

# SCIENTIFIC REPORTS



OPEN

## Housekeeping in Tephritid insects: the best gene choice for expression analyses in the medfly and the olive fly

Received: 22 August 2016

Accepted: 01 March 2017

Published: 03 April 2017

Efthimia Sagri<sup>1</sup>, Panagiota Koskinioti<sup>1</sup>, Maria-Eleni Gregoriou<sup>1</sup>, Konstantina T. Tsoumani<sup>1</sup>, Yiannis C. Bassiakos<sup>2</sup> & Kostas D. Mathiopoulos<sup>1</sup>

Real-time quantitative-PCR has been a priceless tool for gene expression analyses. The reaction, however, needs proper normalization with the use of housekeeping genes (HKGs), whose expression remains stable throughout the experimental conditions. Often, the combination of several genes is required for accurate normalization. Most importantly, there are no universal HKGs which can be used since their expression varies among different organisms, tissues or experimental conditions. In the present study, nine common HKGs (*RPL19*, *tpb*, *ubx*, *GAPDH*,  $\alpha$ -*TUB*,  $\beta$ -*TUB*, *14-3-3zeta*, *RPE* and *actin3*) are evaluated in thirteen different body parts, developmental stages and reproductive and olfactory tissues of two insects of agricultural importance, the medfly and the olive fly. Three software programs based on different algorithms were used (*geNorm*, *NormFinder* and *BestKeeper*) and gave different ranking of HKG stabilities. This confirms once again that the stability of common HKGs should not be taken for granted and demonstrates the caution that is needed in the choice of the appropriate HKGs. Finally, by estimating the average of a standard score of the stability values resulted by the three programs we were able to provide a useful consensus key for the choice of the best HKG combination in various tissues of the two insects.

The Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), and the olive fruit fly, *Bactrocera oleae* (Rossi), belong to the Tephritidae family of insects. As typical fruit flies, females lay their eggs in fruits or vegetables and the emerging larvae feed in the fruit sap, thus destroying the fruit. The medfly is one of the most devastating insects, easily adapting to new environments and hosts, infecting more than 260 species of fruits and vegetables worldwide<sup>1,2</sup>, and causing great economic losses in fruit production and quarantine costs. The olive fruit fly, on the other hand, is a monophagous species, the most important enemy of olive cultivations<sup>3,4</sup>. Whole genome sequencing of both species has been completed<sup>5,6</sup>, offering a holistic view of the entire genomes, allowing the study of any desired gene and thus leading to a profound understanding of the biology of these species. Such understanding is a prerequisite for novel, alternative to insecticides, control approaches.

The study of any gene inevitably goes through detailed and thorough scrutiny of its expression profile in various tissues and under different conditions. An invaluable tool for such expression analysis is RT-qPCR. The same way PCR revolutionized modern day molecular biology, RT-qPCR gave tremendous impetus to studies of gene expression, quantitative genotyping, genetic variation, disease diagnosis, forensics and many more. Due to the simplicity of the reaction, data can be easily collected and published in high impact journals without, necessarily, following good practices of RT-qPCR<sup>7</sup>. One of the most important parameters that should be addressed in order to standardize the reaction and perform a valid RT-qPCR analysis is the selection of suitable reference housekeeping genes. Since the reaction has several limitations as a result of the quality and quantity of starting RNA and the efficiency of its reverse transcription, housekeeping genes are used for the systematic normalization of gene expression data in order to improve the fidelity and accuracy of RT-qPCR<sup>8-10</sup>. Time and again, it has been demonstrated that the use of an unsuitable reference gene can lead to false results of the qPCR data and, consequently,

<sup>1</sup>Department of Biochemistry and Biotechnology, University of Thessaly, Larissa, Greece. <sup>2</sup>Department of Economic Sciences, National and Kapodistrian University of Athens, Athens, 10559, Greece. Correspondence and requests for materials should be addressed to K.D.M. (email: kmathiop@bio.uth.gr)

to erroneous interpretations<sup>11–14</sup>. Most frequently, indeed, more than one housekeeping genes are required for proper normalization of the data<sup>15,16</sup>.

In insects, many articles have been published on the identification and selection of the best reference gene in specific tissues and under different conditions. In the Tephritidae family there are two studies on the oriental fruit fly, *Bactrocera dorsalis*<sup>17,18</sup> and one in the West Indian fruit fly *Anastrepha obliqua*<sup>19</sup>. Among other dipteran species, there are three studies on *Drosophila melanogaster*<sup>15,20,21</sup> and a single one on each of *D. suzukii*<sup>22</sup>, *Musca domestica*<sup>23</sup>, *Lucilia cuprina*<sup>24</sup> and the Calliphoridae family<sup>25</sup>. Interestingly, there are no studies published on any mosquito species. In many mosquito publications, normalization of RT-qPCR is at best performed using a housekeeping gene (HKG) that demonstrates stable expression in microarray or RNAseq results<sup>26,27</sup>. This strategy may seem biologically reasonable, but there is a potential technical artifact considering that microarrays, RNAseq and RT-qPCR constitute quite different methods, with different limitations, requiring different standardization each. Most frequently, however, there is no specific justification regarding the selection of the utilized HKGs<sup>28–34</sup> except, at most, that it may have been used previously in the same<sup>35,36</sup> or related species<sup>37,38</sup>. Furthermore, with regard to published HKG studies on Diptera, the number of HKGs tested varies from as low as six<sup>18,21,25</sup> to over 20<sup>15</sup>. Unfortunately, neither the same genes nor the same tissues and conditions are studied, a fact that makes any effort to compare results practically impossible. Very importantly, these studies hardly ever indicate the use of the same housekeeping gene or gene combination in different tissues of the same insect or in the same tissue of different insects. Since, as mentioned above, the use of improper housekeeping genes for the normalization of the RT-qPCR can lead to erroneous results, this variability necessitates each time, for every organism and every tissue, the search for the proper housekeeping genes. Additionally, given the fact that the available software (such as *geNorm*<sup>9</sup>, *NormFinder*<sup>39</sup>, *BestKeeper*<sup>40</sup> and the web-based *RefFinder* platform<sup>41</sup>) are based on different statistical algorithms, they do not result in the same HKG suggestions for a particular tissue<sup>42,43</sup>.

