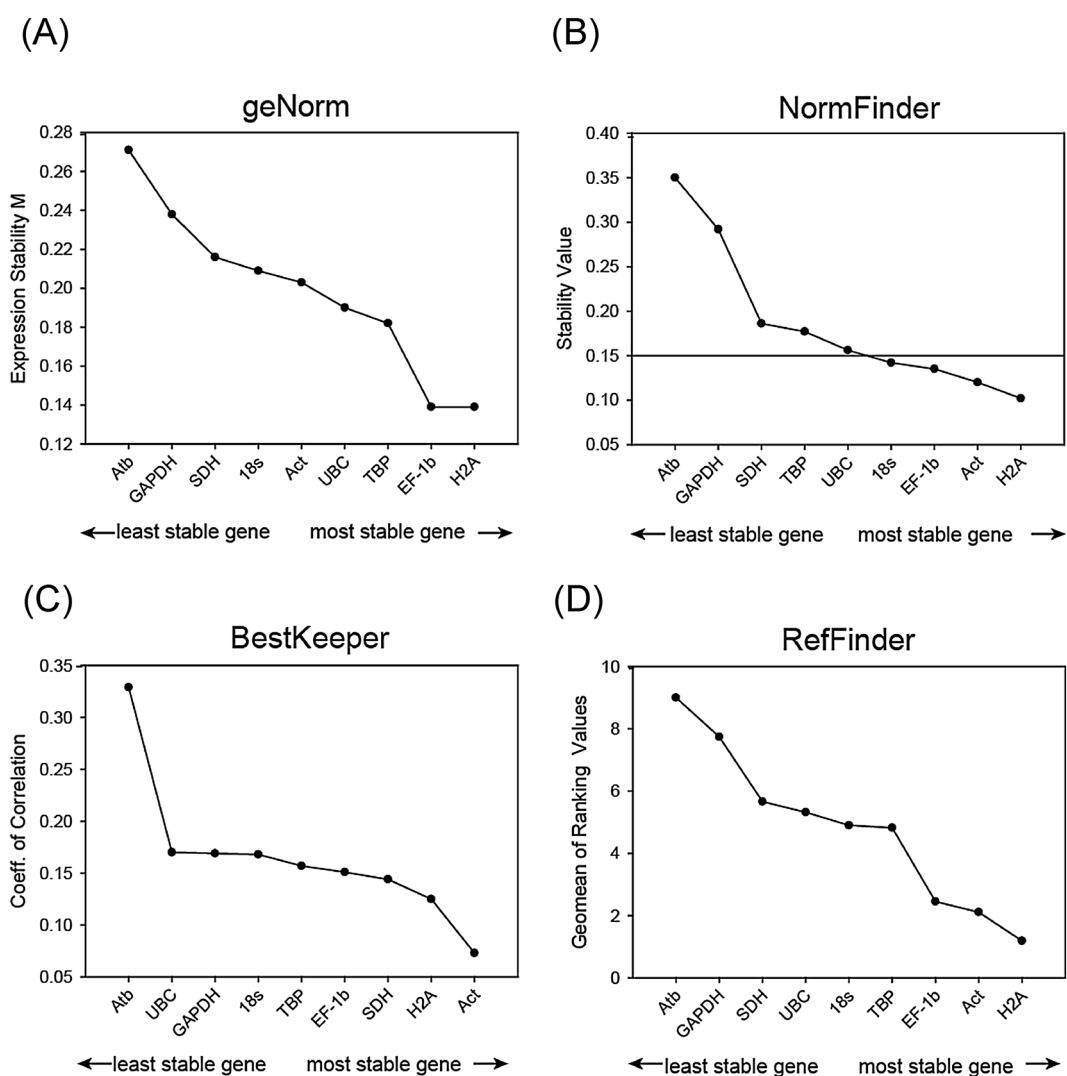


Figure 1. Expression stability of nine candidates for the normalization calculated by different algorithms for three different chemical (B[a]P, BPA, and Hg) exposures. Reference genes are shown in ascending order of expression stability. (A) geNorm (stability M) (B) NormFinder, (C) BestKeeper coefficient of correlation, and (D) RefFinder, a geomean of ranking values integrating other methods.

