

RESEARCH ARTICLE

Identification and validation of reference genes for qRT-PCR studies of the obligate aphid pathogenic fungus *Pandora neoaphidis* during different developmental stages

Shutao Zhang¹, Chun Chen¹*, Tingna Xie¹*, Sudan Ye²

1 China Jiliang University, Zhejiang Provincial Key Laboratory of Biometrology and Inspection & Quarantine, Hangzhou, China, **2** Zhejiang Economic & Trade Polytechnic, Hangzhou, China

* These authors contributed equally to this work.

* aspring@cjl.u.edu.cn



OPEN ACCESS

Citation: Zhang S, Chen C, Xie T, Ye S (2017) Identification and validation of reference genes for qRT-PCR studies of the obligate aphid pathogenic fungus *Pandora neoaphidis* during different developmental stages. PLoS ONE 12(7): e0179930. <https://doi.org/10.1371/journal.pone.0179930>

Editor: Raffaella Balestrini, Institute for Sustainable Plant Protection, C.N.R., ITALY

Received: December 28, 2016

Accepted: June 6, 2017

Published: July 3, 2017

Copyright: © 2017 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was funded by the Natural Science Foundation of China (31461143030), the Science and Technology Planning Project of Zhejiang Province (2013C37087) and the Zhejiang Natural Science Foundation (LR12C03001).

Competing interests: The authors have declared that no competing interests exist.

Abstract

The selection of stable reference genes is a critical step for the accurate quantification of gene expression. To identify and validate the reference genes in *Pandora neoaphidis*—an obligate aphid pathogenic fungus—the expression of 13 classical candidate reference genes were evaluated by quantitative real-time reverse transcriptase polymerase chain reaction (qPCR) at four developmental stages (conidia, conidia with germ tubes, short hyphae and elongated hyphae). Four statistical algorithms, including geNorm, NormFinder, BestKeeper and Delta Ct method were used to rank putative reference genes according to their expression stability and indicate the best reference gene or combination of reference genes for accurate normalization. The analysis of comprehensive ranking revealed that *ACT1* and *18S* was the most stably expressed genes throughout the developmental stages. To further validate the suitability of the reference genes identified in this study, the expression of cell division control protein 25 (*CDC25*) and Chitinase 1 (*CHI1*) genes were used to further confirm the validated candidate reference genes. Our study presented the first systematic study of reference gene(s) selection for *P. neoaphidis* study and provided guidelines to obtain more accurate qPCR results for future developmental efforts.

Introduction

Pandora neoaphidis (Entomophthoromycotina, Entomophthorales) is an obligate aphid pathogenic fungus and has great potential for use in biological control applications [1–3]. Due to the ability to produce forcibly discharged infectious conidia, which germinate and penetrate the aphid through the cuticle, ultimately killing it, *P. neoaphidis* has a great value for the conservation and epizootiological studies [2].

Although there are many sequences of *P. neoaphidis* deposited in the National Center for Biotechnology Information (NCBI) GenBank database, the vast majority of available genomic information is composed of partial gene and intron sequences developed for use in

phylogenetic analyses[4]. Elucidation of the infection processes and pathogenicity by modern genomic tools is desirable because it provides clues for understanding the main elements of host-pathogen evolution and the function of virulence genes in pathogenic fungi[4,5]. However, the function of genes of interest (GOIs), especially gene expression of *P. neoaphidis* during the growth and infection process, is still lacking a detailed molecular understanding nowadays. Generally, a good understanding of the GOIs requires good knowledge of gene expression under a broad set of conditions. A diversity of genes can be up- or down-regulated during various developmental stages as well as pathogen-insect interactions[6]. Therefore, the study of *P. neoaphidis* gene expression in different conditions may provide useful information regarding the molecular mechanisms highlighting the epizootic ability enabling for the development of genetic engineering approaches and field application. Quantitative real-time reverse transcription polymerase chain reaction (qPCR) has been indicated to be a good method for measuring gene expression across different samples because of its advantages, e.g., high throughput capacity, sensitivity and specificity[7,8]. Furthermore, qPCR has been used for this purpose with several fungi species[9–12], even including *P. neoaphidis* which was described previously but using only one reference gene [5]. However, one constant challenge is the selection of suitable reference genes because it requires a species-specific solving process[13,14]. Moreover, it is not acceptable to use a single traditional gene nowadays, and instead the ‘gold standard’ is the normalization of gene expression levels by reference genes. Here, an ideal reference gene set should have constant expression across all the samples to be investigated, regardless of the developmental stage or any other biological or technical variability.

Additionally, the selection of inappropriate reference genes with variable expressions could lead to incorrect calculations of the expression of the target genes and, therefore, incorrect assumptions about the function of those target genes[15]. The identification of suitable reference genes is therefore essential for accurate transcript expression analysis. For this purpose, several statistical algorithms have been developed to identify the most suitable internal controls with the least expression variability; calculating their expression stability based on the qPCR data of candidate reference genes. These algorithms rank the putative reference genes, thereby to indicate the best reference gene or genes for normalization. The four most commonly used algorithms for assessing the appropriateness of reference genes are: geNorm[16], NormFinder [17], BestKeeper[18] and Delta Ct[19], and these algorithms have been widely used[20,21].

In this study, 13 putative housekeeping genes including *18S*(18S ribosomal RNA), *28S*(28S ribosomal RNA), *ACT1* (Actin-1), *ACTN1*(Actinin Alpha 1), *HPRT*(Hypoxanthine-guanine phosphoribosyltransferase), *EF1*(Elongation factor 1), *GAPDH*(Glyceraldehyde-3-phosphate dehydrogenase), *TBCE*(Tubulin-specific chaperone E), *LDHA*(L-lactate dehydrogenase A), *LSM1* (U6 snRNA-associated Sm-like protein LSM1), *HISTH4* (Histone H4), *ALG9*(Alpha-1,2-mannosyltransferase) and *DMA2* (E3 ubiquitin-protein ligase DMA2), were evaluated as potentially reliable reference genes in *P. neoaphidis* using qPCR. As fungal growth and infection of the host usually involve multiple developmental stages[22], the profiling of gene expression in multiple development stages is important for understanding the mechanisms of growth and pathogenesis. With this in mind, we tested four developmental stages, including conidia, conidia with germ tubes, short hyphae and elongated hyphae. The expression stability of each gene in all the samples was analysed by the geNorm, NormFinder, BestKeeper and Delta Ct programs, and the appropriate reference gene set for accurate normalization was then selected. Furthermore, for demonstrating the efficacy of the selected reference genes, we investigated the expression of cell division control protein 25 (*CDC25*), a key protein controlling entry into and progression through various phases of the cell cycle[23], and Chitinase 1 (*CHI1*), responsible for the carbohydrate active enzymes[24]. The expression of these two

genes were significantly different during different developmental stages, as it was already demonstrated in a previous RNA-seq experiment performed by our group (unpublished data).

