Available online at www.sciencedirect.com
ScienceDirect
J. Biochem. Biophys. Methods 70 (2008) 985–991
JOURNAL OF biochemical and biophysical methods
www.elsevier.com/locate/jbbm

Analytical approach for selecting normalizing genes from a cDNA microarray platform to be used in q-RT-PCR assays: A cnidarian case study

Mauricio Rodriguez-Lanetty [a,*], Wendy S. Phillips [b], Sophie Dove [c], Ove Hoegh-Guldberg [c], Virginia M. Weis [b]

[a] *ARC Centre of Excellence for Coral Reef Studies /Centre for Marine Studies, University of Queensland, St. Lucia QLD 4072, Australia*
[b] *Department of Zoology, Oregon State University, Corvallis OR 97331, United States*
[c] *Centre for Marine Studies, University of Queensland, St. Lucia QLD 4072, Australia*

Received 24 April 2007; received in revised form 4 July 2007; accepted 27 August 2007

Abstract

Research in gene function using Quantitative Reverse Transcription PCR (q-RT-PCR) and microarray approaches are emerging and just about to explode in the field of coral and cnidarian biology. These approaches are showing the great potential to significantly advance our understanding of how corals respond to abiotic and biotic stresses, and how host cnidarians/dinoflagellates symbioses are maintained and regulated. With these genomic advances, however, new analytical challenges are also emerging, such as the normalization of gene expression data derived from q-RT-PCR. In this study, an effective analytical method is introduced to identify candidate housekeeping genes (HKG) from a sea anemone (*Anthopleura elegantissima*) cDNA microarray platform that can be used as internal control genes to normalize q-RT-PCR gene expression data. It is shown that the identified HKGs were stable among the experimental conditions tested in this study. The three most stables genes identified, in term of gene expression, were beta-actin, ribosomal protein L12, and a Poly(a) binding protein. The applications of these HKGs in other cnidarian systems are further discussed.
© 2007 Published by Elsevier B.V.

Keywords: Cnidarians; Coral; Housekeeping genes; q-RT-PCR; Microarray; Gene expression; Symbiosis

1. Introduction

As a result of recent advances in genomic technology, a new set of tools has started to be used in addressing central questions in cnidarian and coral reef physiology [1]. Techniques such as gene expression profiling using DNA microarray, quantitative reverse transcription PCR (q-RT-PCR) and differential display assays are showing the great potential to significantly advance our understanding of how corals respond to abiotic and biotic stresses [2,3], and how the complex mutualistic association between host cnidarians and symbiotic dinoflagellates is maintained and regulated [4]. With these genomic advances, however, new analytical and methodological challenges are also emerging, not only for the field of coral biology, but also for the general field of molecular biology.

Gene expression profiling studies using microarray technology to measure messenger RNA (mRNA) expression yields long lists of genes that appear to be differentially expressed. Differential expression of some of these genes is subsequently validated using q-RT-PCR assays, which compared to other methods to assess gene expression, such as northern blot analysis or conventional RT-PCR, are rapid, reproducible, and have a wide dynamic range [5].

q-RT-PCR, a modification of PCR, is used to rapidly measure the quantity of starting amounts of DNA, cDNA, or RNA templates in a particular sample. The technique is based on the detection of a fluorescent reporter molecule that increases as PCR product accumulates with each cycle of amplification. q-RT-PCR exceeds the limitations of traditional end-point PCR methods by allowing either absolute or relative quantification of

* Corresponding author. Tel.: +61 7 33653378; fax: +61 7 33654755.
E-mail address: m.rodriguez-lanetty@uq.edu.au (M. Rodriguez-Lanetty).

0165-022X/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.jbbm.2007.08.005

PCR product at the end of each cycle. This ability has greatly enhanced several areas of research including gene expression analysis and genotyping assays.