Here we present the most extensive study on HKGs, at least in the dipteran order of insects. The study validates nine candidate reference genes in thirteen different tissues of the model tephritid fly, the Mediterranean fruit fly, *C. capitata*, and the olive fruit fly, *B. oleae*. The genes are: *RPL19* (ribosome protein L19), *tbp* (TATA-binding protein), *ubx* (ultrabithorax), *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase),  $\alpha$ -*TUB* ( $\alpha$ -tubulin),  $\beta$ -*TUB* ( $\beta$ -tubulin), *14-3-3zeta*, *RPE* (RNA polymerase II) and *actin3*. The tissues selected for the analysis were mostly tissues from either the reproductive [testes, ovaries, male and female accessory glands (MAGs and FAGs, respectively), ovipositors] or the olfactory (maxillary palps and antennae) systems of the flies. In addition, we analyzed three developmental stages (egg, larva, pupa) and the three sections of the insect body (head, thorax, abdomen), as they are often convenient controls for comparison with other tissues.

## Results

In the present study, the best choice for reference genes for RT-qPCR in thirteen tissues of two insects of the Tephritidae family, the Mediterranean fruit fly, *Ceratitis capitata* and the olive fruit fly, *Bactrocera oleae*, was examined. Three available software programs were used for the analysis and, since each program is based on a different algorithm, an effort was put to generate a consensus of the three programs.

**Gene choice and amplification performance.** Nine different housekeeping genes, commonly used in other dipteran species, were chosen for the analysis. The genes considered were: *RPL19* (ribosome protein L19), *tbp* (TATA-binding protein), *ubx* (ultrabithorax), *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase),  $\alpha$ -*TUB* ( $\alpha$ -tubulin),  $\beta$ -*TUB* ( $\beta$ -tubulin), *14-3-3zeta*, *RPE* (RNA polymerase II) and *actin3*. Gene names and IDs for the two species are presented in Supplementary Table S6.

In all instances, primers were designed by Primer-BLAST<sup>44</sup> in order to get amplicons ranging from 82 to 150 bp, as shown in Supplementary Table S7. Reaction conditions described in the Methods section resulted in one gene-specific peak and the absence of primer dimers peaks (data not shown). The PCR efficiency (E) and the correlation coefficient ( $R^2$ ) characterizing each standard curve are also given in Supplementary Table S7. Efficiencies for all tested genes varied between 90.1% and 106.4%.

All reactions were done in triplicate (three technical replicates). The expression of the reference genes was measured in 8 or 10 biological replicates, as indicated in Table 1. Three negative controls were also used.

**Expression stability by *geNorm*.** *geNorm* is a Visual Basic Application (VBA) for Microsoft Excel that automatically calculates two parameters: the gene-stability measure M and the pairwise variation V. The lower the gene-stability M value indicates the more stably expressed gene. Values of M higher than 1.5 are not considered stable across measurements. The pairwise variation V, on the other hand, indicates the least number of the most stably expressed genes that should be combined for optimal normalization. Additionally, V should be below the cut-off value of 0.15, otherwise, the lowest V should be considered. Using this algorithm, we ranked the nine housekeeping genes in the thirteen tissues tested according to their expression stability (Fig. 1). For *B. oleae* egg, for example, under the cut-off value of 0.150 is V4/5 (0.136, Fig. 1-D) and, therefore, the four most stable genes for the eggs (*ubx* with M = 0.508, *14-3-3zeta* with M = 0.556, *tbp* with M = 0.601 and *RPE* with M = 0.658) should be combined in order to obtain optimal normalization. For the other tissues, the lowest pairwise variation value and the suggested combination of HKGs are presented in Supplementary Table S1. In most cases, *geNorm* suggests the combination of 2–3 HKGs for optimal normalization. In one case (FAGs of *B. oleae*) it suggests the combination of six; and in one other (ovipositor of *B. oleae*) it suggests the combination of seven. For *C. capitata*,  $\alpha$ - and  $\beta$ -*tubulin* are most frequently among the suggested HKGs, while for *B. oleae* *14-3-3zeta* is the winner. *RPE* and *ubx* are never among the suggested HKGs in *C. capitata*.

**Expression stability by *NormFinder*.** *NormFinder* algorithm identifies the optimal normalization gene among a set of candidate genes, providing a stability value for each gene. This value is the estimated expression

Tested tissues	Biological replicates	
Developmental Stages	Egg	10 individuals
	Larva	10 individuals
	Pupa	10 individuals
Body parts	Head	10 individual parts (5 male and 5 female)
	Thorax	10 individual parts (5 male and 5 female)
	Abdomen	10 individual parts (5 male and 5 female)
Reproductive System	MAGs	10 pairs (1 pair of MAGs/fly)
	Testes	10 pairs (1 pair of testes/fly)
	FAGs	10 pairs (1 pair of FAGs/fly)
	Ovaries	10 sets (1 set of ovaries/fly)
	Ovipositors	8 pools (4 flies/pool)
Olfactory System	Maxillary palps	8 pools (4 flies/pool)
	Antennae	8 pools (4 flies/pool)

**Table 1.** The thirteen tested tissues of *C. capitata* and *B. oleae*.

variation if a given gene is used for normalization. Therefore, the candidate genes can be ranked according to their expression stability in the different tissues or experimental conditions<sup>45</sup>. The calculated stability values for each HKG and the according ranking in the thirteen tissues are shown in Supplementary Tables S2A and S2B for *C. capitata* and *B. oleae*, respectively. For *C. capitata*, *GADPH* ranks first in four tissues, *14-3-3zeta* in three, while *ubx* and *actin3* never rank first. For *B. oleae*, *RPE* and *14-3-3zeta* rank first in three tissues each, while *ubx* and *GADPH* never rank first.

**Expression stability by *BestKeeper*.** *BestKeeper* software estimates standard deviation (SD) of the Ct values of all candidate genes. Since the expression levels of suitable HKGs should be highly correlated, the lower the SD the more stable the gene<sup>40</sup>. The disadvantage of *BestKeeper* is that it does not provide a combination of reference genes required for an experiment. The calculated SD values and CV (coefficient of variation) for each HKG in the thirteen tissues are shown in Supplementary Tables S3A and S3B for *C. capitata* and *B. oleae*, respectively. According to *BestKeeper*,  $\alpha$ -*TUB*, *GADPH* and *RPL19* have the least SD values in three different tissues of *C. capitata* each, while *thp* and *ubx* in none. For *B. oleae*,  $\alpha$ -*TUB* and *RPE* have the least SD in four different tissues each, while  $\beta$ -*TUB*, *ubx*, *GADPH* and *actin3* in none.