Materials and methods

Isolate and culture conditions

P. neoaphidis was isolated from an infected green-peach aphid *Myzus persicae* (Sulzer) cadaver from Hangzhou city, China, in 1998 using the 'descending conidia' method[25] and was designated as F98028[1]. It was subcultured on SEMA (Sabouraud dextrose agar [SDA, Oxoid, USA] supplemented with egg yolk and milk)[26] in 9cm diameter Petri dishes for 10d at 20°C in a 12:12 light: dark regime.

P. neoaphidis sample preparations

Preparation of *P. neoaphidis* propagules at different developmental stages were modified slightly as previous description[5]. Primary conidia actively discharged from the mycelial mat during the period of peak sporulation were harvested into 0.01M sterile PBS(pH 7.3)[27]. After the conidia were centrifuged at 3170×g for 10 min, a concentration of 10¹⁰ conidia per ml was then inoculated into GLEN medium(0.4%glucose, 0.5% yeast extract, 0.65% lactalbumin hydrolysate,0.77% NaCl;)[28] in flasks at 20°C and 150 rpm in a light: dark (12:12) for different periods of time[1]. Samples were prepared at ~0h for conidia without morphological changes, after ~6h for conidia with germ tubes, after ~12h for early short hyphae (i.e., the length of the germ tube was no more than 200µm), and after 24h for elongated hyphae (i.e., the length of hyphae exceeded 200µm). Structures of fungal samples were observed microscopically to ensure that they were in the right developmental stage (S1 Fig, Appendix). Three biological replicates were performed for each developmental stage, i.e., 12 flasks in total were established.

Total RNA isolation and cDNA synthesis

Total RNA was prepared and extracted mechanically for each sample using the RNeasy Mini Kit (QIAGEN, Cat. No 74104, Mississauga, Canada) according to the manufacturer's instructions. The quality of the RNA was evaluated by 1% agarose gel electrophoresis to indicate the intact 28S and 18S ribosomal RNA (rRNA) subunits without any contaminating genomic DNA. Nucleic acid concentrations were measured using a Nanodrop 1000 (Thermo Biotech, USA).

First-strand cDNA was synthesised using the PrimeScript™ RT reagent Kit with gDNA Eraser according to the manufacturer's instructions (Takara, Cat.No.RR047A, China). Before the reverse transcription, RNase-free DNase I(Agilent Technologies, USA)was added immediately into the RNA samples in order to avoid potential errors from genomic DNA contamination. One µg total RNA was used for the cDNA synthesis. Subsequently, the cDNA was diluted into 20 ng µl⁻¹ with nuclease-free water(QIAGEN, Canada). The cDNA mixtures were diluted 1:10 with nuclease-free water and stored at -20°C for subsequent qPCR analysis.

Primer design and validation

Nucleotide sequences of the 18S and 28S genes were downloaded from Genbank (<http://www.ncbi.nlm.nih.gov/genbank/>), while the other 11 candidate reference gene sequences, including *ACT1*, *ACTN1*, *HPRT*, *EF1*, *GAPDH*, *TBCE*, *LDHA*, *LSM1*, *HISTH4*, *ALG9* and *DMA2*, were acquired from *P. neoaphidis* RNA-seq dataset(unpublished data). All the Genbank accession numbers are shown in Table 1. Primer pairs for each gene were designed using Beacon

Table 1. Primer and amplicon characteristics of RT-qPCR for the selected candidate reference genes and genes of interest(GOIs).

Gene	Accession No.	Primer sequence	Amplicon size(bp)	R ²	Primers efficiency
<i>18S</i>	HQ677591.1	F:CAAACCCGAGCAATAGTC R:GTAGGAGCCCGATAGTAAA	179	0.996	108.7%
<i>28S</i>	EF392405.1	F:TCTTATCAGTCTCAGCACCTTG R:TCTGTCAATCCTTACTATGTCTGG	181	0.996	104.0%
<i>ACT1</i>	KT715508	F:GACGAAGTTGCTGCTCTTG R:TCCCATACCAACCATGACAC	138	0.996	107.8%
<i>ALG9</i>	KT715507	F:TGAGGATATAAGCCCAAGAACG R:ATTGGGAACCTACCCATTTCG	96	0.998	107.7%
<i>ACTN1</i>	KT715516	F:GCGGAAATACAGCCTTAGC R:CTTCTTCCTGCCGTTCTATC	178	0.998	98.2%
<i>EF1</i>	KT715517	F:CCGATGTCTATGTCTGCTAAGG R:TTGGCGAGGCTCAAATCATG	193	0.999	98.7%
<i>GAPDH</i>	KT715509	F:GGTAGTTATCTCTGCTCCTTCTG R:CAGTGGTAGCGTGACAAGTAG	194	0.996	101.4%
<i>HISTH4</i>	KT71551	F:TGCTATTCGTCGTCGTCGC R:GCATCACGGACAACCTGACTCTA	111	0.998	109.7%
<i>HPRT</i>	KT715515	F:AGGTGGGCTCATAACTTTGG R:GGTCCTCTAGTTGTATGTTGTG	295	0.998	100.7%
<i>LDHA</i>	KT715510	F:CCTGGTTGGTCGTAATCG R:CCGAGTCTCTGAGTATCAAG	178	0.996	100.3%
<i>TBCE</i>	KT715514	F:TGGTAAGCGTCAGCGATACAATGG R:CCGTCTAGCCGTAGCGTATTAAGTG	172	0.997	100.6%
<i>LSM1</i>	KT715513	F:GGGTGAAGTGGACGAAGAAATC R:CCTTGAGAGTGAAGAACATGGTG	141	0.998	108.0%
<i>DMA2</i>	KT715511	F:GCGTTCGGCATTCTTGATG R:CCTGTTGTCGTTCTCTGTTTC	157	0.996	103.1%
<i>CDC25</i> <i>CHI1</i>	KX197113 KX197114	F:CCAAGTCCCTTACAACCTTATCG R:CCACAGTGTGTCGATTGAG F:GCCTGCTTGR:GTCGTATGG R:GTAGAAGGAGGTTTCGCTTGG	185 282	0.997 0.997	101.2% 104.5%

<https://doi.org/10.1371/journal.pone.0179930.t001>

Designer 7 software (Premier Biosoft, Palo Alto, CA, USA) (Table 1). The PCR parameters of each primer pair were evaluated as described in detail previously [5] for all the samples containing the same amount of cDNA from a cDNA mixture (mentioned above) as the template. Specificity of primer pairs for each candidate gene was confirmed using melting curve analysis and agarose gel electrophoresis.

