SCIENTIFIC REPORTS

Reference genes for accessing differential expression among developmental stages and analysis of differential expression of OBP genes in *Anastrepha obliqua*

Aline Minali Nakamura[1], Samira Chahad-Ehlers[1], André Luís A. Lima[1], Cristiane Hayumi Taniguti[1], Iderval Sobrinho Jr.[2], Felipe Rafael Torres[1] & Reinaldo Alves de Brito[1]

The West Indian fruit fly, *Anastrepha obliqua*, is an important agricultural pest in the New World. The use of pesticide-free methods to control invasive species such as this reinforces the search for genes potentially useful in their genetic control. Therefore, the study of chemosensory proteins involved with a range of responses to the chemical environment will help not only on the understanding of the species biology but may also help the development of environmentally friendly pest control strategies. Here we analyzed the expression patterns of three OBP genes, *Obp19d_2*, *Obp56a* and *Obp99c*, across different phases of *A. obliqua* development by *q*PCR. In order to do so, we tested eight and identified three reference genes for data normalization, *rpl17*, *rpl18* and *ef1a*, which displayed stability for the conditions here tested. All OBPs showed differential expression on adults and some differential expression among adult stages. *Obp99c* had an almost exclusive expression in males and *Obp56a* showed high expression in virgin females. Thereby, our results provide relevant data not only for other gene expression studies in this species, as well as for the search of candidate genes that may help in the development of new pest control strategies.

The genus *Anastrepha* has over 235 species[1] most of them endemic to the Neotropics. Only a few of these species, mostly in the *fraterculus* group, are agricultural pests, including the West Indies fruit fly *Anastrepha obliqua* (Diptera, Tephritidae). Because of the vast damage they inflict to several different fruit crops, they are of great economic importance, which elicit the development of several pest management strategies. Genetic control technique's, as the Sterile Insect Technique (SIT) using gamma radiation, have been used to control natural populations of other tephritids such as *Ceratitis capitata*[2]. However, radiation leads to side effects[3,4], decreasing the technique efficiency. Thus, the search of candidate genes that may lead to more competitive transgenic organisms may be an alternative for improvement of SIT programs.

A number of candidate genes has already been explored and promising results are reported for some, such as *Astra* and *Astra-2* genes for the Caribbean fruit fly *Anastrepha suspensa*[5]. This encourages the search to identify genes related to reproduction and species recognition in this genus, which may be a priori good candidate genes for genetic control of pest populations. Because the olfactory system participates in important reproductive aspects, such as choosing a partner for mating and searching for

[1]Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil. [2]Universidade Federal de Goiás, Jataí, Brazil. Correspondence and requests for materials should be addressed to R.D.B. (email: brito@ufscar.br)

oviposition sites[6–10], genes that code for proteins involved in olfactory reception may prove useful in this regard as well.

From chemical stimulus to behavioral response, the olfactory process involves the capture, binding, transport and inactivation of odors; the activation of receptors and signal transduction; and the perception of odor by processing the signals at several levels of the central nervous system[6,7]. Thus, chemoreception participates in a range of responses to the chemical environment for animal survival, such as chemical cues for foraging behavior and food selection, predator avoidance, host plant recognition for oviposition and larval feeding, selection of mating partners, maternal behavior, and kin recognition[11]. In insects, the beginning of the transduction cascade of olfactory signals is the solubilization and transport of the chemical signals through the aqueous lymph of insect's sensilla to the olfactory receptors (ORs). The Odorant Binding Proteins (OBPs) mediate these processes[12], making the interaction between the environment and the insect, and determining the behavioral response. Hence, OBPs have been recognized as good candidates for genetic control programs of insect pests[13,14] and hematophagous insect vectors[15–18].

One important step in the process of prospecting genes for genetic control is the understanding of expression patterns of candidate genes in the target organism. Quantitative polymerase chain reaction ($_q$PCR) has become an invaluable tool in studies of gene expression due to its sensitivity and precision, enabling the detection of different contrasts among RNA samples. Differential expression of some dipterans' OBP family members has been studied by $_q$PCR in different tissues[14] and in different adult life stages[13].

In this study, the goal was to assess by $_q$PCR the expression profile of three OBPs genes, *Obp19d_2*, *Obp56a* and *Obp99c*, at different developmental stages on the West Indies fruit fly. These genes were chosen due to their differential expression between virgin and post-mating males in a previous in silico analysis, suggesting them as potential targets for further studies. However, there is a dearth of information regarding gene expression, as well as good reference genes for data normalization in $_q$PCR expression studies in most animal species, let alone for *A. obliqua*. Thus, in addition to an analysis of OBPs' expression, this study also aimed to determine a set of reliable reference genes for expression studies across different developmental stages for the West Indies fruit fly. In order to do so, we investigate patterns of gene expression for eight potential reference genes using $_q$PCR: *ribosomal protein L18* (*rpl18*), *β-Tubulin* (*btub*), *elongation factor 1α* (*ef1a*), *ribosomal protein S17* (*rps17*), *glyceraldehyde-3-phosphate dehydrogenase* (*gapdh*), *actin* (*act*), *Syntaxin* (*Syx*) and *Troponin C* (*TpnC*) . Our results indicate that the genes *rpl18*, *rps17* and *ef1a* show the highest expression stability throughout development, making them potentially reliable reference genes for gene expression studies in *A. obliqua*. Our results also show that some OBP genes are differentially expressed across different genders and reproductive stages, which make them good candidates for future application in genetic control.

