Assessment of housekeeping genes as internal references in quantitative expression analysis during early development of oyster

Pin Huan¹, Hongxia Wang¹ and Baozhong Liu^{1,2*}

¹Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, 7 Nanhai Road, 266071, Qingdao, China

²Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, 1 Wenhai Road, 266000, Qingdao, China

(Received 2 February 2016, accepted 8 June 2016; J-STAGE Advance published date: 31 August 2016)

The early development of mollusks exhibits important characteristics from the developmental and evolutionary perspective. With the increasing number of genome-wide studies, accurate analyses of quantitative gene expression during development are impeded by the lack of validated reference genes. To improve the situation, in this study, we analyzed the expression stability of seven candidate housekeeping genes during early development of the Pacific oyster *Crassostrea gigas*: *actin*, *glyceraldehyde-3-phosphate dehydrogenase (gapdh)*, *α subunit of elongation factor 1 (elf1α)*, *adp-ribosylation factor 1 (arf1)*, *heterogeneous nuclear ribonucleoprotein q*, *ubiquitin-conjugating enzyme e2d2* and *ribosomal protein s18*. We focused on 11 stages from oocyte to D-veliger, which include crucial developmental processes such as axis determination, gastrulation and shell formation. Gene expression stabilities were assessed with the three commonly used programs geNorm, NormFinder and BestKeeper. Although the results obtained with the three programs varied to some extent, in general, *arf1*, *elf1α* and *gapdh* were highly ranked and *actin* was poorly ranked. This analysis also indicated that multiple genes should be used for normalization, and we concluded that *arf1-elf1α-gapdh* should be used as internal references. The findings of this study will help researchers to obtain accurate results in future quantitative gene expression analysis of development in bivalve mollusks.

Key words: *Crassostrea gigas*, expression stability, internal reference, normalization, real-time PCR

INTRODUCTION

Molluscan development exhibits important characters that are unique to either Lophotrochozoa species or the Mollusca phylum, such as spiral cleavage, a conserved trochophore larval stage, and shell formation (Nielsen, 2004; Mouëza et al., 2006). Current studies on molluscan development include many aspects such as cell fate mapping (Dictus and Damen, 1997; Hejnal et al., 2007), axis determination (Koop et al., 2007), organogenesis (Voronezhskaya et al., 2008; Wang et al., 2008) and shell formation (Nederbragt et al., 2002; Kin et al., 2009). Studies on these processes are essential to understand the development and evolution of mollusks. Despite the

increasing number of reports, however, studies on molluscan development are still in their infancy, especially when compared to model organisms such as the fruit fly and nematode. In recent years, the Pacific oyster *Crassostrea gigas* has emerged as a model mollusk species with the publication of the whole genome sequence (Zhang et al., 2012) and of fundamental studies (Fabioux et al., 2004; Badarotti et al., 2007; Dheilly et al., 2011; Huan et al., 2013). In most of these studies, quantitative gene expression analysis (mainly based on real-time PCR assay) is an important element, providing quantitative gene expression data to study gene function. Nevertheless, it should be noted that accurate evaluation of gene expression using real-time PCR relies on normalization of raw data using a validated internal reference gene(s), and currently there is no such reference gene covering early development of *C. gigas*.

Internal reference genes are stably expressed house-

Edited by Rimantas Kodzius

* Corresponding author. E-mail: bzliu@qdio.ac.cn

DOI: <http://doi.org/10.1266/ggs.16-00007>

keeping genes selected from numerous candidates. In previous research, although their expression stability was not validated, several housekeeping genes of *C. gigas* have been used as internal references; they include α subunit of elongation factor 1 (*elf1 α*) (Fabioux et al., 2004; Badariotti et al., 2007; Gonzalez et al., 2007), *actin* (Zhang et al., 2011), *glyceraldehyde-3-phosphate dehydrogenase* (*gapdh*) (Badariotti et al., 2007; Herpin et al., 2007), 18s rRNA (Meistertzheim et al., 2007) and 28s rRNA (David et al., 2005). However, it is inappropriate to simply assume a gene to be stably expressed without validation, especially given the complex early development of *C. gigas*. As shown in Fig. 1, a fertilized oocyte of *C. gigas* develops to a D-veliger in less than 24 h at 25 °C (Huan et al., 2012). Numerous developmental processes occur in this period, such as axis determination, gastrulation, shell formation and development of muscular and nervous systems. In these processes, expression of housekeeping genes may fluctuate with the dynamic changes of embryonic cells including cell division, migration, differentiation and death. A recent study on zebrafish early development revealed that the results of gene expression analysis could vary considerably when different genes were used for normalization (Hu et al., 2016). It is thus essential to assess the expression stability of candidate reference genes during oyster early development.