Quantitative PCR

The qPCR was performed on an iQ⁵ multicolour real-time PCR detection system (Bio-RAD, USA) using MicroAmp[™] Reaction Tubes sealed with MicroAmp Optical 8-cap Strip (ABI, USA). The reaction mix contained a 20 ng template of cDNA, SYBR Premix Ex Taq[™] II (12.5 μl; Takara, China) and 400 nM each of the forward and reverse primers (Sangon, China) in a final volume of 25 μl. The PCR was accomplished after a 10 min activation/denaturation step at 95°C, followed by 40 cycles of 30 s each at 95°C, 60 s each at 60°C, and 30 s each at 72°C. Fluorescence was detected at each polymerisation step. A cDNA mixture template containing equal amounts of cDNA from all samples was used to evaluate the PCR efficiency of each primer pair. The cDNA mixture template (from 10 ng to 1 pg) was diluted as ten-fold series. After the PCR amplification, a melting curve analysis was performed to check the amplification specificity. Three technical replicates were performed for each sample. Two control

samples, the no-template control (NTC), and the *P. neoaphidis* genomic DNA were included with each primer pair tested.

Statistical analyses

The expression levels of the candidate reference genes were determined from the Cq values (quantification cycle value), i.e., the number of PCR cycles at which the quantity of amplified targets reached a specific threshold level of detection. The expression stability of the reference genes was evaluated by analysing all the raw Cq values using the four programs: geNorm[16], NormFinder[17], BestKeeper[18] and the comparative Delta Ct[19]. geNorm can automatically calculate the gene expression stability measure M which is the mean pair-wise variation for a given gene compared with all other tested genes. Gene with high M values has greater variation in expression, and the rank of tested genes is evaluated via a stepwise exclusion process. NormFinder uses a model-based variation of a gene as its stability value and can rank a set of genes according to their stability values. BestKeeper algorithm evaluates the expression stability of a candidate reference gene by repeated pair-wise correlation analysis of the gene relative to all other candidate reference genes. Further software Delta Ct: a method similar to that described by Vandesompele et al [16], whereby compared 'pairs of genes' for evaluating the expression stability. Besides, in order to balance the ranking of the candidate reference genes across the different algorithms, the expression stability was evaluated with the RefFinder tool (<http://www.leonxie.com/referencegene.php?type=reference>) to determine the overall comprehensive stability value for each candidate reference gene[29]. Additionally, the optimal number of genes needed for accurate normalization—namely the normalization factors NF_n , which is the geometric mean of the n most stable genes—was measured stepwise with the geNorm program[16]. Pairwise variations ($V_{n/n+1}$) between two subsequent normalization factors (NF_n/NF_{n+1}) were calculated to indicate the effect of adding the $(n+1)^{th}$ gene for normalization. Here, if $V_{n/n+1}$ is larger than the recommended threshold of 0.15, that means that adding the $(n+1)^{th}$ gene has a significant effect and thus this target gene is considered as a reliable reference gene.

Validation of reference genes

In order to validate the liability of the selected reference genes as determined by RefFinder, the relative expression profiles of *CDC25*, a cell division control protein 25 involved in controlling entry into and progression through various phases of the cell cycle, and *CH11*, a gene of the carbohydrate active enzymes, were measured and normalized with the most stable and least stable reference genes. The RT-qPCR amplification conditions were the same as described above and primer pairs were shown in Table 1. The relative expression data was calculated according to the $2^{-\Delta\Delta Cq}$ method and presented as fold change [30]. Samples from conidia stage were used as reference samples. In case of statistical significance, the relative quantitation (RQ) results were compared among the more stable and more unstable reference genes with results of RNA-seq analysis by Mann–Whitney U test. Statistical analyses were performed using the computer program DPS System [31].

Results

Quantification and expression levels of the candidate reference genes

18S, 28S and 11 housekeeping genes pre-selected from the RNA-seq database, including *ACT1*, *ACTN1*, *HPRT*, *EF1*, *GAPDH*, *TBCE*, *LDHA*, *LSM1*, *HISTH4*, *ALG9* and *DMA2*, were tested as references for the gene expression studies. There was a single peak in the melting

curve of each qPCR reaction, but there were no signals detected within the controls, indicating that primer dimers or other non-specific amplification products had not appeared (S2A Fig, Appendix). The sizes of the amplicon products generated from specific primers were in the range 96–295 base pairs (Table 1). R^2 presented a good linear relationship between the Cq values and the relevant relative amount of cDNA in all the qPCR reactions. Correspondingly, the amplification efficiencies were estimated to be 98.2–109.7% (Table 1). The specificity of the primer pairs for each candidate reference gene was identified using agarose gel electrophoresis (S2B Fig, Appendix). Thus, the 13 primers were shown to be specific and appropriate for qPCR analysis.

The calculated expression levels (Cq values) of the 13 genes evaluated varied from 10.16 to 35.73, which represents quite a large degree of variation (Fig 1). *18S*, *28S*, *EF1* and *LSM1* showed a compact distribution of Cq values, displaying a relatively low variation among the four developmental stages. Some of these, i.e., *GAPDH*, *HISTH4* and *HPRT*, as well as the *LDHA* gene showed a slightly higher dispersion of their Cq values, thus displaying a relatively high variation among the different developmental stages.