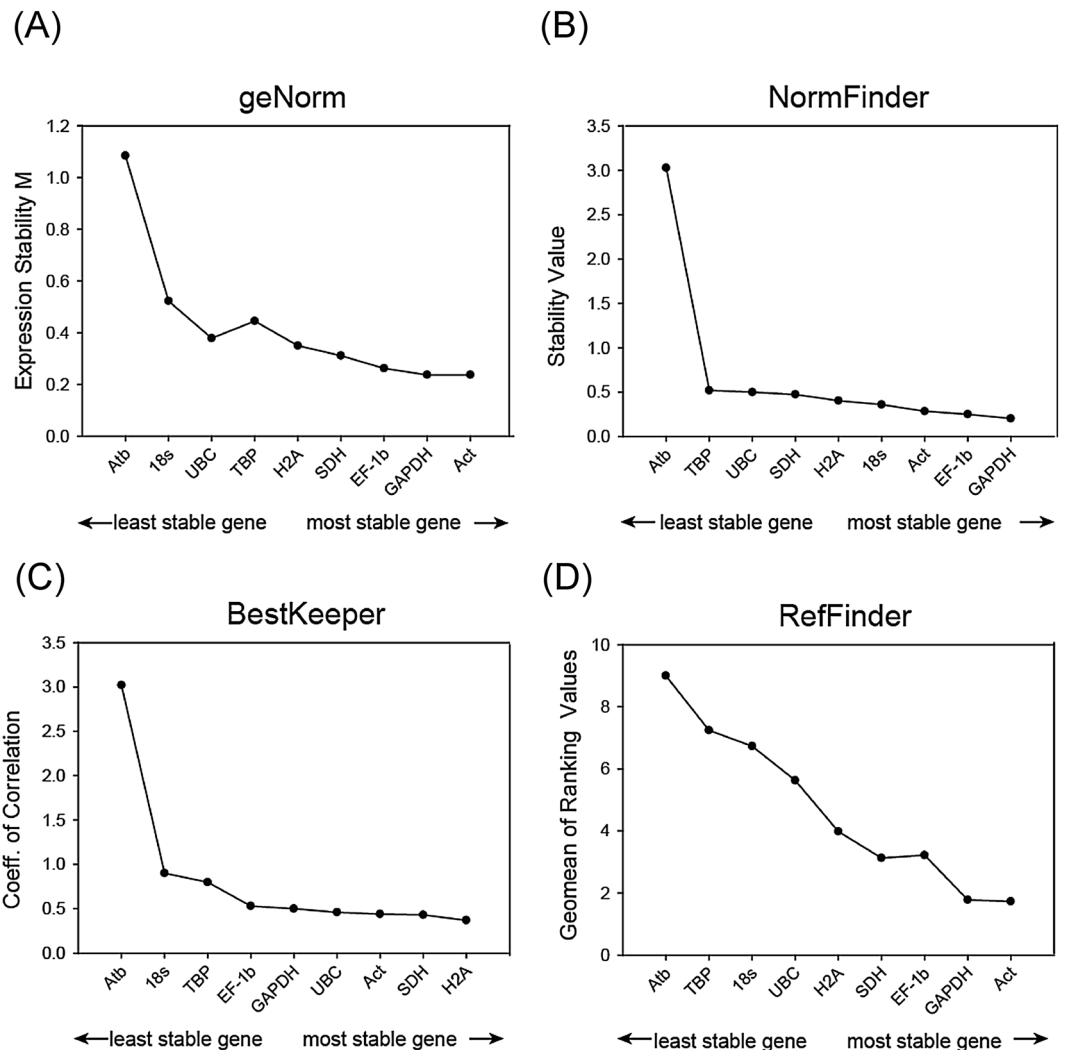


Figure 2. Assessment of nine putative reference genes of *Diaphanosoma celebensis* in different developmental ages (24 h, 4d, 7d, and 10 d) by four different algorithms (A) geNorm, (B) NormFinder, (C) BestKeeper, and (D) RefFinder. Reference genes are shown in ascending order of expression stability.

comprehensive method summarizes *H2A* and *Act* as potential reference genes for the chemical treatment of this species (Fig. 1D and Fig. S5).

Interestingly, in BPA exposure, *Atb*, which showed the most unstable results for the entire sample, was calculated as a stable reference gene as the first (NormFinder) or top ranking in several methods (third stable gene for BestKeeper, second stable gene for RefFinder) only for BPA exposure (Fig. S2-S5). When all three chemicals are considered together, either at the commonly used thresholds (1.5 for M-value for geNorm, SD lower than 1 for BestKeeper) or stability order (RefFinder), *H2A* and *Act* have similarly high stability.

The stability of candidate reference genes under biotic conditions: different ages. Each age group (24 h, 4 d, 7 d and 10 d post hatching) showed different stabilities for the nine reference genes of the brackish water flea *D. celebensis* (Fig. 2, Fig. S6-S9). Interestingly, individuals at post-24 h showed *TBP* to be the most unstable, whereas this gene was relatively stable at other ages of growth (days 4, 7, and 10). Even on day 4, the *TBP* gene was found to be the most stable in the NormFinder method (Fig. S7). All results showed that *Atb* was the most unstable, except at 24 h (Fig. 2 and Fig. S6-S9). In particular, the stability of *Atb* gene expression on day 4 was significantly different from that of the other genes. For example, in geNorm, only *Atb* had an M value greater than 1, and the next unstable gene, *18S*, was 0.5.

NormFinder and BestKeeper also showed that *Atb* was the most unstable gene, and also showed a large difference in stability value with the next unstable genes, *SDH* and *18S*, respectively (Fig. 2 and Fig. S6-S8). NormFinder showed different results from geNorm in the order of stability order: *GAPDH*, *EF-1b*, *Act*, *18S*, *H2A*, *SDH*, *UBC*, *TBP*, and *Atb*, while the three most stable genes were the same. BestKeeper showed that *SDH* is the most stable gene when considering all age samples. The stable results of *18S* shown in geNorm and NormFinder do not apply to the seven-day individuals as coefficients in BestKeeper of *18S* over 1, the threshold of this algorithm.

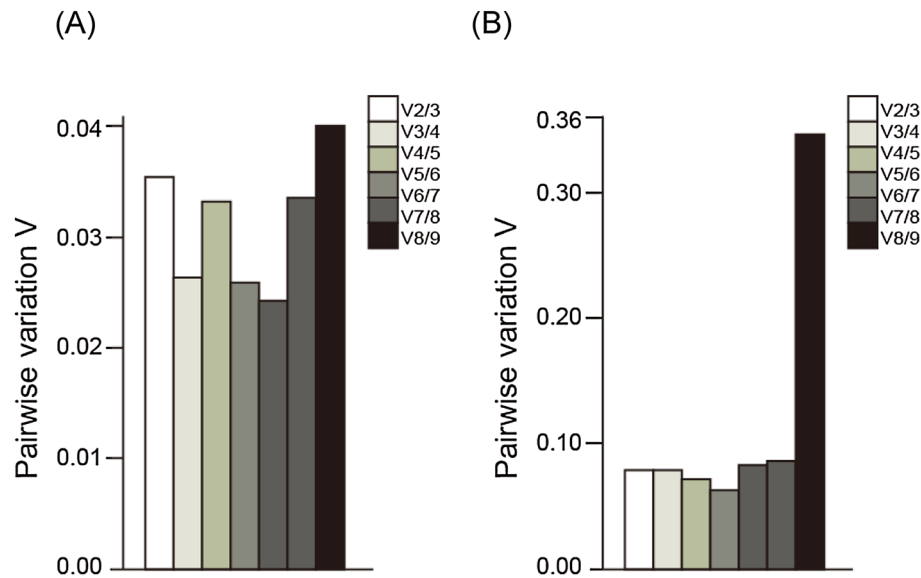


Figure 3. Pairwise variation for evaluating the optimal number of reference genes for accurate normalization (V_n/V_{n+1}) in (A) chemical exposure and (B) ages. The line indicates the cut-off value (0.15), below which the inclusion of an additional reference gene is not required.