To date, two studies have assessed the expression stability of housekeeping genes in *C. gigas*. In 2011, a microarray analysis provided a list of housekeeping genes that are stably expressed in adult tissues and suggested that the *adp-ribosylation factor 1* (*arf1*) gene should be used as the internal reference (Dheilly et al., 2011). More recently, Du et al. (2013) conducted an expression analysis of housekeeping genes in larvae of *C. gigas* and suggested that two ribosomal protein genes be used for normalization. However, Dheilly et al. only analyzed adult tissues, and Du et

al. focused on D-veligers and umbo veligers. Neither of the studies covered developmental processes before the D-veliger stage, such as axis determination, shell formation and neurogenesis (see Fig. 1), which are crucial for studies into molluscan development and evolution. Here, to improve the situation, we analyzed the expression of seven housekeeping genes in the first 24 h of *C. gigas* development (from oocyte to D-veliger), and we propose optimal gene combinations to be used for normalization. Our results should provide fundamental support for accurate quantitative gene expression analysis during early development of this species.

MATERIALS AND METHODS

Sample collection Sexually mature adults of *C. gigas* were obtained from a local market in Qingdao, China. Gametes were dissected from five males and ten females. Proper amounts of sperm were added to oocytes in filtered seawater, and fertilized oocytes were cultured at 25 $\pm$ 0.5 °C with constant aeration. Early development of the oysters was very close to that described in Fig. 1: the first cleavage occurred at about 1.5–2 hours post fertilization (hpf), gastrulation began at around 4 hpf, early trochophores emerged at 9–10 hpf, early D-veligers with developing shell hinges could be discriminated at 11–12 hpf, and fully developed D-veligers were observed at 17–18 hpf. Good developmental synchrony was kept during the first 12 hpf, and was largely sustained thereafter. To obtain a global view of the expression of housekeeping genes during early development, we collected samples at 11 stages, including oocytes and embryos/larvae at 2, 4, 6, 8, 10, 12, 14, 16, 20 and 24 hpf. The samples were immersed in *RNAlater* (Ambion, USA), kept at 4 °C for 24 h with frequent agitation, and stored at –20 °C until RNA extraction. Three samples were collected at each

Fig. 1. Scheme depicting the early development of *C. gigas* in the first 24 h after fertilization at around 25 °C. Black boxes indicate major developmental stages and gray boxes indicate important processes that occur during this period. Note that the boundaries of boxes represent only approximate start or stop time points determined by microscopic observation, and thus do not indicate accurate definition of developmental stages/processes. Asterisks indicate the time points at which samples were collected in the present study. hpf, hours post fertilization.

stage, and in total there were 33 samples representing 11 developmental stages (Fig. 1). After sample collection, the remaining larvae were cultured for an additional 24 h to ensure that they developed normally.

RNA extraction and reverse transcription RNA was extracted using an E.Z.N.A. total RNA kit II (Omega Bio-Tek, USA) and its quality was verified by agarose gel electrophoresis and spectrophotometry. Five hundred nanograms of RNA was reverse-transcribed in a 50- μ l reaction using the Primerscript RT Reagent Kit with gDNA Eraser (Takara, China). To detect any genomic DNA contamination, all RNA samples were pooled to prepare an RT-minus control in which reverse transcriptase was excluded from the reaction mixture. No influence from genomic DNA was confirmed by the fact that there was no amplification when the RT-minus control was used as a template in the subsequent real-time PCR assay.

Selection of candidate housekeeping genes We collected 13 housekeeping genes as candidate internal reference genes (Table 1), including the three commonly used genes *elf1 α* , *gapdh* and *actin*, and the *arf1* gene, which was reported to be stably expressed in adult oyster tissues (Dheilly et al., 2011). The other nine genes were proposed as candidate reference genes by Du et al. (2013). Analysis of these 13 genes based on homology revealed that some belonged to the same functional classes (Table 1). Such genes could be co-regulated and might cause distorted results in the subsequent assessment of gene expression stability. Therefore, we selected only one gene from each functional class, with the exception that we chose two genes from the functional class of translation-related

proteins: *elf1 α* , which is widely used as an internal reference in *C. gigas*, and *rps18*, which was assessed to be one of the most stably expressed genes among different larval samples (Du et al., 2013). The final combination included seven housekeeping genes (shown in bold type in Table 1).