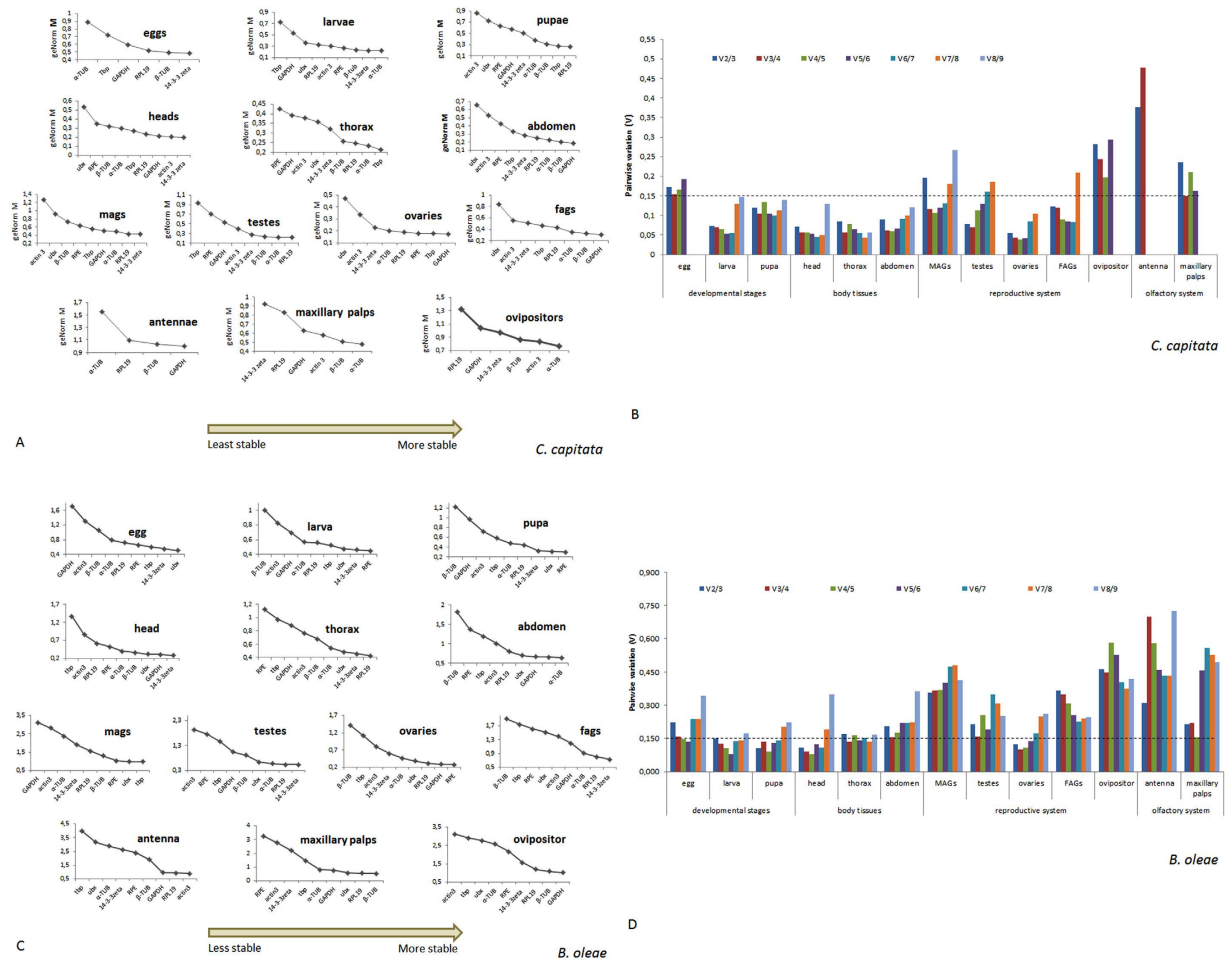
**Seeking consensus.** Since the different software programs use different algorithms to estimate gene expression stability, they rarely reach the same ranking. *RefFinder* software theoretically integrates the results of the previous analyses (by *geNorm*, *Normfinder* and *BestKeeper*). It then assigns an appropriate weight to an individual HKG and calculates the geometric mean of their weights for an overall final ranking<sup>41</sup>. We ran this user-friendly web-based tool as well. However, since the values that *RefFinder* calculated for, e.g., *geNorm* differed from those estimated by *geNorm* itself, we considered *RefFinder* unreliable and we did not use it any further. *RefFinder* results are presented in Supplementary Tables S4A and S4B.

In order to propose a combination of the most stable HKGs that a researcher can use for normalization of gene expression in *C. capitata* and *B. oleae*, we took a different route. We first estimated the average of a standard score (z-score) of the stability values resulted by all three software packages for every single gene and then ranked them according to this new average score. Complete results of this ranking are presented in Supplementary Tables 5A and 5B. The first three genes of this consensus ranking are presented in Table 2. In the medfly, *RPL19* is the HKG that is most often found in the best three ranking genes, followed by  $\beta$ -*TUB*, while *ubx* is never among the top three. Similarly, in the olive fly *14-3-3zeta* is the HKG that is most often found in the best three ranking genes, followed by *GADPH*, while  $\alpha$ -*TUB* is not found at all. To our experience, the combination of at least two HKGs and at most the number of genes suggested by *geNorm*, provides an excellent internal control in all RT-qPCRs.

## Discussion

Several times in the recent years it has been documented that the choice of the right reference gene/s for the standardization of RT-qPCRs is of paramount importance and the possible use of the inappropriate HKGs can lead to incorrect results<sup>11–14</sup>. Common housekeeping genes, that are supposed to be constitutively expressed in order to maintain basic cellular functions, may not have constant and stable expression throughout an experiment. This may be due to the special characteristics of the organism or tissue analyzed or the particular conditions of the experimental design. Good practice of an RT-qPCR experiment requires the establishment of the appropriate HKGs for its standardization<sup>46</sup>, even though good practice is not always observed.