Finally, the integrated approach RefFinder indicates that *Act* is the most stable gene for covering all ages of *D. cebebenensis* (Fig. S8).

Impact of reference genes on real-time qRT-PCR data analysis of target genes. To validate the selected reference genes using algorithms, the expression of target genes was normalized and compared at different conditions. First, the pairwise variation (V) suggested that two reference genes are required for the normalization of gene expression levels (Fig. 3). In fact, all V values evaluated by geNorm were less than 0.15, which is the criterion for selecting a suitable reference gene number for the normalization of gene expression in chemical treatments. However, a test conducted using individuals of different ages showed that increasing the number of reference genes did not always increase reliability (Fig. 3B). For differential ages, *Act* and *GAPDH* were selected using RefFinder as stable reference genes.

The expression levels of *GSTmu* were compared between the most stable and unstable genes; however, no difference was found in the levels and patterns regardless of stable (*H2A*) or not (*Atb*) (Fig. 4). However, studies performed using *D. cebebenensis* of different ages showed significantly different expression patterns of *GSTsigma* and *EcRA* depending on which reference gene was used for normalization (Fig. 5 and Fig. S10). In particular, the result using *Atb* as a reference gene showed superficially different expression patterns of target genes (highly expressed *GSTsigma* and *EcRA* at day 4) unlike other cases using stable reference genes (*GAPDH*, *Act*, and both genes).

Discussion

With its rapid development, qRT-PCR has been widely used to evaluate the expression of target genes in recent years³³. In particular, gene expressions in organisms collected in environments sometimes provide integrated information to assess environmental stress conditions without restricting the number of samples, times, and repeatability³⁴. Recently, RNA transcriptomics has become an irreplaceable approach to measure the expression of the target gene in many fields based on biology in the rapid extraction of large-scale biological samples; however, these advantages, such as fast and mass processing, can cause inevitable errors. Thus, the so-called “housekeeping” genes as internal references have been attracting attention for several years and in all major organisms. Briefly, the absolute expression level of the target gene was calibrated based on that of the reference gene. This indicates that reliable validation of the reference gene is a pivotal step in the use of gene expression for further discussion using qRT-PCR. Furthermore, recent studies have reported that various physiological responses are affected by xenobiotic stress, including persistent organic pollutants and heavy metals^{35,36}. The expression of specific genes associated with these physiological alterations shows biological repercussions and the status of the environment. This is the reason why the comprehensive evaluation of reference gene stability in response to diverse biotic and abiotic factors is a rising issue.

However, no reference gene is truly stably expressed for any type of stress, although many studies are being conducted³⁷. Differential expression of the internal gene under different experimental conditions or at different developmental stages causes errors in normalization of relative expression levels. Furthermore, the variability of the reference gene has been easily overlooked as an advantage of gene expression studies using this approach.

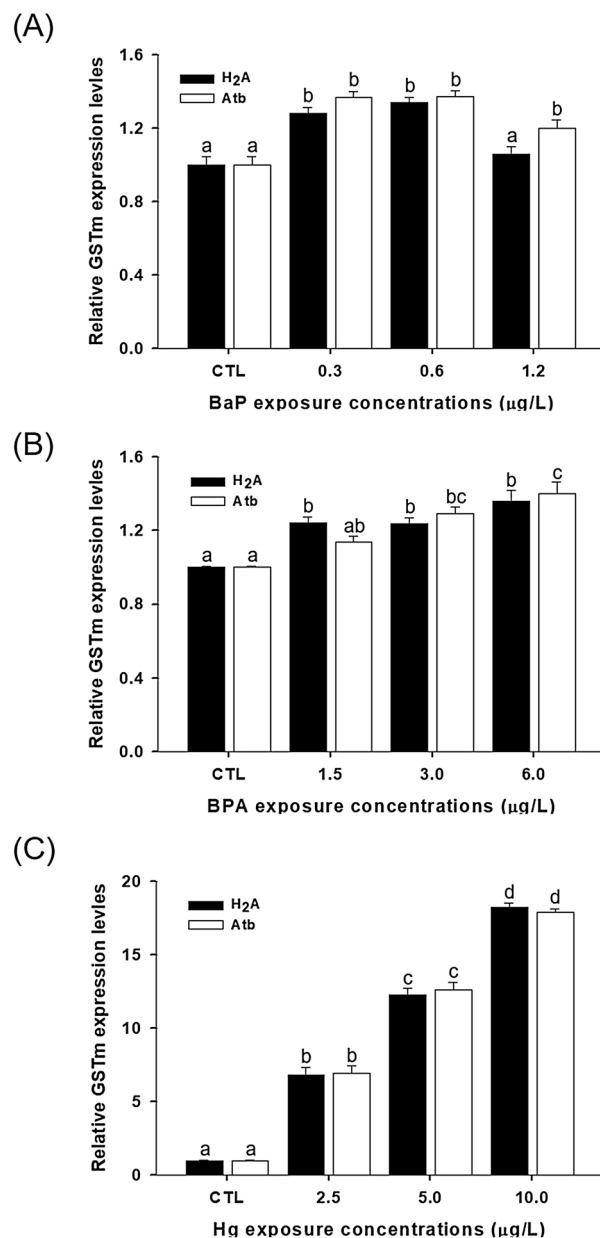


Figure 4. The relative expression level of *GSTmu* was determined using select reference genes, including the most (*H2A*) or least stable (*Atb*) reference genes, for normalization in the brackish water flea *D. celebensis* exposed to (A) B[a]P, (B) BPA, and (C) Hg. Different letters indicate significant differences between groups using different reference genes (ANOVA, Tukey's post-hoc analysis, $p < 0.05$).