Candidate gene expression stability during four developmental stages

Using the algorithms geNorm, NormFinder, BestKeeper and Delta Ct, the results of the stability tests were different (Fig 2). The heat map showing the ranking order indicates the most to the least stable, ranging from red to blue. The stability values calculated by geNorm, NormFinder, BestKeeper and Delta Ct are listed in S1 Table, respectively. The NormFinder and Delta Ct algorithms led to similar classifications, ranking *ACT1*, *ACTN1* and *TBCE* as the three most stable genes. However, the ranking order of stability calculated by the BestKeeper algorithms

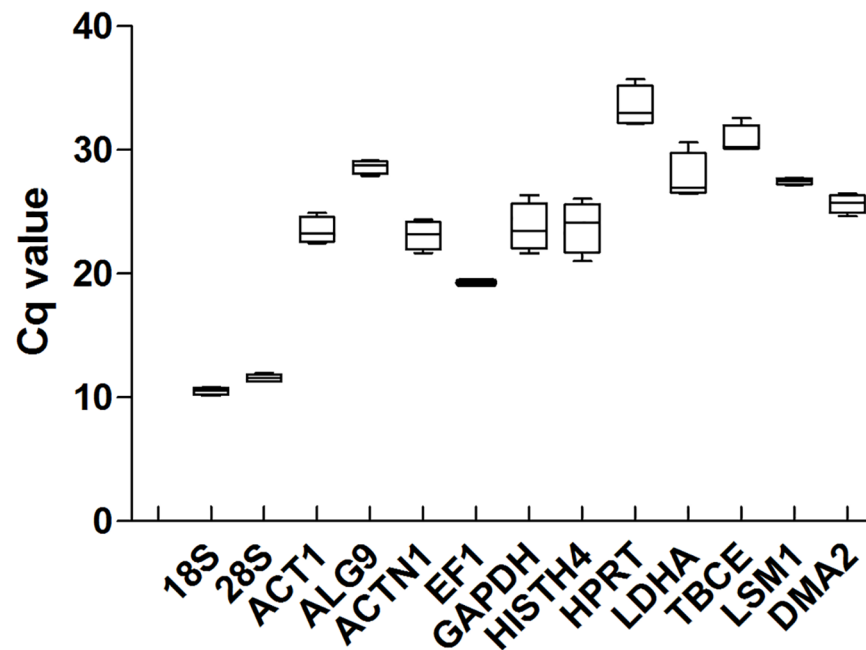


Fig 1. Distribution of the raw Cq (quantification cycle) values of the 13 candidate reference genes at four developmental stages. The lower and upper boundaries of the box (interquartile) represent the 25% and 75% percentiles, respectively. Whiskers indicate the maximum and minimum values. The median is depicted by the line across the box.

<https://doi.org/10.1371/journal.pone.0179930.g001>

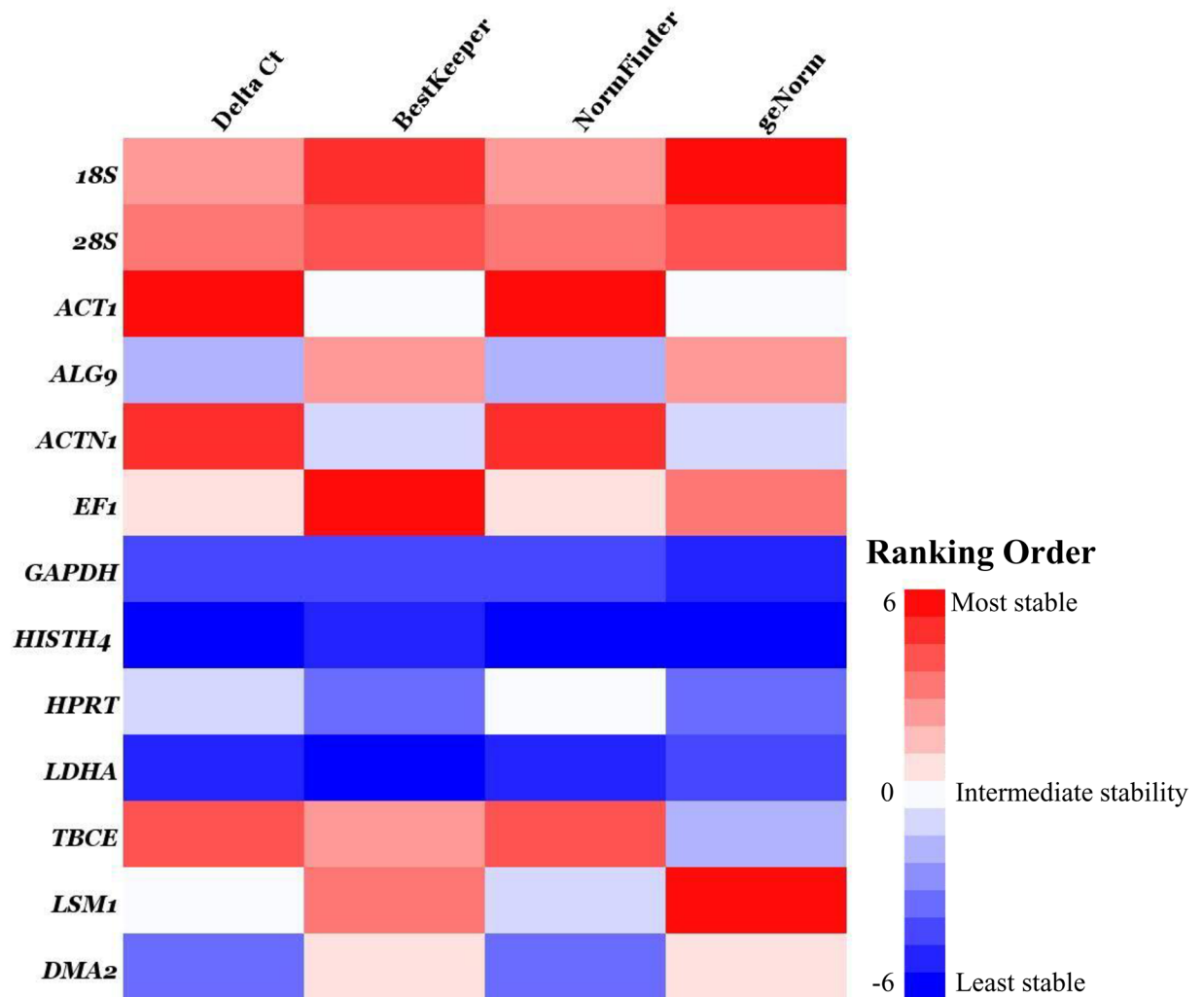


Fig 2. Heat map of the ranking in expression stability calculated by geNorm, NormFinder, BestKeeper and Delta Ct. The analysis was performed separately for each 9 sample group under the conditions examined. The 13 candidate references are coloured based on their stability in different saturations. Red represents the most stable and blue the least stable in ranking order.