Primer design Primers were designed using Primer premier 6.0 with the following criteria: primer T_m, 60 $\pm$ 3 $^{\circ}$ C; primer length, 20–27 bp; expected products, 100–250 bp. Intron-spanning primers were designed for each gene except *gapdh*, which does not contain known introns. Amplification specificity and efficiency of each pair of primers were examined before they were used to estimate gene expression levels. Amplification specificity was certified when only a single peak was observed in melting curve analysis in a real-time PCR assay (Supplementary Fig. S1.). Amplification efficiency was calculated based on the slope of a standard curve generated by plotting Ct values against the relative quantity of templates. Primer sequences were adjusted until all primers had good amplification specificity and had amplification efficiencies of 0.90–1.05. Primer information is provided in Table 2.

Real-time PCR Real-time PCR was conducted in triplicate in a 10- μ l volume containing 1X SuperReal PreMix Plus (Tiangen, China), 0.25 μ M of each primer, and 0.3 μ l of cDNA. Negative controls using the RT-minus control were conducted for each gene to confirm that there was no influence from genomic DNA contamination. The PCR program was 95 $^{\circ}$ C for 15 min (to activate the DNA polymerase, as recommended by the manufacturer's manual), and 40 cycles of 95 $^{\circ}$ C for 15 s, 56 $^{\circ}$ C for 15 s and 72 $^{\circ}$ C

Table 1. The 13 candidate housekeeping genes

Gene name (abbreviation)	Accession no.	Functional class
<i>actin</i> (<i>actin</i>)	AF026063	cytoskeletal protein
<i>glyceraldehyde-3-phosphate dehydrogenase</i> (<i>gapdh</i>)	AJ5448861	glycolysis
<i>adp-ribosylation factor 1</i> (<i>arf1</i>)	HS159238	vesicular trafficking
<i>heterogeneous nuclear ribonucleoprotein q</i> (<i>hnrpq</i>)^a	CGI_10028368	processing of newly synthesized RNA
<i>ubiquitin-conjugating enzyme e2d2</i> (<i>ube2d2</i>)^b	CGI_10007222	protein degradation
<i>α subunit of elongation factor 1</i> (<i>elf1α</i>)^c	AB122066	translation
<i>ribosomal protein s18</i> (<i>rps18</i>)^c	CGI_10008101	translation
<i>heterogeneous nuclear ribonucleoprotein a2/b1</i> ^a	CGI_10028469	processing of newly synthesized RNA
<i>s-phase kinase-associated protein 1</i> ^b	CGI_10028032	protein degradation
<i>elongation factor 2</i> ^c	CGI_10017178	translation
<i>ribosomal protein l27</i> ^c	CGI_10024743	translation
<i>ribosomal protein l36</i> ^c	CGI_10021069	translation
<i>ribosomal protein l7</i> ^c	CGI_10020975	translation

The accession numbers refer to GenBank or OysterBase (<http://www.oysterdb.com>). The names of the seven genes used in the present study are in bold type.

^{a, b, c}: Genes with the same superscript belong to the same functional class.

Table 2. Primers used in this study

Gene	Primers (5'-3')	Intron-spanning	Product length (bp)	E ^a	Coefficient (R ²)
<i>actin</i>	F: CAGAGCTGTGTTTCCCTCCATTGT, R: AGTTGGTGACGATGCCGTGTTC	Yes	155	1.04	0.9982
<i>gapdh</i>	F: AGAACATCGTCAGCAACGCATCC, R: CCTCTACCACCACGCCAATCCT	No ^b	172	0.99	0.9982
<i>arf1</i>	F: TCAGGACAAGATCCGACCACTGT, R: GCAGCGTCTCTAAGTTCATCCTCATTA	Yes	150	0.94	0.9940
<i>hnrpq</i>	F: AGTGCATTCTTGTGTGGCATGATGA, R: ACCGTACTTTCTCTGTCCCGTTGTA	Yes	159	0.99	0.9982
<i>ube2d2</i>	F: ACAGTAATGGAAGCATCTGTCTGGATA, R: CGTCAGGATTTGGGTCAGTTAGCA	Yes	114	0.92	0.9931
<i>rps18</i>	F: ATTGAGCCAGGTCATGTCCAACAC, R: CACCGACAGTTCTTCTCTTCTTCC	Yes	161	0.94	0.9979
<i>elf1α</i>	F: CGAGAAGGAAGCTGCTGAGATGG, R: ACAGTCAGCCTGTGAAGTTCCTGTA	Yes	208	0.96	0.9970
<i>gata2/3</i>	F: TGCCAGAAGAGCAGGAACAAGTTG, R: TTTCTTCATTGTAAGCGGGCGGTTA	Yes	148	1.04	0.9930

^a: Amplification efficiency.

