EUROPEAN JOURNAL OF ENTOMOLOGY
ISSN (online): 1802-8829
http://www.eje.cz
Eur. J. Entomol. **116**: 402–412, 2019
doi: 10.14411/eje.2019.041
ORIGINAL ARTICLE

Screening for stable internal reference genes for quantitative PCR analysis of *Wolbachia*-host interactions in whitefly *Bemisia tabaci* (Homoptera: Aleyrodidae)

Xin-Chao LIU and Zheng-Xi LI*

Department of Entomology and Key Laboratory of Pest Monitoring and Green Management, MOA, College of Plant Protection, China Agricultural University, 2 Yuanmingyuan West Road, Beijing 100193, China; e-mails: zxli@cau.edu.cn, 1696679923@qq.com

Key words. Homoptera, Aleyrodidae, *Bemisia tabaci*, reference gene, quantitative PCR, *Wolbachia* transfection

Abstract. Stable reference genes (RGs) determine the reliability of quantitative polymerase chain reaction (qPCR) analyses and it is recommended that different reference genes are used for different types of DNA and tissues. The present study aimed to screen for stable RGs for the qPCR analysis of the immune responses of the whitefly *Bemisia tabaci* to the *Wolbachia w*Mel strain from *Drosophila melanogaster*. A total of eight candidate RGs were evaluated using five different methods, i.e., Coefficient of Variation analysis, GeNorm, NormFinder, BestKeeper and ΔCt. The stability of these RGs was assessed for both genomic DNA (gDNA) and complementary DNA (cDNA). The results indicate that β-actin (*Actin*) and elongation factor 1 alpha (*EF-1α*) were the most stable RGs for gDNA, whereas 18S rRNA (*18S*) and glyceraldehyde phosphate dehydrogenase (*GAPDH*) were the least stable; in contrast, *Actin* and *GAPDH* were the most stable for cDNA, whereas *RPL29* and *ATPase* were the least stable. The effectiveness of the most stable RGs was then validated against the least stable using qPCR analysis of the titre of *w*Mel (gDNA) and the transcriptional responses of the antimicrobial peptide Alo-3-like and the phosphatidylinositol-bisphosphate 3-kinase catalytic subunit delta isoform (cDNA) to *w*Mel transfection. The results support the notion that reliable RGs are essential for a qPCR analysis of samples of both gDNA and cDNA.

INTRODUCTION

Real-time quantitative PCR (qPCR) is a commonly used technique for gene quantitation at both genomic DNA (gDNA) and complementary DNA (cDNA) levels, which recently has been used in many studies (Artico et al., 2010; Derveaux et al., 2010; Hindson et al., 2013). Nevertheless, the reliability of qPCR is affected by many different factors (Bustin, 2000; Derveaux et al., 2010) and the use of reference genes (RGs) can greatly improve the accuracy of qPCR results by standardizing or normalizing experimental data from different developmental stages, organs or tissues (Artico et al., 2010; Arya et al., 2017). Screening for stable RGs is reported for different species of insects, such as *Aphis glycines* (Bansal et al., 2012), *Nilaparvata lugens* (Yuan et al., 2014), *Bactericera cockerelli* (Ibanez & Tamborindeguy, 2016) and *Lipaphis erysimi* (Koramutla et al., 2016). The RGs have to be carefully evaluated for particular situations, as it is unlikely that the same RGs (non-evaluated) are generally suitable (Thellin et al., 1999; Vandesompele et al., 2002; Ponton et al., 2011; Zhou & Li, 2016). Several algorithms have been developed for evaluating the stability of RGs, including the Coefficient of Variation (Boda et al., 2009), GeNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) and the ΔCt method (Silver et al., 2006). These algorithms have proved to be helpful in the screening for stable RGs in a variety of organisms (Freitag et al., 2018; Meng et al., 2019; Xie et al., 2019).

Whitefly *Bemisia tabaci* is a destructive agricultural pest and causes serious damage to vegetables, ornamental plants and other crops worldwide (De Barro et al., 2011; Navas-Castillo et al., 2011). *B. tabaci* is often infected with *Wolbachia* (Li et al., 2007), a group of maternally inherited intracellular Gram-negative bacteria frequently found in arthropods and filarial nematodes (Hilgenboecker et al., 2008; Zug & Hammerstein, 2012). *Wolbachia* can manipulate host reproduction in different ways, block the transmission of human diseases and control insect pests (Bourtzis & O'Neill, 1998; Stouthamer et al., 1999; Werren et al., 2008). However, the interactions between *Wolbachia* and their hosts are still poorly understood (Lemaitre & Hoffmann, 2007). In our research group, several exog-

* Corresponding author: e-mail: zxli@cau.edu.cn

enous *Wolbachia* strains were successfully transferred into *B. tabaci* by microinjection and transinfected isofemale lines were established in the laboratory (Zhong & Li, 2014; Hu & Li, 2015; Zhou & Li, 2016). Recently, we sequenced the transcriptomes of *B. tabaci* in response to *Wolbachia* transfection using the RNA-seq technique, which is necessary for a detailed investigation of the host-*Wolbachia* interaction, including qPCR analysis of the titre of *Wolbachia* and expression levels of differentially expressed genes (DEGs) and functional analysis of candidate genes via RNA interference (RNAi). All of these studies require the selection of suitable RGs. In a previous study, 15 candidate RGs were evaluated for *B. tabaci* using the algorithms geNorm and Normfinder (Li et al., 2013), which revealed that the stability of RGs was greatly affected by both biotic and abiotic factors, and that different RGs should be used depending on the species and conditions. The purpose of the present study was to screen for suitable RGs for qPCR analysis of both gDNA and cDNA from *B. tabaci* artificially transfected with exogenous *Wolbachia*. Based on our studies, different RGs are needed for determining the titre of *Wolbachia* or level of expression of functional genes in different generations when *B. tabaci* is transfected with the *w*Mel *Wolbachia* strain from the fruit fly *Drosophila melanogaster*.

MATERIAL AND METHODS

Insect rearing and transfection through microinjection

The whitefly *B. tabaci* (Q cryptic species) was collected in tomato greenhouses of Jinan Academy of Agricultural Sciences, Shandong, China, and then maintained on pot-grown plants of cotton in a laboratory (14L : 10D, RH 60–80%, 27 ± 1°C). The 4th-instar nymphs (pseudopupae) of *B. tabaci* were artificially transinfected with the *w*Mel *Wolbachia* strain by microinjection as described by Zhou & Li (2016). The *w*Mel strain was isolated from *D. melanogaster* using the Percoll density-gradient centrifugation method. The primer pair 81F/691R (Table 1) was used to verify the existence of the *w*Mel strain, and on each occasion a volume of 46 nL *Wolbachia* suspension was injected. After transfection, isofemale lines were established and samples were collected from subsequent generations (G).

Extraction of genomic DNA (gDNA) and synthesis of complementary DNA (cDNA)

Twenty fresh whiteflies (adults) were used for extraction of gDNA from samples of G_0, G_2, G_4, G_6 and G_7 individuals using the KAC method as described by Zhong & Li (2013) and the purity and concentration of gDNA were checked using a NanoDrop ND-2000 (Thermo Scientific, Wilmington, DE). Samples of thirty fresh whiteflies were used for the extraction of total RNA in a 1.5-mL centrifuge tube containing TRIZol reagent (TransGen Biotech, Beijing, China), which were from G_0, G_1 and G_4 of the transinfected isofemale lines, with the wild type as the control. The first-strand cDNA was synthesized from 500 ng of total RNA using the PrimeScript™ RT reagent Kit (TaKaRa, Beijing, China) according to the supplier's instructions.

Selection of RGs

Eight RGs were selected: heat shock protein 20 (*HSP20*), heat shock protein 70 (*HSP70*), glyceraldehyde phosphate dehydrogenase (*GAPDH*), 60S ribosomal protein L29 (*RPL29)*, β-actin (*Actin*), 18S rRNA (*18S*), elongation factor 1 alpha (*EF-1a*) and adenosine triphosphate enzyme (*ATPase*) (Table 1). Most of these genes have been used as internal references for real-time qPCR analysis in previous studies (Zhou et al., 1998; Li et al., 2013; Zhou & Li, 2016). These RGs were evaluated against both gDNA and cDNA in the present study, but *HSP20* and *HSP70* were not used for cDNA samples as it is likely they are induced by *Wolbachia* transfection.