When measuring the expression (i.e. amount of mRNA transcripts) of a gene of interest using q-RT-PCR assays, the expression data has to go through a step of normalization to remove or eliminate variation in the amount of starting material between the samples, e.g. caused by sample-to-sample variation, variation in RNA integrity, reverse transcription efficiency differences and cDNA sample loading variation [6–8]. This is especially relevant when the samples have been obtained from different individuals, different tissues and different time courses, and will result in the misinterpretation of the derived expression profile of the target genes. Therefore, normalization of target gene expression levels must be performed to compensate intra-and inter-kinetic RT-PCR variations (sample-to-sample and run-to-run variations) [9]. q-RT-PCR data normalization can be carried out against total cellular DNA or RNA content (molecules/g total DNA/RNA and concentrations/g total DNA/RNA) or against endogenous unregulated reference gene transcript. Normalization according the total cellular RNA content has increasingly been used, but little is known about the total RNA content of cells or even about the mRNA concentrations [10]. The content per cell or per gram tissue may vary in different tissues in vivo, in cell culture (in vitro), between individuals and under different experimental conditions. Alternatively, relative quantification can be determined by normalizing the level of expression of a gene of interest relative to the levels of an endogenous reference gene, also called internal control or most popularly termed housekeeping gene (HKG). This housekeeping gene can be co-amplified in the same tube in a multiplex assay or can be amplified in a separate tube. Therefore, relative quantification using HKGs does not require standards with known concentrations and the reference can be any transcript, as long as its sequence is known. The normalization values for different samples can be then directly compared [10].

When using housekeeping genes to normalize gene expression data, the underlying assumption is that they are expressed at constant levels across the samples and that their expression does not vary in response to the experimental manipulation. Because the expression of the target gene is measured relative to the housekeeping gene, it is critical that these assumptions hold. However, it has become clear that no single housekeeping gene can be used for all studies and that the choice of the control gene depends on the tissue type and the experimental conditions [11–13]. Furthermore, it has been shown that normalization based on a single housekeeping gene led to erroneous quantifications of up to 3-fold in 25% of cases and 6.4-fold in 10% of cases, with sporadic cases showing errors greater than 20-fold [14]. It is recommended, instead normalizing the gene expression of gene of interests to the geometric mean of several housekeeping genes [14]. Nevertheless, finding multiple HKGs to be used as normalizing internal controls is challenging and a difficult task.

Since the selection of internal control genes is tissue and experimental condition dependent, and there is a need for several of them for q-RT-PCR normalization, we suggest that potential HKG candidates for normalizing q-RT-PCR data could be attained from a cDNA microarray platform under similar experimental settings. The justification of this approach lies on the fact that only small proportions of genes from large, comprehensive cDNA microarrays tend to be differentially expressed under any experimental design, leaving the gene expression of most of the genes on the microarrays unalterable. Because of this, microarray experiments can be self-normalized using the data from the thousands of genes in the array that do not change expression [15]. Thus, microarrays offer a great advantage in mining for candidate HKGs.

In this study, we propose a statistical and analytical approach to select candidate “Housekeeping Genes” from microarray experiments, which then can be used as normalizing genes in q-RT-PCR assays. The gene expression stability among these candidate genes from q-RT-PCR data is further tested. For this study, we used a (cnidarian) sea anemone (*Anthopleura elegantissima*) cDNA microarray platform, where two treatments or conditions were compared: symbiont-free anemones (aposymbiotic) and symbiotic anemones (see [4], for further details regarding the experimental design). After statistical and screening analyses, we found six candidate genes out of 10,368 cDNAs. We discuss the potential for these HKGs to be tested and used in other cnidarian systems.

2. Materials and methods

2.1. Selection of candidate housekeeping genes from a microarray platform

The gene expression data from the microarray experiments published by Rodriguez-Lanetty et al. [4] and deposited in Gene Expression Omnibus (GSE3858) was used. In this study, the gene transcription profile from sea anemones hosting symbionts were compared to anemones free of symbionts (referred to as aposymbiotic).

Three criteria were established for selecting candidate HKGs from the microarray platform; (1) these genes should show low gene expression variability among samples within and between treatments; (2) They should show a ratio of expression not statistically different than 1 between treatments, and (3) they should be abundant transcripts.

In order to filter out genes whose expressions are very variable among samples, we ranked all the genes on the array based on their coefficient of variation (CV) of the array-normalized Cy5/Cy3 ratios and pre-selected those genes that had a coefficient of variation lower than 5%. From this list of pre-selected genes, we tested whether the Cy5/Cy3 ratios were not significantly different to ratio = 1 (null hypothesis), using a one-sample *t*-test (which compares the mean of a group to a hypothetical value), filtering out those genes whose ratios were significantly different from one ($p < 0.05$). In the case where microarray data includes more than two treatments or groups, we recommend that ANOVA analysis be performed and only those significant genes among the treatment be filtered out.