^b: There is no known intron in this gene.

for 20 s. A melting curve analysis was performed after each reaction to confirm specificity. Reactions were conducted on an Eppendorf Mastercycler ep realplex thermocycler. Baseline values were automatically determined using the Mastercycler ep realplex software V2.2. The drift correction option (provided by the software) was activated to correct drift in the baseline. To ensure the comparability of results from different experimental plates, several common samples were included in different plates, and the thresholds used to determine Ct values were adjusted slightly to obtain the same Ct values for the common samples from different reactions.

Gene expression stability analysis Gene expression stability was evaluated using three commonly used programs, geNorm V3.5 (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004) and BestKeeper V1 (Pfaffl et al., 2004). Raw Ct values and amplification efficiency were directly submitted to BestKeeper. For geNorm and NormFinder, data were transformed to relative expression levels by normalizing to the sample with the highest Ct value using the formula: relative expression level = $(1 + E)^{\Delta Ct}$, where ΔCt = the highest Ct value – Ct value of the sample, and E represents the amplification efficiency.

Quantification of a target gene using different normalization strategies The *gata2/3* gene, which has dynamic expression during early development (Liu et al., 2015), was selected as the target gene to test the effect of using different normalization strategies. The primers were designed as described above and the information is provided in Table 2. The relative expression of

gata2/3 was calculated through the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) using either *actin* or the gene combination *arf1-elf1α-gapdh* as reference. When multiple genes were used for normalization, the geometric mean of their Ct values was applied as recommended by Vandesompele et al. (2002). Student's *t*-test was used to determine whether the relative expression levels (at each time point) derived from the two normalization strategies were different.

RESULTS

Expression of housekeeping genes A boxplot graph depicting the distributions of Ct values is shown in Fig. 2, and all Ct values are provided in Supplementary Table S1 (n = 33 for each gene). The Ct values ranged from 19.43 to 30.13, with the majority ranging from 21 to 28. The Ct values of *elf1α* and *rps18* were generally lower than those of the other five genes, indicating that these two genes have relatively higher expression levels than the others, which exhibited similar expression levels.

Gene expression stability analysis Expression stability of the seven genes was evaluated using the three commonly used programs geNorm, NormFinder and BestKeeper, which are based on different algorithms (Vandesompele et al., 2002; Andersen et al., 2004; Pfaffl et al., 2004). The geNorm program calculates M values of genes to represent the gene expression stability, where lower M values indicate more stable expression. The M values of the seven genes ranged from 0.885 to 1.230. The two genes with lowest M values were *gapdh* and *arf1*,

Fig. 2. Boxplot graph depicting the distributions of Ct values of the seven candidate housekeeping genes. The lower/upper quartiles and median are indicated by the two ends of boxes and the bands within boxes, respectively. The two ends of whiskers represent the lowest and highest values still within 1.5 inter-quartile ranges of the lower/upper quartiles. The black dots represent outliers.

Fig. 3. M values of genes calculated by geNorm.

Table 3. Results obtained from the three programs

Gene	geNorm		NormFinder		BestKeeper		
	M value	Rank	Stability value	Rank	Coeff. of corr. [r]	p value	Rank
<i>arf1</i>	0.885	1	10.674	7	0.926	0.001	1
<i>gapdh</i>	0.885	1	8.844	5	0.900	0.001	3
<i>hnrpq</i>	1.016	3	7.306	2	0.837	0.001	4
<i>elf1α</i>	1.034	4	7.165	1	0.901	0.001	2
<i>ube2d2</i>	1.085	5	9.556	6	0.827	0.001	5
<i>actin</i>	1.154	6	8.767	4	0.820	0.001	6
<i>rps18</i>	1.230	7	8.323	3	0.685	0.001	7

followed by *heterogeneous nuclear ribonucleoprotein q* (*hnrpq*) (Table 3 and Fig. 3). NormFinder calculates gene expression stability values, which revealed that *elf1α* was the most stably expressed gene (stability value 7.165) (Table 3 and Fig. 4). NormFinder also suggested the best

Fig. 4. Stability values of genes calculated by Normfinder.

Fig. 5. Assessment of optimal number of reference genes for normalization. The $V_{n/n+1}$ values were calculated by geNorm.

combination of two genes to be *elf1α* and *arf1* (stability value 3.911). BestKeeper calculates a BestKeeper Index and evaluates gene expression stability based on correlation analysis between a gene and the BestKeeper Index (reflected by the Pearson's correlation coefficient (r) and p values) (Pfaffl et al., 2004). Genes with high r values and low p values are considered to be stably expressed. The top three ranked genes were *arf1* ($r = 0.926$), *elf1α* ($r = 0.901$) and *gapdh* ($r = 0.900$) (Table 3).