Results

Efficiency of candidate reference genes. A quantitative study of relative gene expression requires normalization of data for target genes against a set of reference genes that are used as internal standards and are evenly expressed among tested conditions. Therefore, we investigated gene expression patterns of eight potential reference genes in different *A. obliqua* developmental stages. To determine the best primer concentrations, we tested three different concentrations in several primer combinations. The optimal concentration was $0.3\,\mu$M compared to $0.1\,\mu$M and $0.15\,\mu$M. For sensitivity assays, all genes tested, with the exception of *syx* and *TpnC*, showed efficiency values between 95% to 105% and standard curve correlation coefficients of 0.99 (Table 1; Supplementary Fig. S1 online; Supplementary Table S1 online), meaning that the amount of PCR product approximately doubles each cycle[19]. For all genes, the melting curves showed one single peak, indicating that assays led to specific amplifications (Supplementary Fig. S1 online). Because *syx* and *TpnC* failed to show acceptable amplification efficiency ($>$105%), they were disregarded as reference genes, leaving six genes that were further evaluated.

Candidate genes with relative expression stability across life stages. The C_q (quantitative cycle) values generated from the six candidates for reference genes were used to estimate stability of gene expression across different life stages. Figure 1A,B reveals different patterns of expression and variation among candidate genes. The difference in average C_qs among the candidate genes was only 2.18 cycles, with the lowest C_qs values (19.23) for *ef1a* gene and the highest (21.41) for *rpl18* gene (Fig. 1A). The investigation of variation across different life stages revealed two different groups (Fig. 1B): one group consisted of *rpl18*, *rps17*, *ef1a* and *btub* with standard deviation (SD) $<$ 1.0; and another group consisted of *gapdh* e *act* with SD $>$ 1.0. There are several different algorithms that investigate patterns of expression to identify the best reference genes which are available for $_q$PCR studies. In this study, by using three tools (NormFinder, BestKeeper and RefFinder) available online, we identified the most stable genes as being *rpl18*, *rps17* and *ef1a* (Table 2). It is worth noting that all three algorithms indicated that *act* exhibited the largest variation and lowest stability, and is therefore not suitable as a reference gene for the experimental conditions used in this study.

Investigation on target genes involved in reproduction. We used the results generated from differential expression of transcripts of virgin mature and post-mating males of *Anastrepha fraterculus*

Symbol	Gene Name	GeneBank Acession Number	Function	Ta (°C)	PCR Efficiency(%)	Correlation Coefficient	Forward primer	Reverse primer
act	Actin	L12254.1	structural constituent of cytoskeleton	59	95.3	0.994	TGACGATGAGGTTGCTGCTT	TTGTCCCATACCGACCATCA
btub	β-Tubulin	EU980443.1	structural constituent of cytoskeleton	59	101.3	0.995	CCCTCACCCAAAGTATCAGAC	TGGTCAATTTCAGAGTGCGG
ef1a	Elongation factor-1α	GU339154.1	translation elongation factor activity	59	100.0	0.999	TTGCCTTCGTACCAATCTCTG	AACCTTCCATCCCTTGAACC
gapdh	Glyceraldehyde 3 phosphate dehydrogenase	GU269901.1	(NAD+) (phosphorylating) activity	58	98.7	0.995	AAACTGTGGCGTGATGGAC	GACTGAGACATTGGGTGTGG
rpL18	Ribosomal proteinL18	NM139834.2	structural constituent of ribosome	58	99.5	0.996	GTGCCTAAGATGACCATTTGC	GACCTCCAGCCTTAATGATACG
rpS17	Ribosomal proteinS17	NM079278.2	structural constituent of ribosome	58	103.4	0.993	GCCGCAAAGGTCATAATTGAG	GCTACTTCTTCACAAATACGCT
syx	Syntaxin	NM168905	SNAP receptor activity	58	107.2	0.935	GAAATCTACAAGAAACTCGGTGC	CACACTTGTTTGCTCCACAG
tpnC	TroponinC	EZ423429.1	calcium ion binding	58	109.1	0.998	TCCTGAAAGAGTTAGACGAT	CTCTCCTGTCATCATTTCCAT

Table 1. Eight candidates for reference gene in expression studies by qPCR in *Anastrepha obliqua*.

transcriptome to select some differentially expressed genes. The rationale was that this difference indicates genes potentially involved in the reproductive process. Of the differentially expressed genes obtained in our in silico analysis, three OBPs were selected to represent a gene family already associated with the reproduction process[8,20]. Table 3 shows the differentially expressed genes selected from the *Anastrepha* transcriptome. Based on the fold-change differences, two OBP genes were chosen for being supposedly up-regulated (*Obp19d_2* and *Obp99c*) and one down-regulated (*Obp56a*) by mating.

Expression profiles of OBPs in different life stages. Despite the important role of OBPs in the reproductive behavior of Diptera, we still lack greater knowledge about the function and expression patterns of these proteins at different developmental stages of insects. Thus, we used $_q$PCR to determine the expression profiles of the three selected OBPs in *A. obliqua* larvae, pupae, and mature virgin and post-mating adults of both sexes. In order to generate OBP sequences from *A. obliqua*, we used *A. fraterculus* contig sequences to create primers for cloning and sequencing these genes from *A. obliqua*. This procedure generated full OBP sequences from *A. obliqua* from which we designed qPCR primers to study *A. obliqua* expression. Sequences from *A. obliqua* were very similar to OBPs from *A. fraterculus*, sharing 100% similarity for *Obp19d_2*, 97% for *Obp56a* and 93% for *Obp99c* at the amino acid level. Furthermore, the region we used to derive the qPCR primers had 100% similarity between the species, so it is likely that these primers could be used to investigate expression studies in both species, though we have only tested them for *A. obliqua*. The sensitivity of each primer set determined by standard curves with seven dilutions is shown in Table 4. These primers showed efficiency close to 100%, and standard curve correlation coefficients of 0.99 (Supplementary Fig. S2 online; Supplementary Table S2 online). The specificity of each primer set was confirmed by single melting peaks (Supplementary Fig. S2 online). For gene expression analysis, reactions were performed with three technical replicates of three biological replicates for each life stage. The relative normalized expression was calculated by $2^{-\Delta\Delta C_q}$ method[21] in contrast with three validated reference genes expression (*rpl18*, *rps17* and *ef1a*).

