Fish & Shellfish Immunology
Volume 29, Issue 6, December 2010, Pages 937-945
http://dx.doi.org/10.1016/j.fsi.2010.07.028

Identification of genes from flat oyster *Ostrea edulis* as suitable housekeeping genes for quantitative real time PCR

Benjamin Morga, Isabelle Arzul*, Nicole Faury, Tristan Renault

IFREMER, Laboratoire de Génétique et Pathologie (LGP); La Tremblade, France

Corresponding author. Mailing address: IFREMER, Laboratoire de Génétique et Pathologie, Avenue de Mus de Loup, 17390 La Tremblade, France. Phone: +33 5 46 76 26 10. Fax: +33 5 46 76 26 11. E-mail: iarzul@ifremer.fr

Abstract:

Bonamia ostreae is an intrahaemocytic protozoan affecting *Ostrea edulis*. The parasite multiplies within haemocytes without being degraded and involves changes in cellular activities. Studies aiming at better understanding host response to a pathogen at the transcriptome levels are frequently based on the use of real time PCR assays, which require some reference genes. However, very few sequence data is available for *Ostrea edulis* in public databases.
Subtracted cDNA libraries were constructed from the *Ostrea edulis* haemocytes in order to identify genes involved in host reactions against the parasite and quantitative real time PCR assays were developed to study expression of these genes. In this context, identification of reference genes and study of their relative expression stability were required for quantitative real time PCR normalization. The expression of 5 potential candidate reference genes from *Ostrea edulis* (ie elongation factor 1 alpha (EF1-α), 60S ribosomal protein L5 (L5), glyceraldehyde 3-phosphate-dehydrogenase (GAPDH), polyubiquitin (Ubiq) and β-actin (ACT)) was studied using RNAs extracted from pools of haemocytes in contact with the parasite *Bonamia ostreae* and haemocytes alone. Gene expression was quantified by real time PCR and expression stability was analysed with two analytical approaches *GeNorm* and *NormFinder*. GAPDH and EF1-α were identified as the most stable genes with the *GeNorm* analysis. Whatever were the tested conditions, EF1-α was also found as the most stable gene using *Normfinder*. The less stable gene was β-actin although this gene is commonly used as housekeeping gene in many studies.
Our results suggest using GAPDH and EF1-α combined as reference genes when studying expression levels in haemocytes of *Ostrea edulis*. In addition, the complete ORF of these two genes was characterized.

Keywords: Real time PCR, Housekeeping genes, Haemocytes, *Ostrea edulis*, *Bonamia ostreae*, EF1-α, GAPDH

1. Introduction

The parasite *Bonamia ostreae* was first reported in France in 1979 infecting the European flat oyster, *Ostrea edulis* [1, 2]. Despite 30 years of research on this protozoan pathogen, host parasite interactions are poorly known and few sequence data is available for *O. edulis* in public databases. A recent study has been conducted to investigate the transcriptome during an *in vitro* infection of *O. edulis* haemocytes with purified parasites. cDNA libraries allowed obtaining a large number of expressed sequence tags (ESTs). Based on sequence homologies, some ESTs were selected for expression assays and the quantification of transcription levels in haemocytes was undertaken using quantitative real time PCR.

Quantitative real time PCR is commonly used to study gene expression due to high sensitivity, specificity, reproducibility and broad dynamic range [3-5]. Level of gene expression is determined by quantifying the relative amount of target mRNA between different conditions, using the comparative Ct ($2^{-\Delta\Delta ct}$), or Pfaffl method including the adjustment by the amplification efficiencies [6]. The relative quantification method requires an internal control or a housekeeping gene. The internal control is an endogenous control that allows correction of experimental variations caused by the quantity and quality of starting material, inhibitory compounds, pipetting errors, biological variability, the reverse transcription efficiency and the experimental design [7]. Reference genes are thus compulsory but they must be carefully validated as pointed out in the MIQE guideline. This guideline suggests a minimum set of information that researcher should provide for their quantitative real time PCR data [8]. The ideal housekeeping gene should present a stable mRNA expression and should be minimally regulated under experimental conditions [9-12].

18S ribosomal RNA, glyceraldehyde-3 phosphate-dehydrogenase and β-actin have frequently been used as reference genes in bivalves [13-15]. These genes are involved in ribosome structure, metabolism and cytoskeleton respectively and their expression level is assumed to be constant. However, some studies have shown that expression of these genes can vary depending on development stages or in response to an external treatment [10, 16]. Relative quantification by real time PCR requires first a study of the suitability and stability of one or more housekeeping genes in selected experimental conditions [17]. Some mathematical methods have been developed to evaluate relative stability of gene expression in order to identify the most suitable housekeeping genes [18, 19]. Two Visual Basic Applications for Microsoft Excel are currently widely used to determine reference gene suitability: *GeNorm* [18] and *NormFinder* [19]. *GeNorm* is a mathematical method based on the principle that the expression ratio of two ideal control genes should be identical in all samples and experimental conditions. It calculates gene expression stability (M), which is the mean pair-wise variation between an individual gene and all other tested reference genes [18]. *NormFinder* is another Excel applet based on an algorithm for identifying the optimal normalization gene(s) among a set of candidates [19]. It ranks the candidate genes according to their mRNA expression stability value in a given sample set and a given experimental design. This approach combines the intra-group and inter-group expression variation to a stability value that enables the ranking of genes by mRNA expression stability.

In this context, the aim of the present study was to identify from ESTs database previously obtained, optimal reference genes that could be used for expression analysis in flat oyster haemocytes experimentally infected by the parasite *Bonamia ostreae*. The five tested reference genes (Table 1) were housekeeping genes commonly used in quantitative real time PCR studies: elongation factor 1 alpha (EF1-α), 60S ribosomal protein L5 (L5), glyceraldehyde 3-phosphate-dehydrogenase (GAPDH), polyubiquitin (Ubiq) and the β-actin (ACT) [10, 16, 20]. In addition, the ORF of EF1-α and GAPDH, identified as the most suitable housekeeping genes in our study, were fully characterized.

2. Material and methods

2.1. Biological material

2.1.1. Oysters

Two-year-old flat oysters *Ostrea edulis* were collected from Quiberon Bay (France) in November 2008 and acclimatized in the quarantine facilities of IFREMER laboratory in La Tremblade (Charente–Maritime, France) over 30 days. Flat oysters were maintained in 120 l raceways supplied with a constant flow of seawater enriched in phytoplankton (*Skeletonema costatum, Isochrisis galbana, Chaetoceros gracialis* and *Tetraselmis succica*).

2.1.2. Haemolymph collection

Haemolymph was collected from the adductor muscle sinus of 40 oysters using a 1 ml syringe equipped with a needle (0.40 mm x 90 mm). To eliminate cell aggregates, the haemolymph samples were filtered through 60 µm nylon mesh and held on ice to prevent cell aggregation. Approximately 1 ml of haemolymph was collected from each oyster. Samples were pooled, and haemocytes counted using a Malassez-cell. The haemocyte concentration was adjusted to 1.10^6 cells ml^{-1} using 0.22 µm filtered seawater (FSW).

2.2. In vitro challenge of Ostrea edulis haemocytes with Bonamia ostreae

Haemocytes (1.10^6 cells ml-1) were incubated with live purified parasites at a ratio of five parasites per one haemocyte for two hours at 15°C. *Bonamia ostreae* was purified according to the protocol developed by Mialhe et al., (1988) [21]. The control consisted of haemocytes alone suspended in FSW (1.10^6 cells ml^{-1}). Four challenges were performed using four different pools of haemolymph, each pool of haemolymph was divided in two parts, one part for the control and a second part for in vitro infection. Two hours post challenge, haemocytes were recovered by centrifugation (1500 rcf, 4°C for 15 min).