From the list of non-differentially expressed genes, we then filtered out those genes that were below a set threshold of fluorescent intensity [measured by the Photo-Multiplier Tube (PMT) detector] per array-scanned channel. Only those pre-

Table 1
Candidate housekeeping genes, ranked in order of their expression stability, and their forward and reverse primers used in q-RT-PCR assays

Gene	Rank	Primers	Product size
Beta-actin (b-actin)	1	For: CTG ATG GAC AGG TCA TCA CCA T Rev: CTC GTG GAT ACC AGC AGA TTC C	200 bp
Ribosomal protein L12 (RibpL12)	2	For: CCT AAG TTC GAC CCC AAT GA Rev: TTT TGG ATG GTC AGT TGG AC	191 bp
Poly(a) binding protein (Pabp)	3	For: AAT GGC GTC TCT ATA CGT TGG Rev: CGT GAG CTG GTT GTT GGA A	167 bp
Ribosomal protein L10 (RibpL10)	4	For: TGT GTG GCA AAG AAG GAT GA Rev: GTA GCA GGG ACG ATT TCA GG	212 bp
CCAAT/enhancer-binding protein (C/EBP)	5	For: CGC TTC AAG ACC AAA CAA AA Rev: TCG GAG TAC GTT GAG CTC TTT	120 bp
Senescence-associated protein (Sap)	6	For: TCG CAC CAT TGG CTT TTC Rev: CCT ATC ATT GCG AAG CAG AA	204 bp

selected genes with PMT values higher than 30% of the saturation PMT values were chosen (PMT > 20,000). The rationale for this last criterion of selection is to choose common HKGs based on the assumption that transcript abundance correlates with PMT intensity values.

2.2. Stability of the selected candidate housekeeping genes from the microarray platform

2.2.1. Primer design and q-RT-PCR assays

Specific primers amplifying approximately 100–200 bp PCR products were designed for the chosen candidate HKGs, and specific amplification of appropriately sized bands was checked (Table 1). The selected range of PCR product length was based on the recommendations given by Fleige et al. [16]. The candidate HKGs were tested on new sea anemone samples from the same two conditions, one infected with symbionts and the other one free of symbionts. cDNAs of five samples from each of the two treatments or conditions were synthesized from 100 ng of mRNA and then diluted to a final volume of 330 µl. Two µl of cDNA was used in triplicate 20 µl quantitative RT-PCR reactions with 250 nM primers and iQ SYBR Green Supermix (BioRad) for a total of 40 cycles. The comparative delta CT method corrected for the actual PCR efficiency was used to determine relative quantities of mRNA transcripts from each sampled anemone. The PCR efficiency was determined using Lin-RegPCR [17].

2.2.2. Gene-stability measure of candidate housekeeping genes

We used the method described by Vandesompele et al. [14] to determine the expression stability (*M* value) of control HKGs on the basis of non-normalized expression data from q-RT-PCR assays. The "*M*-value" measure relies on the principle that the expression ratio of two hypothetical ideal internal control genes is identical in all samples, regardless of the experimental condition or treatment. In this way, variation of the expression ratios of two actual housekeeping genes reflects the fact that one (or both) of the genes is (are) not constantly expressed, with increasing variation in ratio corresponding to decreasing expression stability (*M* value). For every candidate housekeeping gene, we determined the pairwise variation with all other candidate control genes as the standard variation of the logarithmically transformed expression ratios, which was defined as the internal control gene-stability measure "*M* value", the average pairwise variation of a particular gene with all other control genes. Genes with the lowest *M* values have the most stable expression and therefore would be selected as the ideal housekeeping genes.

The calculation of the *M* values was performed using the software GeNorm created by Vandesompele et al. [14]. This program automatically calculates the gene-stability measure *M* for all control genes in a given set of samples, and enables elimination of the worst-scoring housekeeping gene (which is the one with the highest *M* value).

2.2.3. Testing ideal housekeeping genes: comparison of array and q-RT-PCR expression data

A normalization factor based on the expression level of the best-performing or the most stable housekeeping genes was calculated by computing the geometric mean among these genes. The reason to use a geometric average instead of an arithmetic average relies on the fact that the former is less sensitive to outlying values and abundance differences between the different genes [14]. This normalization factor was used to normalize the expression q-RT-PCR data of four target genes of interest (Table 2) known to be differentially expressed from a previous microarray study [4]. The expression values of array data and q-RT-PCR were compared. The sample of ratios (symbiotic vs aposymbiotic state) from the Q-RT-PCR data for each gene was compared to the absolute ratio value from the microarray data (Table 2) using a one-sample *t*-test ($\alpha = 0.05$).