Quantitative PCR analysis

The specificity of the primers for the eight RGs (Table 1) was examined using PCR in a total reaction volume of 25 µL containing 2 µL DNA template, 1 µL of each primer (10 µM), 2.5 µL 10 × PCR buffer (TransGen Biotech, Beijing, China), 2 µL dNTPs (2.5 mM), 0.5 µL *Taq* DNA polymerase (5.0 U/µL) and 16 µL ddH_2O. The thermocycling program was: 94°C for 5 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 30 s and a final 10 min extension at 72°C. PCR products were analysed using 1% agarose gel electrophoresis. For the qPCR analysis, the DNA templates (gDNA or cDNA) were successively diluted by 5^0, 5^1, 5^2, 5^3 and 5^4 times for construction of standard curves and melting curves. A suitable pair of primers for qPCR analysis should have a correlation coefficient (R^2) > 0.99, an amplification efficiency (E) > 90% and a unimodal melting curve (Livak & Schmittgen, 2001). The reaction was performed in a total volume of 20 µL containing 10 µL AceQ® SYBR® Green Master Mix (Vazyme, Nanjing, China), 0.4 µL of each primer (10 µM), 1 µL DNA template and 8.2 µL ddH_2O on ABI 7500 platform (Applied Biosystems, Foster City, California, USA); the primer pair *wsp*Q384/*wsp*Q513 (Table 1) was used to detect the titre of *w*Mel strain in *B. tabaci* after transfection. The thermocycling program was 50°C for 2 min, 95°C for 5 min, 40 cycles of 95°C for 10 s and 60°C for 30 s. The program used for determining the melting curve was 95°C for 15 s, 60°C for 60 s and 95°C for 15 s. DNase/RNase-free water was used as the negative control. Each treatment was performed in triplicate.

Stability of RGs

The cycle threshold (*Ct*) values of RGs for different generations were compared in order to evaluate their stability. The working concentration of gDNA for the qPCR analysis was 20 ng/µL and that of cDNA was 100 ng/µL. The stability of RGs was evaluated using four different algorithms: GeNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004) and ΔCt method (Silver et al., 2006). The four algorithms rated the stability according to different variables: GeNorm provides an *M*-value based on the average paired expression ratio. The lower the *M*-value, the more stable the expression: $M \leq 1.5$ was considered to be stable. NormFinder calculates stability based on the inter- and intra-group variance of each candidate RG; the lower the NormFinder value, the more stable the RG. BestKeeper determines the stability of RG according to the standard deviation (SD): The higher the SD, the more unstable the expression. ΔCt method performs pairwise multiple comparisons between the expression levels and identifies the most stable RGs. Each algorithm sorted RGs according to their stability and assigned the RGs a series of consecutive integers (starting at one) and the geometric average (geomean) of the weights based on the four algorithms was then calculated for each RG. The RG with the lowest geomean value was considered to be the most stable and the RGs were ranked accordingly. Subsequently, the Coefficient of Variation (CV) was used to measure the population variance of each gene as it is the only method that is not affected by other factors (Boda et al., 2009). Here a threshold of CV = 50% was defined: genes with a CV value above this threshold were

considered highly variable and excluded from further analyses. Finally, a corrected ranking of the RGs was obtained.

Validation of RGs

The reliability of candidate RGs was validated by a qPCR analysis of the titre of *w*Mel (gDNA) and the transcriptional response of two genes (antimicrobial peptide Alo-3-like; phosphatidylinositol-bisphosphate 3-kinase catalytic subunit delta isoform, *Pl3K*) to *w*Mel transfection (cDNA) (the primer sequences are included in Table 1). Alo-3-like is an antimicrobial peptide involved in immune response of the host to *w*Mel transfection, and *Pl3K* is involved in the immune-related signalling pathway in *B. tabaci* based on our transcriptome sequencing data (Tables S1 and S2; Figs S1 and S2). Here we used the two most stable and two least stable RGs based on their geomean values to normalize the qPCR analysis before the results were compared. The procedures for qPCR analysis were the same as described above.

Data analysis

The statistical differences were analysed using One-way ANOVA followed by Student Newman Keuls (SNK) test at 0.05, 0.01 and 0.001 levels of probability on SPSS v.20.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

Specificity, amplification efficiency and cycle thresholds of RG primers

The specificity of the RG primers (Table 1) was determined using PCR and agarose gel electrophoresis, which indicated that all primer pairs resulted in single specific bands of expected size. In addition, the melting curves of all the primer pairs had single peaks, and the correlation coefficients and amplification efficiencies of the standard curves were all within the range of $R^2 > 0.99$ and $90\% < E < 110\%$ (Table 1; Fig. S3). Moreover, the determination of RG using qPCR revealed that the *Ct*-value of *18S* was the lowest (the log of starting template copies) for both gDNA and cDNA templates, while that of *GAPDH* was the highest for cDNA (Fig. 1) and that of *RPL29* was the most variable (Fig. 1A).

Stability of candidate RGs

Based on qPCR analysis of gDNA and cDNA samples taken from transinfected *B. tabaci* adults and the different algorithms (GeNorm, NormFinder, BestKeeper or ΔCt method), stabilities were assigned to each RG and the geomeans of these stabilities calculated, which were used to determine their combined ranking (Table 2). The results indicated that *ATPase* and *EF-1α* were the two most stable RGs for gDNA templates and *RPL29* and *GAPDH* the two least stable RGs; in contrast, *18S* and *Actin* were the two most stable RGs for cDNA templates, whereas *EF-1α* and *ATPase* were the two least stable RGs.

The CV analysis was the most appropriate method for identifying the most variable RGs, which were excluded from further assessments. The analysis showed that *RPL29* and *EF-1α* were the most variable RGs for both gDNA and cDNA samples (Table 3). After removing *RPL29* and *EF-1α* from the ranking analysis, a corrected stability ranking value was generated for each RG using the different algorithms and finally all RGs were ranked according to their geomean values (Table 4). The results indicate that *Actin* and *EF-1α* were the two most stable RGs for gDNA samples, whereas *Actin* and *GAPDH* were the two most stable for cDNA samples. Interestingly, *18S*/*GAPDH* and *18S*/*RPL29*/*ATPase* were identified as the least stable RGs for both gDNA and cDNA samples.

Table 1. Primer sequences used in this study.

Gene name	GenBank acc. no.	Primer sequence (5'→3')	Expected size (bp)	gDNA		cDNA		References
				E%	R^2	*E*%	R^2	
HSP20	EU934239	F: AAGAAGTCAGCGTGAAAGTCG R: GTACCTCCTAGTGAAAGATCGG	107	103.8	0.9986	—	—	Li et al., 2013
HSP70	EU934240	F: AGCACTCCGGCGTCTACG R: CGAACCTGGCACGGGACAC	134	108.4	0.9966	—	—	Li et al., 2013
GAPDH	JU470454	F: GGACACGGAAAGCCATACCAG R: ACCACCGCTACCCAAAAGACC	166	102.7	0.9971	109	0.9965	Li et al., 2013
RPL29	EE596314	F: TCGGAAAATTACCGTGAG R: GAACTTGTGATCTACTCCTCTCGTG	144	107.4	0.9984	92.9	0.9939	Li et al., 2013
Actin	AF071908	F: TCTTCCAGCCATCCTTCTTG R: CGGTGATTTCCTTCTGCATT	174	96.75	0.996	102.6	0.9999	Zhou & Li, 2016
18S	U20401	F: CGGCTACCACATCCAAGGAA R: GCTGGAATTACCGCGGCT	187	94.62	0.9941	104.9	0.9913	Li et al., 2013
EF-1a	EE600682	F: ATGCCATGGTCAAGGGATGG R: ACATCTGGAGTGGCAATCGG	134	97.69	0.9951	110	0.9971	Designed by the authors
ATPase	JU470453	F: CGTTACTCCCCTCTTGGCTG R: CAGAAGACGGCGATTCGAGA	122	108.5	0.9943	107.7	0.996	Designed by the authors
wsp	KU870673	81F: TGGTCCAATAAGTGATGAAGAAAC 691R: AAAAATTAAACGCTACTCCA	632	—	—	—	—	Zhou et al., 1998
wspQ384 *wspQ513*	KU870673	F: TGGAACCCGCTGTGAATGAT R: GCACCATAAGAACCGAAATAACG	130	93.76	0.9976	95.46	0.9957	Zhou & Li, 2016
Pl3K	LOC109034225	F: TGTTGCAACGTATGTGCTTGG R: TAAGAATTGCCGCCTGGACT	146	97.51	0.9987	93.14	0.9989	Designed by the authors
AMP	LOC109033344	F: TTCCAAGCCAAAACCAACCAC R: CCGTCTGAGAGGCAATTCGAT	110	106.5	0.9961	108.2	0.9916	Designed by the authors