2.3. Total RNA extraction and cDNA synthesis

Total RNA was extracted using TRIZOL reagent (Invitrogen) following the manufacturer's instructions. Briefly, cells were lysed in TRIZOL, RNA was separated from DNA and proteins using phenolchloroform. Total RNAs were precipitated using isopropanol. The RNA pellet was washed with 75% ethanol solution and finally eluted in 22 µl. Total RNAs were treated with RQ1 RNAse-free DNAse (Promega) to remove remaining genomic DNA. The RNA concentration was determined using a spectrophotometer at 260 nm and RNA quality was assessed on a 1% agarose gel.
First strand cDNA was synthesized using the oligo (dT) anchor primer (5'-GAC CACGCGTATCGATGTCGACT(16)V-3') and Moloney murine leukaemia virus (M-MLV) reverse transcriptase SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen).

2.4. Real time quantitative PCR

ESTs of five candidate housekeeping genes were obtained from a subtracted cDNA library from *Ostrea edulis*. For each of these genes forward and reverse primers were designed using primer 3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (Table 2) and synthesized by Eurogentec. Real time quantitative PCR was carried out in duplicate, all amplification reactions were performed in a total volume of 25 µL using a Mx3000 Thermocycler sequence detector (Stratagene) in 96-microwell plates. Each well (25 µL) contained 5 µL of cDNA dilution (1/30), 12.5 µL of Brilliant® SYBR® Green II PCR Master Mix (Stratagene), 2.5 µL of each diluted primer (3 µM) and 2.5 µL of distilled water. Thermal cycling conditions were: 1 cycle of activation of Hot start Taq polymerase at 95 °C for 10 min; 40 cycles of amplification at 95°C for 10 s, 60°C for 30s; and melting temperature curve analysis at 95°C to 60°C by 0.5°C decrease of the temperature every 10s. In all cases, negative controls (without cDNA) were included to rule out DNA contamination.
For each candidate housekeeping gene, melting curve, gel picture and sequences were analysed in order to verify the specificity of the amplified products and that at a single PCR product was amplified. Efficacy (E) was calculated for each primer pair by determining the slopes of standard curves. Efficacies were calculated according to the following equation $E=10^{(1-slope)}$ [6].

2.5. Rapid amplification of cDNA ends and sequence analysis

In order to obtain the complete open reading frame (ORF) of EF1-α and GAPDH, RACE (Rapid Amplification of cDNA ends) reactions were carried out using SMART RACE cDNA Amplification Kit from Clontech according to the manufacturer's instructions. 5' and 3' primers were designed using primer 3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (3'elongrace GTAATCAAGAGCGGAGATGC, 5'elongrace GTCCAATAAATGACCTGCTGTGGTC and 3'GAPDHrace GCAATGCTTCGTGCACCACTAACTGT, 5'GAPDHrace GCGGTGTATGCATGAACCGTTGTC) and synthesized by Eurogentec. After ligation and cloning in

TOPO Vector System (Invitrogen) and transformation in Top 10 competent bacteria (Invitrogen), several clones were sequenced from both ends with TOPO forward and reverse primers using BigDye terminator Cycle Sequencing Ready Reaction Kit and an automated DNA sequencer ABI 3130 XL. Chromatograms were analyzed with Chromas 231 software (Technelysium). NCBI-BLAST (http://www.ncbi.nlm.nih.gov/blast/) was used for detection of known orthologs and for comparison of the obtained cDNAs and deduced polypeptides.

2.6. Phylogenic analysis of EF1-α and GAPDH

The sequence of the complete ORF of EF1-α was aligned with homologous genes from GenBank database (*Mus musculus* (L26479.1), *Homo sapiens* (AY043301.1), *Bos taurus* (XM001254483.2), *Salmo salar* (BT059777.1), *Solea senegalensis* (AB326302.1), *Sparus aurata* (AF184170), *Mytilus galloprovincialis* (AB162021.1), *Crassostrea gigas* (AB122066.1), *Strongylocentrotus purpuratus* (XM 0011764), *Danio rerio* (NM 131263.1), *Argopecten purpuratus* (ES469321), *Pecten maximus* (DN794050) and *Chlamys farreri* (DT716075).

The sequence of the complete ORF of GAPDH was aligned with homologous genes from GenBank database (*Homo sapiens* (NP002037.2), *Mus musculus* (NP032110.1), *Bos taurus* (NP001029206.1), *Sparus aurata* (ABG23666.1), *Xenopus lavevis* (NP001080567.1), *Crassostrea gigas* (ABU84874.1), *Crassostrea virginica* (CAD677717.1) and *Tribolium castaneum* (XP974943.2).

Alignments were performed using the Clustal W [22] including in MEGA 4 [23]. Phylogenetic trees based on amino acid sequences were performed using the Neighbour-Joining [24] algorithm with the MEGA 4 software program. Statistical confidence on the inferred phylogenetic relationships was assessed by bootstrap of 1000 replicates.

2.7. Statistical analysis

Cts were calculated with the stratagene Mxpro software 4.0. Data obtained for each experiment was analyzed using two excel applets, *GeNorm* [18] and *NormFinder* [19].
GeNorm was used to calculate the gene expression normalization factor and to determine the most stable internal controls. Determining the average pairwise variation of a particular gene with all other genes assessed the gene stability measurement M. Moreover, the optimal number of housekeeping genes required for normalization was determined using *GeNorm*. Inter-group and intra-group variations were determined using *NormFinder* and allowed ranking and estimating reference gene stability.

3. Results

3.1. Real time PCR amplification of candidate housekeeping genes

The relative expression level of five candidate housekeeping genes (L5, Ubiq, ACT, GAPDH and EF1-α) was studied during four different *in vitro* infections of haemocytes with the parasite *Bonamia ostreae*. Ct values were distributed between 16.1 (±1.5) for ACT and 21 (±1.2) for GAPDH. Ct value between replicates was less than one. Standard curves exhibited correlation coefficient (R^2) higher than 0.99 and PCR efficiencies ranged from 95% to 100% (Table 3). The amplification specificity was confirmed by melting curve analysis (Fig 1). The melting curves for all genes demonstrated a single peak confirming gene-specific amplification. Agarose gel analyses showed a single band for all amplified genes (data not shown).

3.2. Expression stability of candidate housekeeping genes

Expression stability (M) of each candidate housekeeping gene through different tested RNA samples was calculated using the *GeNorm* program. A low M value translates a highly stable expression in the tested conditions. The rank of the candidate housekeeping genes based on their M values is shown in Fig 2a. ACT (M=0.516) Ubiq (M=0.469) and L5 (M=0.451) appeared less stable while, GAPDH (M=0.187) and EF1-α (M=0.187) were the most stable genes. The best combination of two genes appeared as GAPDH and EF1-α (M=0.187).
Gene expression stability was also evaluated using another software, *NormFinder*. Stability values for the tested genes from the least to the most stable were: ACT (M=0.130), Ubiq (M=0.119), L5

(M=0.112), GAPDH (M=0.049) and EF1-α (M=0.046) (Fig 2b). The best combination for the most stable was GAPDH and EF1-α (0.034).

3.3. Optimum number of housekeeping genes

In addition, *GeNorm* method can determine the optimum number of housekeeping genes recommended in an experiment using a pairwise number variation Vn/n+1 analysis (Fig 3), Vandesompele et al., (2002) [18] proposed 0.15 as a cut-off value, below which the inclusion of an additional reference gene is not required. In our experiment the $V_{2/3}$ and $V_{3/4}$ values were 0.192 and 0.112, respectively suggesting that the normalization factor should preferably contain 3 housekeeping genes.