<https://doi.org/10.1371/journal.pone.0179930.g002>

was: *EF1*, *18S*, *28S*, *LSM1*, *ALG9*, *DMA2*, *ACT1*, *ACTN1*, *TBCE*, *HPRT*, *GAPDH*, *HISTH4*, *LDHA*. Additionally, the ranking order of stability calculated by the geNorm algorithms was: *18S*, *LSM1*, *28S*, *EF1*, *ALG9*, *DMA2*, *ACT1*, *ACTN1*, *TBCE*, *HPRT*, *LDHA*, *GAPDH*, *HISTH4*. Unexpectedly, as calculated by geNorm, the *M* values of these 13 genes were below the recommended cut-off value of 1.5, even for the *GAPDH* gene, but also for *LDHA* and *HISTH4*, both of which ranked among the least stable genes, consistent with the behaviour calculated by the other three algorithms.

To reinforce this result, RefFinder, a web-based tool that integrates four algorithms: geNorm, NormFinder, BestKeeper and the comparative Delta Ct method—was also employed to balance the ranking of the candidate genes. As can be seen in Fig 3, the overall stability was calculated as the geometric mean of their weights. The ranking order from the most to least stable was: *ACT1*, *18S*, *28S*, *EF1*, *LSM1*, *ACTN1*, *TBCE*, *ALG9*, *DMA2*, *HPRT*, *GAPDH*, *LDHA*, *HISTH4*.

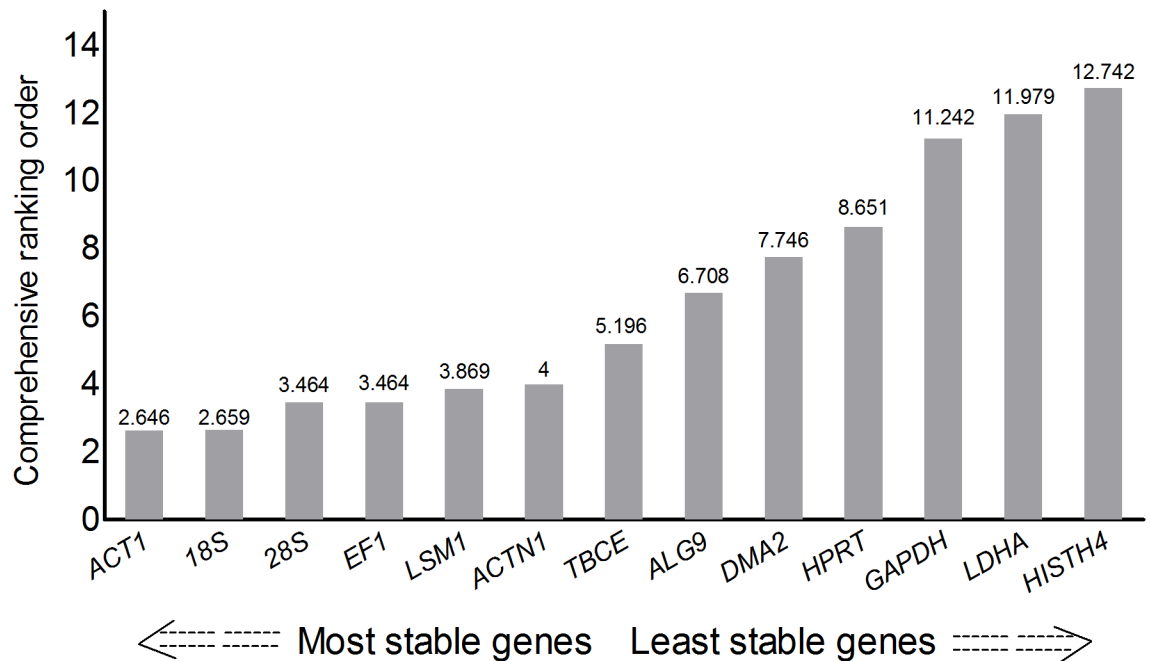


Fig 3. Comprehensive ranking of the candidate reference genes. The expression stability was evaluated with the RefFinder tool (<http://www.leonxie.com/referencegene.php?type=reference>) to determine the overall comprehensive ranking order for each candidate reference gene.

<https://doi.org/10.1371/journal.pone.0179930.g003>

Determination of the optimal number of reference genes

The optimal number of reference genes was evaluated as required for accurate normalization in the four developmental stages using the geNorm program. As shown in Fig 4, a combination of the three most stable genes, *ACT1*, *18S* and *28S*, has the lowest pairwise variation value ($V_{3/4}$) of 0.062, which is lower than the threshold of 0.15. If *28S* is excluded, $V_{2/3} = 0.068$ is higher than $V_{3/4}$, but is still lower than the threshold of 0.15. Therefore, a set with two reference genes (*ACT1* and *18S*) is sufficient for accurate normalization of the transcript levels in all of the samples.

Validation of reference genes

The quantification was realized using normalization by the combination of the two most stable genes as determined above and one of the least stable genes in developmental stages, i.e., “*ACT1+18S*” and *HISTH4*, respectively. As shown in Fig 5, we observed that *CDC25* gene was indeed up-regulated and *CH11* gene was down-regulated regardless using the set of “*ACT1+18S*” reference genes or *HISTH4* reference gene to normalize. The relative expression data of *CDC25* and *CH11* were calculated according to the $2^{-\Delta\Delta Cq}$ method based on the qPCR data and presented as fold change. Furthermore, the relative quantification (RQ) results obtained for *CDC25* and *CH11* using the “*ACT1+18S*” gene set for the normalization were not significantly different from that obtained from our RNA-seq database ($u = 0.72$, $p > 0.05$; $u = -0.37$, $p > 0.05$, respectively). In contrast, using *HISTH4* for expression normalization gave rise to very different results compared to that from RNA-seq ($u = 8.23$, $p < 0.01$; $u = 2.38$, $p < 0.05$, respectively). These results confirm the suitability of the “*ACT1+18S*” gene set for normalization purposes.