This means that the use of a reference gene for calculating the expression of target genes under specific conditions is always required when new species are used in ecotoxicological studies¹⁴.

Reference gene validation has been performed in many other aquatic invertebrates, such as cladocera *Daphnia magna*¹³, amphipod *Gammarus fossarum*¹⁴, and bivalves *Pecten maximus* L.³⁸ and *Ruditapes philippinarum*³⁹. In particular, Volland et al.³⁹ demonstrated that many invalidated uses of genes still introduce artefactual variance, causing misinterpretations of gene expression. These studies insist that preliminary investigations, such as conducting basic research to find reference genes, are important for studies using gene expression as well as for toxicogenomics. Furthermore, the results of studies on reference genes have been helpful in using these species in ecotoxicological studies, followed by the suggestion of specific genes^{40,41}. In the brackish water flea, *D. celebensis* is also one of the potential species that might be used in several studies on pollution in many tropical and coastal areas, either in lab-based work or fields^{26,42,43}. The nine candidate reference genes, *Atb*, *Act*, *18S*, *GAPDH*, *EF-1b*, *UBC*, *H2A*, *TBP*, and *SDH* were used to identify adequate and reliable reference genes for brackish water flea exposed to waterborne chemicals. In our study, however, no reference gene was excluded from analysis because all primer pairs yielded sufficient efficiencies in RT-qPCR (Ct values < 30).

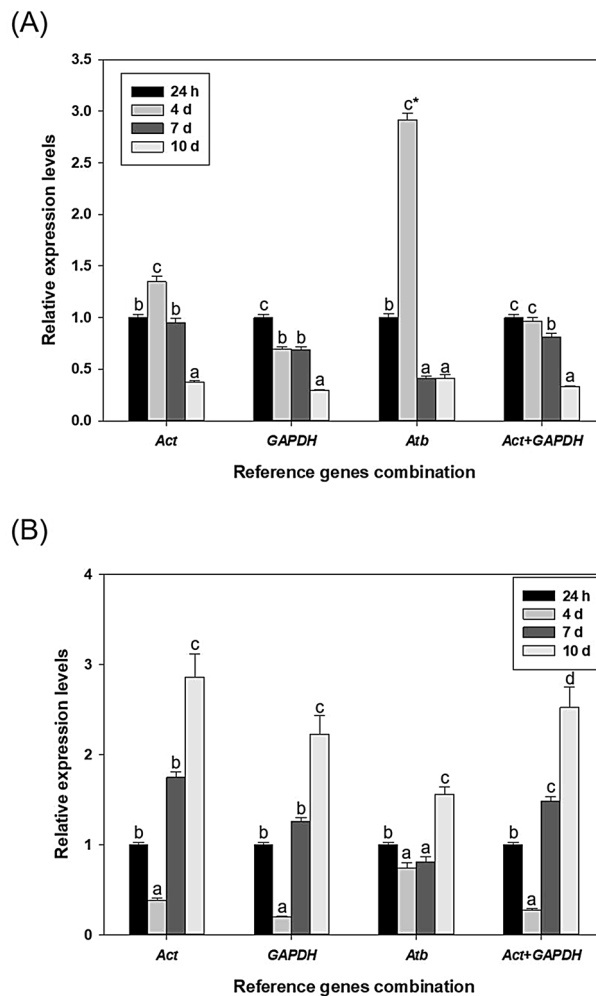


Figure 5. Relative quantification of (A) *GSTsigma* and (B) *EcRA* expression in *D. celebensis* at different ages, 24 h, 4 d, 7 d, and 10 d. The error bars represent the standard errors of the means of three biological replicates. Different letters and asterisks indicate significantly different expressions compared to the 24 h sample (ANOVA, $p < 0.05$) and differences under different normalizations of the reference gene at expression levels for the same age, respectively (t-test, $p < 0.05$).

In the four different approaches for stability testing, each method selected the stability of the potential reference genes with different orders (Table S1). When considering all samples for the three chemical exposures, NormFinder showed different results in the order (*H2A*, *Act*, *EF-1b*, *18S*) of stability compared to those collected by geNorm (*H2A*, *EF-1b*, and *TBP*). Many studies insist that these two methods yield similar results, but the discrepancy between geNorm and NormFinder in this study might be explained by the high expression shown in *18S* (approximately two-fold Ct value smaller than any other gene). The BestKeeper, which indicates the standard deviation (SD) of Ct is considered stable when a value less than 1 leads to similar results from NormFinder for the two most stable genes.

To validate the effect of the reference gene, we compared the target genes under different conditions of the reference gene. *GST*, selected as a target gene, is a phase II detoxifying enzyme that involves the removal of electrophilic substances using reduced glutathione (GSH) as a substrate in most organisms⁴⁴. In particular, it is well established as a biomarker for diagnosis of pollution as *GST* mediates the binding between reactive metabolites generated by phase I enzymes and the thiol group of GSH, thereby converting it to a nontoxic form⁴⁵. Our recent study also showed that the *GST* isoform, particularly *GSTmu* of *D. celebensis* normalized by *18S rRNA*, was similarly upregulated in the presence of heavy metals and BPA exposure^{47,48}. The increased patterns of *GSTmu* observed in the chemical exposures demonstrated that the calculated results based on reference genes in this study are reasonable. Furthermore, the results in which chemical exposures did not elicit comparable patterns in target genes suggested that all reference genes could be candidates when using individuals at least by day 4 of these three chemical exposures (Fig. 4). These results were expected, as all genes used in this batch had M values at geNorm significantly lower than the cut-off, 1.5 (Fig. 1A).