3.4. Characterization of Ostrea edulis EF1-α

ESTs with high similarity to EF1-α have been identified in subtracted cDNA libraries. RACE reactions allowed obtaining the full cDNA of *Ostrea edulis* EF1-α. The length of the EF1-α was 1698 nucleotides (nt). Complete sequence was deposited in GenBank and assigned under the accession number EU651798. The sequence contained a short 5'-untranslated region (1-57 nt) followed by the ORF of 1385 nt (461 codons: 58-1443 nt) and finished by an 3'-untranslated region (1444-1698 nt) (Fig 4). The ORF included three domains: elongation factor I 82-768 nt, elongation factor II 778-1050 nt and elongation factor III 1057-1368 nt (Fig 5).
The ORF of the EF1-α presented characteristic regions G1 to G4 critical in GDP/GTP exchange, GTP-induced conformational change and GTP hydrolysis (Fig 5). The amino acid sequence was submitted in PROSITE database (http://www.expasy.ch/prosite/) and a conserved motif DKLKAERERGITIDIA was identified at position 61 to 76. This motif is a signature of the GTP-binding elongation factor (Fig 5).
The nucleotide sequence of EF1-α from *Ostrea edulis* presented 73% to 86% of identity with other species. *Crassostrea gigas* presented the highest percentage of identity (86%).
The alignment of the selected amino acid EF1-α sequences and phylogenetic tree showed that *Ostrea edulis* EF1-α clustered with other invertebrate EF1-α (Fig 6).

3.5. Characterization of Ostrea edulis GAPDH

Ostrea edulis cDNA GAPDH showed a 1309 nucleotide (nt) sequence. Complete sequence was deposited in GenBank and assigned under the accession number GQ150762. The sequence contained a short 5'-untranslated region (1-70 nt) followed by an ORF of 1008 nt (335 codons: 69-1077 nt) and finished by a 3'-untranslated region (1078-1309 nt) (Fig 7). The complete ORF had a putative molecular mass at 36 143.34 Da and a predicted isoelectric point at 7.58. The analysis of the ORF revealed the presence of two domains, NAD binding domain (3-150 aa) and C terminal domain (155-312) (Fig. 7). ORF of GAPDH had a typical eukaryotic GAPDH signature (ASCTTNCL; positions from 148 to 155) [25].
The nucleotide sequence of GAPDH from *Ostrea edulis* presented 74% to 91% of identity with other species. *Crassostrea gigas* presented the highest percentage of identity (91%).
The phylogenic tree based on amino acid sequences of GAPDH from different species showed that *Ostrea edulis* GAPDH belong to the "mollusc cluster" with *Crassotrea gigas* and *Crassostrea virginica* (Fig 8).

4. Discussion

The quantitative real time PCR is a powerful tool to measure gene expression [6]. This method requires an internal control (ie reference gene or housekeeping gene) in order to normalize gene expression and to integrate experimental variations. The expression of housekeeping genes must be stable during the experiment. Thus, studying the expression stability of housekeeping genes is recommended before quantifying genes of interest by quantitative real time PCR [8, 17].
In our study, five candidate housekeeping genes identified from subtracted libraries have been selected and expression monitored by real-time quantitative PCR during *in vitro* infections of *Ostrea edulis* haemocytes with the parasite *Bonamia ostreae* to check their stability. Selected housekeeping

genes are involved in ribosomal metabolism (L5), cytoskeleton structure (ACT), protein metabolism (Ubiq and GAPDH) and elongation process (EF1-α).
Two different analytical approaches based on *GeNorm* and *NormFinder* algorithms were used to identify the most suitable housekeeping genes. Their expression was compared between exposed and non-exposed haemocytes during an *in vitro* infection of *Ostrea edulis* cells with the parasite *Bonamia ostreae*. *GeNorm* and *NormFinder* allowed ranking candidate housekeeping genes according to their expression stability.
Both approaches gave concordant results. EF1-α and GAPDH genes appeared as the most suitable housekeeping genes in the present experiment. Both genes are widely used as housekeeping genes. EF1-α is a member of the G protein family, and one of the four subunits that compose the eukaryotic elongation factor 1 [26, 27]. EF1-α plays a key role in protein translation. Araya et al., (2008) [10] demonstrated that EF-1α is a suitable reference gene during an *in vitro* challenge of *Mya arenaria* haemocytes with *Vibrio splendidus*. In Atlantic salmon, EF1-α was identified as the most stable gene in both healthy and virus infected tissues [28]. This gene was also reported as the best housekeeping gene during the larval development of the flatfish *Solea senegalensis* [29] and during the development of the European sea bass *Dicentrarchus labrax* [30].
The GAPDH gene is frequently used as a housekeeping gene in human [31] or in plants [32]. Although the expression of this gene appeared stable in our experiment, it was shown unsuitable in other studies because of its involvement in different functions. GAPDH is implicated in nuclear RNA export, DNA replication, DNA repair, exocytotic membrane fusion, cytoskeleton organisation and phosphotransferase activity [33]. Moreover, Cho et al., (2008) [34] showed that the expression of 2 GAPDH isoforms from the marine teleost fish *Oplegnathus fasciatus* was modulated in response to bacterial and viral infections.
The less stable genes identified in the present study were Ubiq and ACT.
Ubiq has previously been used as housekeeping gene in the study of *Mya arenaria* responses to infection with *Vibrio splendidus* [10]. However, in our study, the expression of this gene was found unstable between the tested conditions. This instability could be explained by the involvement of polyubiquin in protein degradation through the proteasome. Although β-actin is commonly used as housekeeping genes in quantitative real time PCR assays, our results suggest that the parasite *Bonamia ostreae* affects the expression of ACT. β-actin is involved in the cytoskeleton structure which plays a pivotal role in the phagocytosis and encapsulation. Intracellular parasites such as apicomplexan parasites require polymerisation of host actin in order to penetrate into host cells [35]. A similar mechanism might be developed by the parasite *Bonamia ostreae*. Chagot et al., 1992 [36] demonstrated that parasites were not engulfed when haemocytes were treated with cytochalasin B prior parasite contact. These results support an active participation of host cells in the internalisation of the parasite.

The *GeNorm* analysis also defined the optimal number of housekeeping genes in the present study as 3. Vandesompele et al., (2002) [18] suggested that in most cases three housekeeping genes is the optimal number for a valid normalization strategy.
The molecular characterization of the EF1-α and GAPDH in *Ostrea edulis* allowed studying the sequence of the gene and identifying conserved domains and signatures present in other species. The phylogenetic analysis assigned the flat oyster EF1-α and GAPDH inside the mollusc clusters.

This study is the first, to our knowledge, that describes the evaluation of flat oyster genes as housekeeping genes for real time quantitative PCR. The most appropriate housekeeping gene in flat oyster haemocytes is the EF1-α. In this respect it is now possible to evaluate the level of expression of genes of interest in order to improve the knowledge on interactions between the flat oyster and the parasite *Bonamia ostreae*. This study has also shown the importance to determine statistically the stability of references genes in every experiment because of the impact of the experimental conditions used on gene expression.

5. Acknowledgement

The authors are very grateful to anonymous reviewers for revising the manuscript. The Région of Poitou Charentes supported in part this research. The authors thank Raphael Brizard for technical assistance for the maintenance of oysters in IFREMER La Tremblade quarantine facilities.

6. References

[1] Pichot Y, Comps M, Tige G, Grizel H, Rabouin MA. Research on *Bonamia ostreae* gen. n., sp. n., a new parasite of the flat oyster *Ostrea edulis* (L.) Rev Trav Inst Pêches Marit 1979; 43: 131-140.

[2] Comps M, Tige G, Grizel H. Etude ultrastructurale d'un protiste parasite de l'huître *Ostrea edulis* (L.). *C. R*. Acad Science Paris 1980 ; 290:383-385.

[3] Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K et al. The real-time polymerase chain reaction. Mol Aspects Med 2006; 27:95-125.

[4] Bustin SA, Benes V, Nolan T, Pfaffl MW. Quantitative real-time RT-PCR-a perspective. J Mol Endocrinol 2005; 34:597-601.