Table 2
Comparison of ratios of expression of four selected genes involved in symbiosis from microarray and q-RT-PCR data

Gene	Gene expression (microarray) [a]	Gene expression (q-RT-PCR) [b]	GenBank accession ID
Death-associated protein 1 (DAP-1)	↑ 1.39	↑ 1.51	DQ309530
Calcium/calmodulin-dependent protein kinase IV	↓ 1.57	↓ 1.33	DQ309544
Plasma glutamate carboxypeptidase	↑ 1.70	↑ 1.77	DQ309521
Collagen alpha-1 chain, type XII	↓ 1.94	↓ 2.25	DQ309546

Arrows indicate the direction of expression as a function of symbiosis: up-or down-regulation.

[a] Data from Rodriguez-Lanetty et al. [4].

[b] Data from this study.

3. Results

3.1. Selection of candidate housekeeping genes

Out of 10,368 genes on the microarray, 18 genes satisfied the three stringent criteria of selection for housekeeping genes, which were low gene expression variability among all samples; no significant difference in gene expression between treatments ($\alpha=0.05$), and high abundance of transcripts (Fig. 1).

After completing sequencing and BLAST analyses, six unique genes out of the 18 candidate genes showed significant homology to genes deposited in GenBank. These genes were beta-actin, the ribosomal protein L12 and L10, a Poly(a)-binding protein gene, a CCAAT/enhancer-binding protein gene and a putative senescence-associated protein gene (Table 1), which were then further tested as potential housekeeping genes in q-RT-PCR assays. The genes that did not show any significant hit with the GenBank database were not further tested as potential candidate HKGs. The reason to remove the unknown genes from the analyses is because we are not able in the current stage to satisfy one of the important criteria when selecting multiple housekeeping genes to be used for normalizing Q-RT-PCR data, which is the selected HKGs should not be co-expressed or belonging to the same functional group. So because we could not determine the actual function of some of the genes from gene database comparisons, we decided not to include them in our analyses.

3.2. Analysis of gene-stability of candidate housekeeping genes from q-RT-PCR assays

The gene-stability expression measures (M) for the six candidate housekeeping genes on the basis of non-normalized expression levels from q-RT-PCR assays are shown in Fig. 2. From a stepwise exclusion of the least stable control gene, we determined that beta-actin and the ribosomal protein L12 mRNA were the two most stable control genes, with an average expression stability value of 0.43, followed by the Poly(a)-binding protein gene (Fig. 2; Table 1). The least stable gene within these six candidates was the putative senescence-associated protein gene which increased the average expression stability to 0.59.

3.3. Determination of the optimal number of control genes for q-RT-PCR normalization

In order to measure expression levels accurately in q-RT-PCR assays, normalization by multiple housekeeping genes are required, as discussed by Vandesompele et al. [14]. The number of genes used for calculating a geometric average normalization factor should be a trade-off between practical considerations and accuracy. It would be impractical to quantify the six candidate control genes to obtain a normalization factor when only a few target genes would need to be studied. Furthermore, it would be a waste of resources to quantify more housekeeping genes than necessary if all genes are relatively stably expressed and if the normalization factor does not significantly change whether or not more genes are included. It has been recommended that the minimal use of the three most stable internal control genes for calculation of an q-RT-PCR normalization factor, and stepwise inclusion of more control genes until the $(n+1)$th gene has no significant contribution to the newly calculated normalization factor [14].

From our analysis, the pairwise variation of the normalization factor for the three most stable (i.e. best performing) control genes was less than 0.15, which is already a very low pairwise variation, suggesting that these three most stable genes might be enough for calculating an adequate normalization factor (Fig. 3). However, we observed that the inclusion of the 4th ranked stable gene slightly reduced the pairwise variation of the new normalization