The relative transcript abundance of OBPs in larvae II, larvae III, early pupae, late pupae, mature virgin and post-mating female and mature virgin and post-mating male is exhibited in Fig. 2. All three OBP genes showed low expression during immature stages (pupae and larvae) compared to adult phases. We failed to detect significant differences in *Obp19d_2* expression between virgin and post-mating male or female adults. However, we observed a higher expression of this OBP in post-mating males when compared to post-mating females p (<0.05). As for *Obp56a*, a 2-fold decrease in abundance (unpaired t test, p < 0.01) was observed in post-mating females compared to virgin females, whereas no significant change was detected in males among different profiles. The levels of expression of this OBP in virgin and post-mating males were as low as that observed for post-mating females, but significant difference was observed only between post-mating males and females (p < 0.05). On the other hand, *Obp99c* had low expression values for mature virgin and post-mating females, comparable to the expression observed in the early stages of development. However, we observed a significantly higher expression in virgin and post-mating males in comparison to virgin and post-mating females (2-fold and 3-fold increase, respectively) (unpaired t test, p < 0.001).

Discussion

Normalization is an essential step for the quantitative study of relative gene expression, which means that the relative expression of a target gene must be determined against a reference gene, whose expression
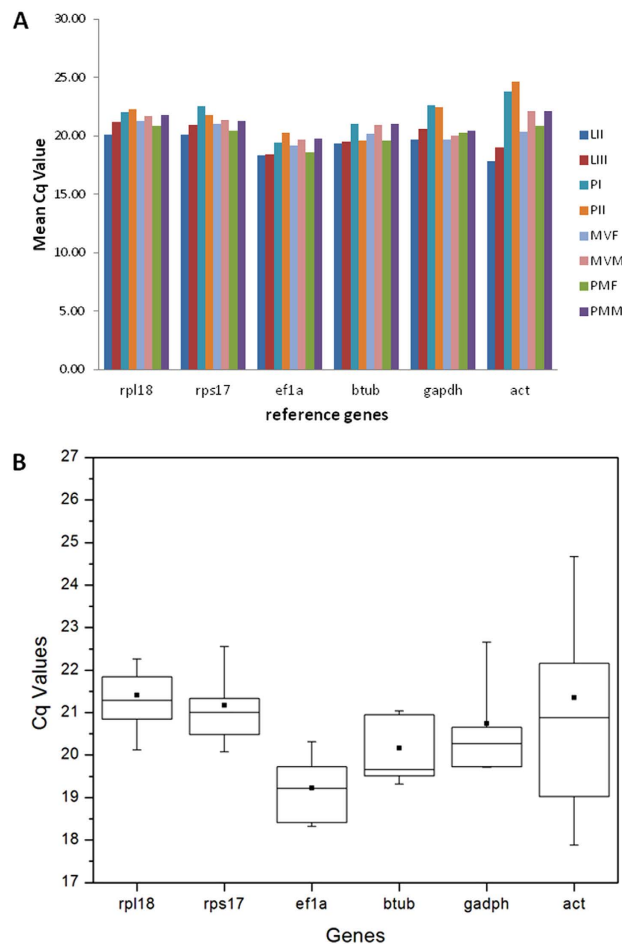
Figure 1. Gene expression stability of candidate reference genes. (**A**) Gene expression across different *Anastrepha obliqua* life stages. The bars show mean C_q values for each life stage from three biological replicates. (**B**) Variation in expression levels of candidate reference genes. Box-plots show medians (lines), means (dots), 95th percentile (boxes) and ranges (whiskers); (n = 24). The variation in expression level for each candidate gene was estimated across eight different life stages from *Anastrepha*, each stage with three biological replicates. The genes and their respective acronyms are: *ribosomal protein L18* (*rpl*18), *β-Tubulin* (*btub*), *elongation factor 1α* (*ef1a*), *ribosomal protein S17* (*rps*17), *glyceraldehyde 3-phosphate dehydrogenase* (*gapdh*) and *actin* (*act*). LII - larval stage II; LIII - larval stage III; PI - pupal stage I; PII - pupal stage II; MVF - mature virgin female; MVM - mature virgin male; PMF - post-mating female; PMM - post-mating male.

Gene	NormFinder	BestKeeper	RefFinder
*rpl*18	0.10 (1)	0.54 (1)	1.00 (1)
*rps*17	0.14 (2)	0.55 (2)	1.68 (2)
ef1a	0.40 (3)	0.59 (3)	3.00 (3)
btub	0.83 (5)	0.63 (4)	4.23 (4)
gapdh	0.73 (4)	0.91 (5)	4.73 (5)
act	1.65 (6)	1.82 (6)	6.00 (6)

Table 2. Expression stability values of the candidate reference genes through different life stages of *A. obliqua*, according to three algorithms (ranking in parentheses).

Gene	Symbol	logFC[a]	logCPM[b]	Most expressed in
Odorant-binding protein 19d	*Obp19d_2*	−3.47735	6.501031	post-mating male
Odorant-binding protein 56a	*Obp56a*	4.199486	8.609884	mature virgin male
Odorant-binding protein 99c	*Obp99c*	−4.62665	11.25094	post-mating male

Table 3. Differentially expressed genes selected as candidates for expression studies. [a]logFC: $\log_2$ fold-change; ratio of virgin and post-mating expression values. [b]logCPM: $\log_2$ counts-per-million.

levels should be stable and not influenced by the experimental procedure[22]. Since levels of expression in reference genes may vary depending on the species being studied, or even on environmental or developmental conditions, no reference gene should be considered universal[23]. Therefore, every $_q$PCR experiment should begin with the identification and validation of a set of reference genes that can be reliably used according to a specific condition.