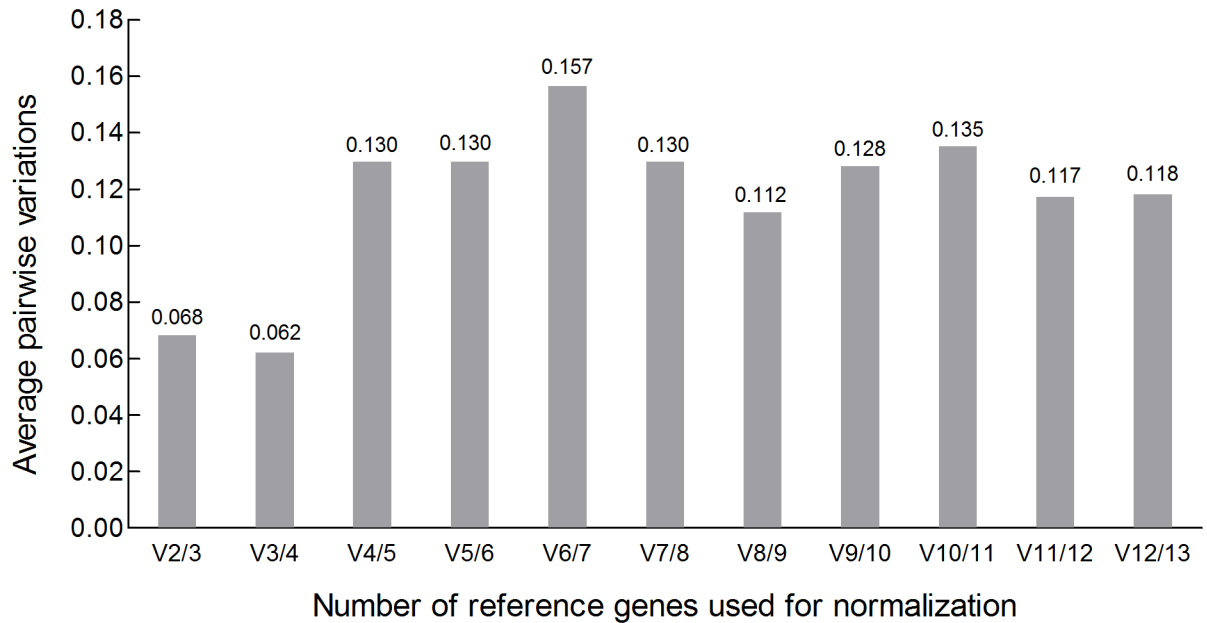


Fig 4. Optimal number of reference genes in different developmental stages of *P. neoaphidis* for normalization as evaluated by geNorm. Pairwise variations ($V_{n/n+1}$) between two subsequent normalization factors NF_n (geometric mean of the expression levels of n reference genes) and NF_{n+1} in the course of a stepwise inclusion of reference genes were calculated to indicate the effect of adding the $(n+1)^{th}$ gene for normalization.

<https://doi.org/10.1371/journal.pone.0179930.g004>

Discussion

Fungal growth and infection of the host usually involves multiple developmental stages[22]. In this study, we evaluated the profiling of gene expression in four development stages for understanding the mechanisms of growth and pathogenesis. Using RNA-seq databases of *P. neoaphidis*, we carried out a pre-selection of the stable genes with a fold change(FC) equal to one between hyphae RNA-seq database and conidia RNA-seq database. Thirteen classical

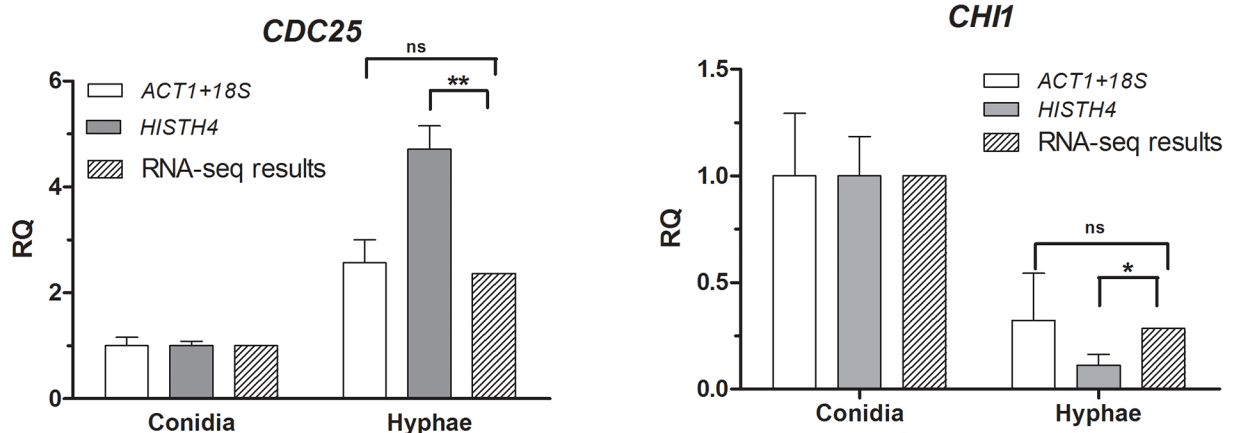


Fig 5. Expression profile of *CDC25* and *CHI1* in *P. neoaphidis* under conidia stage and elongated hyphae stage. The best stable combination of reference genes (*ACT1* and *18S*, white bars) and the least stable reference gene (*HISTH4*, black bars) were used to normalize the expression data. The relative quantification (RQ) of *CDC25* and *CHI1* (conidia stage VS elongated hyphae stage) obtained from RNA-seq analysis were marked as grey bars and chosen as standard for identifying their stability. Samples from the conidia stage were used as reference samples. The data show the mean expression \pm standard deviation calculated from three biological replicates.

<https://doi.org/10.1371/journal.pone.0179930.g005>

housekeeping genes were used as candidate reference genes for evaluating stability studies in *P. neoaphidis* (Table 1).