[a] *E*% – amplification efficiency; R^2 – correlation coefficient.

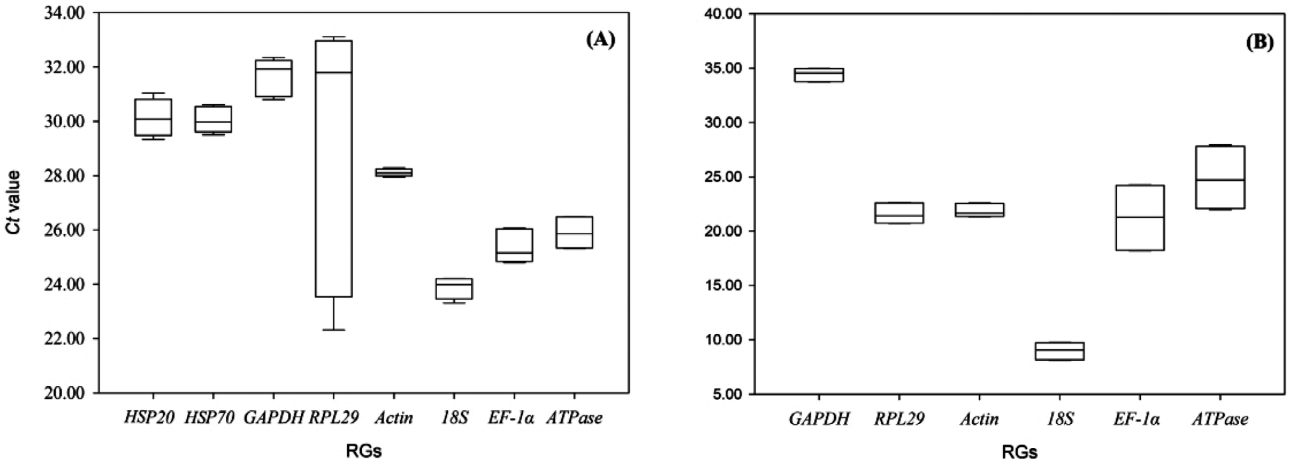


Fig. 1. A boxplot of the amplification efficiencies of the reference genes (RGs) based on the cycle threshold (*Ct*) of the qPCR analysis. The *Ct* values are those for gDNA (A) recorded in G_0, G_2, G_4, G_6 and G_7 of the transinfected isolines, and for cDNA (B) in G_0, G_1, and G_4 of the transinfected isolines and of the wild type. The line in the middle of the box represents the median value. The box indicates the 25% and 75% percentiles, and the extremities indicate the maximum and minimum values. The bigger the box, the greater the variation.

Validation of stable RGs

The qPCR analysis of the relative titre (gDNA) of the *Wolbachia w*Mel strain was normalized by using a combination of the two most stable (*Actin* and *EF-1α*) and two least stable (*18S* and *GAPDH*) RGs. The results revealed that the measurement of *Wolbachia* titre was indeed affected by the nature of the RG in all of the five generations assessed in this study (Fig. 2). Specifically, the relative titres normalized by unstable RGs varied dramatically between generations, especially in G_6 for which the titre was several-fold higher than in that recorded in other generations. In contrast, the titres normalized by stable RGs looked more normal.

The qPCR analysis of the relative expression (cDNA) of two immune-related genes (*AMP* and *Pl3K*) was also normalized using a combination of the two most stable (*Actin* and *GAPDH*) and two least stable (*RPL29*/*ATPase*) RGs. The results indicate that the relative level of expression of *AMP* (Alo-3-like gene) when normalized by unstable RGs was very high in the wild type, while there were no significant differences between G_0 and G_4 and the control (G_1) (Fig. 3A). In the case of *Pl3K*, the relative level of expression normalized by unstable RGs were exceptionally high in the wild type and in G_0 of the transinfected lines (Fig. 3B). Moreover, the relative levels of expression of both the Alo-3-like gene and *Pl3K* when normalized using stable RGs were generally consistent with the transcriptome data, whereas those normalized using unstable RGs were not (Tables S1 and S2).

DISCUSSION

Our studies confirmed that the reference genes significantly affected the accuracy of qPCR analysis of both gDNA and cDNA samples. Moreover, the selection of candidate RGs based on the algorithms used in this study seemed very effective: the scoring and ranking procedures ensured the selection of suitable RGs for the qPCR analysis of both gDNA and cDNA templates. Our results support

Table 2. Rankings of selected internal reference genes based on different algorithms and a comprehensive calculation (geomean)[a].

Ref. gene	GeNorm *M*-value	NormFinder Stability	BestKeeper SD ± CP	ΔCt method	Geomean	Ranking
			gDNA			
ATPase	0.253(1)	0.0883(2)	0.458(5)	0.986(1)	1.77	1
EF-1α	0.253(1)	0.0882(1)	0.503(6)	0.988(2)	1.86	2
HSP20	0.346(2)	0.121(3)	0.393(4)	0.998(3)	2.91	3
Actin	0.528(4)	0.475(5)	0.074(1)	1.05(5)	3.16	4
HSP70	0.474(3)	0.154(4)	0.317(3)	1.01(4)	3.46	5
18S	0.613(6)	0.608(7)	0.246(2)	1.19(7)	4.92	6
GAPDH	0.578(5)	0.555(6)	0.504(7)	1.14(6)	5.95	7
RPL29	1.56(7)	3.042(8)	3.70(8)	3.94(8)	7.37	8
			cDNA			
18S	0.636(2)	0.194(2)	0.623(4)	1.131(1)	1.86	1
Actin	0.296(1)	0.621(4)	0.493(2)	1.29(3)	2.21	2
GAPDH	0.296(1)	0.847(5)	0.373(1)	1.25(5)	2.23	3
RPL29	0.841(3)	0.194(1)	0.604(3)	1.01(2)	2.37	4
ATPase	1.49(4)	0.528(3)	0.805(5)	1.21(4)	3.94	5
EF-1α	1.70(5)	1.78(6)	2.55(6)	2.29(6)	5.73	6

[a] The gDNA and cDNA samples are taken from transinfected *B. tabaci* adults. The final rankings are determined by geomean values calculated based on the algorithms GeNorm, NormFinder, BestKeeper and ΔCt method. The numbers in brackets indicate the ranking values based on individual algorithms.

Table 3. Stability of internal reference genes based on a Coefficient of Variation (CV) analysis.

Ref. gene	CV%	Ranking
	gDNA	
Actin	11.71	1
HSP70	26.26	2
18S	28.99	3
EF-1α	33.5	4
ATPase	33.56	5
HSP20	35.4	6
GAPDH	39.89	7
RPL29	191.27	8
	cDNA	
Actin	28.27	1
GAPDH	34.17	2
RPL29	43.62	3
ATPase	43.77	4
18S	43.91	5
EF-1α	106.11	6

the notion that the use of suitable RGs is essential for a qPCR analysis.