[5] Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse transcription polymerase chain reaction. J Biomol Tech 2004; 15:155-166.

[6] Pfall MW. A new mathematical model for relative quantificationin real-time RT-PCR. Nucleic Acid Res 2001; 29: 2002-2007.

[7] Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. Biotechniques 2004; 37:112-119.

[8] Bustin SA. Why the need for qPCR publication guidelines? -The case for MIQE. Method 2010; 50: 217-226.

[9] Nicot N, Hausman JF, Hoffmann L, Evers D. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot 2005; 56: 2907–2914.

[10] Araya M.T, Siah A, Mateo D, Markham F, McKenna P, Johnson G et al. Selection and evaluation of housekeeping genes for haemocytes of soft-shell clams *(Mya arenaria)* challenged with *Vibrio splendidus*. J Invertebr Pathol 2008; 99: 326–331.

[11] De Boer ME, De Boer T , Mariën J, Timmermans MJ, Nota B, Van Straalen NM et al. Reference genes for QRT-PCR tested under various stress conditions in *Folsomia candida* and *Orchesella cincta* (Insecta, Collembola). BMC Mol Biol 2009; 10:54.

[12] Paolacci AR, Tanzarella OA, Porceddu E and Ciaffi M. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol Biol 2009; 10:11.

[13] Huvet A, Daniela JY, Quéré C, Dubois S, Prudence M, Van Wormhoudt et al. Tissue expression of two a-amylase genes in the Pacific oyster *Crassostrea gigas*. Effects of two different food rations. Aquaculture 2003; 228:321–333.

[14] Veldhoen N, Loweb CJ, Davis C, Mazumder A, Helbing CC. Gene expression profiling in the deep water horse mussel *Modiolus modiolus* (L.) located near a marine municipal wastewater outfall. Aquatic Toxicology 2009; 93:116–124.

[15] Dondero F, Dagnino A, Jonsson H, Capri F, Gastaldi L, Viarengo A. Assessing the occurrence of a stress syndrome in mussels (*Mytilus edulis*) using a combined biomarker/gene expression approach. Aquatic Toxicology 2006; 78S:S13–S24.

[16] Olsvik PA, Lie KK, Jordal AEO, Nilsen TO, Hordvik I. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. BMC Mol Biol 2005; 6:21.

[17] Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA et al. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. Anal Biochem 2005; 344:141-143.

[18] 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:RESEARCH0034.

[19] Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004; 64:5245-5250.

[20] Siah A, Dohoo C, Mckenna P, Delaporte M, Berthe, FCJ. Selecting a set of housekeeping genes for quantitative real-time PCR in normal and tetraploid haemocytes of soft-shell clams, *Mya arenaria*. Fish Shellfish Immunol 2008; 25:202–207.

[21] Mialhe E, Bachere E, Chagot D, Grizel H. Isolation and purification of the protozoan *Bonamia ostreae* (Pichot et *al.* 1980), a parasite affecting the flat oyster *Ostrea edulis* (L.). Aquaculture 1988; 71:293-299.

[22] Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality tools. Nucleic Acids Res 1997; 24:4876-4882.

[23] Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 2007; 24:1596-1599.

[24] Saitou N & Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 1987; 4:406-425.

[25] Sirover MA. New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. Biochim Biochys Acta 1999; 1432:159–84.

[26] Browne GJ, Proud CG. Regulation of peptide-chain elongation in mammalian cells. Eur J Biochem 2002; 269:5360-5368.

[27] Ejiri S. Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization. Biosci Biotechnol Biochem 2002; 66:1-21.

[28] Jorgensen SM, Kleveland EJ, Grimholt U, Gjoen, T. Validation of reference genes for real-time polymerase chain reaction studies in Atlantic salmon. Mar. Biotechnol 2006; 8: 398–408.

[29] Infante C, Matsuoka MP, Asensio E, Cañavate JP, Reith M, Manchado M. Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. BMC Mol Biol 2008; 9:28.

[30] Mitter K, Kotoulas G, Magoulas A, Mulero V, Sepulcre P, Figueras A, Novoa B, Sarropoulou E. Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (Dicentrarchus labrax). Comp Biochem Physiol B Biochem Mol Biol. 2009; 4: 340-347

[31] Barber RD, Harmer DW, Coleman RA, and Clark BJ. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiol Genomics 2005; 21: 389–395.

[32] Barsalobres-Cavallari CF, Severino FE, Maluf MP and Maia IG. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. BMC Mol Biol 2009; 10:1

[33] Bustin, SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR) trends and problems. J Mol Endocrinol 2002; 29: 23–39.

[34] Cho YS, Lee SY, Kim KH, Nam YK. Differential modulations of two glyceraldehydes 3-phosphate dehydrogenase mRNAs in response to bacterial and viral challenges in a marine teleost *Oplegnathus fasciatus* (Perciformes). Fish Shellfish Immunol 2008; 25: 472–476.

[35] Gonzalez V, Combe A, David V, Malmquist NA, Delorme V, Leroy C et al. Host cell entry by apicomplexa parasites requires actin polymerisation in the host cell. Cell Host & Microbe 2009; 5:259-272.

[36] Chagot D, Boulo V, Hervio D, Mialhe E, Bachere E, Mourton C et al. Interactions between *Bonamia ostreae* (Protozoa: Ascetospora) and hemocytes of *Ostrea edulis* and *Crassostrea gigas (*Mollusca: Bivalvia): Entry mechanisms. J Invertebr Pathol 1992; 5: 241-249.

Figure captions

Figure 1 Dissociation curves of (A) ACT, (B) EF1-α, (C) GAPDH, (D) L5 and (E) Ubiq.

Figure 2 Determination of the most stable gene during an *in-vitro* infection of haemocytes from flat oyster *Ostrea edulis* with parasites *Bonamia ostreae* using the *GeNorm* and *NormFinder* methods. (a) Average expression stability value M of the five candidate reference genes using *GeNorm*. (b) Stability index of the five candidate reference genes using *NormFinder*.

Figure 3 Determination of the optimal number of housekeeping genes required for accurate normalization based on pairwise variation (V_n/V_{n+1}) between candidate genes using *GeNorm* analysis.

Figure 4 Complete ORF nucleotide and deduced amino acid sequences of the flat oyster EF1-α. ORF contains three domains: elongation factor 1 (82-768 nt) in black, elongation factor II (778-1050 nt) in grey and elongation factor III (1057-1368 nt) in white. Start and stop codons are highlighted in grey.

Figure 5 Comparison of the EF1-α amino acid sequences of *Ostrea edulis* and three other species (*Crassostrea gigas*, *Mytilus galloprovincialis* and *Homo sapiens*). Dots represent identity of amino acid between species and dashes represent gaps. G1 to G4 indicate the critical regions involved in GDP/GTP exchange and GTP hydrolysis. The consensus sequence composed of the three consensus elements GXXXXGK (G^{14}-K^{20}), DXXG (D^{91}-G^{94}), and NKXD (N^{153}-D^{156}) present in the GTP-binding domain is shaded in grey. The GTP-binding elongation factor signature corresponding to amino acids 61 to 76 is boxed.

Figure 6 Un-rooted phylogenetic tree on the selected EF1-α amino acid sequences. The tree was constructed using the neighbour-joining algorithm in the Mega 4.0 program. Bootstrap values (shown) are based on 1 000 resamplings of the data. The scale for branch length (0.05 substitutions/site) is shown below the tree.

Figure 7 Comparison of the GAPDH amino acid sequences of *Ostrea edulis* and three other species (*Crassostrea gigas*, *Sparus aurata* and *Homo sapiens*). Dots represent identity of amino acid between species and dashes represent gaps. The NAD binding domain (3-150 aa) is underlined in black and the C terminal domain (155-312) is underlined in grey. The typical eukaryotic GAPDH signature (ASCTTNCL) corresponding to amino acids 148 to 155 is boxed.