The selection of stable reference genes is a critical influence on the accuracy of qPCR[32]. Through the stability analyses, we identified that the *ACT1* and *18S* reference genes were the most suitable for normalization of the qPCR data from the multiple developmental stages. Those encoding rRNA genes are most conserved and constitute about 85–90% of the total cellular RNA in eukaryotes[33]. The *18S* gene—a classic housekeeping gene—has been widely used as a reference gene in the analysis of gene expression, e.g., in the entomopathogenic fungus *Beauveria bassiana*[9,34], *P. neoaphidis*[5] and other species[35,36]. Additionally, the *ACT1* gene encodes the single essential gene for actin, which is a ubiquitous, conserved cytoskeletal element critical for many cellular processes[37]. Although *ACT1* was not always encourage to use as reference gene since stationary phase samples showed a more significant drop of *ACT1* mRNA[38], they are still assumed to be constitutively expressed and widely used as reference genes[39,40]. In this study, *ACT1* is appropriate reference gene for evaluating expression stability during four developmental stages in *P. neoaphidis*.

Furthermore, by using simply a single algorithm to analyse the biological stability of expression in candidate reference genes, we often have difficult to evaluate and chose the best reference gene exhaustively [41]. Therefore, in this study, we adopted a comprehensive online tool based on four algorithms involved geNorm, NormFinder, BestKeeper and Delta Ct. Although the stability of the 13 candidate reference genes differed among the four software programs, reliable reference genes were identified for use in gene expression analyses. The overall ranking order—from the most to least stable—was *ACT1*, *18S*, *28S*, *EF1*, *LSM1*, *ACTN1*, *TBCE*, *ALG9*, *DMA2*, *HPRT*, *GAPDH*, *LDHA*, *HISTH4*, as highlighted by the comprehensive RefFinder tool. These algorithms did not always return identical results, especially for geNorm and BestKeeper compared to the NormFinder and Delta Ct methods. In addition, we obtained an unexpectedly identical stability of the candidate reference genes in different culture nutrition, indicating that they are all suitable for normalization (unpublished data). This may also be the reason for the different algorithms giving different normalization results. Furthermore, to evaluate the profiling of gene expression in growth period and pathogenesis stage, the relative quantitation of two relative genes including *CDC25* and *CHI1*, were measured using the set of “*ACT1+18S*” reference genes and *HISTH4* reference gene under conidia stage and hyphae stage. Compared with the RNA-seq database, results showed that the use of unsuitable references leads to differences in the relative expression profile. These results further confirmed the importance of validating reference genes prior to experimental applications.

Conclusions

In this study, our results suggest that the *ACT1* and *18S* gene set would be the most reliable for normalizing the expression levels in multiple stages of the entomophthoralean fungus *P. neoaphidis* with the accurate and widespread method: qPCR. However, a further complication is present in that fungus-insect interactions involved fungal propagules and the host. Thus, further studies are required to assess and validate the identified reference genes suitable for the expression of GOIs during fungus-insect interactions.

Supporting information

S1 Table. Stability values calculated by geNorm, NormFinder, BestKeeper and Delta Ct programs.

(DOCX)

S1 Fig. Image of *P. neoaphidis* propagules.

(DOCX)

S2 Fig. Confirmation of primer specificity and amplicon length for thirteen candidate reference genes.

(DOCX)

Acknowledgments

We thank Prof. M.G. Feng (Zhejiang University, Hangzhou) for providing *Pandora neoaphidis* F98028 strain. This study was funded by the Natural Science Foundation of China (31461143030), the Science and Technology Planning Project of Zhejiang Province (2013C37087) and the Zhejiang Natural Science Foundation (LR12C03001).

Author Contributions

Data curation: Shutao Zhang, Tingna Xie.

Project administration: Chun Chen.

Software: Tingna Xie, Sudan Ye.

Writing – original draft: Chun Chen.

Writing – review & editing: Chun Chen.

References

1. Chen C, Feng MG. Experimental simulation of transmission of an obligate aphid pathogen with aphid flight dispersal. *Environ Microbiol.* 2006; 8: 69–76. <https://doi.org/10.1111/j.1462-2920.2005.00869.x> PMID: 16343323
2. Pell JK, Eilenberg J, Hajek AE. Biology, ecology and pest management potential of Entomophthorales. In: Butt TM, Jackson C & Magan N, (eds.) *Fungal biological control agents: progress, problems and potential*, CABI Publishing, Oxord. 2001, pp. 71–153
3. Jensen AB, Hansen LM, Eilenberg J. Grain aphid population structure: no effect of fungal infections in a 2-year field study in Denmark. *Agri Forest Entomol.* 2008; 10: 279–290.
4. De Fine Licht HH, Hajek AE, Eilenberg J, Jensen AB. Utilizing genomics to study entomopathogenicity in the fungal phylum Entomophthoromycota: a review of current genetic resources. *Adv Genet.* 2016; 94:41–65. <https://doi.org/10.1016/bs.adgen.2016.01.003> PMID: 27131322
5. Chen C, Xie T, Ye S, Jensen AB, Eilenberg J. Selection of reference genes for expression analysis in the entomophthoralean fungus *Pandora neoaphidis*. *Braz J Microbiol.* 2016; 47(1):259–65 <https://doi.org/10.1016/j.bjm.2015.11.031> PMID: 26887253
6. Grell MN, Jensen AB, Olsen PB, Eilenberg J, Lange L. Secretome of fungus-infected aphids documents high pathogen activity and weak host response. *Fungal Genet Biol.* 2011; 48: 343–352. <https://doi.org/10.1016/j.fgb.2010.12.003> PMID: 21156213
7. Bustin SA, Benes V, Nolan T, Pfaffl MW. Quantitative real-time RT-PCR—a perspective. *J Mol Endocrinol.* 2005; 34: 597–601. <https://doi.org/10.1677/jme.1.01755> PMID: 15956331
8. Ginzinger DG. Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Exp Hematol.* 2002; 30: 503–512. PMID: 12063017
9. Fang W, Bidochka MJ. Expression of genes involved in germination, conidiogenesis and pathogenesis in *Metarhizium anisopliae* using quantitative real-time RT-PCR. *Mycol Res.* 2006; 110: 1165–1171. <https://doi.org/10.1016/j.mycres.2006.04.014> PMID: 17010593
10. Lian TT, Yang T, Liu GJ, Sun J, Dong C. Reliable reference gene selection for *Cordyceps militaris* gene expression studies under different developmental stages and media. *FEMS Microbiol Lett.* 2014; 356: 97–104. <https://doi.org/10.1111/1574-6968.12492> PMID: 24953133
11. Llanos A, François JM, Parrou JL. Tracking the best reference genes for RT-qPCR data normalization in filamentous fungi. *BMC Genomics.* 2015; 16:71 <https://doi.org/10.1186/s12864-015-1224-y> PMID: 25757610