The *w*Mel *Wolbachia* strain from *D. melanogaster* is able to establish and induce strong cytoplasmic incompatibility (CI) in *B. tabaci* (Zhou & Li, 2016), an important worldwide agricultural pest, which indicates it might be possible to use it to control this pest. Nevertheless, the titre of *w*Mel fluctuated during transgenerational transmission in *B. tabaci*. For instance, the titre of *w*Mel was extremely low after transfection in $G_{1–2}$ and G_2 and higher in $G_{4–6}$. This observation prompted us to speculate that the titre of *Wolbachia* might be modulated by the interaction between the host and *Wolbachia*. Obviously, clarifying the quantitative relationships between the titre of *Wolbachia* and the levels of expression of the candidate genes involved in the host-*Wolbachia* interaction would help us understand the mechanisms underlying the fluctuations in the *Wolbachia* titre, especially when the titre is potentially related to the pest control capability of *Wolbachia* (Breeuwer & Werren, 1993; Noda et al., 2001). Therefore, it is essential to quantitatively measure both the titre of *Wolbachia* (gDNA) and the levels of expression (cDNA) of functional genes, and thus the screening for stable RGs for both gDNA and cDNA samples is a prerequisite for a reliable qPCR analysis. In the present study, a combination of algorithms was used to evaluate the candidate RGs, which circumvented the drawbacks of using a single algorithm, and the results indicate that use of the stable RGs rather than unstable RGs could result in more reliable qPCR data. Indeed, the stable RGs specifically identified in this study were successfully used in the qPCR analysis of both gDNA and cDNA samples from *B. tabaci* artificially transfected with *Wolbachia*, which will facilitate the study of the host insect-*Wolbachia* interaction.

Different algorithms were developed for selecting RGs in the past, but as mentioned above, each algorithm has advantages and disadvantages, which potentially generate biased results. The use of a combination of algorithms may hopefully counteract any bias. For instance, GeNorm and BestKeeper are based on paired comparisons, and the selection of the most appropriate RGs are based on the change in the expression of genes, which does not eliminate the influence of co-regulation; on the other hand, NormFinder and the ΔCt method can neutralize this influence. Here we used geomeans to aggregate the RGs and generated an integrated ranking for the candidate RGs and the most variable RGs based on the CV analysis were not used in this study. Our corrected results indicate that the combined analysis worked well in selecting stable RGs.

The overall design of our experiments was based on the results of our previous study of the *w*Mel *Wolbachia* strain after transfection by microinjection, which revealed the titre of *w*Mel differed greatly in the different generations. As noted above, the change in titre was most marked from G_0 to G_6; hence, the samples (both gDNA and cDNA) were collected during this period of time. Hopefully, the change in the expression of the functional genes selected and the titre of *Wolbachia* were effectively detected. In addition, the selection of immune-related genes for validating the RGs was based on those genes that were identified by the transcriptome sequencing, which revealed that an active immune response was induced after infection with *Wolbachia*, and thus they will be functionally analysed in the next step. The antimicrobial peptide Alo-3 was greatly

Table 4. Corrected rankings of reference genes.

Ref. gene[a]	GeNorm *M*-value	NormFinder Stability	BestKeeper SD ± CP	ΔCt method	Geomean	Ranking
			gDNA			
Actin	0.528(4)	0.231(1)	0.138(1)	0.536(4)	2	1
EF-1α	0.253(1)	0.271(3)	0.503(6)	0.518(1)	2.06	2
HSP20	0.346(2)	0.268(2)	0.394(4)	0.534(3)	2.63	3
ATPase	0.253(1)	0.279(5)	0.458(5)	0.532(2)	2.66	4
HSP70	0.474(3)	0.276(4)	0.317(3)	0.549(5)	3.66	5
18S	0.631(6)	0.462(7)	0.298(2)	0.685(7)	4.92	6
GAPDH	0.578(5)	0.402(6)	0.504(7)	0.637(6)	5.95	7
			cDNA			
Actin	0.297(1)	0.444(3)	0.394(2)	0.721(2)	1.86	1
GAPDH	0.297(1)	0.614(4)	0.373(1)	0.923(4)	2	2
18S	0.636(2)	0.158(1)	0.704(4)	0.812(3)	2.21	3
RPL29	0.774(3)	0.358(2)	0.604(3)	0.716(1)	2.21	3
ATPase	0.841(4)	0.738(5)	0.805(5)	0.973(5)	4.4	4

[a] *RPL29* and *EF-1α* are excluded from gDNA and cDNA samples, respectively, in the corrected ranking analysis based on their highest CV values (Table 3).

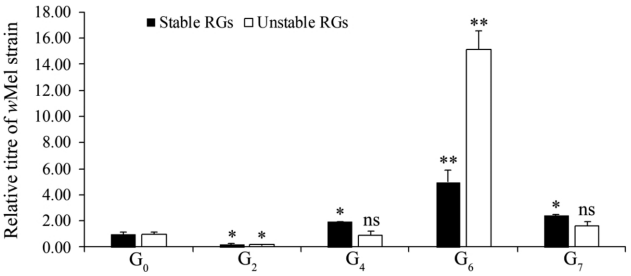


Fig. 2. Validation of stable RGs based on the qPCR analysis of the relative titre (gDNA) of *w*Mel normalized by a combination of the two most stable RGs (*Actin* and *EF-1α*) and two least stable RGs (*18S* and *GAPDH*). The data are the means ± standard deviation of the three replicates. G_0 is used as the control in the statistical analysis using One-way ANOVA followed by SNK test at 0.05, 0.01 and 0.001 levels of probability (ns – not significant; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$). G_0 – adults that were microinjected with the *w*Mel strain; G_2 – 2nd generation after transfection; G_4 – 4th generation after transfection; G_6 – 6th generation after transfection; G_7 – 7th generation after transfection.

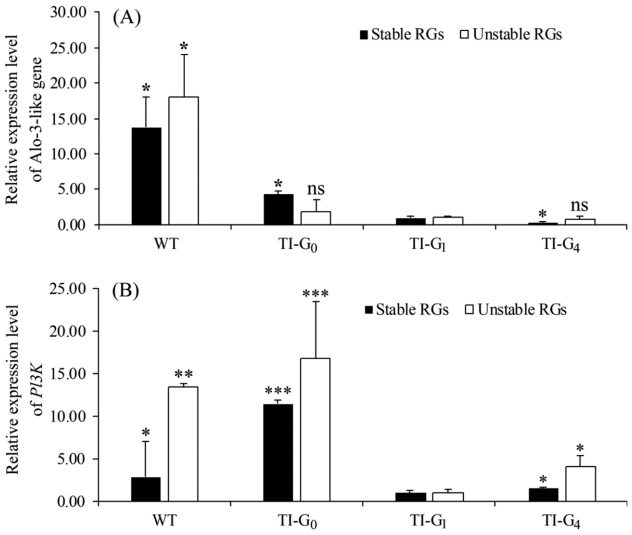


Fig. 3. Validation of stable RGs by qPCR analysis of the relative expression levels of *AMP*/*PI3K* normalized by the combination of the two most stable RGs (*Actin* and *GAPDH*) and two least stable RGs (*RPL29* and *ATPase*). (A) *AMP*, antimicrobial peptide Alo-3-like; (B) *PI3K*, phosphatidylinositol-bisphosphate 3-kinase catalytic subunit delta isoform. The data are means ± standard deviations of three replicates. $TI\text{-}G_1$ is used as the control in the statistical analysis using One-way ANOVA followed by SNK test at 0.05, 0.01 and 0.001 levels of probability (ns – not significant; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$). WT – wild-type; TI – transfected isoline; G_0 – adults emerged after microinjection with the *w*Mel strain; G_1 – 1st generation; G_4 – 4th generation.

down-regulated, while the *PL3K* involved in the immune-related signalling pathway was significantly up-regulated (Tables S1 and 2), indicating that these immune genes might play a role in the modulation of the titre of *w*Mel and in the immune reaction of the host to the *Wolbachia*. The results obtained in the present study confirmed the reliability of the transcriptome sequencing.