To our knowledge, this is the first study for suitable reference genes for gene expression studies in the genus *Anastrepha* across different life stages. Eight potential reference genes were analyzed and ranked according to $_q$PCR efficiency and expression stability. From them, three (*rpl18*, *rps17* and *ef1a*) were reliable as reference genes because they showed appropriate efficiency values and constitutive expression levels across the different stages we contrasted. In fact, the efficiency and the expression stability of these three genes have already been reported for other insects. The gene *rpl18* was the most stable for the bed bug *Cimex lectularius* and the Colorado potato beetle *Leptinotarsa decemlineata*[24,25]. For the mosquito *Aedes aegypti*, *rps17* was the most stable gene across different ages[26] and other experimental conditions[27–29]. On the other hand, *rps17* was not reliable as reference gene in three Calliphorid flies[30]. The *ef1a* gene was analyzed on *Bombus terrestris*, *Bombus lucorum*, *Bactrocera dorsalis*, *Cimex lectularius* and *Leptinotarsa decemlineata*[24,25,31–33] and among these species *ef1a* was reported as a good reference gene only for *B. lucorum*. A range of mixed results for the same potential reference gene in different studies and experimental conditions reinforces the importance of search and normalization of suitable reference genes for $_q$PCR studies. Indeed, even species from the same genus as *B. terrestris* and *B. lucorum*, have required different reference genes for $_q$PCR studies[31], which emphasizes the need of normalization even for closely related species.

The wide range of expression levels shown by *act* in our experiments with *A. obliqua* also highlights the importance of normalization and species-specific studies for $_q$PCR. This gene is widely used as reference in $_q$PCR experiments in a range of organisms, because it is assumed to be constitutively expressed. For instance, *act* is considered one of the most stable gene in dipterans such as the tephritid *B. dorsalis*[33], and *D. melanogaster*[34]. However, our data showed that, among the potential reference genes tested, *act* had the widest range in expression values across different stages of development in *A. obliqua*, making it inappropriate as a reference gene in the conditions we considered. We should emphasize that even though *rpl18*, *rps17* and *ef1a* have been confirmed as appropriate reference genes for *A. obliqua* across the developmental stages we considered, any condition other than the ones tested here may require new validations.

An attribute that makes insects a group of great success is their sophisticated olfactory system, which is paramount to their survival and reproduction. Since we still know very little of the molecular events affecting behavior and biology of the West Indies fruit fly, information on the expression profiling of OBPs across different developmental stages and pre- and post-mating stages contributes relevant information about the role of these proteins on reproduction. Such knowledge might help towards the development of new and more environmentally friendly pest control strategies. *Obp19d_2*, *Obp56a* and *Obp99c* showed different levels of expression across development, even in immature stages (larvae and pupae), despite the very low relative transcript abundance compared to adult stages. Although several of the different expression levels here found are significant, some profiles show higher standard deviation, particularly virgin and post-mating adults in *Obp19d_2* and virgin adults in *Obp56a* (Fig. 2). Because these deviations are consistent across technical replicates, we believe that they are due to variation in expression among individuals. OBPs belong to an epistatic network of pleiotropic genes that are plastic, dynamic and modulated by several factors, such as sex, physiological state, social relationships, as well as by the environment[35]. Though we may be able to control several of these aspects in the lab, we cannot control for population variation, since we performed our study in non-inbred populations. Variation at the population level for OBP genes and their expression might be adaptive, since it might enable different response to important environmental cues. The combination of population variation, epistatic interactions, and environmental effects makes it a daunting task to investigate OBP expression, but our results still show relevant biological associations.

In *D. melanogaster*, microarray gene expression studies showed that *Obp19d* had an adult-biased expression, as observed in our data, whereas *Obp56a* had a larval-biased expression, which differs from our results. As for *Obp99c*, high expression levels were observed in both larvae and adults of *D. melanogaster*[11], which differs from *A. obliqua* since the expression was very low in immature stages as well as in female adults. Overall, the low levels of expression of these OBPs in immature stages of *A. obliqua* lead

Symbol	GenBank accession number	Tm (°C)	PCR efficiency (%)	Correlation coefficient	Forward Primer	Reverse Primer
Obp19d_2	KP939313	60	100.5	0.999	CAAATGACAAAGCTAAGTGCCTG	CATCACCATCAGCAACTGC
Obp56a	KP939314	60	100.1	0.994	TGGCGACAGTTCTACTCTG	GACGATAGGCAATCCAAGC
Obp99c	KP939315	60	100.0	0.994	CGGATGTACCAGAGGTAAACTC	ATCTTCAGCATTATTTGGGTCC

Table 4. OBP genes studied in *Anastrepha obliqua*. Efficiency and correlation coefficient determined by standard curves for primers of the selected genes.
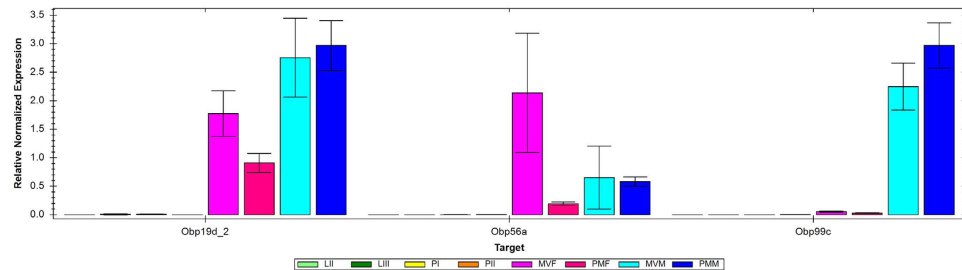
Figure 2. Relative normalized expression of *Obp*19d_2, *Obp*56a and *Obp*99c through different *Anastrepha* life stages. LII - larval stage II; LIII - larval stage III; PI - pupae stage I; PII - pupae stage II; MVF - mature virgin female; MVM - mature virgin male; PMF - post-mating female; PMM - post-mating male.

us to speculate that, at least in larvae, these genes seem not be involved with some basic biological needs such as nutrient sensing. Furthermore, the adult-biased expression observed for these OBP genes in *A. obliqua* may be related to the search for partners and oviposition sites.