On the other hand, the different ages cause differential expression of reference genes, even though these nine genes are believed to be suitable for reference and showed stable results in other studies, including our first

batch of experiments conducted under chemical exposure^{27–30}. First, the alterations of *GSTsigma* at different ages suggested that age may have different abilities to metabolize, resulting in different gene expressions. In *D. celebensis*, 4-day old release the first brood of neonates²³ and 7-day old are in the molting period (unpublished data). This indicates that different age groups may lead to large alterations in the expression of some reference genes. As expected, the result calculated from the use of an unstable reference (*Atb*) causes a distinct pattern of *GSTsigma*, as if the 4-day individuals were overestimated.

The expression patterns of another selected target gene, ecdysteroid receptor A (*EcRA*), was also similar to *GSTsigma*. *EcRA* is a key marker that is involved in the transport of ecdysteroid, a molting hormone in arthropods, into the nucleus, and thus plays a key role in the modulation of growth, reproduction, and development⁴⁹. Differently modulated expression according to age with the molting process suggested that these genes can be markers of *D. celebensis* to evaluate development²¹. However, the selection of *Atb* resulted in significantly different results in target gene expression. This indicates that the Ct value of *Atb* highly fluctuates according to age. This can be a falsification to report that *Atb* levels remain stable and good candidates as a reference gene for this organism as day 4 is the tipping point of life table for the growth of *D. celebensis*. In the case of both *GAPDH* and *Act*, there was no difference in the expression level in each group, but the expression level of target genes was significantly upregulated in a normalization with *Atb* (t-test, $p < 0.05$). Previous studies have also pointed out that some reference genes can be differently suggested according to the developmental stage of the plant^{50,51}. Thus, the results obtained from *D. celebensis* show evidence of the importance of reference gene selection to avoid misinterpretation in specific organisms that have distinctive growth and reproductive periods in life tables.

In addition, the *18S rRNA* gene was classified as stable by the four different algorithms despite showing a significantly lower Ct value. This is interesting, as *18S rRNA* has been considered to be one of most popular reference genes, as it is expressed in most cell types, but several identified problems suggest that *18S rRNA* may lead to unexpected analytical errors in the cDNA synthesis step, and the ribosomal fraction may not actually represent the whole cellular mRNA population⁵². These results suggest that additional research and validation under various environmental and physiological conditions are required to validate the use of *18S rRNA*.

Conclusion

In our study, nine candidate reference genes were selected for real-time qRT-PCR standardization and evaluated using four different algorithms. The different analytical methods result in comparable values and ranks of the candidates, but the integrated application of these results leads to more reliable information on which gene must be sorted out. The results showed that *H2A* and *Act* were the most stable reference genes for considering 4th day individuals of brackish water flea for chemical exposure, while *Atb* was the most unstable reference gene.

The results for the unstable gene were the same in terms of age. However, the selection of the reference gene results in a significantly different pattern in the target gene when prominent development or alterations, such as molting or reproduction, are observed. This is of great importance in marine environments, as most hypotrophic organisms have a short life cycle, rapid sexual maturation, and prominent features such as idiosyncratic growth, such as molting.

The stable reference genes identified in this study will enhance the accuracy of qRT-PCR-based analysis of target gene expression and can be used to study related ecotoxicological applications using this species in estuarine environments. In this study, normalization with the most stable or with a combination of two carefully chosen reference genes led to different patterns in the expressions that were calculated using unsuitable reference genes. Thus, careful evaluation of the reference genes prior to expression studies is essential. For future directions in environmental studies, more studies are required as the selection of appropriate reference genes from field conditions might be more complex than controlled environmental studies. However, the results accumulated in many studies will be useful for extracting candidates for future studies.

Materials and methods

Reagents. All chemical reagents used in this study are listed in Table S3. Briefly, all reagents for molecular studies and chemical exposure were of ultrapure grade for molecular biology and analytical grade (HPLC grade). Oligonucleotide synthesis and DNA sequencing analyses were performed by Bionics Co., Ltd. (Seoul, South Korea).

Culture and maintenance. The cladoceran *D. celebensis* strain was maintained at the Molecular Toxicology Laboratory of Sangmyung University, South Korea, since they were obtained from Dr. Kyun-Woo Lee of the Korea Institute of Ocean Science & Technology (KIOST; Busan, South Korea) in 2015. The culture medium used was 0.2 µm-filtered artificial seawater (15 psu) using Instant Ocean (Aquarium system, France). Culture conditions were maintained at a photoperiod of 12 h:12 h light/dark and 25 ± 1 °C. The food source was *Chlorella vulgaris* (4 × 10⁷ cells/L) cultured in Jaworski's medium and supplied once every 2 days.

Waterborne exposure to chemicals. Three chemicals (BPA, B[a]P, and Hg), which have been extensively studied in aquatic environments, were selected as representative toxicants. Stock solutions of BPA (2,2-bis-(4-hydroxyphenyl) propane; 3 mg/ml), B[a]P (5 mg/ml) were prepared by dissolving in dimethyl sulfoxide (DMSO), and those of Hg (HgCl₂) were dissolved in distilled water at a concentration of 1 mg/ml. Working solutions were prepared by adding stock solution to 200 ml of 15 psu sea water. To examine the concentration-dependent effects of chemicals, *D. celebensis* (4 days old, ~200 individuals/each concentration) was exposed to BPA (1.5, 3, and 6 mg/L), B[a]P (0.3, 0.6 and 1.2 mg/L), and Hg (2.5, 5, and 10 µg/L) for 48 h in a 250 ml beaker. The concentrations of each chemical were designed based on an acute toxicity test that was conducted as a preliminary study (Table S2). In addition, to check the developmental stage-dependent effects, 24 h, 4 d, 7 d, and

10 d old *D. celebensis* (~200 individuals/each concentration) were used. The final DMSO concentration was less than 0.05%, and no mortality was observed. All examinations were performed in triplicates.