While the results from different programs varied, they shared many similarities (especially for geNorm and BestKeeper; see Table 3). *elf1α* was generally highly ranked by all programs (a top-two gene by Normfinder and Bestkeeper, and ranked fourth by geNorm). *arf1* and *gapdh* were both highly ranked by geNorm and BestKeeper, but poorly ranked by NormFinder. However, NormFinder recommended *arf1* and *elf1α* as the best gene combination for normalization. Taken together, these findings suggested that the optimal gene combination for normalization would be *arf1-elf1α-gapdh*.

Optimal reference gene number for normalization Using multiple genes is preferred for normalization because there is probably no ideal single gene in many cases (Vandesompele et al., 2002). The geNorm program calculated pairwise variation ($V_{n/n+1}$) to assess the optimal number of reference genes for normalization; n genes should be enough for normalization when $V_{n/n+1}$ is

Fig. 6. Comparison of different normalization strategies.

less than 0.15 (Vandesompele et al., 2002). As shown in Fig. 5, the $V_{n/n+1}$ values decreased when n increased from two to six. There was no $V_{n/n+1}$ value below 0.15; the two lowest values were $V_{5/6}$ (0.187) and $V_{6/7}$ (0.175). On the other hand, it should be noted that when gene number decreased from six to three, the $V_{n/n+1}$ value only increased from 0.175 to 0.220 (for comparison, when using two genes, $V_{n/n+1}$ increased markedly to 0.330).

Comparison of different normalization strategies

To test whether different normalization strategies might influence the quantification results of gene expression, we analyzed the expression of the *gata2/3* gene in early development, using either *actin* or the gene combination *arf1-elf1α-gapdh* for normalization. As shown in Fig. 6, the expression patterns of *gata2/3* obtained from the two normalization strategies were similar, comprising low expression before 4 hpf, high expression from 4 to 12 hpf (except 10 hpf), and low expression from 14 to 24 hpf. Statistical analysis revealed no significant differences at any time point between the two normalization strategies (Student's *t*-test, $P > 0.05$).

DISCUSSION

Quantitative gene expression analysis is reliable only when proper reference genes are used for normalization. In the present study, we revealed varied expression stabilities of seven housekeeping genes during early development of *C. gigas* and thus highlighted the need to evaluate the expression stability of housekeeping genes for normalization.

To develop a strategy to assess expression stabilities of housekeeping genes, researchers have to solve a circular problem: how to evaluate the expression stability of a gene while no reliable measure of normalization is available (Vandesompele et al., 2002; Andersen et al., 2004; Pfaffl et al., 2004). Current methods generally rely on the relative expression of multiple housekeeping genes. Under

these conditions, co-regulated genes should be avoided to the maximum extent. After assembling a list of 13 candidate genes, we examined their potential function based on homology and selected only one gene from each functional class (except for *elf1α* and *rps21*, which are both involved in translation). Although 18S and 28S rRNA are widely used as internal references, we excluded them, both because they also function in translation and because their ultra-high abundance in cDNA would cause false-positive results when evaluating expression stability (Vandesompele et al., 2002).

As mentioned above, co-regulated genes can cause distorted results when evaluating gene expression stability. Therefore, if it is not possible to exclude some genes belonging to the same functional category, more attention should be paid when analyzing data. The top three genes ranked by geNorm and BestKeeper (*arf1*, *gapdh* and *hnnpq/elf1α*) belong to different functional classes. On the other hand, two of the top three genes ranked by NormFinder are both translation-related genes (*elf1α* and *rps18*). In addition, a transcriptomic analysis revealed that the other gene, *hnnpq* (ranked third), which is involved in processing of newly synthesized RNA (pre-mRNA), had a similar expression trend to *rps18* during development (Du et al., 2013). Therefore, it can be argued that the top three genes ranked by NormFinder (*elf1α*, *hnnpq* and *rps18*) might be co-regulated. However, the results of geNorm counter this argument. The algorithm of geNorm is based on pairwise correlation analysis (Andersen et al., 2004) and it should therefore be very sensitive to co-regulated genes. If the three genes are co-regulated, they should be highly ranked by geNorm. However, geNorm did not rank these three genes highly, and indeed assessed *rps18* as the least stably expressed gene (Table 3). Therefore, the top three genes ranked by NormFinder are unlikely to be co-regulated. On the other hand, to be cautious, we included only *elf1α* in the gene combination that we suggested for normalization.