In adults, chemically-induced changes on females due to mating, mainly related to seminal fluid proteins, have been well characterized in *D. melanogaster*[36–38]; as well as mating consequences in males, like immunological sexual dimorphism[39] and stimulation of male accessory glands proteins[40]. For *A. obliqua*, we found a difference in expression levels between virgin and post-mating adults only for *Obp56a* in females, with a significant lower level of expression before mating. This difference may indicate possible regulation by mating and it could contribute to changes on female's behavior as it happens in *D. melanogaster* females that, after sperm contact due to mating, experience a decrease in attractiveness and receptiveness to other males[37]. In fact, *Obp56a* is regulated on females by sperm contact in *D. melanogaster*[36]. In *C. capitata*, females experience a switch on certain olfactory-mediated behaviors after mating: virgin females choose the odor of male pheromone instead of volatile host plant emissions and maintain this preference even without contact with males, whereas mated females have a stronger preference for host plant odors to oviposit[41]. Thereby, changes in chemoreception genes may be expected due to mating and, although the function of *Obp56a* is not known, the decrease of its abundance in *A. obliqua* post-mating females may suggest that this protein would be more related to perception of male pheromones, or other aspects related to mating, instead of the search for oviposition sites.

Furthermore, *Obp56a* also showed sexual dimorphism in expression, since it had a higher expression on post-mating males than on females. The same was observed on mated males for *Obp19d_2* and *Obp99c*. This differential expression may be a consequence of the different gender response to the chemical environment. This sexual dimorphism is expected since males and females may have very different chemical cues. The higher expression of OBPs on post-mating males would suggest that these genes could be related to seeking partners, responding to female's pheromone, while post-mating females gear their efforts to seeking oviposition sites, and therefore may be more responsive to volatile host plant emissions. *Obp19d* is also known as pheromone binding protein-related protein (*PBPRP2*), indicating its role with reproduction in *D. melanogaster*[42]. On the other hand, *Obp99c* was significantly more expressed on virgin and mated males suggesting that this protein may be related to functions restricted to this gender. Genes in the *Obp99* class have been associated with the smell of benzaldehyde in *D. melanogaster*[43,44], which is common to several plants, especially in flowers and fruits[45], making it a possible class of genes associated with host preference. Furthermore, benzaldehyde increased attraction of males to sex pheromone lures in *Spodoptera exigua* by over 100%[46], making it possible that *OBP99c* is a gene involved with host and mating preference. A widespread sexual dimorphism in chemosensory genes expression, including OBPs, was documented in *D. melanogaster*, in response to different environmental, physiological and social conditions[11]. Indeed, responses to different odorants have been observed as sexually dimorphic[47].

The results here obtained make *Obp99c* a good candidate for the application in techniques that use genetic manipulation like the Sterile Insect Technique (SIT), since its relative expression was almost exclusive to males. Likewise, the high expression of *Obp56a* on virgin females may be an important factor involved in their search for males and or mating sites. Some OBPs like *Obp19d_2*, *Obp56a* and *Obp99c*,

may regulate parts of sexual behavior such as the search for hosts and sexual partners, therefore, an improvement on the knowledge about these proteins in agricultural pests like *A. obliqua* impacts not only the understanding of their biology, but may contribute to future research on the use and manipulation of OBPs for the development of new environmentally friendly pest control strategies. Furthermore, this knowledge may aid on the development of novel species-specific attractants and repellents, lures and traps for pest control programs.

Methods

Sampling. Flies were obtained from an *A. obliqua* population maintained in natural photoperiod at $28 °C \pm 2.0 °C$ in the Population Genetics and Evolution laboratory, at the Federal University of São Carlos, Brazil. Different stages of development were used: 2nd and 3rd larval instars (LII and LIII) taken from mangoes offered to the population for oviposition, using mean standard length of each instar (3–5 mm for 2nd instar and around 8.6 mm for 3rd instar)[48]; early pupal stage (PI), taken around 24 hours after pupation, later pupal stage (PII), taken 10 days after pupation; sexually mature virgin males (MVM) and females (MVF), taken 10 days after adult emergence; and post-mating males (PMM) and females (PMF), taken 24 hours after mating. For each stage, three independent biological replicates were produced from pools of five different individuals.

RNA isolation and cDNA synthesis. Total RNA was extracted using the TRIzol/chloroform protocol[49]. RNA quality was visualized in an agarose gel electrophoresis for integrity, absorbance was measured in Nanodrop (Thermo Scientific^{TM}) and then RNA was quantified using a Qubit® 2.0 Fluorometer (Invitrogen^{TM}). The extracted RNA was treated with DNase I Amplification Grade (Invitrogen^{TM}) prior to reverse transcription. One μg of treated RNA was converted into cDNA using iScript^{TM} cDNA Synthesis Kit (Bio-Rad). To perform the assays, all cDNA samples were diluted to 4 ng/μl. The samples were stored at $-80 °C$ until used.