Total RNA extraction and cDNA synthesis. *D. celebensis* was collected 48 h after exposure to chemicals or at different developmental stages and homogenized in five volumes of TRIzol reagent (Thermo Fisher Scientific Inc., USA). Total RNA was extracted according to the manufacturer's instructions and kept at -80 °C until use. The quality and quantity of total RNA were checked using agarose gel electrophoresis and a Nano drop (Maestrogen Nano Drop, Taiwan). Only RNA samples with a 260/280 nm ratio of 1.8–2.0 were used. cDNA was synthesized from 500 ng of total RNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific Inc., USA).

Polymerase chain reaction (PCR) amplification and sequence analysis. The cDNA sequences of nine reference genes (*Atb*, *Act*, *18S*, *GAPDH*, *EF-1b*, *UBC*, *H2A*, *TBP*, and *SDH*) were obtained from the local database of *D. celebensis* transcriptome (Molecular Toxicology Laboratory, Sangmyung University). The PCR protocol was as follows: 95 °C for 2 min; 35 cycles of 95 °C for 30 s, 55 °C for 30 s, 72 °C for 2 min; and an extension step at 72 °C for 10 min. The PCR product was visualized on a 1.4% agarose gel and purified using the AccuPrep® Gel Purification Kit (Bioneer, South Korea) for sequencing. Basic Local Alignment Search Tool (BLAST) homology searches of databases were performed using Blast + 2.8.1 from the NCBI to identify the gene. The sequences of each gene were deposited in GenBank (Table 1).

Quantitative real-time RT-PCR. To validate the reference genes, real-time qRT-PCR was performed using a CFX96™ real-time PCR system (Bio-Rad, USA). Each PCR reaction involved 2 µl cDNA and 2 µl of 10 pmol primer set (Table 1). The PCR cycle conditions were as follows: 95 °C for 10 min, followed by 33 cycles of 95 °C for 15 s, 60 °C for 1 min. To check a specific product after PCR, melting curves were analyzed under the following conditions: 95 °C for 15 s and 60 °C for 1 min with a 0.5 °C increase per second (Text S1). All experiments used SYBR master mix (KAPA Bioassay System, USA) and performed in triplicate. The threshold cycle (Ct) from each experiment was used to compare the stability of the reference genes. All examinations were performed in triplicates.

Expression stability analysis of candidate reference genes. The Ct values obtained from qRT-PCR were analyzed using geNorm (Vandesompele et al. 2002), NormFinder⁵⁵, BestKeeper⁵⁶, and RefFinder (<https://www.heartcure.com.au/reffinder/>) algorithms. For geNorm, the expression stability of nine reference genes was evaluated using statistical analysis and the average expression stability (M value) using geNorm analysis⁵⁴ in qbase PLUS 3.2 (Biogazelle, Ghent, Belgium). Similar to geNorm, NormFinder also calculated the average expression stability values based on the $2^{-\Delta Ct}$ value, and low values were stable. BestKeeper, a method using the Ct value directly without any conversion step, showed stability based on the coefficient of variance (CV) and standard deviation (SD), and RefFinder integrating all the results from the three software mentioned above, was used to calculate the ranking index for showing stability in this study. In addition, the optimal number of reference genes required for reliable normalization was also evaluated by pairwise variation (V value) through geNorm analysis⁵⁴.

Normalization with reference genes. From the stability analysis, the less stable and the most stable reference genes were used to normalize the expression of one target gene (glutathione S-transferase mu, *gst-mu*) in 4-day old *D. celebensis* exposed to chemicals for 48 h. On the other hand, one less stable and two more stable reference genes were used to normalize the expression of two target genes (ecdysone receptor A, *EcRA*; glutathione S-transferase, *GST-sigma*, and *GST-mu*) in *D. celebensis* of four different ages. These genes were selected as target genes because they showed the most significant changes under chemical exposure conditions in our previous studies^{44,47–49}. The fold change values were calculated using the $2^{\Delta\Delta Ct}$ method^{57,58} using formulas (1) and (2):

$$\text{Average CT of reference genes (RG)} : CT_{RG} = (CT_{RG1} + CT_{RG2} + \dots + CT_{RGn})/n \quad (1)$$

$$\text{Gene expression level} : G = 2^{-(CT_{\text{gene}} - CT_{RG})} \quad (2)$$

Statistical analysis. Data from all experiments are represented as the mean ± standard deviation (SD) of the three replicates. One-way analysis of variance (one-way ANOVA) and Student's t-test followed by Tukey's test were used for statistical analysis among samples using the PASW statistics 18.0 program (SPSS Inc., Chicago, IL, USA). Statistical significance was set at $p < 0.05$.