Figure 8 Un-rooted phylogenetic tree on the selected GAPDH amino acid sequences. The tree was constructed using the neighbour-joining algorithm in the Mega 4.0 program. Bootstrap values (shown) are based on 1 000 resamplings of the data. The scale for branch length (0.05 substitutions/site) is shown below the tree.

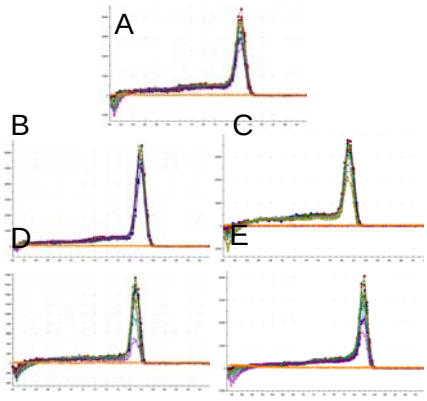

Figure 1 Dissociation curves of (A) ACT, (B) EF1-α, (C) GAPDH, (D) L5 and (E) Ubiq.

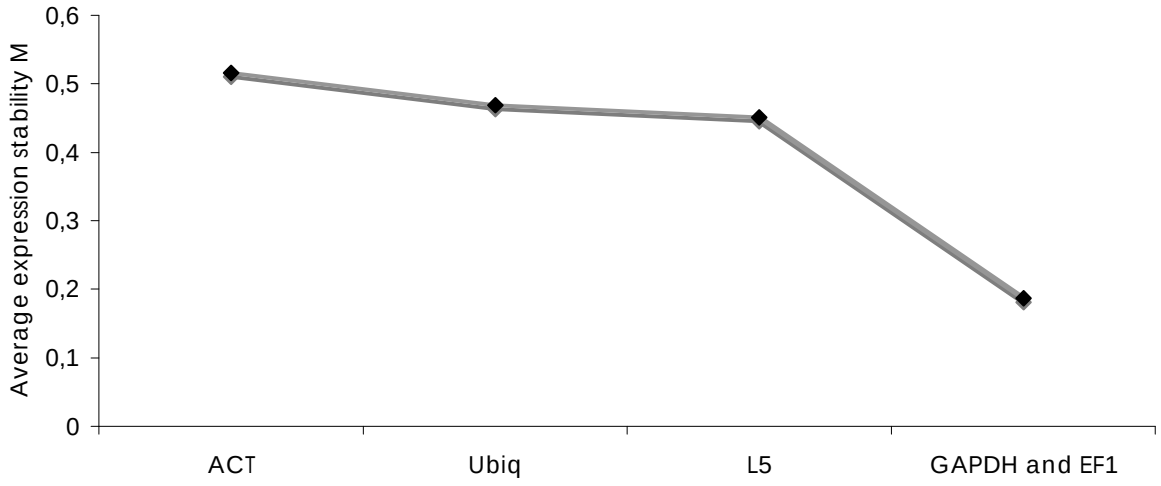

Fig. 2a

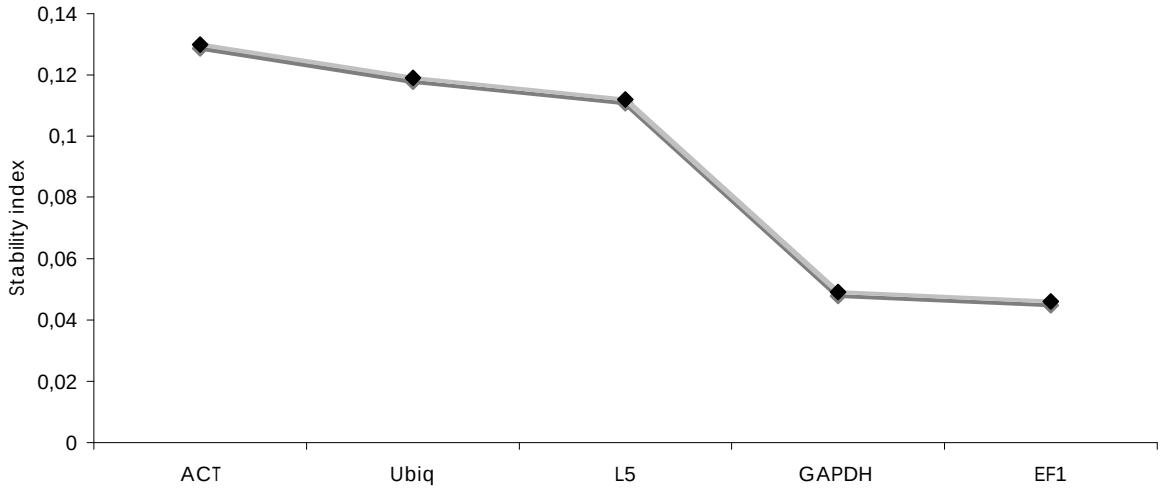

Fig. 2b

Figure 2 Determination of the most stable gene during an *in-vitro* infection of haemocytes from flat oyster *Ostrea edulis* with parasites *Bonamia ostreae* using the *GeNorm* and *NormFinder* methods. (a) Average expression stability value M of the five candidate reference genes using *GeNorm*. (b) Stability index of the five candidate reference genes using *NormFinder*.