Candidate genes and primer design. For the validation of reference genes, eight constitutive genes were chosen, most of them already used as reference for qPCR experiments in insects[24,31–33,50,51]. In order to obtain DNA sequences for *A. obliqua* for these genes, *D. melanogaster's* sequences from the following genes were selected from GenBank database: *rpl18* (NM_139834.2), *btub* (EU980443.1), *ef1a* (GU339154.1), *rps17* (NM_079278.2), *gapdh* (GU269901.1), *syx* (NM_168905.1), *tpnC* (EZ423429.1) and *act* (L12254.1). These sequences were used to BLAST against an *A. fraterculus* transcriptome (Rezende, V.B.; Congrains, C.; Lima, A.L.; Campanini, E.B.; Nakamura, A.M.; Oliveira, J.L.; Chahad-Ehlers, S.; Sobrinho-Jr., I.; de Brito, R.A.; unpublished data) that was the only transcriptome data available at that time in our lab. We did not consider this to be much of a problem, because *A. fraterculus* and *A. obliqua* are closely related species and these genes are evolutionarily very conservative. We performed a differential expression analysis between mature virgin male and post-mating male in the *A. fraterculus* transcriptome data set using the software edgeR following default specifications[52] and identified three differentially expressed OBP genes (*Obp19d_2*, *Obp56a* and *Obp99c*) between these profiles. These genes were chosen following the premise that the difference in expression indicates that the gene may play an important role in reproduction. We used inferred amino acid sequences for each of these OBPs to perform a BLAST search on GenBank with the best close match options to *C. capitata*, which identified the sequences *Obp19d_2* (XM_004524978.1), *Obp56a* (XM_004517843) and *Obp99c* (XM_004521129.1). We aligned these sequences to their orthologous copies obtained from the *A. fraterculus* transcriptome using MAFFT[53] and designed primers to amplify conserved regions of each gene in *A. obliqua*. These PCR products were PEG purified and sent to a specialized company (MACROGEN INC., South Korea) to carry out the sequencing. Based on the Sanger sequences, $_q$PCR primers were designed in regions with 100% of similarity between each *A. obliqua* sequence and *A. fraterculus* contig, using in flanking intron-exon junctions if possible, and checked for duplex formation and secondary structures with the OligoAnalyzer tool (http://www.idtdna.com/analyzer/applications/oligoanalyzer/).

$_q$PCR. For $_q$PCR assays, *SsoFast^{TM}* Evagreen® Supermix (Bio-Rad) was used with final primer concentration of 0.3 μM in a total reaction of 10 μl. Reactions were performed in a CFX96 Touch^{TM} Real-Time PCR Detection Systems (Bio-Rad) using the following protocol: 95° for 30 seconds followed by 40 cycles of dissociation at 95 °C for 10 seconds and annealing and extension at 60 °C for 1 min. For melt curve analysis we used a protocol with temperatures that varied from 65 °C to 95 °C with increments of 0.5 °C for 5 seconds and continuous fluorescent measurements. The efficiency of the reactions at high and low concentrations were assessed to ensure that primer efficiency would not vary significantly due to variations in cDNA concentration among samples. Therefore, efficiency of $_q$PCR (E) and the correlation coefficients (R^2) were calculated for each pair of primer using standard curves of correlation between cDNA concentration and Cq values. They were generated with a serial dilution (7x) of a pool with all cDNA samples with an initial concentration of 4 ng/μl, in triplicate. Moreover, to check for the uniform expression of reference genes across life stages, which is necessary for the gene stability calculation and to calculate the relative normalized expression for OBPs, assays were performed in triplicate, using each biological replicate at a concentration of 4 ng/μl. No-template control (NTC) was used in all $_q$PCR experiments to confirm that there was no contamination with exogenous material.

Gene stability analyses of reference genes. The C_q average values calculated for the candidate reference genes in CFX Manager^{TM} software (Bio-Rad) were analyzed by three different descriptive statistic tools freely available online: NormFinder software (http://moma.dk/normfinder-software), BestKeeper software (http://www.gene-quantification.de/bestkeeper.html) and RefFinder web-based tool (http://www.leonxie.com/referencegene.php). NormFinder calculates not only the overall variation of the candidate normalization genes but also the variation between sample subgroups of the sample set[54]. The software BestKeeper estimates the correlation of expression levels of all candidates and the genes more stable are combined in an index. Three main indicators were used to determinate the best candidate: standard deviation, covariance and correlation analysis[55]. RefFinder ranks all candidates of reference genes based on the main algorithms currently used, including Normfinder and BestKeeper, as well as the geNorm algorithm[56] and the Comparative ΔC_t Method[57]. Thus, RefFinder allows the choice of the best candidates based on the final ranking results of different programs[58–62].

Relative normalized expression calculation of OBPs. Like the reference genes, data analysis for OBPs was also performed using CFX Manager Software^{TM} (Bio-Rad). The values of relative normalized expression were calculated for each gene using the $2^{\Delta\Delta C_t}$ method[21], in contrast with the three validated reference genes expression. Since experiments were performed in different plates, inter-run calibrators were used and a common threshold was set. Data were statistically compared by ANOVA and Unpaired 2-tailed t test using Prism 5.01 software (GraphPad Software, San Diego, CA, USA).