Received: 26 August 2021; Accepted: 29 November 2021

Published online: 08 December 2021

References

1. Chouvelon, T. *et al.* Chemical contaminants (trace metals, persistent organic pollutants) in albacore tuna from western Indian and south-eastern Atlantic Oceans: Trophic influence and potential as tracers of populations. *Sci. Total Environ.* **596–597**, 481–495 (2017).
2. Ali, H., Khan, E. & Ilahi, I. Environmental chemistry and ecotoxicology of hazardous heavy metals: Environmental persistence, toxicity, and bioaccumulation. *J. Chem.* Article ID 6730305. (2019).
3. Leppänen, J. J. An overview of Cladoceran studies conducted in mine water impacted lakes. *Int. Aquat. Res.* **10**, 207–221 (2018).
4. Nolan, T., Hands, R. E. & Bustin, S. A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–1582. <https://doi.org/10.1038/nprot.2006.236> (2006).
5. Valones, M. A. *et al.* Principles and applications of polymerase chain reaction in medical diagnostic fields: A review. *Braz. J. Microbiol.* **40**, 1–11 (2009).
6. Dheda, K. *et al.* Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* **37**, 112–119 (2004).
7. Chendler, G. T. & Green, A. S. Developmental stage-specific life-cycle bioassay for assessment of sediment-associated toxicant effects on benthic copepod production. *Environ. Toxicol. Chem.* **20**, 171–180 (2001).
8. Martins, P., Mafra, V. & de Souza, W. Selection of reliable reference genes for RT-qPCR analysis during developmental stages and abiotic stress in *Setaria viridis*. *Sci. Rep.* **6**, 28348 (2016).
9. Luo, X. *et al.* Stable reference gene selection for RT-qPCR analysis in *Synechococcus elongatus* PCC 7942 under abiotic stresses. *Biomed. Res. Int.* **3**, 1–15 (2019).
10. Cottin, D. *et al.* Comparison of heat-shock responses between the hydrothermal vent shrimp *Rimicaris exoculata* and the related coastal shrimp *Palaemonetes varians*. *J. Exp. Mar. Biol. Ecol.* **393**, 9–16 (2010).
11. Leelatanawit, R., Klanchui, A., Uawisetwathana, U., & Karoonuthaisiri, N. Validation of reference genes for real-time PCR of reproductive system in the black tiger shrimp. *PLoS ONE*, **7**(12), e52677 (2012).
12. Jaramillo, M. L. *et al.* Identification and evaluation of reference genes for expression studies by RT-qPCR during embryonic development of the emerging model organism *Macrobrachium olfersii*. *Gene* **598**, 97–106 (2017).
13. Heckmann, L. H. *et al.* Expression of target and reference genes in *Daphnia magna* exposed to ibuprofen. *BMC Genom.* **7**, 175 (2006).
14. Mehennaoui, K. *et al.* Identification of reference genes for RT-qPCR data normalization in *Gammarus fossarum* (Crustacea Amphipoda). *Sci. Rep.* **8**, 15225 (2018).
15. Tarrant, A. M., Nilsson, B. & Hansen, B. W. Molecular physiology of copepods—From biomarkers to transcriptomes and back again. *Comp. Biochem. Physiol. D* **30**, 230–247 (2019).
16. Suzuki, T., Higgins, P. J. & Crawford, D. R. Control selection for RNA quantitation. *Biotechniques* **29**, 332–337 (2000).
17. Köhler, M., Leitsch, D., Müller, N. & Walochnik, J. Validation of reference genes for the normalization of RT-qPCR gene expression in *Acanthamoeba* spp.. *Sci. Rep.* **10**, 10362 (2020).
18. Bustin, S. A. *et al.* The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622 (2009).
19. Nicot, N., Hausman, J.-F., Hoffmann, L. & Evers, D. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J. Exp. Bot.* **56**, 2907–2914 (2005).
20. Sarma, S. S. & Nandini, S. Review of recent ecotoxicological studies on cladocerans. *J. Environ. Sci. Health B* **41**, 1417–1430 (2006).
21. Kwok, W. H. K. P., Souissi, S., Dur, G., Won, E.-J. & Lee, J.-S. Copepods as Reference Species in Estuarine and Marine Waters. In *Aquatic Ecotoxicology: Advancing Tools for Dealing with Emerging Risks* (pp. 281–308). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-800949-9.00012-7> (2015).
22. Marcial, H. S. & Hagiwara, A. Multigenerational effects of 17 β -estradiol and nonylphenol on euryhaline cladoceran *Diaphanosoma celebensis*. *Fish. Sci.* **73**, 324–330 (2007).
23. In, S. *et al.* Acute toxicity of bisphenol A and its structural analogues and transcriptional modulation of the ecdysone-mediated pathway in the brackish water flea *Diaphanosoma celebensis*. *Ecotoxol. Environ. Safe* **179**, 310–317 (2019).
24. In, S., Cho, H., Lee, K.-W., Won, E.-J. & Lee, Y.-M. Cloning and molecular characterization of estrogen-related receptor (ERR) and vitellogenin genes in the brackish water flea *Diaphanosoma celebensis* exposed to bisphenol A and its structural analogues, *Mar. Pollut. Bull.* **154**, 111063 (2020).
25. Kim, B.-M. *et al.* De novo transcriptome assembly of brackish water flea *Diaphanosoma celebensis* based on short-term cadmium and benzo[a]pyrene exposure experiments. *Hereditas* **155**, Article number 36 (2018).
26. Kim, D.-H. *et al.* The genome of the marine water flea *Diaphanosoma celebensis*: Identification of phase I, II, and III detoxification genes and potential applications in marine molecular ecotoxicology. *Comp. Biochem. Physiol. D* **37**, 100787 (2021).
27. Barber, R. D., Harmer, D. W., Coleman, R. A. & Clark, B. J. GAPDH as a housekeeping gene: Analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol. Genom.* **21**, 389–395 (2015).
28. Kuchipudi, S. V. *et al.* 18S rRNA is a reliable normalisation gene for real time PCR based on influenza virus infected cells. *Virol. J.* **9**, 230 (2012).
29. McCulloch, R. S., Ashwell, M. S. & O’Nan, A. T. Identification of stable normalization genes for quantitative real-time PCR in porcine articular cartilage. *J. Anim. Sci. Biotechnol.* **3**, 36 (2012).
30. Zhang, G. *et al.* Characterization of reference genes for quantitative real-time PCR analysis in various tissues of *Anoectochilus roxburghii*. *Mol. Biol. Rep.* **39**, 5905–5912 (2012).
31. Soetaert, A. *et al.* *Daphnia magna* and ecotoxicogenomics: Gene expression profiles of the anti-ecdysteroidal fungicide fenarimol using energy-, molting- and life stage-related cDNA libraries. *Chemosphere* **67**(1), 60–71 (2007).
32. Hwang, D.-S. *et al.* BDE-47 causes developmental retardation with down-regulated expression profiles of ecdysteroid signaling pathway-involved nuclear receptor (NR) genes in the copepod *Tigriopus japonicus*. *Aquat. Toxicol.* **177**, 285–294 (2016).
33. Deepak, S. A. *et al.* Real-time PCR: Revolutionizing detection and expression analysis of genes. *Curr. Genom.* **8**, 234–251 (2007).
34. VanGuilder, H. D., Vrana, K. E. & Freeman, W. M. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques* **44**, 619–626 (2008).
35. Canesi, L. & Fabbri, E. Environmental effects of BPA: Focus on aquatic species. *Dose Response* **13**, 1–14 (2015).
36. Hudspith, M., Reichelt-Brushett, A. & Harrison, P. L. Factors affecting the toxicity of trace metals to fertilization success in broadcast spawning marine invertebrates: A review. *Aquat. Toxicol.* **184**, 1–13 (2017).
37. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. R. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* **139**(1), 5–17 (2005).
38. Mauriz, O., Maneiro, V., Pérez-Parallé, M. L., Sanchez, J. L. & Pazos, A. J. Selection of reference genes for quantitative RT-PCR studies on the gonad of the bivalve mollusc *Pecten maximus* L. *Aquaculture* **370–371**, 158–165 (2012).
39. Volland, M., Blasco, J. & Hampel, M. Validation of reference genes for RT-qPCR in marine bivalve ecotoxicology: Systematic review and case study using copper treated primary *Ruditapes philippinarum* hemocytes. *Aquat. Toxicol.* **185**, 86–94 (2018).