We set out to address the above question for the medfly, *C. capitata*, and the olive fly, *B. oleae*, both very important agricultural pests. Particularly, the medfly is a cosmopolitan pest and due to its great importance in the cultivation and export of more than 260 fruits and vegetables<sup>1,2</sup>, it has turned out to be a model organism in the Tephritidae family of insects and beyond, for studies ranging from classical genetics to genomics<sup>5,47–52</sup>, as well as area-wide control practices<sup>53,54</sup>. The olive fly, on the other hand, is a strictly monophagous cousin of the medfly,



**Figure 1.** (A) Stability values of the reference genes in the 13 *C. capitata* tissues under study as generated by the *geNorm* algorithm. The average expression stability values from least stable (left) to most stable (right) for the egg, larva, pupa, head, thorax, abdomen, MAGs, testes, ovaries, FAGs, antennae, maxillary palps and ovipositor of the Mediterranean fruit fly. (B) Pairwise variation ( $V_n/V_{n+1}$ ) analysis determines the optimal number of reference genes for all of the tissues under study. (C) Stability values of the reference genes in the 13 *B. oleae* tissues as generated by the *geNorm* algorithm. The average expression stability values from least stable (left) to most stable (right) for the egg, larva, pupa, head, thorax, abdomen, MAGs, testes, ovaries, FAGs, antennae, maxillary palps and ovipositor of the olive fruit fly. (D) Pairwise variation ( $V_n/V_{n+1}$ ) analysis determines the optimal number of reference genes for all of the tissues under study.

of particular interest in the olive producing areas of the world<sup>4</sup>. Recent development of molecular and genomics tools have made it focus of active research, with renewed interest in its control<sup>55,56</sup>. The tissues selected for the analysis were mostly tissues from either the reproductive (testes, ovaries, male and female accessory glands, ovipositors) or the olfactory (maxillary palps and antennae) systems of the flies. The reproductive system is involved in the successful mating and egg development while the olfactory system plays a crucial role in insect survival and reproductive success, mediating responses to food, mates and oviposition. Beyond their general interest, such systems can serve as targets for alternative control approaches, such as the Sterile Insect Technique and its alternatives<sup>52,57–59</sup> and, therefore, are currently under scrutiny in the scientific community. In addition, we analyzed three developmental stages (egg, larva, pupa), as they are useful in order to obtain the expression profile throughout the life cycle of an insect. Finally, we included the three sections of the insect body (head, thorax, abdomen), as they are often convenient controls for comparison with other tissues.

In order to determine the best combination of HKGs, we performed our analyses with the three most popular software programs, *geNorm*<sup>9</sup>, *NormFinder*<sup>39</sup> and *BestKeeper*<sup>40</sup>. As anticipated, results were largely inconsistent among them, as they are based on different algorithms. A fourth user-friendly web-based software, *RefFinder*<sup>41</sup>, that is supposed to integrate the results of the previous three programs, gave inconsistent results with the programs themselves and so it was deemed untrustworthy. Instead, we decided to take a different route: we first transformed the raw scores of the three programs into standard scores, then calculated the average of the standard scores and finally ranked them. The use of the average of scores is based on the underlying idea of producing a

Tested tissues		The best ranking reference genes in <i>C. capitata</i>			The best ranking reference genes in <i>B. oleae</i>		
Developmental stages	Egg	<b>14-3-3zeta</b>	<b>RPL19</b>	$\beta$ -TUB	RPE	<b>14-3-3zeta</b>	<b>RPL19</b>
	Larva	<b>RPE</b>	<i>actin3</i>	RPL19	14-3-3zeta	<b>RPE</b>	GAPDH
	Pupa	<i>tbp</i>	<b>RPL19</b>	$\beta$ -TUB	<b>RPL19</b>	14-3-3zeta	RPE
Body tissues	Head	<b>14-3-3zeta</b>	<b>RPL19</b>	<b>actin3</b>	<b>14-3-3zeta</b>	<b>RPL19</b>	<b>actin3</b>
	Thorax	RPL19	<b>GAPDH</b>	<b>14-3-3zeta</b>	<b>14-3-3zeta</b>	<b>GAPDH</b>	$\beta$ -TUB
	Abdomen	$\alpha$ -TUB	<b>GAPDH</b>	$\beta$ -TUB	14-3-3zeta	<b>GAPDH</b>	<i>ubx</i>
Reproductive system	Testes	<b>actin3</b>	RPL19	$\alpha$ -TUB	14-3-3zeta	<b>actin3</b>	RPE
	MAGs	<b>RPL19</b>	$\alpha$ -TUB	<b>GAPDH</b>	<b>RPL19</b>	<i>actin3</i>	<b>GAPDH</b>
	Ovaries	<b>GAPDH</b>	$\alpha$ -TUB	RPE	<i>actin3</i>	<b>GAPDH</b>	RPL19
	FAGs	$\beta$ -TUB	<i>tbp</i>	<b>RPE</b>	GAPDH	<b>RPE</b>	<i>actin3</i>
	Ovipositor	$\beta$ -TUB	$\alpha$ -TUB	<b>14-3-3zeta</b>	<i>tbp</i>	<b>14-3-3zeta</b>	RPL19
Olfactory system	Antennae	<b>14-3-3zeta</b>	$\beta$ -TUB	<b>GAPDH</b>	<b>14-3-3zeta</b>	<i>actin3</i>	<b>GAPDH</b>
	Maxillary palps	$\alpha$ -TUB	RPL19	$\beta$ -TUB	<i>ubx</i>	GAPDH	<i>actin3</i>

**Table 2.** Consensus ranking of tested *Ceratitis capitata* and *Bactrocera oleae* housekeeping genes according to the mean of the z-scores of their stability values obtained by *geNorm*, *NormFinder* and *BestKeeper*. Only the first three genes are indicated, listed from the most stable (left) to the least stable (right) gene order. Genes in bold contain highly ranked HKGs that are common in both *C. capitata* and *B. oleae*.

composite score using a linear combination of the individual score values. This practice is common in Statistics, e.g., Principal Components Analysis, Factor Analysis and other multivariate<sup>60–62</sup> methods. In all these methods the individual variables do not have to be similar in derivation nor do they have to measure the same quantity. Instead, they measure different facets of the same concept, in many instances using different measurement tools. This is the case in our work. Averaging is the simplest form of linear combination (all scores have the same coefficient). The major issue in this case is not how the individual scores are derived, but if their values are in a similar range. When this is not true the score with the larger values would dominate the composite score. We resolved this problem by score standardization.