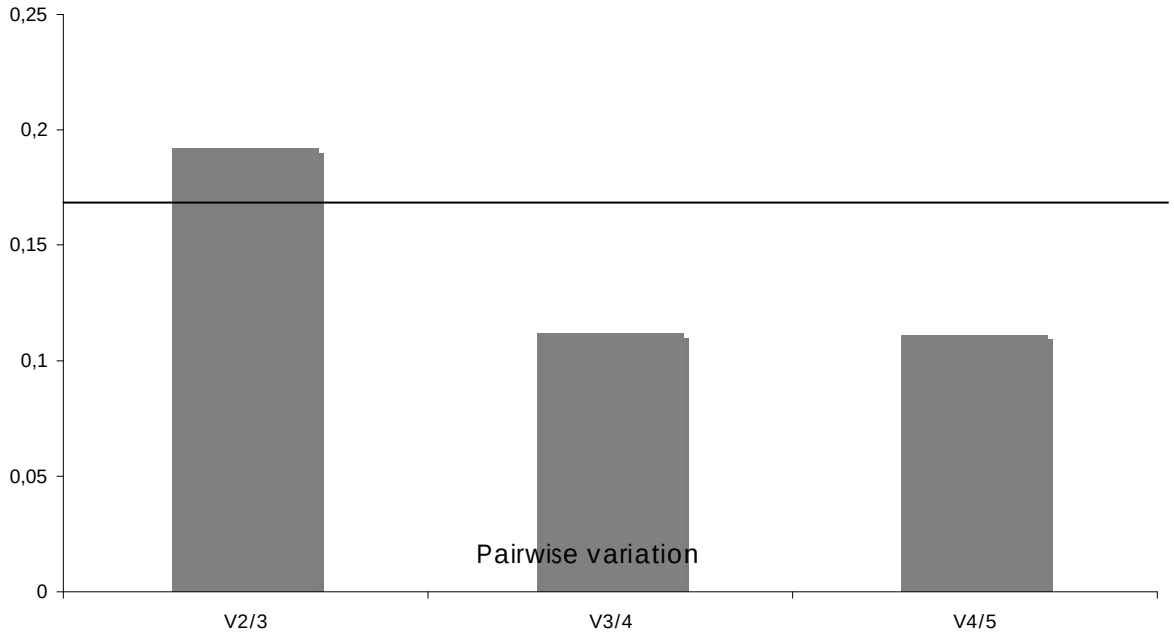

Figure 3 Determination of the optimal number of housekeeping genes required for accurate normalization based on pairwise variation (V_n/V_{n+1}) between candidate genes using *GeNorm* analysis.

```
1     GCC TTT TTG CTG TGG CTG ACT CTG GGA TAT CAA GAG TAT TAA ATA   45
1                                                                  15

46    CTA GGA GCA ACG ATG GTT AAA GAA AAG ATT CAC ATC AAC ATT GTC   90
16                    M   V   K   E   K   I   H   I   N   I   V   30

91    GTC ATT GGT CAT GTC GAC TCT GGA AAA TCG ACC ACT ACT GGA CAT   135
31     V   I   G   H   V   D   S   G   K   S   T   T   T   G   H   45

136   CTT ATC TAC AAA TGT GGT GGA ATT GAC GAG AGA ACC ATT GCC AAG   180
46     L   I   Y   K   C   G   G   I   D   E   R   T   I   A   K   60

181   TTT GAA AAG GAA GCT GCT GAG ATG GGC AAG GGT TCC TTC AAG TAC   225
61     F   E   K   E   A   A   E   M   G   K   G   S   F   K   Y   75

226   GCC TGG GTG TTG GAC AAA CTG AAG GCT GAG CGT GAA CGT GGA ATC   270
76     A   W   V   L   D   K   L   K   A   E   R   E   R   G   I   90

271   ACC ATT GAC ATT GCC CTC TGG AAG TTT GAG ACG GCA AAG TAC CAC   315
91     T   I   D   I   A   L   W   K   F   E   T   A   K   Y   H   105

316   ATC ACA ATT ATT GAT GCT CCA GGC CAC AGA GAT TTC ATC AAA AAC   360
106    I   T   I   I   D   A   P   G   H   R   D   F   I   K   N   120

361   ATG ATT ACA GGA ACA TCT CAG GCT GAT TGT GCT GTG TTG ATC GTA   405
121    M   I   T   G   T   S   Q   A   D   C   A   V   L   I   V   135

406   GCA GCA GGT GTT GGG GAA TTT GAA GCA GGT ATC TCC GCC AAT GGA   450
136    A   A   G   V   G   E   F   E   A   G   I   S   A   N   G   150

451   CAA ACT CGT GAG CAC GCT CTA TTG GCT TTC ACC TTG GGA GTC AAA   495
151    Q   T   R   E   H   A   L   L   A   F   T   L   G   V   K   165

496   CAG CTC ATT GTC GGA GTC AAC AAA ATG GAC AGC ACG GAA AAA CCA   540
166    Q   L   I   V   G   V   N   K   M   D   S   T   E   K   P   180

541   TAC AGT GAG ACA CGT TTT GAG AAT ATC AAA GGG GAG GTG GAG AAG   585
181    Y   S   E   T   R   F   E   N   I   K   G   E   V   E   K   195

586   TAC ATC AAG AAG ATT GGA TAT AAT CCA AAG ACT GTT GCC TTT GTA   630
196    Y   I   K   K   I   G   Y   N   P   K   T   V   A   F   V   210

631   CCT ATC TCT GGT TGG CAT GGT GAC AAC ATG ATT GAA CAG TCA AAA   675
211    P   I   S   G   W   H   G   D   N   M   I   E   Q   S   K   225

676   AAT ATG TCT TGG TTC AGG GGT TGG AAT GTG GAG AGG AAA GAA GGA   720
226    N   M   S   W   F   R   G   W   N   V   E   R   K   E   G   240

721   AAT GCC AGC GGA TTT ACT CTC TTG CAA GCT TTG GAT TCC ATC CTG   765
241    N   A   S   G   F   T   L   L   Q   A   L   D   S   I   L   255

766   CCA CCA AAG AGA CCC ACA GAT TTG GCC CTT CGT CTT CCT CTG CAG   810
256    P   P   K   R   P   T   D   L   A   L   R   L   P   L   Q   270

811   GAT GTA TAC AAA ATT GGA GGT ATT GGA ACT GTG CCA GTA GGA AGA   855
271    D   V   Y   K   I   G   G   I   G   T   V   P   V   G   R   285

856   GTT GAA ACT GGT ATT ATG AAG CCT GGT ATG GTT ATC ACC TTC GCT   900
286    V   E   T   G   I   M   K   P   G   M   V   I   T   F   A   300

901   CCA CCC AAC ATC ACC ACT GAG GTC AAA TCA GTG GAA ATG CAT CAC   945
301    P   P   N   I   T   T   E   V   K   S   V   E   M   H   H   315

946   GAG TCG CTC ACA GAA GCT GTA CCA GGA GAC AAT GTT GGC TTC AAT   990
316    E   S   L   T   E   A   V   P   G   D   N   V   G   F   N   330

991   ATA AAG AAC GTG TCT GTT AAG GAA ATC CGT CGA GGA AAT GTC TGT   1035
331    I   K   N   V   S   V   K   E   I   R   R   G   N   V   C   345

1036  GGA GAC AGC AAG AAT GAC CCT CCC AAG GGT GCC AAG AAC TTT ATT   1080
346    G   D   S   K   N   D   P   P   K   G   A   K   N   F   I   360

1081  GCC CAG GTC ATC ATC TTG AAC CAC CCT GGT GAA ATC AAG AAT GGA   1125
361    A   Q   V   I   I   L   N   H   P   G   E   I   K   N   G   375
```

```
1126  TAT GCG CCT GTC CTC GAT TGC CAC ACT GCT CAT ATT GCC TGC AAA  1170
376    Y   A   P   V   L   D   C   H   T   A   H   I   A   C   K   390

1171  TTT GTC GAA ATC AGA GAG AAA TGC GAT CGT CGT AGT GGA AAA GTT  1215
391    F   V   E   I   R   E   K   C   D   R   R   S   G   K   V   405

1216  TTG GAA GAA GCC CCT AAA GTA ATC AAG AGC GGA GAT GCT GCT ATG  1260
406    L   E   E   A   P   K   V   I   K   S   G   D   A   A   M   420

1261  GTC CTC ATG GTG CCC AGC AAC CCA ATG TGT GTT GAA CAA TTC TCT  1305
421    V   L   M   V   P   S   N   P   M   C   V   E   Q   F   S   435

1306  AAG TAC GCA CCA TTG GGA CGT TTT GCT GTC CGT GAC ATG AGG CAA  1350
436    K   Y   A   P   L   G   R   F   A   V   R   D   M   R   Q   450

1351  ACT GTA GCT GTT GGT GTC ATC AAA GAG GTT GAG AAA AAT GAA CCA  1395
451    T   V   A   V   G   V   I   K   E   V   E   K   N   E   P   465

1396  AGT CAA GGC AAA GTG ACC AAG GCT GCT CAG AAG GCT GGA AAG AAG  1440
466    S   Q   G   K   V   T   K   A   A   Q   K   A   G   K   K   480

1441  TGA AAA TTA ATG AAC TTT GGA ACC CTT CAA TCC CAT CAG CTT AAT  1485
481    *                                                         495

1486  TAT GAT TTT TAT TTA AAG AGA AAT AAT GAC ATC ATG CCA CTA TTT  1530
496                                                             510

1531  AAT TTT GAA AGA CCA CAG CAG GTC ATT TAT TGG ACA ATT TGG TTT  1575
511                                                             525

1576  TTG AAG AAA AAT ATC CAT CGC AAA GAA GTC CGT CGG AAA AAA CGG  1620
526                                                             540

1621  ATT TGA AGT AGC AAG TAC TTG AAT TAT GTC GAA TAA ATT TGG ACA  1665
541                                                             555

1666  TTT TAA AAA AAA AAA AAA AAA AAA AAA AAA AAA   1698
556
```

Figure 4 Complete ORF nucleotide and deduced amino acid sequences of the flat oyster EF1-α. ORF contains three domains: elongation factor 1 (82-768 nt) in black, elongation factor II (778-1050 nt) in grey and elongation factor III (1057-1368 nt) in white. Start and stop codons are highlighted in grey.