The RGs selected here are extensively used in other studies. Taking *Actin* as an example, it encodes a major component of the protein scaffold that supports the cell and determines its shape. *Actin* is well expressed in most types of cells and is widely used as an RG in the qPCR analysis of many different insects, including whitefly (Zhou & Li, 2016), desert locust (Van Hiel et al., 2009), European honey bee (Scharlaken et al., 2008) and two species of Collembola (Van Hiel et al., 2009). Our results indicate that *Actin* and *EF-1α* are the most stable RGs for gDNA samples of *B. tabaci*, while *Actin* and *GAPDH* were the top two RGs for cDNA samples. Similarly, *EF-1α* is used as an RG for salmon (Olsvik et al., 2005), humans (Shen et al., 2010; Silver et al., 2006), Orthoptera (Van et al., 2009) and Hymenoptera (Horňáková et al., 2010). Nevertheless, *GAPDH*, the commonly used RG in many species is reported to be unstable in some studies (Tong et al., 2009; Martins et al., 2016). It is also noteworthy that RPL29 is a recommended reference gene for several conditions, including different host plants, TYLCV infection, different developmental stages and tissues (Li et al., 2013). However, it is not recommended for use in studies on the artificial transfection of exogenous *Wolbachia* into *B. tabaci*, which further supports the notion that different reference genes should be used in different systems.

CONCLUSIONS

We developed a set of suitable RGs for studying the host insect-*Wolbachia* interaction using the agricultural pest *B. tabaci* artificially transfected with an exogenous strain of *Wolbachia*. The RGs are suitable for both gDNA and cDNA templates. A more reliable qPCR analysis of the titre of *Wolbachia* and the expression of functional genes will increase our understanding of the infection dynamics of *Wolbachia* in this pest insect, which might provide a scientific basis for the development of a CI-based control strategy for this pest.

ACKNOWLEDGEMENTS. This work is supported by the National Natural Science Foundation of China (Grant nos. 31772169 and 31972267).

REFERENCES

Andersen C.L., Jensen J.L. & Ørntoft T.F. 2004: Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. — *Cancer Res*. **64**: 5245–5250.

Artico S., Nardeli S.M., Brilhante O., Grossi-de-Sa M.F. & Alves-Ferreira M. 2010: Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. — *BMC Plant Biol*. **10**: 49, 12 pp.

Arya S.K., Jain G., Upadhyay S.K., Singh H., Dixit S. & Verma P.C. 2017: Reference genes validation in *Phenacoccus solenopsis* under various biotic and abiotic stress conditions. — *Sci. Rep*. **7**: 13520, 12 pp.

Bansal R., Mamidala P., Mian M.R., Mittapalli O. & Michel A.P. 2012: Validation of reference genes for gene expression studies in *Aphis glycines* (Hemiptera: Aphididae). — *J. Econ. Entomol*. **105**: 1432–1438.

Boda E., Pini A., Hoxha E., Parolisi R. & Tempia F. 2009: Selection of reference genes for quantitative real-time RT-PCR studies in mouse brain. — *J. Mol. Neurosci*. **37**: 238–253.

Bourtzis K. & O'Neill S. 1998: "*Wolbachia*" Infections and arthropod reproduction. — *Bioscience* **48**: 287–293.

Breeuwer J.A. & Werren J.H. 1993: Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. — *Genetics* **135**: 565–574.

Bustin S.A. 2000: Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. — *J. Mol. Endocrinol*. **25**: 169–193.

De Barro P.J., Liu S.S., Boykin L.M. & Dinsdale A.B. 2011: *Bemisia tabaci*: a statement of species status. — *Annu. Rev. Entomol*. **56**: 1–19.

Derveaux S., Vandesompele J. & Hellemans J. 2010: How to do successful gene expression analysis using real-time PCR. — *Methods* **50**: 227–230.

Freitag D., Koch A., McLean A.L., Kalff R. & Walter J. 2018: Validation of reference genes for expression studies in human meningiomas under different experimental settings. — *Mol. Neurobiol*. **55**: 5787–5797.

Hilgenboecker K., Hammerstein P., Schlattmann P., Telschow A. & Werren J.H. 2008: How many species are infected with *Wolbachia*? A statistical analysis of current data. — *FEMS Microbiol. Lett.* **281**: 215–220.

Hindson C.M., Chevillet J.R., Briggs H.A., Gallichotte E.N., Ruf I.K., Hindson B.J. & Tewari M. 2013: Absolute quantification by droplet digital PCR versus analog real-time PCR. — *Nat. Methods* **10**: 1003.

Horňáková D., Matoušková P., Kindl J., Valterová I. & Pichová I. 2010: Selection of reference genes for real-time polymerase chain reaction analysis in tissues from *Bombus terrestris* and *Bombus lucorum* of different ages. — *Anal. Biochem.* **397**: 118–120.

Hu H.Y. & Li Z.X. 2015: A novel *Wolbachia* strain from the rice moth *Corcyra cephalonica* induces reproductive incompatibility in the whitefly *Bemisia tabaci*: sequence typing combined with phenotypic evidence. — *Environ. Microbiol. Rep.* **7**: 508–515.

Ibanez F. & Tamborindeguy C. 2016: Selection of reference genes for expression analysis in the potato psyllid, *Bactericera cockerelli*. — *Insect Mol. Biol.* **25**: 227–238.

Koramutla M.K., Aminedi R. & Bhattacharya R. 2016: Comprehensive evaluation of candidate reference genes for qRT-PCR studies of gene expression in mustard aphid, *Lipaphis erysimi* (Kalt). — *Sci. Rep.* **6**: 25883, 10 pp.

Lemaitre B. & Hoffmann J. 2007: The host defense of *Drosophila melanogaster*. — *Annu. Rev. Immunol.* **25**: 697–743.

Li Z.X., Lin H.Z. & Guo X.P. 2007: Prevalence of *Wolbachia* infection in *Bemisia tabaci*. — *Curr. Microbiol.* **54**: 467–471.

Li R., Xie W., Wang S., Wu Q., Yang N., Yang X. & Zhou X. 2013: Reference gene selection for qRT-PCR analysis in the sweetpotato whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae). — *PLoS ONE* **8**: e53006, 8 pp.

Livak K.J. & Schmittgen T.D. 2001: Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. — *Methods* **25**: 402–408.

Martins P.K., Mafra V., De Souza W.R., Ribeiro A.P., Vinecky F., Basso M.F. & Molinari H.B.C. 2016: Selection of reliable reference genes for RT-qPCR analysis during developmental stages and abiotic stress in *Setaria viridis*. — *Sci. Rep*. **6**: 28348, 10 pp.

Meng H., Yang Y., Gao Z.H. & Wei J.H. 2019: Selection and validation of reference genes for gene expression studies by RT-PCR in *Dalbergia odorifera*. — *Sci. Rep*. **9**: 3341, 10 pp.

Navas-Castillo J., Fiallo-Olivé E. & Sánchez-Campos S. 2011: Emerging virus diseases transmitted by whiteflies. — *Ann. Rev. Phytopathol*. **49**: 219–248.

Noda H., Koizumi Y., Zhang Q. & Deng K. 2001: Infection density of *Wolbachia* and incompatibility level in two planthopper species, *Laodelphax striatellus* and *Sogatella furcifera*. — *Insect Biochem. Mol. Biol*. **31**: 727–737.

Olsvik P.A., Lie K.K., Jordal A.E.O., Nilsen T.O. & Hordvik I. 2005: Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. — *BMC Mol. Biol*. **6**: 21, 9 pp.

Pfaffl M.W., Tichopad A., Prgomet C. & Neuvians T.P. 2004: Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. — *Biotechnol. Lett.* **26**: 509–515.

Ponton F., Chapuis M.P., Pernice M., Sword G.A. & Simpson S.J. 2011: Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. — *J. Insect Physiol.* **57**: 840–850.