The aforementioned approach resulted in a useful consensus key (Table 2) for the choice of the best HKG combination in various tissues of the medfly and the olive fly. A few qualified comments based on Table 2 are worth making. First, the most common genes found in the top three choices for both the medfly and the olive fly (i.e., found five times or more in both organisms in Table 2) are *14-3-3zeta*, *RPL19* and *GAPDH*, while the least common (two times or less) are, *tbp* and *ubx*. Curiously,  $\alpha$ - and  $\beta$ -tubulins are quite frequently found in the medfly (6 and 7 times, respectively), while only  $\beta$ -tubulin is found only once in the olive fly. Secondly, in quite a few occasions (indicated by the genes in bold in Table 2) the same HKGs are found in the top three genes in the same tissue of both insects. For example, in eggs *14-3-3zeta* and *RPL19* are ranked in the top three HKGs in both the medfly and the olive fly. All things being equal, the probability of finding one particular gene out of the nine tested HKGs among three selected genes is  $\frac{1}{3}$ ; the probability of finding two particular genes out of nine tested HKGs among three selected genes is  $\frac{1}{12}$ ; while the probability of finding three particular genes out of nine tested HKGs among three selected genes is  $\frac{1}{84}$  (calculation based on hypergeometric probabilities). Furthermore, the probability of finding the *same two* of the nine tested HKGs among three genes of both organisms (independent selections from a probability point of view) acquires the statistically significant value of 0.0069 ( $\frac{1}{12} \times \frac{1}{12}$ ), while finding the *same three* of the nine tested HKGs among three genes of both organisms acquires the statistically significant value of 0.00014 ( $\frac{1}{84} \times \frac{1}{84}$ ). This observation may suggest a biological explanation for the stability of *14-3-3zeta*, *RPL19* and *actin3* in heads or the stability of *14-3-3zeta* and *RPL19* in eggs of both species. A similar situation is detected in thoraces, MAGs and antennae, but not in FAGs or maxillary palps. Therefore, one can imagine that similar patterns of neuronal development in heads or embryonic development in eggs of both species would require similar expression of HKGs; on the contrary, differences in the female reproductive system (FAGs) or different diets (perceived by the maxillary palps) between the two insects would be reflected in the expression of different HKGs. More analyses are needed to substantiate such claims that are beyond the scope of this article.

Closing, we should iterate once again that the stability of common HKGs should not be taken for granted and that a lot of caution is needed in the choice of the appropriate HKGs. In fact, there is a need to validate the use of the proper HKG more often than practically encountered in recent literature. Even though we consider that our analysis offers a useful tool in the medfly and olive fly research community, we do encourage researchers to check these HKGs on their own subjects before use in a particular expression study.

## Methods

**Fly strains.** The ‘Benakeion’ medfly and the ‘Demokritos’ olive fly strains were used in the experiments. The ‘Benakeion’ strain was originally established at the Benakeion Institute of Phytopathology, Athens, Greece, and has been kindly provided by Prof Nikos Papadopoulos at the Department of Agriculture Crop Production and Rural Environment, University of Thessaly, Greece. The ‘Demokritos’ strain originally comes from the Nuclear Research Centre in Athens, Greece, and has been reared in our laboratory for over 15 years. Both strains are maintained in wooden, nylon-screened, holding cages (30 × 30 × 30 cm) under an LD 14:10 h photocycle at 25 ± 1 °C and 60 ± 10% relative humidity. Olive fly rearing conditions are described in refs 63–65, while medfly conditions are described by Boller<sup>66</sup>.

**RNA isolation from specific tissues of *Ceratitis capitata* and *Bactrocera oleae*.** Thirteen specific tissues at different developmental stages were used, as shown in Table 1. Eggs were collected from adult females 15 minutes after being laid. Larvae were 2<sup>nd</sup> stage and pupae were harvested 12 hours after pupation. All dissected tissues (heads, thoraces and abdomens, as well as reproductive and olfactory) were from 5 day-old adult male and female insects.

Total RNA was isolated with the use of TRIsure™ (Bioline) following the instructions of the manufacturer with minor modifications. RNA extraction was followed by an additional DNA removal using the TURBO DNA-free Kit (Ambion-Invitrogen), according to manufacturer's instructions. The integrity of RNA was assessed in a 1% agarose gel electrophoresis and quantified by Qubit® 2.0 Fluorometer (Thermo Fisher Scientific).

The RNA extracted from: a single larva, a single pupa, a single head, a single thorax, a single abdomen and a set of ovaries from a single female, was quantified by Qubit and the amount of 2 µg was used for cDNA preparation. The entire RNA amount extracted from: a single egg, one pair of testes, one pair of FAGs, and maxillary palps, antennae or ovipositors from a pool of 4 individual flies, was used for cDNA preparation, since the amount of RNA was undetectable.

DNA-free total RNA was converted into cDNA using 300 ng Random hexamer primers (equimolar mix of N<sub>5</sub>A, N<sub>5</sub>G, N<sub>5</sub>C and N<sub>5</sub>T), 200 units MMLV Reverse Transcriptase (Bioline), 10× reaction buffer, 40 mM dNTP mix and 40 units RNase Inhibitor (Bioline) according to the manufacturer's instructions.

**Expression stability of candidate reference genes in *C. capitata* and *B. oleae*.** 82 to 150 bp amplicons from nine different housekeeping genes commonly used in other dipteran species were analyzed. The genes considered were: *RPL19* (ribosome protein L19), *tbp* (TATA-binding protein), *ubx* (ultrabithorax), *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase),  $\alpha$ -*TUB* ( $\alpha$ -tubulin),  $\beta$ -*TUB* ( $\beta$ -tubulin), *14-3-3zeta*, *RPE* (RNA polymerase II) and *actin3* (Supplementary Table S6). For the medfly, primers were based on sequences retrieved in the NCBI database. For the olive fly, primers were based on the sequences obtained during the transcriptome analysis of *B. oleae*<sup>55,67</sup>. Specific primers for the amplification of these HKGs were designed by Primer-BLAST<sup>44</sup> (Supplementary Table S7). Each primer was also evaluated using OligoAnalyzer 3.1 tool<sup>68</sup> in order to avoid hairpin formation and self-/hetero-dimerization of the oligonucleotides.