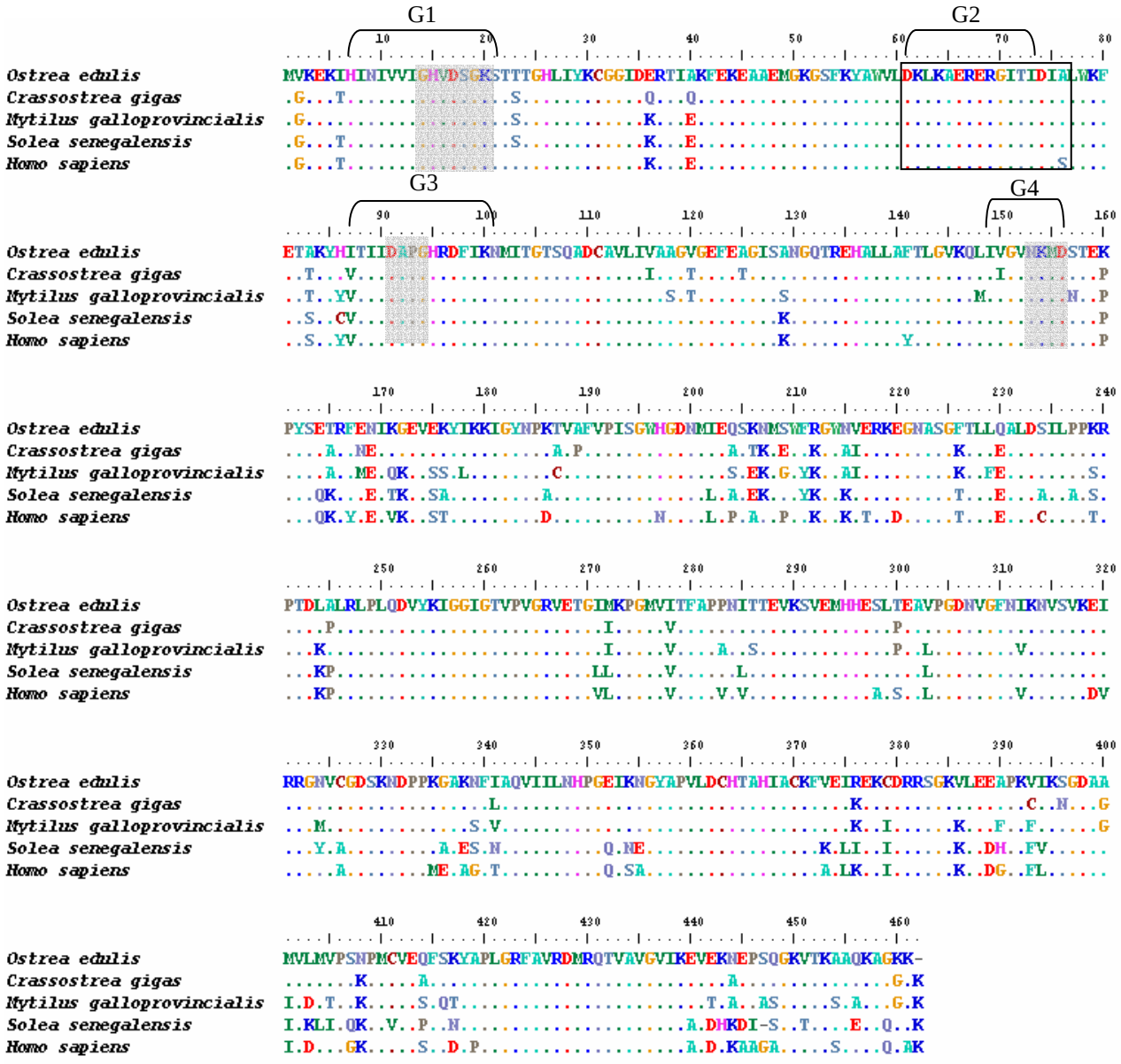

Figure 5 Comparison of the EF1-α amino acid sequences of *Ostrea edulis* and three other species (*Crassostrea gigas*, *Mytilus galloprovincialis* and *Homo sapiens*). Dots represent identity of amino acid between species and dashes represent gaps. G1 to G4 indicate the critical regions involved in GDP/GTP exchange and GTP hydrolysis. The consensus sequence composed of the three consensus elements GXXXXGK (G^{14}-K^{20}), DXXG (D^{91}-G^{94}), and NKXD (N^{153}-D^{156}) present in the GTP-binding domain is shaded in grey. The GTP-binding elongation factor signature corresponding to amino acids 61 to 76 is boxed.

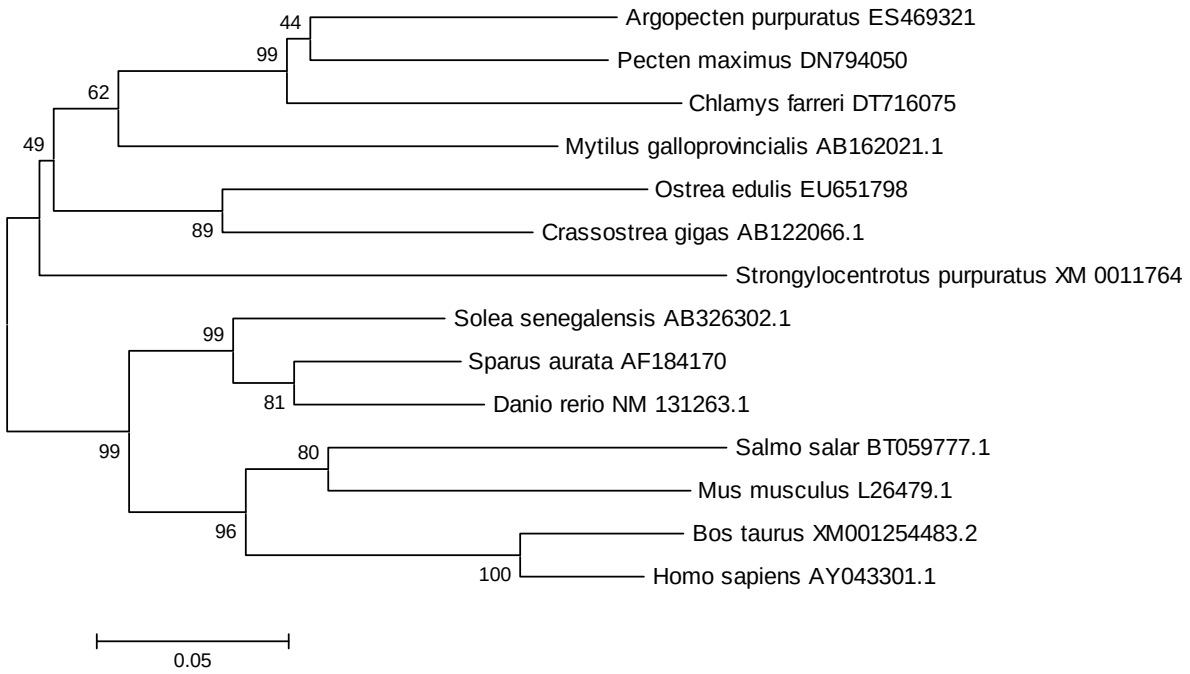

Figure 6 Un-rooted phylogenetic tree on the selected EF1-α amino acid sequences. The tree was constructed using the neighbour-joining algorithm in the Mega 4.0 program. Bootstrap values (shown) are based on 1 000 resamplings of the data. The scale for branch length (0.05 substitutions/site) is shown below the tree.