The dynamic changes of cells in embryos and larvae

of *C. gigas* during early development should involve changes in the expression of various housekeeping genes that participate in metabolism, the cytoskeleton, translation, etc. Therefore, it is not surprising that all $V_{n/n+1}$ values were higher than 0.15. This indicates that all seven genes showed some variation in expression, and that even all seven are not sufficient for normalization according to the criterion suggested by Vandesompele et al. (2002) ($V_{n/n+1} < 0.15$). In fact, similar to our results, high $V_{n/n+1}$ values (all over 0.15) were also observed in a study on the gonad of another bivalve species (Mauriz et al., 2012), although different candidate housekeeping genes were used. Obtaining relatively high $V_{n/n+1}$ values may be a common result in analyses using molluscan samples. However, the number of genes used for normalization is a trade-off between practical considerations and accuracy (Vandesompele et al., 2002). Many reference genes are unlikely to be included if only few target genes are to be investigated, or if only limited amounts of samples are available. We noticed that the $V_{n/n+1}$ value did not increase substantially (by 0.045) when n decreased from six to three. Therefore, if we take a looser criterion (e.g., $V_{n/n+1} \leq 0.22$), using our three selected genes for normalization would be acceptable. Taking all these considerations together, we decided to use the three genes *arf1*, *elf1 α* and *gapdh* for normalization when studying early development of *C. gigas*.

Among these three selected reference genes, *arf1* encodes a conserved guanine nucleotide-binding protein that has a central role in intra-Golgi transport (Serafini et al., 1991; Li et al., 2004). Although it has been used as reference gene in human (Eisenberg and Levanon, 2003), *arf1* of *C. gigas* was only revealed to be stably expressed in adult tissues in a recent microarray study (Dheilly et al., 2011). *gapdh* encodes a key enzyme in glucose metabolism and is one of the most commonly used reference genes in quantitative gene expression analysis (Huggett et al., 2005). In the previous study by Dheilly et al. (2011), *gapdh* was also found to be stably expressed in most adult tissues, except adductor muscle. Therefore, it is likely that *arf1* and *gapdh* are stably expressed in early development as well as adult tissues. The third gene, *elf1 α* , encodes the α subunit of eukaryotic elongation factor 1, which mediates the entry of the aminoacyl tRNA into a free site of the ribosome. *elf1 α* is a commonly used reference gene, and has been recommended to be used to study gene expression during development of several fishes (Tang et al., 2007; Infante et al., 2008; McCurley and Callard, 2008). In *C. gigas*, however, *elf1 α* was revealed as an unsuitable reference gene when analyzing adult tissues (Dheilly et al., 2011). It is likely that *elf1 α* has different expression patterns in embryonic, larval and adult stages. In previous studies on the development of *C. gigas*, both *elf1 α* and *gapdh* have been used as internal references (Fabioux et al., 2004; Badariotti et al.,

2007; Gonzalez et al., 2007; Herpin et al., 2007), including our studies (Huan et al., 2012, 2014). Here, by revealing *elf1 α* and *gapdh* to be stably expressed genes, we confirm the reliability of previous studies. On the other hand, *actin* was not stably expressed, and should not be used in future studies.

To date, there are only a few studies on the expression of housekeeping genes during development of bivalve mollusks. In a recent report on the D-veliger and umbo veliger of *C. gigas*, two ribosomal protein genes were recommended as internal references (Du et al., 2013). The differences between their study and ours may be attributable to different developmental stages (D-umbo veligers in that study vs. development from oocyte to D-veliger here) and also the different candidate gene sets (14 genes in that study vs. seven genes here). In another bivalve species, Yesso scallop (*Patinopecten yessoensis*), the expression of 12 housekeeping genes was analyzed in early embryos/larvae (Feng et al., 2013). Those authors also used samples from fertilized oocyte to D-veliger, but the sampling points differed considerably from ours (five stages in that study vs. 11 stages here). The candidate genes were also very different, with only two common genes (*actin* and *gapdh*). Despite the different experimental designs, however, *gapdh* was highly ranked and *actin* was poorly ranked in both studies, indicating some common characters of gene expression among different bivalve species.