Relative quantitation was used to analyze changes in expression levels of the selected genes using a quantitative real-time PCR approach. The RT-qPCR conditions were: polymerase activation and DNA denaturation step at 95 °C for 4 min, followed by 40 cycles of denaturation at 95 °C for 30 s, annealing/extension and plate read at 56 °C (for all the tested housekeeping genes) and 60 °C (only for the reference genes *RPE* and *actin3*) for 30 s and finally, a step of melting curve analysis at a gradual increase of temperature over the range 55 °C → 95 °C. In this step, the detection of one gene specific peak and the absence of primer dimer peaks were assured. Each reaction was performed in a total volume of 15 µl, containing 5 µl from a 1:10 dilution of the cDNA template, 1 × iTaq Universal SYBR Green Supermix (Bio-Rad) and 400 nM of each primer. The reactions were carried out on Bio-Rad Real-Time thermal cycler CFX96 (Bio-Rad, Hercules, CA, USA) and data analyzed using the CFX Manager™ software. The expression of the reference genes was measured in 8 or 10 biological replicates, as indicated in Table 1. Three negative controls were also used. All reactions were done in triplicate (three technical replicates). The amplification efficiency of the reactions was calculated by the CFX Manager™ software (Bio-Rad). The PCR efficiency (E) and the correlation coefficient (R<sup>2</sup>) characterizing each standard curve are given in Supplementary Table S7. Efficiencies for all tested genes varied from 90.1% to 106.4%. The 2<sup>- $\Delta\Delta C_t$</sup>  method was used for the analysis of relative gene expression<sup>69</sup>.

**geNorm analysis.** The expression stability of the nine reference genes was assessed using the *geNorm* software. This algorithm is based on the principle that the logarithmically transformed expression ratio between two genes should be constant if both genes are stably expressed in a given sample set. The candidate reference genes were ranked by *geNorm* based on the expression stability value M, which is calculated for all genes under study. The lower the M value, the higher the gene's expression stability. Furthermore, *geNorm* performs a stepwise calculation of the pairwise variation ( $V_n/V_{n+1}$ ) between sequential normalization factors (NF<sub>n</sub> and NF<sub>n+1</sub>) to determine the optimal number of reference genes required for accurate normalization<sup>9</sup>. Results are presented in Fig. 1 and Supplementary Table S1.

*Normfinder*<sup>39</sup> is an algorithm for identifying the optimal normalization gene among a set of candidate genes. This software is based on a mathematical model of gene expression that enables estimation not only of the overall variation of the candidate normalization genes but also of the variation between samples subgroups of the sample set<sup>39</sup>. Results are presented in Supplementary Table S2A and S2B for *C. capitata* and *B. oleae*, respectively.

*BestKeeper* determines the most stably expressed genes based on the coefficient of correlation (r) to the *BestKeeper* Index (BI), which is the geometric mean of the candidate reference gene Cq values. Additionally, it calculates the standard deviation (SD) and the coefficient of variation (CV) based on the Cq values of all candidate reference genes<sup>40</sup>. Reference genes are identified as the most stable genes, i.e. those that exhibit the lowest coefficient of variance and standard deviation<sup>70</sup>. Results are presented in Supplementary Table S3A and S3B for *C. capitata* and *B. oleae*, respectively.

The *RefFinder* tool ranks all the potential reference genes according to the gene expression stability based on the rankings from *geNorm*, *Normfinder*, *BestKeeper* and the comparative  $\Delta\Delta C_t$  method programs. Also, this program assigns an appropriate weight to an individual gene and calculates the geometric mean of their weights for the overall final ranking<sup>71</sup>. Results are presented in Supplementary Table S4A and S4B for *C. capitata* and *B. oleae*, respectively.

**Statistical Analysis.** Four different types of Microsoft Excel-based software, *geNorm*<sup>9</sup>, *NormFinder*<sup>39</sup>, *BestKeeper*<sup>40</sup> and *RefFinder*<sup>41</sup> were used to rank the expression stability of reference genes for all the experimental sets in the specific tissues of the medfly and the olive fruit fly. Relative quantities were used for *geNorm* and *NormFinder*, while *BestKeeper* analyses and the web-based program *refFinder* were based on untransformed Cq values. All four software packages were used according to the manufacturer's instructions.

The consensus rank of the reference genes was estimated by the combination of the stability measurements obtained by *geNorm*, *Normfinder* and *BestKeeper*. More specifically, the raw scores calculated by these three software (M value by *geNorm*, stability value by *Normfinder* and SD by *BestKeeper*) were transformed into standard scores (z-score) for each housekeeping gene separately. The average of the three z-scores was subsequently calculated and the final rank was computed using the RANK function in Excel software. The above measurements were produced for every single reference gene in each one of the insect tissues under study. Thus, a consensus ranking of all nine genes was estimated for each one of the 13 tissues separately. Results are presented in Supplementary Table S5A and S5B for *C. capitata* and *B. oleae*, respectively.

**Ethics statement.** The study was carried out on laboratory reared olive flies and medflies. No specific permissions are required for these experiments or collections, since these studies did not involve endangered or protected species.