Scharlaken B., de Graaf D.C., Goossens K., Brunain M., Peelman L.J. & Jacobs F.J. 2008: Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial challenge. — *J. Insect Sci*. **8**: 33, 10 pp.

Shen Y., Li Y., Ye F., Wang F., Lu W. & Xie X. 2010: Identification of suitable reference genes for measurement of gene expression in human cervical tissues. — *Anal. Biochem.* **405**: 224–229.

Silver N., Best S., Jiang J. & Thein S.L. 2006: Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. — *BMC Mol. Biol.* **7**: 33, 9 pp.

Stouthamer R., Breeuwer J.A. & Hurst G.D. 1999: *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. — *Annu. Rev. Microbiol*. **53**: 71–102.

Thellin O., Zorzi W., Lakaye B., De Borman B., Coumans B., Hennen G. & Heinen E. 1999: Housekeeping genes as internal standards: use and limits. — *J. Biotechnol*. **75**: 291–295.

Tong Z., Gao Z., Wang F., Zhou J. & Zhang Z. 2009: Selection of reliable reference genes for gene expression studies in peach using real-time PCR. — *BMC Mol. Biol*. **10**: 71, 13 pp.

Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. & Speleman F. 2002: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. — *Genome Biol*. **3**: research0034-1, 12 pp.

Van Hiel M.B., Van Wielendaele P., Temmerman L., Van Soest S., Vuerinckx K., Huybrechts R. & Simonet G. 2009: Identification and validation of housekeeping genes in brains of the desert locust *Schistocerca gregaria* under different developmental conditions. — *BMC Mol. Biol.* **10**: 56, 10 pp.

Werren J.H., Baldo L. & Clark M.E. 2008: *Wolbachia*: master manipulators of invertebrate biology. — *Nat. Rev. Microbiol.* **6**: 741–751.

Xie L.H., Quan X., Zhang J., Yang Y.Y., Sun R.H., Xia M.C. & Yang L.R. 2019: Selection of reference genes for real-time quantitative PCR normalization in the process of *Gaeumannomyces graminis* var. *tritici* infecting wheat. — *Plant Pathol. J.* **35**: 11–18.

Yuan M., Lu Y., Zhu X., Wan H., Shakeel M., Zhan S. & Li J. 2014: Selection and evaluation of potential reference genes for gene expression analysis in the brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae) using reverse-transcription quantitative PCR. — *PLoS ONE* **9**: e86503, 10 pp.

Zhong Y. & Li Z.X. 2013: Influences of tetracycline on the reproduction of the B biotype of *Bemisia tabaci* (Homoptera: Aleyrodidae). — *Appl. Entomol. Zool*. **48**: 241–246.

Zhong Y. & Li Z.X. 2014: Bidirectional cytoplasmic incompatibility induced by cross-order transfection of *Wolbachia*: Implications for control of the host population. — *Microb. Ecol.* **68**: 463–471.

Zhou X.F. & Li Z.X. 2016: Establishment of the cytoplasmic incompatibility-inducing *Wolbachia* strain *w*Mel in an important agricultural pest insect. — *Sci. Rep.* **6**: 39200, 9 pp.

Zhou W., Rousset F. & O'Neill S. 1998: Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. — *Proc. Biol. Sci.* **265**: 509–515.

Zug R. & Hammerstein P. 2012: Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. — *PLoS ONE* **7**: e38544, 3 pp.

Received August 28, 2019; revised and accepted November 4, 2019
Published online November 15, 2019

Table S1. DEGs involved in the immune-related signalling pathways.

Pathway / Gene ID	Annotated function	GenBank acc. no.	$\log_2$(FC)[a]
Recognition			
gene5108	Peptidoglycan recognition protein 2-like (PGRP)	LOC109032651	–1.96
Imd			
gene1970	RING-box protein 1A	LOC109044484	–1.22
gene7989	RING-box protein 2	LOC109035809	–1.96
gene1881	RING finger protein 207-like	LOC109044343	2.00
gene15250	Mitogen-activated protein kinase kinase kinase (MAP3K)	LOC109043842	–1.03
gene13482	Mitogen-activated protein kinase kinase kinase (MAP3K)	LOC109041956	2.26
SP			
gene3222	Putative serine protease K12H4.7	LOC109030575	–2.15
gene8224	Putative serine protease F56F10.1	LOC109036101	–2.15
gene9459	Venom serine protease Bi-VSP-like	LOC109037519	–1.71
gene6008	Transmembrane protease, serine 9	LOC109033755	–2.13
gene10506	Serine protease nudel	LOC109038630	3.83
gene3265	Serpin B3-like	LOC109030665	–3.16
Toll			
gene1030	Protein spätzle-like	LOC109040362	–3.84
gene6128	Toll-like receptor 7	LOC109033743	–1.47
gene13805	Toll-like receptor	LOC109042320	1.03
PI3K/AMPK/mTOR			
gene6501	**Phosphatidylinositol-bisphosphate 3-kinase catalytic subunit delta isoform (PI3K)**[b]	LOC109034225	2.83
gene12334	5'-AMP-activated protein kinase subunit gamma-2 (AMPK)	LOC109040764	1.09
gene14490	DNA-dependent protein kinase catalytic subunit (mTOR)	LOC109043096	1.19
gene9417	Ribosomal protein S6 kinase beta-1	LOC109037417	1.42

[a] Only the genes with *q* value < 0.005 and $|\log_2(FC)|>1$ (significantly regulated) are listed. FC – fold change. The same below. [b] The gene in bold was used in the validation of the RGs.

Table S2. DEGs involved in the immune responses of the host.

Category/Gene ID	Annotated function	GenBank acc. no.	$\log_2$(FC)
AMP			
gene553	**Antimicrobial peptide Alo-3-like**[a]	LOC109033344	–3.02
Lysozyme			
gene12425	Lysozyme-like	LOC109040865	2.11
gene13644	Lysozyme C, milk isozyme-like	LOC109042207	–6.47
Phagocytosis			
gene3382	Down syndrome cell adhesion molecule-like protein Dscam2	LOC109030752	1.38
gene8209	Down syndrome cell adhesion molecule-like protein Dscam2	LOC109036103	1.11
gene9357	cdc42 homolog	LOC109037334	1.28
gene14561	cdc42 homolog	LOC109043068	1.17
gene6552	Partitioning defective 3 homolog	LOC109034237	1.38
gene3743	Epidermal growth factor receptor (Efgr)	LOC109031082	1.44
gene2039	Epidermal growth factor receptor substrate 15-like	LOC109044512	1.33
gene11461	Dynamin	LOC109039701	1.70
gene6591	Myc box-dependent-interacting protein 1	LOC109034291	1.65
gene9406	G protein-coupled receptor kinase 1	LOC109037372	1.26
gene9298	Protein disabled (Dab1)	LOC109037388	1.43
gene15013	Tyrosine-protein kinase Src64B	LOC109043668	1.19
gene1358	Ras-like GTP-binding protein Rho1	LOC109040974	1.02
gene9020	Ras-related protein Rad-11A	LOC109037003	1.04
gene9786	EH domain-containing protein	LOC109037843	1.20
gene12339	EH domain-containing protein 1-like	LOC109040805	1.26
gene4313	Rab11 family-interacting protein	LOC109031749	1.06
Encapsulation			
gene7278	Integrin alpha-PS2-like	LOC109035063	1.24
gene4193	Integrin beta-PS	LOC109031639	1.47
gene7161	Rho GTPase-activating protein gacZ-like	LOC109034894	1.95
gene7940	Rho GTPase-activating protein 10-like	LOC109035840	1.16

[a] The gene in bold was used in the validation of the RGs.