To test the effect of reference gene choice on gene expression quantification, we analyzed the expression of *gata2/3* using two normalization strategies. Although they were below the threshold of significance, differences in relative *gata2/3* expression could be observed at several time points (e.g., at 4, 12, 20 and 24 hpf; see Fig. 6), indicating that normalization strategy likely influences quantification results. Combined with the results of gene stability analysis mentioned above, we recommend that the three genes be used for normalization in future studies on early development of *C. gigas*. Expression of the *gata2/3* gene during early development has been investigated using whole mount *in situ* hybridization (WMISH) in our previous work (Liu et al., 2015). The results obtained from real-time PCR are generally consistent with those obtained using WMISH, including initiation of expression during early gastrulation (4–6 hpf), relatively high expression in gastrulas and trochophores (6–12 hpf), and decreased expression after the trochophore stage (14–24 hpf). Analysis of the results based on both real-time PCR and WMISH also reveals some indications on *gata2/3* function. For instance, the decreased expression beyond 14 hpf revealed by real-time PCR may reflect decreased expression surrounding shell plates in this period, given that expression in the presumed hematopoietic tissue was relatively constant (Liu et al., 2015).

In conclusion, our results revealed the different expres-

sion stabilities of housekeeping genes during early development of *C. gigas*. We propose using the multiple gene combination *arf1-elf1 α -gapdh* for normalization. By applying the normalization strategy provided here, accurate results can be obtained in quantitative gene expression studies, which will promote research on early development of *C. gigas*.

This work was financially supported by the National Natural Science Foundation of China (31472265) and the Scientific and Technological Innovation Project from Qingdao National Laboratory for Marine Science and Technology (No. 2015ASKJ02). The authors declare no conflict of interest.

REFERENCES

- Andersen, C. L., Jensen, J. L., and Orntoft, T. F. (2004) 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.
- Badariotti, F., Thuau, R., Lelong, C., Dubos, M. P., and Favrel, P. (2007) Characterization of an atypical family 18 chitinase from the oyster *Crassostrea gigas*: evidence for a role in early development and immunity. *Dev. Comp. Immunol.* **31**, 559–570.
- David, E., Tanguy, A., Pichavant, K., and Moraga, D. (2005) Response of the Pacific oyster *Crassostrea gigas* to hypoxia exposure under experimental conditions. *FEBS J.* **272**, 5635–5652.
- Dheilly, N. M., Lelong, C., Huvet, A., and Favrel, P. (2011) Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns. *BMC Genomics* **12**, 468.
- Dictus, W. J., and Damen, P. (1997) Cell-lineage and clonal-contribution map of the trochophore larva of *Patella vulgata* (Mollusca). *Mech. Develop.* **62**, 213–226.
- Du, Y., Zhang, L., Xu, F., Huang, B., Zhang, G., and Li, L. (2013) Validation of housekeeping genes as internal controls for studying gene expression during Pacific oyster (*Crassostrea gigas*) development by quantitative real-time PCR. *Fish Shellfish Immunol.* **34**, 939–945.
- Eisenberg, E., and Levanon, E. Y. (2003) Human housekeeping genes are compact. *Trends Genet.* **19**, 362–365.
- Fabioux, C., Huvet, A., Lelong, C., Robert, R., Pouvreau, S., Daniel, J. Y., Minguant, C., and Le Pennec, M. (2004) Oyster *vasa*-like gene as a marker of the germline cell development in *Crassostrea gigas*. *Biochem. Biophys. Res. Commun.* **320**, 592–598.
- Feng, L., Yu, Q., Li, X., Ning, X., Wang, J., Zou, J., Zhang, L., Wang, S., Hu, J., Hu, X., and Bao, Z. (2013) Identification of reference genes for qRT-PCR analysis in Yesso scallop *Patinopekten yessoensis*. *PloS one* **8**, e75609.
- Gonzalez, M., Gueguen, Y., Desserre, G., de Lorgeril, J., Rometand, B., and Bachere, E. (2007) Molecular characterization of two isoforms of defensin from hemocytes of the oyster *Crassostrea gigas*. *Dev. Comp. Immunol.* **31**, 332–339.
- Hejnal, A., Martindale, M. Q., and Henry, J. Q. (2007) High-resolution fate map of the snail *Crepidula fornicata*: The origins of ciliary bands, nervous system, and muscular elements. *Dev. Biol.* **305**, 63–76.
- Herpin, A., Lelong, C., Becker, T., Favrel, P., and Cunningham, C. (2007) A tolloid homologue from the Pacific oyster *Crassostrea gigas*. *Gene Expr. Patterns* **7**, 700–708.
- Hu, Y., Xie, S., and Yao, J. (2016) Identification of novel reference genes suitable for qRT-PCR normalization with respect to the zebrafish developmental stage. *PLoS One* **11**, e0149277.
- Huan, P., Wang, H., Dong, B., and Liu, B. (2012) Identification of differentially expressed proteins involved in the early larval development of the Pacific oyster *Crassostrea gigas*. *J. Proteomics* **75**, 3855–3865.
- Huan, P., Liu, G., Wang, H., and Liu, B. (2013) Identification of a tyrosinase gene potentially involved in early larval shell biogenesis of the Pacific oyster *Crassostrea gigas*. *Dev. Genes Evol.* **223**, 389–394.
- Huan, P., Liu, G., Wang, H., and Liu, B. (2014) Multiple ferritin subunit genes of the Pacific oyster *Crassostrea gigas* and their distinct expression patterns during early development. *Gene* **546**, 80–88.
- Huggett, J., Dheda, K., Bustin, S., and Zumla, A. (2005) Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* **6**, 279–284.
- Infante, C., Matsuoka, M., Asensio, E., Canavate, J., Reith, M., and Manchado, M. (2008) Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. *BMC Mol. Biol.* **9**, 28.
- Kin, K., Kakoi, S., and Wada, H. (2009) A novel role for *dpp* in the shaping of bivalve shells revealed in a conserved molluscan developmental program. *Dev. Biol.* **329**, 152–166.
- Koop, D., Richards, G. S., Wanninger, A., Gunter, H. M., and Degnan, B. M. (2007) The role of MAPK signaling in patterning and establishing axial symmetry in the gastropod *Haliotis asinina*. *Dev. Biol.* **311**, 200–212.
- Li, Y., Kelly, W. G., Logsdon, J. M., Schurko, A. M., Harfe, B. D., Hill-Harfe, K. L., and Kahn, R. A. (2004) Functional genomic analysis of the ADP-ribosylation factor family of GTPases: phylogeny among diverse eukaryotes and function in *C. elegans*. *FASEB J.* **18**, 1834–1850.
- Liu, G., Huan, P., and Liu, B. (2015) A GATA2/3 gene potentially involved in larval shell formation of the Pacific oyster *Crassostrea gigas*. *Dev. Genes Evol.* **225**, 253–257.
- Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2^{- $\Delta\Delta$ CT} method. *Methods* **25**, 402–408.
- Mauriz, O., Maneiro, V., Pérez-Parallé, M. L., Sánchez, J. L., and Pazos, A. J. (2012) Selection of reference genes for quantitative RT-PCR studies on the gonad of the bivalve mollusc *Pecten maximus* L. *Aquaculture* **370–371**, 158–165.
- McCurley, A., and Callard, G. (2008) Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Mol. Biol.* **9**, 102.
- Meistertzheim, A. L., Tanguy, A., Moraga, D., and Thebault, M. T. (2007) Identification of differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged thermal stress. *FEBS J.* **274**, 6392–6402.
- Mouëza, M., Gros, O., and Frenkiel, L. (2006) Embryonic development and shell differentiation in *Chione cancellata* (Bivalvia, Veneridae): an ultrastructural analysis. *Invertebr. Biol.* **125**, 21–33.
- Nederbragt, A. J., van Loon, A. E., and Dictus, W. J. A. G. (2002) Expression of *Patella vulgata* orthologs of *engrailed* and *dpp-BMP2/4* in adjacent domains during molluscan shell development suggests a conserved compartment boundary mechanism. *Dev. Biol.* **246**, 341–355.
- Nielsen, C. (2004) Trochophora larvae: cell-lineages, ciliary bands, and body regions. 1. Annelida and Mollusca. *J. Exp. Zool. B Mol. Dev. Evol.* **302**, 35–68.