## References

1. Khoo, K. C., Ooi, P. A. C. & Ho, C. T. *Crop pests and their management in Malaysia*. Malaysia: Tropical Press SDN. BHD at <http://www.cabi.org/isc/abstract/19941105253> (1991).
2. Liquido, N. J., Shinoda, L. A. & Cunningham, R. T. Host plants of the Mediterranean fruit fly (Diptera, Tephritidae). An annotated world list. *Ann. Entomol. Soc. Am.* **77**, 1–57 (1991).
3. Mazomenos, B. E. *Estimates of the crop losses caused by Dacus oleae (Gmel.) (Diptera, Tephritidae) in Crete, in Fruit Flies of Economic Importance*. (Elsevier Science Publishers B.V., Amsterdam., 1989).
4. Daane, K. M. & Johnson, M. W. Olive fruit fly: managing an ancient pest in modern times. *Annu. Rev. Entomol.* **55**, 151–69 (2010).
5. Papanicolaou, A. *et al.* The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* **17**, 192 (2016).
6. The olive fly genome. at [https://15k.nal.usda.gov/Bactrocera\\_oleae](https://15k.nal.usda.gov/Bactrocera_oleae) (November 30, 2016)
7. Bustin, S. A. *et al.* The need for transparency and good practices in the qPCR literature. *Nat. Methods* **10**, 1063–1067 (2013).
8. Huggett, J., Dheda, K., Bustin, S. & Zumla, A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* **6**, 279–84 (2005).
9. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).
10. Nolan, T., Hands, R. E. & Bustin, S. A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–82 (2006).
11. Tricarico, C. *et al.* Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal. Biochem.* **309**, 293–300 (2002).
12. Bas, A., Forsberg, G., Hammarström, S. & Hammarström, M.-L. Utility of the housekeeping genes 18S rRNA, beta-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. *Scand. J. Immunol.* **59**, 566–73 (2004).
13. Babij, C. *et al.* STK33 Kinase Activity Is Nonessential in KRAS-Dependent Cancer Cells. *Cancer Res.* **71**, 5818–5826 (2011).
14. Scholl, C. *et al.* Synthetic Lethal Interaction between Oncogenic KRAS Dependency and STK33 Suppression in Human Cancer Cells. *Cell* **137**, 821–834 (2009).
15. Ling, D. & Salvaterra, P. M. Robust RT-qPCR data normalization: validation and selection of internal reference genes during post-experimental data analysis. *PLoS One* **6**, e17762 (2011).
16. Xiao, X. *et al.* Validation of suitable reference genes for gene expression analysis in the halophyte *Salicornia europaea* by real-time quantitative PCR. *Front. Plant Sci.* **5**, 788 (2014).
17. Shen, G.-M., Jiang, H.-B., Wang, X.-N. & Wang, J.-J. Evaluation of endogenous references for gene expression profiling in different tissues of the oriental fruit fly *Bactrocera dorsalis* (Diptera: Tephritidae). *BMC Mol. Biol.* **11**, 76 (2010).
18. Shen, A. G., Huang, Y., Jiang, X. & Dou, W. Effect of  $\beta$ -Cypermethrin Exposure on the Stability of Nine Housekeeping Genes in *Bactrocera dorsalis* (Diptera: Tephritidae). *Florida Entomol.* **96**, 442–450 (2013).
19. Nakamura, A. M. *et al.* Reference genes for accessing differential expression among developmental stages and analysis of differential expression of OBP genes in *Anastrepha obliqua*. *Sci. Rep.* **6**, 17480 (2016).
20. Ponton, F., Chapuis, M.-P., Pernice, M., Sword, G. A. & Simpson, S. J. Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. *J. Insect Physiol.* **57**, 840–850 (2011).
21. Matta, B. P., Bitner-Mathé, B. C. & Alves-Ferreira, M. Getting real with real-time qPCR: a case study of reference gene selection for morphological variation in *Drosophila melanogaster* wings. *Dev. Genes Evol.* **221**, 49–57 (2011).
22. Zhai, Y. *et al.* Identification and Validation of Reference Genes for Quantitative Real-Time PCR in *Drosophila suzukii* (Diptera: Drosophilidae). *PLoS One* **9**, e106800 (2014).
23. Zhong, M. *et al.* Selection of reference genes for quantitative gene expression studies in the house fly (*Musca domestica* L.) using reverse transcription quantitative real-time PCR. *Acta Biochim. Biophys. Sin. (Shanghai)*. **45**, 1069–1073 (2013).
24. BAGNALL, N. H. & KOTZE, A. C. Evaluation of reference genes for real-time PCR quantification of gene expression in the Australian sheep blowfly, *Lucilia cuprina*. *Med. Vet. Entomol.* **24**, 176–181 (2010).
25. Cardoso, G. A., Matioli, C. C., de Azeredo-Espin, A. M. L. & Torres, T. T. Selection and validation of reference genes for functional studies in the Calliphoridae family. *J. Insect Sci.* **14**, 2 (2014).
26. Sanders, H. R., Evans, A. M., Ross, L. S. & Gill, S. S. Blood meal induces global changes in midgut gene expression in the disease vector, *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **33**, 1105–22 (2003).
27. Faucon, F. *et al.* Identifying genomic changes associated with insecticide resistance in the dengue mosquito *Aedes aegypti* by deep targeted sequencing. *Genome Res.* **25**, 1347–59 (2015).
28. Zhao, L., Pridgeon, J. W., Becnel, J. J., Clark, G. G. & Linthicum, K. J. Mitochondrial gene cytochrome b developmental and environmental expression in *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **46**, 1361–9 (2009).
29. Bariami, V., Jones, C. M., Poupardin, R., Vontas, J. & Ranson, H. Gene Amplification, ABC Transporters and Cytochrome P450s: Unraveling the Molecular Basis of Pyrethroid Resistance in the Dengue Vector, *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **6**, e1692 (2012).
30. Zink, S., Van Slyke, G., Palumbo, M., Kramer, L. & Ciota, A. Exposure to West Nile Virus Increases Bacterial Diversity and Immune Gene Expression in *Culex pipiens*. *Viruses* **7**, 5619–5631 (2015).
31. Liu, H. *et al.* Functional analysis of Orco and odorant receptors in odor recognition in *Aedes albopictus*. *Parasit. Vectors* **9**, 363 (2016).