References

1. Marsaro Júnior, A., Deus, E., Ronchi-Teles, R., Adaime, B. & Silva Júnior, R. Species of *Anastrepha* (diptera: Tephritidae) captured in a guava orchard (*Psidium guajava* l., myrtaceae) in Boa Vista, Roraima, Brazil. *Brazilian Journal of Biology.* **73,** 879–886 (2013).
2. Bjelis, M., Ljubetic, V. & Novosel, N. Symposium on fruit flies of economic importance. *Proceedings of the 7ᵗʰ International Symposium* **7th,** 255–259 (2007).
3. Heath, R. R., Epsky, N. D., Dueben, B. D., Guzman, A. & Andrade, L. E. Gamma radiation effect on production of four pheromonal components of male mediterranean fruit flies (Diptera: Tephritidae). *Journal of Economic Entomology* **87,** 904–909 (1994).
4. Calcagno, G. E., Manso, F. & Vilardi, J. C. Comparison of mating performance of medfly (diptera, tephritidae Diptera, Tephritidae) genetic sexing and wild type strains: field cage and video recording experiments. *Florida Entomologist* **85,** 41–50 (2002).
5. Schetelig, M. F., Milano, A., Saccone, G. & Handler, A. M. Male only progeny in *Anastrepha suspensa* by rnai-induced sex reversion of chromosomal females. *Insect Biochemistry and Molecular Biology* **42,** 51–57 (2012).
6. Benton, R. On the origin of smell: odorant receptors in insects. *Cell.Mol.Life Sci* **63,** 1579–1585 (2006).
7. Matsuo, T., Sugaya, S., Yasukawa, J., Aigaki, T. & Fuyama, Y. Odorant-binding proteins *OBP57d* and *OBP57e* affect taste perception and host-plant preference in *Drosophila sechellia*. *PLoS Biol* **5,** e118 (2007).
8. Pelletier, J., Guidolin, A., Syed, Z., Cornel, A. J. & Leal, W. Knockdown of a mosquito odorant-binding protein involved in the sensitive detection of oviposition attractants. *J Chem Ecol* **36,** 245–248 (2010).
9. Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annual Review of Entomology* **58,** 373–391 (2013).
10. Yi, X., Zhao, H., Wang, P., Hu, M. & Zhong, G. Bdor/orco is important for oviposition-deterring behavior induced by both the volatile and non-volatile repellents in *Bactrocera dorsalis* (*Diptera*: Tephritidae). *Journal of Insect Physiology* **65,** 51–56 (2014).
11. Zhou, S., Stone, E. A., Mackay, T. F. C. & Anholt, R. R. H. Plasticity of the chemoreceptor repertoire in *Drosophila melanogaster*. *PLoS Genet* **5,** e1000681 (2009).
12. Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nature* **293,** 161–163 (1981).
13. Zheng, W. *et al.* Identification and expression profile analysis of odorant binding proteins in the oriental fruit fly *Bactrocera dorsalis*. *International Journal of Molecular Science* **14,** 14936–14949 (2103).
14. Siciliano, P. *et al.* Sniffing out chemosensory genes from the mediterranean fruit fly. *Ceratitis capitata*. *PLoS ONE* **9,** e85523 (2014).
15. Bohbot, J. & Vogt, R. G. Antennal expressed genes of the yellow fever mosquito (*Aedes aegypti* l.); characterization of odorant-binding protein 10 and takeout. *Insect Biochemistry and Molecular Biology* **35,** 961–979 (2005).
16. Leal, W. S. *et al.* Reverse and conventional chemical ecology approaches for the development of oviposition attractants for culex mosquitoes. *PLoS ONE* **3,** e3045 (2008).
17. Lagarde, A. *et al.* The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. *Journal of Molecular Biology* **414,** 401–412 (2011).
18. Tsitsanou, K. E. *et al. Anopheles gambiae* odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cell.Mol.Life Sci.* **69,** 283–297 (2012).
19. Bustin, S. A. *et al.* The miqe guidelines: Minimum information for publication of quantitative real-time pcr experiments. *Clinical Chemistry* **55,** 611–622 (2009).
20. Li, S. *et al.* Multiple functions of an odorant-binding protein in the mosquito *Aedes aegypti*. *Biochemical and Biophysical Research Communications* **372,** 464–468 (2008).
21. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative pcr and the 2(-delta delta ct) method. *Method. Methods* **25,** 402–408 (2001).
22. Radonic, A. *et al.* Guideline to reference gene selection for quantitative real-time pcr. *Biochemical and Biophysical Research Communications* **313,** 856–862 (2004).
23. Chen, D., Pan, X., Xiao, P., Farwell, M. A. & Zhang, B. Evaluation and identification of reliable reference genes for pharmacogenomics, toxicogenomics, and small rna expression analysis. *Journal of Cellular Physiology* **226,** 2469–2477 (2011).
24. Mamidala, P., Rajarapu, S. P., Jones, S. C. & Mittapalli, O. Identification and validation of reference genes for quantitative real-time polymerase chain reaction in *Cimex lectularius*. *Journal of Medical Entomology* **48,** 947–951 (2011).
25. Shi, X. Q. *et al.* Validation of reference genes for expression analysis by quantitative real-time pcr in *Leptinotarsa decemlineata* (say). *BMC Research Notes* **6,** 93 (2013).
26. Cook, P. E. *et al.* The use of transcriptional profiles to predict adult mosquito age under field conditions. *Proceedings of the National Academy of Sciences* **103,** 18060–18065 (2006).
27. Hugo, L. E. *et al.* Field validation of a transcriptional assay for the prediction of age of uncaged *Aedes aegypti* mosquitoes in northern australia. *PLoS Negl Trop Dis* **4,** e608 (2010).

28. 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. *Journal of Insect Physiology* **57,** 1220–1226 (2011).

29. Coggins, S. A., Estévez-Lao, T. Y. & Hillyer, J. F. Increased survivorship following bacterial infection by the mosquito *Aedes aegypti* as compared to *Anopheles gambiae* correlates with increased transcriptional induction of antimicrobial peptides. *Developmental and Comparative Immunology* **37,** 390–401 (2012).

30. Cardoso, G. A., Matioli, C. C., Azeredo-Espin, A. M. L. & Torres, T. T. Selection and validation of reference genes for functional studies in the Calliphoridae family. *Journal of Insect Science* **14,** 1–15 (2014).

31. Hornáková, D., Matoušková, P., Jirí, K., Valterová, I. & Pichová, I. Selection of reference genes for real-time polymerase chain reaction analysis in tissues from *Bombus terrestris* and *Bombus lucorum* of different ages. *Analytical Biochemistry* **397,** 118–120 (2010).

32. 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 Molecular Biology* **11,** 76 (2010).

33. 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).

34. 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*. *Journal of Insect Physiology* **57,** 840–850 (2011).