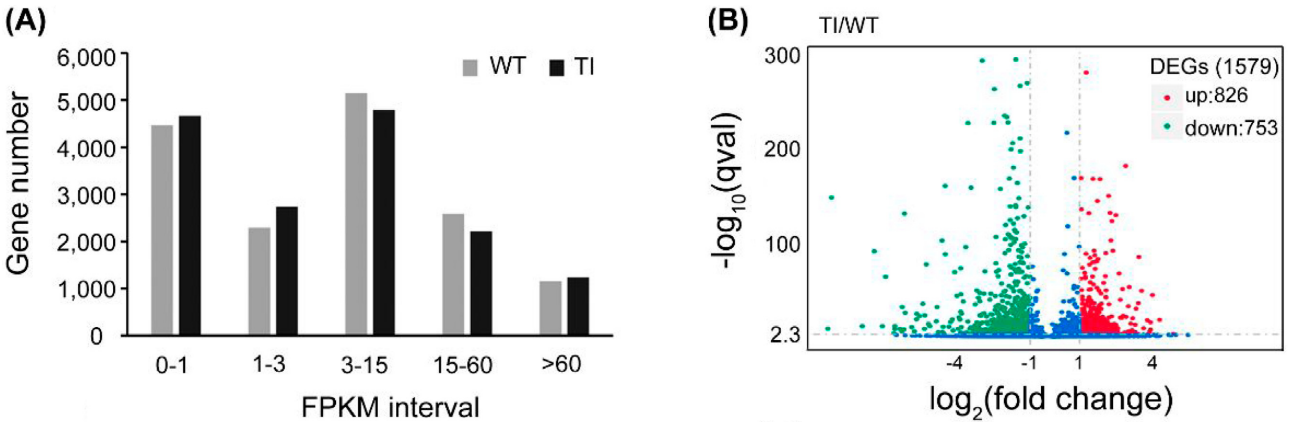


Fig. S1. Distribution of UniGenes per FPKM interval (A) and Volcanoplot of DEGs (B). The up- and down-regulated genes are indicated by red and green dots, respectively. Blue dots indicate no significant difference.

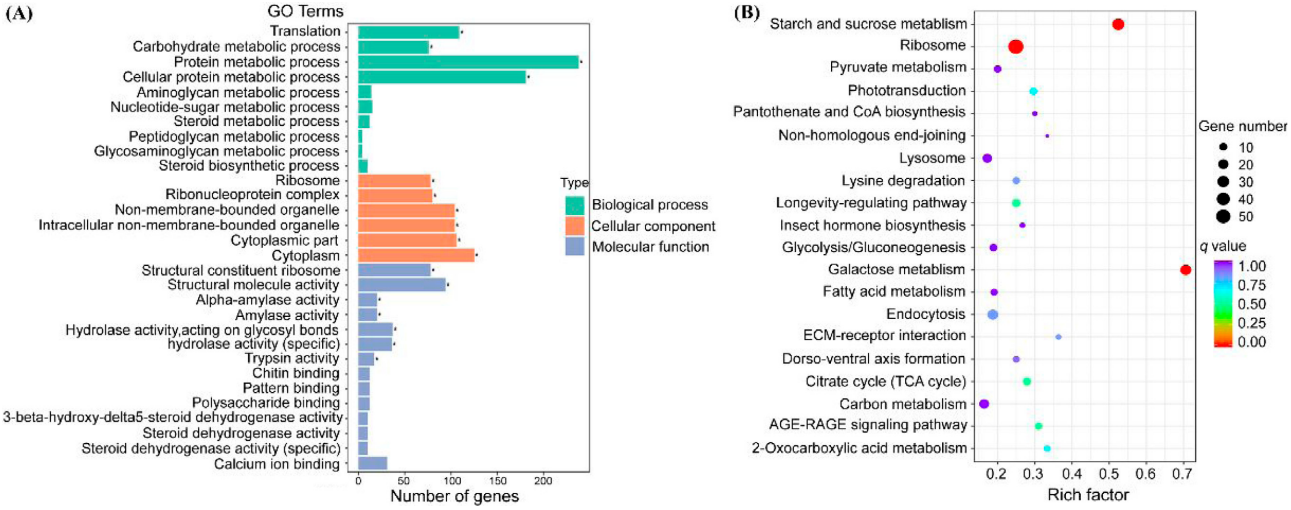
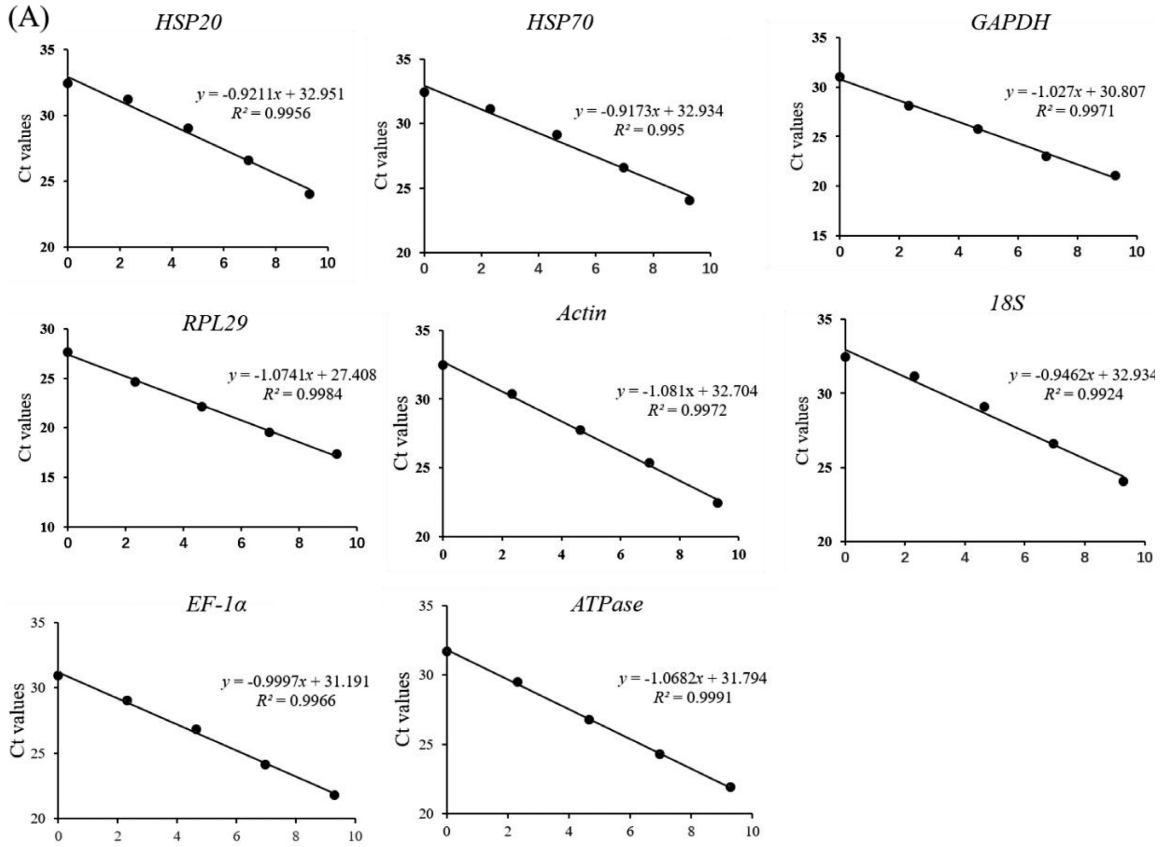
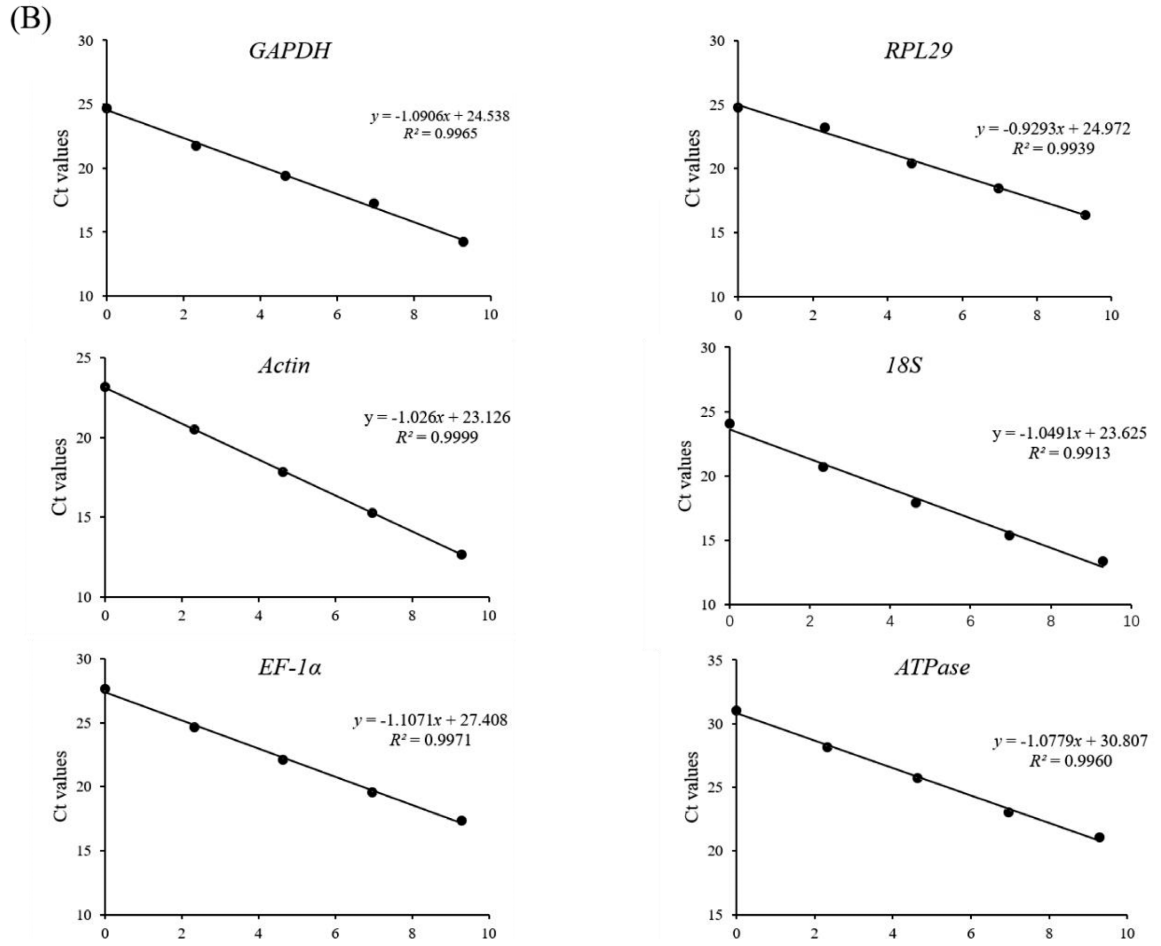