32. Yang, L. & Piermarini, P. M. Molecular expression of aquaporin mRNAs in the northern house mosquito, *Culex pipiens*. *J. Insect Physiol.* **96**, 35–44 (2016).
33. Kang, D. S., Cotten, M. A., Denlinger, D. L. & Sim, C. Comparative Transcriptomics Reveals Key Gene Expression Differences between Diapausing and Non-Diapausing Adults of *Culex pipiens*. *PLoS One* **11**, e0154892 (2016).
34. Alfonso-Parra, C. *et al.* Mating-Induced Transcriptome Changes in the Reproductive Tract of Female *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **10**, e0004451 (2016).
35. Shin, D., Jin, L., Lobo, N. F. & Severson, D. W. Transcript profiling of the meiotic drive phenotype in testis of *Aedes aegypti* using suppressive subtractive hybridization. *J. Insect Physiol.* **57**, 1220–1226 (2011).
36. Cassone, B. J. *et al.* Differential gene expression in incipient species of *Anopheles gambiae*. *Mol. Ecol.* **17**, 2491–2504 (2008).
37. Pelletier, J. & Leal, W. S. Characterization of olfactory genes in the antennae of the Southern house mosquito, *Culex quinquefasciatus*. *J. Insect Physiol.* **57**, 915–929 (2011).
38. Lv, Y. *et al.* Comparative transcriptome analyses of deltamethrin-susceptible and -resistant *Culex pipiens pallens* by RNA-seq. *Mol. Genet. Genomics* **291**, 309–321 (2016).
39. Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. 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**, 5245–50 (2004).
40. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–15 (2004).
41. 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.* **80**, 75–84 (2012).
42. Mallona, I., Lischewski, S., Weiss, J., Hause, B. & Egea-Cortines, M. Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. *BMC Plant Biol.* **10**, 4 (2010).
43. Mafra, V. *et al.* Reference genes for accurate transcript normalization in citrus genotypes under different experimental conditions. *PLoS One* **7**, e31263 (2012).
44. Primer-BLAST. at <http://www.ncbi.nlm.nih.gov/tools/primer-blast> (November 30, 2016)
45. Zhong, H.-Y. *et al.* Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions. *Plant Cell Rep.* **30**, 641–53 (2011).
46. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55**, 611–622 (2009).
47. Robinson, A. S. Genetic sexing strains in medfly, *Ceratitis capitata*, sterile insect technique programmes. *Genetica* **116**, 5–13 (2002).
48. Delprat, M. A., Stolar, C. E., Manso, F. C. & Cladera, J. L. Genetic stability of sexing strains based on the locus sw of *Ceratitis capitata*. *Genetica* **116**, 85–95 (2002).
49. Gasperi, G. *et al.* Genetic differentiation, gene flow and the origin of infestations of the medfly, *Ceratitis capitata*. *Genetica* **116**, 125–35 (2002).
50. Scolari, F. *et al.* How functional genomics will impact fruit fly pest control: the example of the Mediterranean fruit fly, *Ceratitis capitata*. *BMC Genet.* **15**, S11 (2014).
51. Loukeris, T. G., Livadaras, I., Arcà, B., Zabalou, S. & Savakis, C. Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. *Science* **270**, 2002–5 (1995).
52. Schetelig, M. F., Caceres, C., Zacharopoulou, A., Franz, G. & Wimmer, E. A. Conditional embryonic lethality to improve the sterile insect technique in *Ceratitis capitata* (Diptera: Tephritidae). *BMC Biol.* **7**, 4 (2009).
53. Hendrichs, J., Robinson, A. S., Cayol, J. P. & Enkerlin, W. Medfly Area-wide Sterile Insect Technique Programmes for Prevention, Suppression or Eradication: The Importance of Mating Behavior Studies. *Florida Entomol.* **85**, 1–13 (2002).
54. Sterile Insect Technique Principles and Practice in Area-Wide Integrated Pest Management (eds Dyck, V.A., Hendrichs, J., & Robinson, A.S.), (Springer, 2005).
55. Sagri, E. *et al.* The molecular biology of the olive fly comes of age. *BMC Genet.* **15** Suppl 2, S8 (2014).
56. Estes, A. M. *et al.* A basis for the renewal of sterile insect technique for the olive fly, *Bactrocera oleae* (Rossi). *J. Appl. Entomol.* **136**, 1–16 (2011).
57. Gong, P. *et al.* A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. *Nat. Biotechnol.* **23**, 453–6 (2005).
58. Fu, G. *et al.* Female-specific insect lethality engineered using alternative splicing. *Nat. Biotechnol.* **25**, 353–357 (2007).
59. Scolari, F. *et al.* Fluorescent sperm marking to improve the fight against the pest insect *Ceratitis capitata* (Wiedemann; Diptera: Tephritidae). *N. Biotechnol.* **25**, 76–84 (2008).
60. Meng, J., Chen, H.-I., Zhang, J., Chen, Y. & Huang, Y. Uncover cooperative gene regulations by microRNAs and transcription factors in glioblastoma using a nonnegative hybrid factor model. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 6012–6015 (IEEE, 2011). doi:10.1109/ICASSP.2011.5947732
61. Wang, W., Mo, J., Cheng, J., Zhuang, P. & Tang, Z. Selection and characterization of spinosad resistance in *Spodoptera exigua* (Hübner) (Lepidoptera: Noctuidae). *Pestic. Biochem. Physiol.* **84**, 180–187 (2006).
62. Child, D. *The Essentials of factor analysis*. (Universitas Negeri Malang, 1975).
63. Tzanakakis, M., Economopoulos, A. P. & Tsitsipis, J. The importance of conditions during the adult stage in evaluating an artificial food for larvae of *Dacus oleae* (Gmel.) (Diptera, Tephritidae). *Z. Angew. Entomol.* **59**, 127–130 (1967).
64. Tsitsipis, J. Development of a caging and eggging system for mass rearing the olive fruit fly, *Dacus oleae* (Gmel.) (Diptera, Tephritidae). *Ann. Zool. Ecol. Anim* **9**, 133–139 (1977).
65. Tsitsipis, J. A. & Kontos, A. Improved solid adult diet for the olive fruit fly *Dacus oleae*. *Entomol. Hell.* **1**, 24–29 (1983).
66. Boller, E. *Rhagoletis cerasi* and *Ceratitis capitata*. In *Handbook of insect rearing* (eds Sing, P. & Moore, R.) 135–144 (The Netherlands: Elsevier, 1985).
67. Sagri, E. *et al.* Olive fly transcriptomics analysis implicates energy metabolism genes in spinosad resistance. *BMC Genomics* **15**, 714 (2014).
68. OligoAnalyzer 3.1 tool. at <http://eu.idtdna.com/calc/analyser> (November 30, 2016)
69. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods* **25**, 402–8 (2001).
70. Chang, E. *et al.* Selection of reference genes for quantitative gene expression studies in *Platyclusus orientalis* (Cupressaceae) Using real-time PCR. *PLoS One* **7**, e33278 (2012).
71. Yuan, M. *et al.* 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 (2014).

## Acknowledgements

This research has been co-financed by: the European Union (ESF) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework - Research Funding Program: Heracleitus II, “Investing in knowledge society through the European Social Fund”; State of California Specialty Crops Block Grant Program award SCB10037; and the two postgraduate programs



of the Department of Biochemistry and Biotechnology of the University of Thessaly (“Biotechnology - Nutrition and Environment” and “Molecular Biology and Genetics applications”).

### Author Contributions

E.S. maintained the laboratory strains, isolated the tissues egg, larva, pupa, head, thorax and abdomen, performed the functional analyses for *Bactrocera oleae* and the bioinformatics analysis and designed part of the study; P.K. isolated all the tissues for the medfly, performed the functional analyses and the bioinformatics analysis for *Ceratitis capitata*; M.G. isolated the tissues MAGs, FAGs, testes, ovaries and ovipositor for *Bactrocera oleae*; K.T. isolated the tissues antennae and maxillary palps for *Bactrocera oleae*; Y.C.B. guided the calculation of the consensus ranking of the three software programs used; K.D.M. designed and coordinated the study. All authors participated in drafting the manuscript and read and approved the final document.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Sagri, E. *et al.* Housekeeping in Tephritid insects: the best gene choice for expression analyses in the medfly and the olive fly. *Sci. Rep.* 7, 45634; doi: 10.1038/srep45634 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017