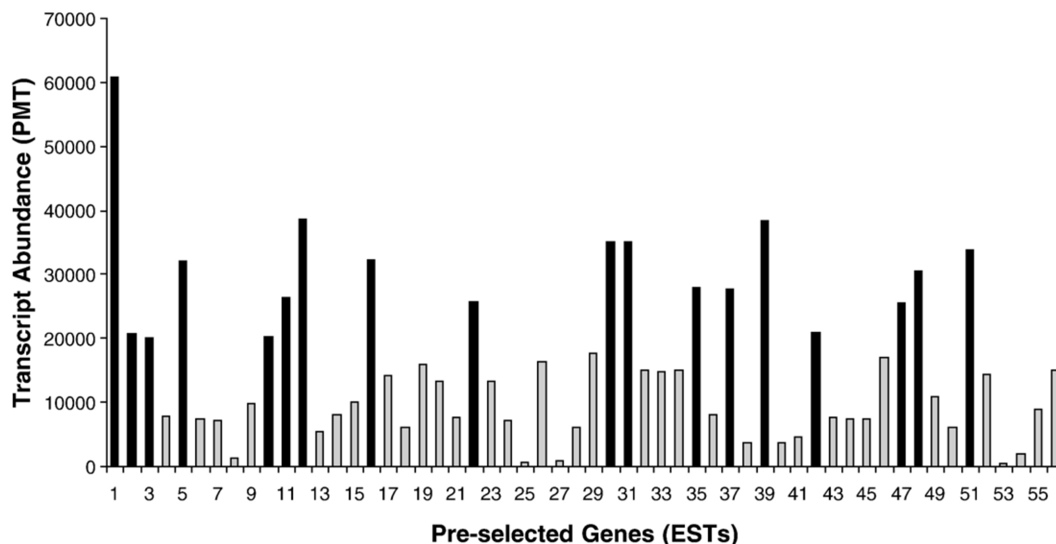

Fig. 1. Transcript abundance of pre-selected genes (ESTs) that did not show significant differences between treatments ($\alpha<0.05$), and which expression variation among all samples were less than 5%. The ESTs that show transcript abundance ($\geq$20,000 PMT values) [bars in black] were finally chosen as candidate housekeeping genes and sequenced.

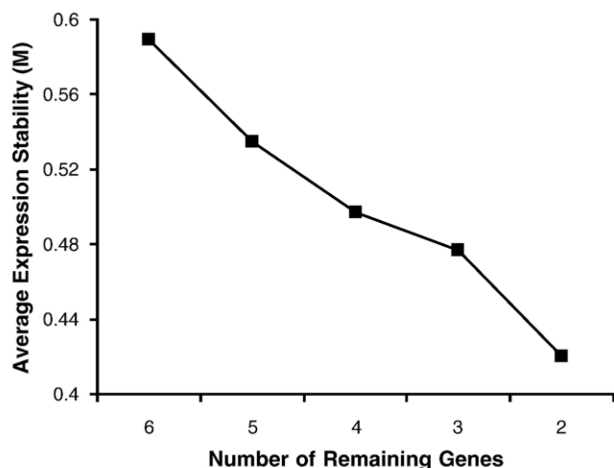

Fig. 2. Average variation of the expression stability value (M) for the six final candidate housekeeping genes during stepwise exclusion of the least stable control gene (from the left to the right on the X-axis).

factor to 0.11, advocating that this extra housekeeping gene might also be included in the calculation of the normalization factor. The inclusion of any further candidate housekeeping gene did not affect the pairwise variation of the subsequent normalization factors.

3.4. Testing housekeeping genes: comparison of microarray and q-RT-PCR expression data

In the previous section, we determined that the following four genes; beta-actin, Poly(a)-binding protein gene, the ribosomal protein L12 and L10 were the most stable housekeeping genes to be used for normalization of q-RT-PCR expression data (Table 1). However, we decided to exclude one of the two ribosomal protein genes from the normalization factor calculation as these two genes might be co-regulated since they belong to the same functional class. Thus, we omitted the ribosomal L10 as this was slightly less stable than the ribosomal L12 gene.

We then compared the expression of four other specific genes, which are particularly involved in symbiosis regulation [4], measured by q-RT-PCR assays and normalized with the geometric mean of the three best-performing housekeeping genes determined previously, with the expression of the same genes measured by the microarray assays. Although there was slight variation in the mean of gene expression, no significant differences ($p>0.1$; one-sample t-test) were detected between the q-RT-PCR and microarray data (Table 2). Moreover, the patterns of gene expression as a function of symbiosis were the same. These results validate the gene expression data for these four genes measured previously using microarrays, but never previously validated with other gene expression assays [4]. These results also indicate that the normalization of the q-RT-PCR data by our three best-performing housekeeping genes was adequate. It is important to mention that the gene expression as function of symbiosis for few other genes has also been validated using the stable HKGs reported here [4]. Rodriguez-Lanetty et al. [4] reported the senescence-associated protein (Sap) [6th stable HKG] was included in the normalization factor in that study, but that was a typo mistake as it was indeed the Poly(a) binding protein (Pabp) used instead along with Beta-actin and Ribosomal protein L12 (RibpL12).