Fig. S2. GO (A) and KEGG (B) pathway enrichment analyses of DEGs. *q* value: the corrected *P* value, indicating a significant difference between WT and TI. Rich factor: The ratio of DEG numbers annotated in a given pathway to all the genes annotated in that pathway.


(A)
HSP20
y = -0.9211x + 32.951
R² = 0.9956
HSP70
y = -0.9173x + 32.934
R² = 0.995
GAPDH
y = -1.027x + 30.807
R² = 0.9971
RPL29
y = -1.0741x + 27.408
R² = 0.9984
Actin
y = -1.081x + 32.704
R² = 0.9972
18S
y = -0.9462x + 32.934
R² = 0.9924
EF-1α
y = -0.9997x + 31.191
R² = 0.9966
ATPase
y = -1.0682x + 31.794
R² = 0.9991
Ct values


(B)
GAPDH
y = -1.0906x + 24.538
R² = 0.9965
RPL29
y = -0.9293x + 24.972
R² = 0.9939
Actin
y = -1.026x + 23.126
R² = 0.9999
18S
y = -1.0491x + 23.625
R² = 0.9913
EF-1α
y = -1.1071x + 27.408
R² = 0.9971
ATPase
y = -1.0779x + 30.807
R² = 0.9960
Ct values

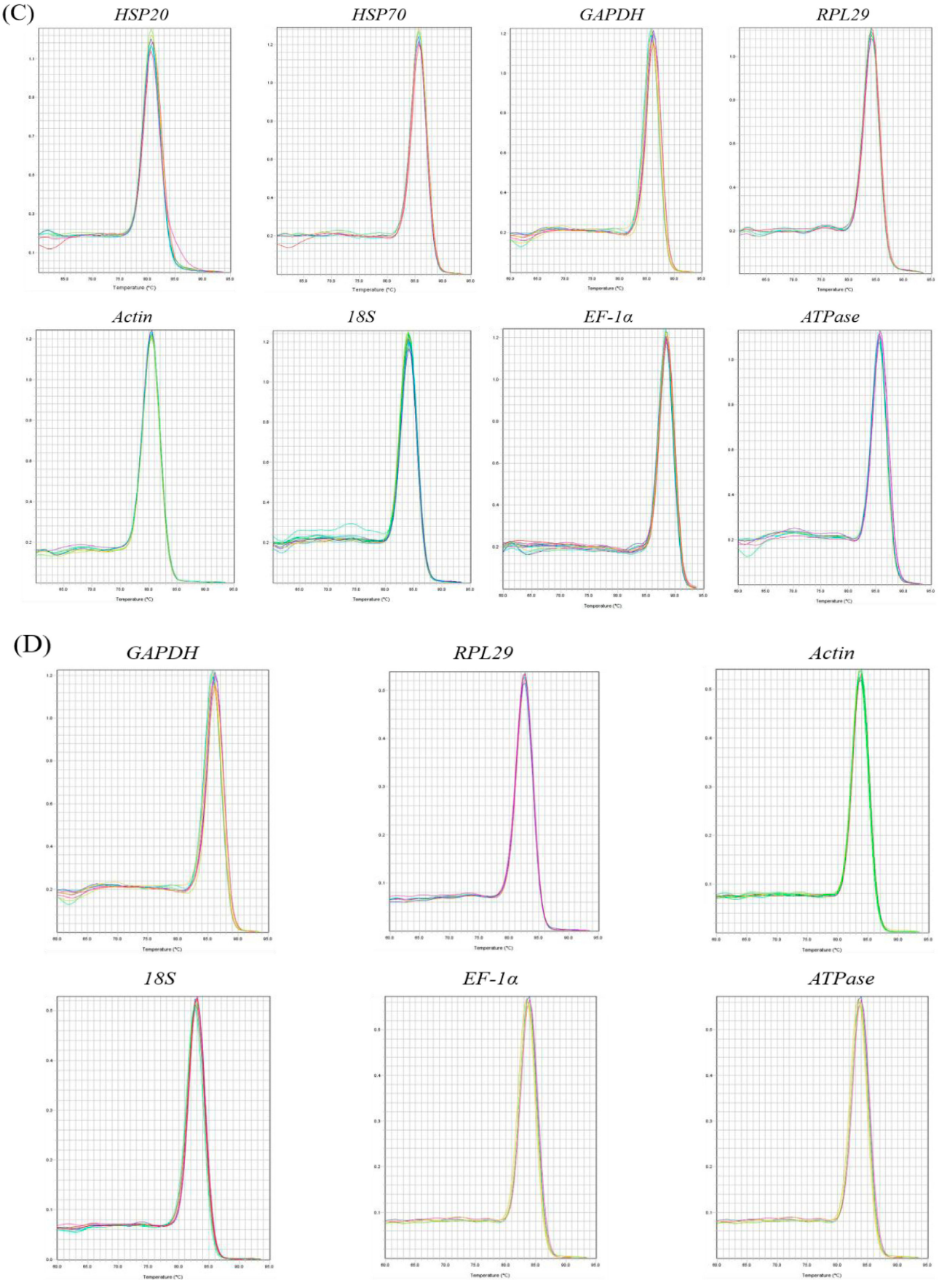


Fig. S3. Standard curves and melting curves of the reference genes. (A) Standard curves for gDNA templates; (B) Standard curves for cDNA templates; (C) Melting curves for gDNA templates and (D) Melting curves for cDNA templates.