40. Solberg, M.F., Kvamme, B.O., Nilsen, F. & Glover, K.A. Effects of environmental stress on mRNA expression levels of seven genes related to oxidative stress and growth in Atlantic salmon *Salmo salar* L. of farmed, hybrid and wild origin. *BMC Res. Notes*. **5**, 672 (2012).
41. Ventoso, P. *et al.* Transcriptional response in the digestive gland of the King Scallop (*Pecten maximus*) after the injection of domoic acid. *Toxins* **13**, 339 (2021).
42. Isabel, L. *et al.* In situ assays with tropical cladocerans to evaluate edge-of-field pesticide runoff toxicity. *Chemosphere* **67**, 2250–2256 (2007).
43. Yoo, J. *et al.* Effects of polystyrene in the brackish water flea *Diaphanosoma celebensis*: Size-dependent acute toxicity, ingestion, egestion, and antioxidant response. *Aquat. Toxicol.* **235**, 105821 (2021).
44. Sheehan, D., Meade, G., Foley, V. M. & Dowd, C. A. Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* **360**, 1–16 (2001).
45. Regoli, F. & Giuliani, M. E. Oxidative pathways of chemical toxicity and oxidative stress biomarkers in marine organisms. *Mar. Environ. Res.* **93**, 106–117 (2014).
46. Bae, C. & Lee, Y.-M. Antioxidant responses in brackish water flea *Diaphanosoma celebensis*-exposed to mercury. *J. Mar. Life Sci.* **3(2)**, 74–80 (Korean) (2018).
47. Yoo, J., Cho, H. & Lee, Y.-M. Modulation of glutathione S-transferase and superoxide dismutase in response to heavy metals in brackish water flea *Diaphanosoma celebensis*. *J. Toxicol. Environ. Biol.* **12**, 83–90 (2020).
48. Yoo, J., Cha, J., Kim, H., Pyo, J. & Lee, Y.-M. Modulation of antioxidant defense system in the brackish water flea *Diaphanosoma celebensis* exposed to bisphenol A. *Korean J. Environ. Biol.* **37**, 72–81 (2019).
49. Nakagawa, Y. & Henrich, V. C. Arthropod nuclear receptors and their role in molting. *FEBS J.* **276**, 6128–6157 (2009).
50. Wan, Q. *et al.* Stability evaluation of reference genes for gene expression analysis by RT-qPCR in soybean under different conditions. *PLoS ONE* **12(12)**, e0189405 (2017).
51. Wang, Y., Zhang, Y. & Liu, Q. Selection and validation of appropriate reference genes for RT-qPCR analysis of flowering stages and different genotypes of *Iris germanica* L. *Sci. Rep.* **11**, 9901 (2021).
52. Zhu, L. J. & Altmann, S. W. mRNA and 18S-RNA coapplication-reverse transcription for quantitative gene expression analysis. *Anal. Biochem.* **345**, 102–109 (2005).
53. Riemer, A. B., Keskin, D. B. & Reinherz, E. L. Identification and validation of reference genes for expression studies in human keratinocyte cell lines treated with and without interferon- γ —A method for qRT-PCR reference gene determination. *Exp. Dermatol.* **21**, 625–629 (2012).
54. Vandesompele, J., De Preter, K. & Pattyn, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, 0034.1 (2002).
55. Andersen, C. L., Jensen, J. L. & Orntoft, T. F. 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.* **64**, 5245–5250 (2004).
56. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–515 (2004).
57. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-DD CT-method. *Methods* **25**, 402–408 (2001).
58. Riedel, G. *et al.* An extended Δ CT-Method facilitating normalisation with multiple reference genes suited for quantitative RT-PCR analyses of human hepatocyte-like cells. *PLoS ONE* **9(3)**, e93031 (2014).

Acknowledgements

This study was funded by the National Research Foundation of Korea (2020R1F1A1069736) to Young-mi Lee and also supported by the National Research Foundation of Korea (2021R1C1C2005922) to Eun-Ji Won.

Author contributions

Y.-M.L., Conceptualization, writing, editing, funding; H.Y.C. and S.I., experiments, imaging, statistical analysis; R.-O.K., experiments, analysis; S.-J.K., data analysis; E.-J.W., Conceptualization, writing, editing, funding. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03098-x>.

Correspondence and requests for materials should be addressed to Y.-M.L. or E.-J.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021