4. Discussion

We have introduced here an effective and original analytical approach to identify candidate housekeeping genes from a microarray platform that can be used as internal control genes to normalize q-RT-PCR data. Based on q-RT-PCR data, we have then tested and shown that the identified HKGs were stable among the experimental conditions of our study. Furthermore, comparison of q-RT-PCR and microarray data showed that the final three best-performing HKGs identified from our method were adequate to be used as normalizing genes in gene expression assays.

4.1. The best-performing housekeeping genes

Based on the experimental conditions from our study, we determined the best-performing housekeeping genes in q-RT-PCR assays among the six candidates genes discovered from a microarray platform containing 10,368 ESTs. It is important, however, to mention that the rank based on performance of these HKGs could change depending on the experimental conditions of the samples. Thus we suggest that, if other researchers plan to use the HKGs described in this study on other cnidarian organisms, including anemones and corals, the stability and ranked performance should be established *a priori* from the new experimental conditions. The cellular function of the three best-performing (most stable) housekeeping genes is further discussed.

4.2. Beta-actin

Actin is a ubiquitous protein involved in the formation of filaments that are major components of the cytoskeleton. These filaments interact with myosin to produce a sliding effect, which is the basis of muscular contraction and many aspects of cell motility, including cytokinesis [18]. The three major groups of

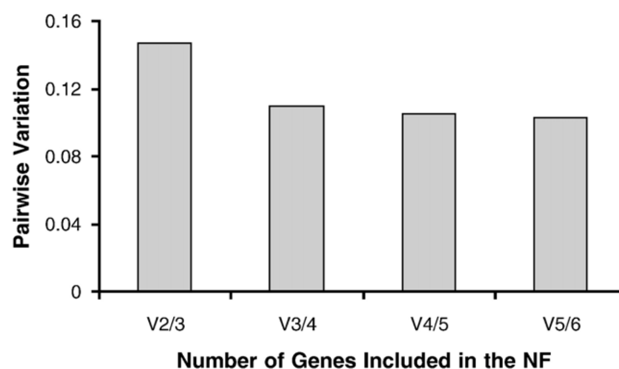

Fig. 3. Pairwise variation ($V_{n/n+1}$) analysis between the normalization factor NF_n and NF_{n+1} to determine the number of control genes required for accurate normalization. N= number of the gene, so gene 1 to 6 refers from the most stable to least ones (ranked genes from Table 1).

actin isoforms are the alpha, beta and gamma. The alpha actins are found in muscle tissues and are a major constituent of the contractile apparatus. The beta and gamma actins co-exist in most cell types as components of the cytoskeleton and as mediators of internal cell motility [18].

Over the last 8 years, beta-actin has become the most common internal control gene in gene expression studies; however many studies have recently shown that caution should be taken when using beta-actin or any other common internal control gene, such as GAPDH, as their gene expression can vary based on the experimental conditions [19]. This does not imply that beta-actin or any other previous HKG cannot be used at all as normalizing or internal control gene. But it is becoming widely accepted that HKGs need careful validation before their use in quantitative mRNA assays. For instance, Murphy et al. [20] have shown that beta-actin among a few other control genes may be used as HKGs to study gene expression in human muscle in experiments employing short-term creatine supplementation.

From our study, we have also shown that beta-actin is not only an adequate housekeeping gene, but the best-performing (most stable) gene along with the ribosomal protein L12 gene when used to address questions on cnidarian/dinoflagellate symbiosis by comparing aposymbiotic with symbiotic host anemones.

4.3. Ribosomal protein L12

From our study, ribosomal protein L12 was equally ranked with beta-actin as the best-performing (stable) gene among the tested housekeeping genes.

Ribosomal proteins are structural proteins that in conjunction with rRNA, make up the ribosomal subunits involved in the cellular process of DNA translation. The L12, along with L7 polypeptides, however, are among the few protein components of the translational machinery that is involved in translation factor binding and in the ribosomal GTPase center [21]. It has been well documented that ribosomes depleted of L7/L12 protein cores inefficiently support protein synthesis and exhibit a higher error frequency [22,23]. The proteins interact with initiation-, elongation-, and release factors, and are responsible for the factor-dependent GTP hydrolysis [24,25]. All these finding suggest that ribosomal protein L7/L12 exerts well-defined functional roles to keep the homeostasis of the DNA translational processes in the cell, and thus to be considered as a good candidate housekeeping gene, as we have shown in this study.