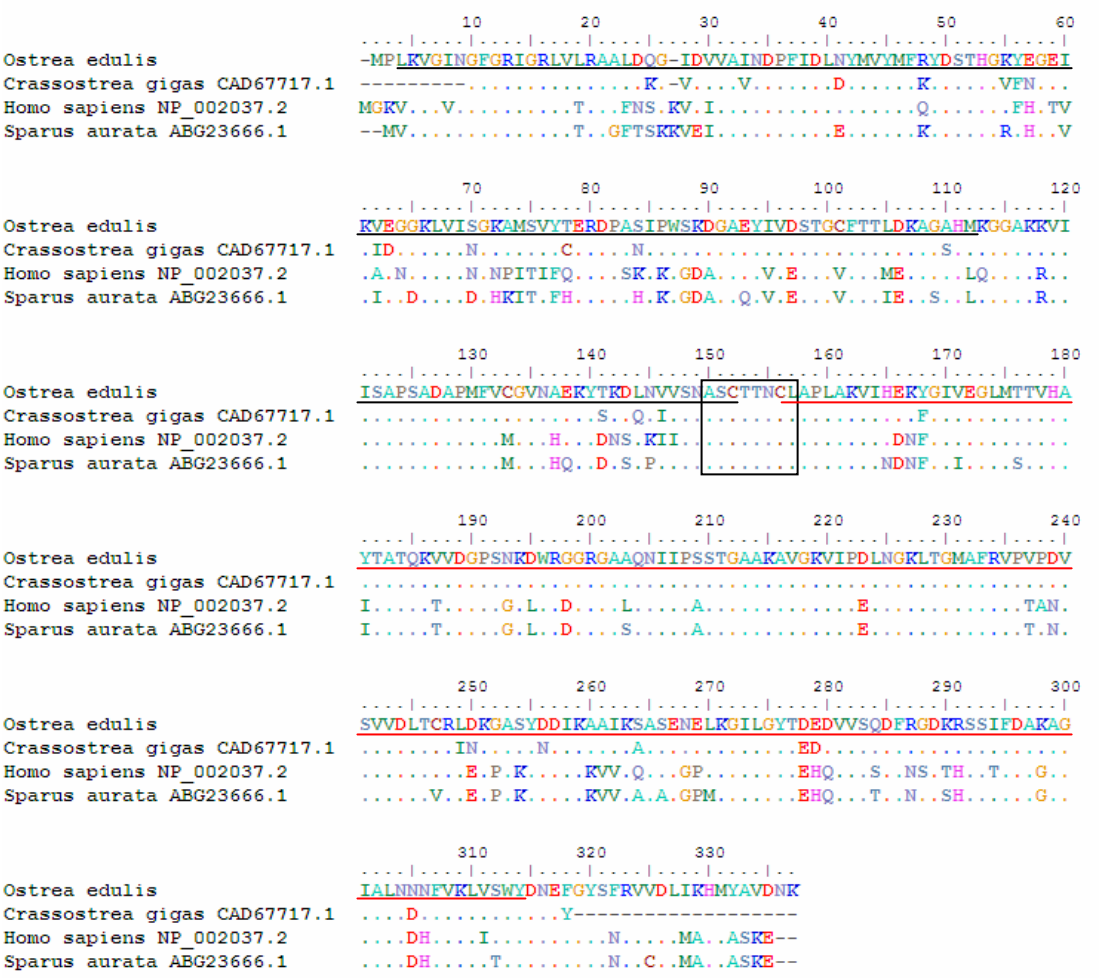

Figure 7 Comparison of the GAPDH amino acid sequences of *Ostrea edulis* and three other species (*Crassostrea gigas*, *Sparus aurata* and *Homo sapiens*). Dots represent identity of amino acid between species and dashes represent gaps. The NAD binding domain (3-150 aa) is underlined in black and the C terminal domain (155-312) is underlined in grey. The typical eukaryotic GAPDH signature (ASCTTNCL) corresponding to amino acids 148 to 155 is boxed.

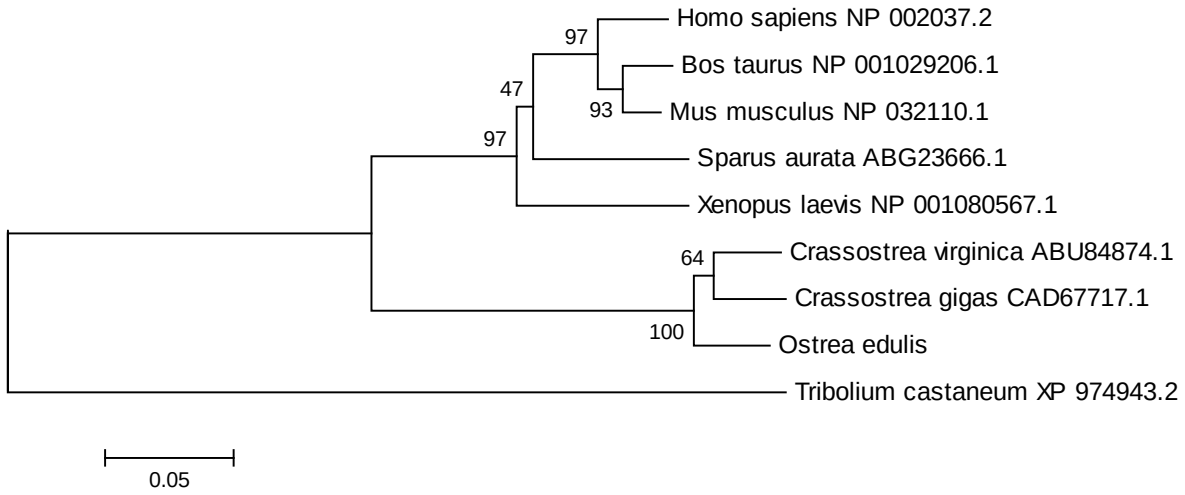

Figure 8 Un-rooted phylogenetic tree on the selected GAPDH amino acid sequences. The tree was constructed using the neighbour-joining algorithm in the Mega 4.0 program. Bootstrap values (shown) are based on 1 000 resamplings of the data. The scale for branch length (0.05 substitutions/site) is shown below the tree.

Table 1

Genes	Reference gene name	Cellular functions
ACT	β-Actin	Cytoskeleton structure protein
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Oxydoreductase in glycolysis and gluconeogenesis
Ubiq	Polyubiquitin	Protein degradation
EF-1α	Elongation factor alpha	Translation eukaryotic factor
L5	Ribosomal L5	Ribosome submit

Table 1 Candidate genes evaluated for reference genes during an *in vitro* infection between haemocytes from flat oyster *Ostrea edulis* and the parasite *Bonamia ostreae.*

Table 2

Genes	Sequences 5'-3'	Amplicon length (bp)	Annealing temperature (°C)
ACT	Forward:ACCAGTTGTACGACCGGAAG Reverse:CACGGTATCGTCACCAACTG	234	60
GAPDH	Forward: TCCCGCTAGCATTCCTTG Reverse: TTGGCGCCTCCTTTCATA	108	60
Ubiq	Forward:ACCAAATGAAGCGTGGATTC Reverse:TGAGGTCGAACCATCAGACA	165	60
EF-1α	Forward:GTCGCTCACAGAAGCTGTACC Reverse:CCAGGGTGGTTCAAGATGAT	162	60
L5	Forward: TCAGTGCAGAGGTTCACAGG Reverse: TAGCAGCATGGCACTTTTTG	171	60

Table 2 Housekeeping gene primers and characteristic for quantitative real time PCR.

Table 3

Genes	GAPDH	L5	EF-1	ACT	Ubiq
Ct	21	19,9	19,1	16,1	18
sd	1,2	1,2	1,4	1,5	0,9
E	96	95	100	96	95

Table 3 Descriptive statistics of threshold cycles (Ct) of 5 housekeeping genes. (E) Quantitative PCR efficiency of the candidate housekeeping genes. (Sd) corresponds to standard deviation (n=8).