35. Arya, A. L. & Weber, G. H. *et al.* Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*. *Genetics* **186,** 1475–1485 (2010).

36. Wolfner, M. F. Tokens of love: Functions and regulation of *Drosophila* male accessory gland products. *Insect Biochemistry and Molecular Biology* **27,** 179–192 (1997).

37. Tram, U. & Wolfner, M. F. Seminal fluid regulation of female sexual attractiveness in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences* **95,** 4051–4054 (1998).

38. McGraw, L. A., Gibson, G., Clark, A. G. & Wolfner, M. F. Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*. *Current Biology* **14,** 1509–1514 (2004).

39. McKean, K. A. & Nunney, L. Sexual selection and immune function in *Drosophila melanogaster*. *Evolution* **62,** 386–400 (2008).

40. DiBenedetto, A. J., Harada, H. A. & Wolfner, M. F. Structure, cell-specific expression, and mating-induced regulation of a *Drosophila melanogaster* male accessory gland gene. *Developmental Biology* **139,** 134–148 (1990).

41. Jang, E. B. Effects of mating and accessory gland injections on olfactory-mediated behavior in the female Mediterranean fruit fly, *Ceratitis capitata Journal of Insect Physiology* **41,** 705–710 (1995).

42. Menuz, K., Larter, N. K., Park, J. & Carlson, J. R. An rna-seq screen of the *Drosophila* antenna identifies a transporter necessary for ammonia detection. *PLOS Genetics* **10,** e1004810 (2014).

43. Wang, P., Lyman, R. F., Shabalina, S. A., Mackay, T. F. C. & Anholt, R. R. H. Association of polymorphisms in odorantbinding protein genes with variation in olfactory response to benzaldehyde in *Drosophila*. *Genetics* **177,** 1655–1665 (2007).

44. Wang, P., Lyman, R. F., Mackay, T. F. C. & Anholt, R. R. H. Natural variation in odorant recognition among odorantbinding proteins in *Drosophila melanogaster*. *Genetics* **184,** 759–767 (2010).

45. Rodríguez, A., Alquézar, B. & Peña, L. Fruit aromas in mature fleshy fruits as signals of readiness for predation and seed dispersal. *New Phytologist* **197,** 36–48 (2013).

46. Deng, J. Y., Wei, H. Y., Huang, H. P. & Du, J. W. Enhancement of attraction to sex pheromones of *Spodoptera exigua* by volatile compounds produced by host plant. *Journal of Chemical Ecology* **30,** 2037–2045 (2004).

47. Swarup, S., Williams, T. I. & Anholt, R. R. H. Functional dissection of odorant binding protein genes in *Drosophila melanogaster*. *Genes, Brain and Behavior* **10,** 648–657 (2011).

48. DaCruz, I. B. M., Nascimento, J. C., Taufer, M. & Oliveira, A. K. in. *Moscas-das-frutas de importância econômica no Brasil: conhecimento básico e aplicado. Holos Editora* 6, 55–66.

49. Chomczynski, P. & Mackey, K. Modification of the tri reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *Biotechniques* **19,** 942–945 (1995).

50. Bagnall, N. & Kotze, A. Evaluation of reference genes for real-time PCR quantification of gene expression in the Australian sheep blowfly. *Lucilia cuprina. Medical and Veterinary Entomology* **24,** 176–181 (2010).

51. Majerowicz, D. *et al.* Looking for reference genes for real- time quantitative PCR experiments in *Rhodnius prolixus* (hemiptera: Reduviidae). *Insect Molecular Biology* **20,** 713–722 (2011).

52. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **25,** 139–140 (2010).

53. Katoh, K. & Standley, D. M. Mafft multiple sequence alignment software version 7: improvements in performance and usability. molecular biology and evolution. *Molecular Biology and Evolution* **30,** 772–780 (2013).

54. Andersen, C. L., Jensen, J. L. & Orntoft, 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 Research* **64,** 5245–5250 (2004).

55. 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. *Biotechnology* **26,** 509–515 (2004).

56. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3, research0034–research0034.0011 (2002).

57. Silver, N., Best, S., Jiang, J. & Thein, S. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology* **7,** 33 (2006).

58. Ovesná, J. *et al.* Validation of the *b-amy1* transcription profiling assay and selection of reference genes suited for a RT-qPCR assay in developing barley caryopsis. *PLoS ONE* **7,** e41886 (2012).

59. Chao, W. S., Dogramaci, M., Foley, M. E., Horvath, D. P. & Anderson, J. V. Selection and validation of endogenous reference genes for qRT-PCR analysis in leafy spurge (*Euphorbia esula*). *PLoS ONE* **7,** e42839 (2012).

60. Jacob, F. *et al.* Careful selection of reference genes is required for reliable performance of RT-qPCR in human normal and cancer cell lines. *PLoS ONE* **8,** e59180 (2013).

61. Cheng, D., Zhang, Z., He, X. & Liang, G. Validation of reference genes in *Solenopsis invicta* in different developmental stages, castes and tissues. *PLoS ONE* **8,** e57718 (2013).

62. Wang, Y. *et al.* Identification of endogenous reference genes for RT-qPCR analysis of plasma micrornas levels in rats with acetaminophen-induced hepatotoxicity. *Journal of Applied Toxicology* **33,** 1330–1336 (2013).

Acknowledgements

Author Contributions

A.M.N., S.C.E., A.L.A.L., C.H.T., I.S.J., F.R.T. and R.A.B. planned the experiments. A.M.N., S.C.E., C.H.T. and F.R.T. performed the experiments. A.M.N., S.C.E., F.R.T. and R.A.B. analysed the results and prepared the manuscript. All authors reviewed the manuscript.

Additional Information

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

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

How to cite this article: 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; doi: 10.1038/srep17480 (2016).