4.4. Poly(a) binding protein

In our study, the Poly(a) binding protein was the third best ranked housekeeping gene. Poly(a) binding proteins play a role in the synthesis and mediating the biological function of Poly(a) tails, which are tails permanently attached to almost all eukaryotic mRNAs [26]. Two types of conserved eukaryotic Poly(a) binding proteins have been described so far in detail, one being localized in the cytoplasm, which is present in all eukaryotes and plays important roles in at least two processes within nucleic acid metabolism, translation and mRNA decay; and another type localized in the nucleus, which was discovered through its role in polyadenylation of mRNA precursors, and is also ubiquitous in eukaryotic organisms [26]. This group of proteins has been implicated in global gene regulation, both stimulating translation initiation [27] and by enhancing mRNA stability and function [28]. Therefore they play a significant role in keeping steady the healthy levels of nucleic acid metabolism, which make them good candidate housekeeping genes to be used as control gene for normalization of gene expression data. We have indeed shown in this study that this is the case for a cnidarian organism, as the gene expression of one of these proteins show high stability among two different experimental settings.

5. Simplified description of the method and its (future) applications

The analytical method introduced in this study allows the identification of candidate housekeeping genes from a microarray platform which then can be used as internal control genes to normalize q-RT-PCR data. The approach encompasses three key steps for selecting candidate HKGs from a microarray platform; (1) the pre-selection of genes showing low gene expression variability among samples within and between treatments; (2) the test of whether the ratio of expression of these genes is not statistically different than 1 between treatments; and (3) The pre-selection of abundant transcripts. Finally, the gene-stability of the selected candidate housekeeping genes is tested in q-RT-PCR assays. The approach proposed in this study to identify HKGs can be applied to any comprehensive cDNA microarray platform generated from any biological system.

5.1. Applications for other cnidarian studies

Research in gene function using q-RT-PCR and microarray approaches are emerging and just about to explode in the field of coral and cnidarian biology. Because of the evolutionary importance of cnidarian organisms, as a basal metazoan, more attention has started to grow and more research in evolutionary developmental biology has began to be conducted in cnidarian species [29]. An entire genome has been already sequenced from a sea anemone (*Nematostella vectensis*) [30], and several ESTs project are in progress for different corals species [31,32]. All these genomics tools would provide molecular platforms to start studying the molecular basis of very important and significant biological phenomena, such as coral bleaching, the effect of ocean acidification on reef corals and host/symbiont regulation. However, in order to make sense with the gene expression data very important technological issues involving normalization in q-RT-PCR need careful attention.

Currently, there is no data available in regard what genes can be used as housekeeping genes to normalize gene expression data for cnidarian systems. We are proposing in this study several housekeeping genes that were successfully used to normalize gene expression data in cnidarian/algal symbiosis experiments. These genes may also be tested as normalizing genes in other coral/cnidarian gene expression studies. This is currently very feasible, as these genes have also already been

pulled out from the *Nemastostella* genome [31] and from the *Acropora millepora* [32] coral EST project. Furthermore, the gene expression of Beta-actin (our most stable housekeeping gene) has also been shown to be stable in the coral *A. millepora* under different thermal stresses (Seneca et al. in preparation). These findings are promising and we then encourage other researchers to test, under other different experimental scenarios of interest, not only the three best-performing HKGs showed here but as well the other three normalizing genes reported in this study.

Acknowledgements

We would like to thank to the members of Hoegh-Guldberg/Dove's lab in the University of Queensland for meaningful discussions and comments on early versions of this manuscript. We also are grateful for the comments and suggestions by two anonymous reviewers which have helped to improve this manuscript. This research was funded by the National Science Foundation (IBN: 0342585) to V. M. W; and by the Australian Research Council Centre of Excellence for Coral Reef Studies.