12. Vieira A, Talhinhos P, Loureiro A, Duplessis S, Fernandez D, Silva Mdo C, et al. Validation of RT-qPCR reference genes for in planta expression studies in *Hemileia vastatrix*, the causal agent of coffee leaf rust. *Fungal Biol-UK*.2011; 115: 891–901.
13. Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot*.2009; 60: 487–493. <https://doi.org/10.1093/jxb/ern305> PMID: 19264760
14. Rocha-Martins M, Njaine B, Silveira MS. Avoiding Pitfalls of Internal Controls: Validation of reference genes for analysis by qRT-PCR and western blot throughout rat retinal development. *PLOS One*.2012; 7: 9.
15. 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. <https://doi.org/10.1016/j.ab.2005.05.022> PMID: 16054107
16. 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: 12.
17. 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. <https://doi.org/10.1158/0008-5472.CAN-04-0496> PMID: 15289330
18. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. *Biotechnol Lett*.2004; 26: 509–515. PMID: 15127793
19. Silver N, Best S, Jiang J, Thein SL. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol*.2006; 7: 33. <https://doi.org/10.1186/1471-2199-7-33> PMID: 17026756
20. Yin RF, Liu XX, Liu C. Systematic selection of housekeeping genes for gene expression normalization in chicken embryo fibroblasts infected with Newcastle disease virus. *Biochem Bioph Res Co*.2011; 413: 537–540.
21. Kianianmomeni A, Hallmann A. Validation of reference genes for quantitative gene expression studies in *Volvox carteri* using real-time RT-PCR. *Mol Biol Rep*.2013; 40: 6691–6699. <https://doi.org/10.1007/s11033-013-2784-z> PMID: 24057254
22. Humber RA. Synopsis of a revised classification for the Entomophthrales (Zygomycotina). *Mycotaxon*.1989; 34: 441–460.
23. Morgan D. *The Cell Cycle: Principles of Control*. New Science Press, London. 2007, 96–98.
24. Battaglia E, Benoit I, van den Brink J, Wiebenga A, Coutinho PM, Henrissat B, et al. Carbohydrate-active enzymes from the zygomycete fungus *Rhizopus oryzae*: a highly specialized approach to carbohydrate degradation depicted at genome level. *BMC Genomics*.2011; 12:38. <https://doi.org/10.1186/1471-2164-12-38> PMID: 21241472
25. Papierok B. On the occurrence of Entomophthorales (Zygomycetes) in Finland.1. species attacking aphids (Homoptera, Aphididae). *Annal Entomolo Fenn*.1989; 55(2):63–9.
26. Wilding N, B P J. Effects of fungicides on the development of Entomophthora aphidis. *Trans Brit Mycol Soc*.1980; 75: 279–302.
27. Chen C, Ye SD, Wang DQ, Hatting JL, Yu XP. Alginate embedding and subsequent sporulation of in vitro-produced *Conidiobolus thomboides* hyphae using a pressurised air-extrusion method. *Biol Control*.2014; 69: 52–58.
28. Beauvais A, Latgé JP. A simple medium for growing Entomophthoralean protoplasts. *J Invertebr Pathol*.1988; 51: 175–178.
29. Xie F, Xiao P, Chen D, Xu L, Zhang B. miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol Biol*.2012; 80 (1): 75–84.
30. Livak KJ and Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods*, 2001; 25: 402–408. <https://doi.org/10.1006/meth.2001.1262> PMID: 11846609
31. Tang QY, Feng MG. DPS data processing system for practical statistics. Science Press, Beijing. 2002:648.
32. Robledo D, Hernandez-Urcera J, Cal RM, Pardo BG, Sánchez L, Martínez P, et al. Analysis of qPCR reference gene stability determination methods and a practical approach for efficiency calculation on a turbot (*Scophthalmus maximus*) gonad dataset. *BMC Genomics*.2014; 15: 648–663. <https://doi.org/10.1186/1471-2164-15-648> PMID: 25091330

33. Eickbush TH, Eickbush DG. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 2007; 175: 477–485.
34. Jin K, Zhang YJ, Fang WG, Luo Z, Zhou Y, Pei Y. Carboxylate Transporter Gene *JEN1* from the Entomopathogenic fungus *Beauveria bassiana* is involved in conidiation and virulence. *Appl Environ Microbiol.* 2010; 76: 254–263. <https://doi.org/10.1128/AEM.00882-09> PMID: 19854926
35. de Almeida MR, Ruedell CM, Ricachenevsky FK, Sperotto RA, Pasquali G, Fett-Neto AG. Reference gene selection for quantitative reverse transcription-polymerase chain reaction normalization during in vitro adventitious rooting in *Eucalyptus globulus* Labill. *BMC Mol Biol.* 2010; 11: 12.
36. Foldager CB, Munir S, Ulrik-Vinther M, Søballe K, Bünger C, Lind M. Validation of suitable house keeping genes for hypoxia-cultured human chondrocytes. *BMC Mol Biol.* 2009; 10: 8.
37. Botstein D. The yeast cytoskeleton. In: Pringle JR, Broach JR and Jones EW (eds). *The Molecular and Cellular Biology of the Yeast Saccharomyces: Cell Cycle and Cell Biology*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1997, 1–90
38. Teste MA, Duquenne M, François JM, Parrou JL. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Mol Biol.* 2009; 10: 99. <https://doi.org/10.1186/1471-2199-10-99> PMID: 19874630
39. Maroniche G, Sagadín M, Mongelli VC, Truol GA, del Vas M. Reference gene selection for gene expression studies using RT-qPCR in virus-infected planthoppers. *Virology* 2011; 8: 308. <https://doi.org/10.1186/1743-422X-8-308> PMID: 21679431
40. Nailis H, Coenye T, Van Nieuwerburgh F, Deforce D, Nelis HJ. Development and evaluation of different normalization strategies for gene expression studies in *Candida albicans* biofilms by real-time PCR. *BMC Mol Biol.* 2006; 7: 25. <https://doi.org/10.1186/1471-2199-7-25> PMID: 16889665
41. Mascia T, Santovito E, Gallitelli D, Cillo F. Evaluation of reference genes for quantitative reverse-transcription polymerase chain reaction normalization in infected tomato plants. *Mol Plant Pathol.* 2010; 11: 805–816. <https://doi.org/10.1111/j.1364-3703.2010.00646.x> PMID: 21029324