- Pfaffl, M. W., Tichopad, A., Prgomet, C., and Neuvians, T. P. (2004) 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.
- Serafini, T., Orci, L., Amherdt, M., Brunner, M., Kahn, R. A., and Rothman, J. E. (1991) ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell* **67**, 239–253.
- Tang, R., Dodd, A., Lai, D., McNabb, W. C., and Love, D. R. (2007) Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. *Acta Biochim. Biophys. Sin.* **39**, 384–390.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, DOI:10.1186/gb-2002-3-7-research0034.
- Voronezhskaya, E., Nezlin, L., Odintsova, N., Plummer, J., and Croll, R. (2008) Neuronal development in larval mussel *Mytilus trossulus*; (Mollusca: Bivalvia). *Zoomorphology* **127**, 97–110.
- Wang, X., Liu, B., Wang, G., Tang, B., and Xiang, J. (2008) Molecular cloning and functional analysis of cathepsin B in nutrient metabolism during larval development in *Meretrix meretrix*. *Aquaculture* **282**, 41–46.
- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., et al. (2012) The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**, 49–54.
- Zhang, Y., He, X., and Yu, Z. (2011) Two homologues of inhibitor of NF-kappa B (IkappaB) are involved in the immune defense of the Pacific oyster, *Crassostrea gigas*. *Fish Shellfish Immunol.* **30**, 1354–1361.