References

[1] Hofmann GE, Burnaford JL, Fielman KT. Genomics-fueled approaches to current challenges in marine ecology. TREE 2005;20:305–11.
[2] Edge SE, Morgan MB, Gleason DF, Snell TW. Development of a coral cDNA array to examine gene expression profiles in *Montastraea faveolata* exposed to environmental stress. Mar Pollut Bul 2005;51:507–23.
[3] Morgan MB, Edge SE, Snell TW. Profiling differential gene expression of corals along a transect of waters adjacent to the Bermuda municipal dump. Mar Pollut Bull 2005;51:524–33.
[4] Rodriguez-Lanetty M, Phillips WS, Weis VM. Transcriptome analysis of a cnidarian — dinoflagellate mutualism reveals complex modulation of host gene expression. BMC Genomics 2006;7:23.
[5] Bustin SA. Quantification of nucleic acids by PCR. In: Bustin SA, editor. A–Z of Quantification PCR. IUL Biotechnology Series, La Jolla; 2004. p. 1–29.
[6] Stahlberg A, Aman P, Ridell B, Mostad P, Kubista M. Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of kappa and lambda immunoglobulin light chain expression. Clin Chem 2003;49:51–9.
[7] Stahlberg A, Kubista M, Pfaffl M. Comparison of reverse transcriptases in gene expression analysis. Clin Chem 2004;50:1678–80.
[8] Stahlberg A, Hakansson J, Xian XJ, Semb H, Kubista M. Properties of the reverse transcription reaction in mRNA quantification. Clin Chem 2004;50:509–15.
[9] Pfaffl MW, Hageleit M. Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. Biotechnol Lett 2001;23:275–82.
[10] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.
[11] 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;75:291–5.
[12] Schmittgen TD, Zakrajsek BA. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. J Biochem Biophys Methods 2000;46:69–81.
[13] Schmid H, Cohen CD, Henger A, Irrgang S, Schlondorff D, Kretzler M. Validation of endogenous controls for gene expression analysis in microdissected human renal biopsies. Kidney Int 2003;64:356–60.
[14] 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;3:1–11.
[15] Yang YH, Dudoit S, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method adressing single and multiple slide systematic variation. Nucleic Acids Res 2002;30:1–e15.
[16] Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW. Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. Biotechnol Lett 2006;28:1601–13.
[17] Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 2003;339:62–6.
[18] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. Garland Science. New York: Taylor & Francis Group; 2002.
[19] Glare EM, Divjak M, Bailey MJ, Walters EH. beta-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalizing mRNA levels. Thorax 2002;57:765–70.
[20] Murphy RM, Watt KKO, Cameron-Smith D, Gibbons CJ, Snow RJ. Effects of creatine supplementation on housekeeping genes in human skeletal muscle using real-time RT-PCR. Physiol Genomics 2003;12:163–74.
[21] Markus CW, Gleb P, Bourenkov HD, Bartunik RH. Flexibility, conformational diversity and two dimerization modes in complexes of ribosomal protein L12. EMBO J 2000;19:174–86.
[22] Hamel E, Koka M, Nakamoto T. Requirement of an *Escherichia coli* 50S ribosomal protein component for effective interaction of the ribosome with T and G factors and with guanosine triphosphate. J Biol Chem 1972;247:805–14.
[23] Pettersson I, Kurland CG. Ribosomal protein L7/L12 is required for optimal translation. Proc Natl Acad Sci USA 1980;77:4007–10.
[24] Kischa K, Möller W, Stöffler G. Reconstitution of a GTPase activity by a 50S ribosomal protein from E. coli. Nat New Biol 1971;233:62–3.
[25] Möller W, Schrier PI, Maassen JA, Zantema A, Schop E, Reinalda H, Cremers AF, Mellema JE. Ribosomal proteins L7/L12 of *Escherichia coli*: localization and possible molecular mechanism in translation. J Mol Biol 1983;3:553–73.
[26] Kuhn U, Wahle E. Structure and function of poly(A) binding proteins. Biochim Biophys Acta 2004;1678:67–84.
[27] Sachs AB, Davis RW. The Poly(A) binding protein is required for Poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell 1989;58:857–67.
[28] Wang Z, Kiledjian M. The Poly(A)-binding protein and an MrNA stability protein jointly regulate an endoribonuclease activity. Mol Cell Biol 2000;20:6334–41.
[29] Miller DJ, Ball EE. The coral *Acropora*: what it can contribute to our knowledge of metazoan evolution and the evolution of developmental processes. BioEssays 2000;22:291–6.
[30] Sullivan JC, Ryan JF, Watson JA, Webb J, Mullikin JC, Rokhsar D, et al. StellaBase: The *Nematostella vectensis* Genomics Database. Nucleic Acids Res 2006;34:D495–9.
[31] Kortschak RD, Samuel G, Saint R, Miller DJ. EST analysis of the Cnidarian *Acropora* millepora reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr Biol 2003;13:2190–5.
[32] Schwarz JA, Brokstein P, Manohar C, Coffroth MA, Szmant A, Medina MA. Coral reef genomics: developing tools for functional genomics of coral symbiosis. Proc 10th Int Coral Reef Symp 2006;1:274–81.