

# SCIENTIFIC REPORTS



OPEN

## Identification and preliminary characterization of chemosensory perception-associated proteins in the melon fly *Bactrocera cucurbitae* using RNA-seq

Received: 18 June 2015  
Accepted: 26 November 2015  
Published: 11 January 2016

Samia Elfekih<sup>1</sup>, Chien-Yu Chen<sup>2</sup>, Ju-Chun Hsu<sup>3</sup>, Mahdi Belcaid<sup>4</sup> & David Haymer<sup>5</sup>

An investigation into proteins involved in chemosensory perception in the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae) is described here using a newly generated transcriptome dataset. The melon fly is a major agricultural pest, widely distributed in the Asia-Pacific region and some parts of Africa. For this study, a transcriptome dataset was generated using RNA extracted from 4-day-old adult specimens of the melon fly. The dataset was assembled and annotated via Gene Ontology (GO) analysis. Based on this and similarity searches to data from other species, a number of protein sequences putatively involved in chemosensory reception were identified and characterized in the melon fly. This included the highly conserved “Orco” along with a number of other less conserved odorant binding protein sequences. In addition, several sequences representing putative ionotropic and gustatory receptors were also identified. This study provides a foundation for future functional studies of chemosensory proteins in the melon fly and for making more detailed comparisons to other species. In the long term, this will ultimately help in the development of improved tools for pest management.

Sophisticated mechanisms of chemosensory perception are essential to many aspects of insect behavior and survival. Insects must be able to detect the presence of specific chemicals in their environment over relatively long distances to find food sources and potential mates. In addition to these more universal aspects of chemosensory perception, the evolutionary trajectories followed by individual species to exploit different aspects of their environment, such as the presence of specific host plants and/or encounters with conspecific mates, are most likely reflected in the presence of defined sets of species-specific genes and proteins involved in chemosensory perception<sup>1,2</sup>.

Beginning primarily with work on *Drosophila melanogaster*, a considerable amount of information about genes and proteins involved in chemosensory perception specific to insects has been developed<sup>3</sup>. At the genome level in particular, the *Drosophila* studies have provided tremendously valuable reference sets of information to facilitate mining for similar datasets in a wide range of other insects<sup>4–6</sup>. Among these references, are sets of protein sequences, predicted mainly from transcriptome studies, involved in critical aspects of chemosensory perception such the odorant receptors and odorant binding proteins (OBPs), the ionotropic receptor (IR) proteins, and the taste or gustatory receptor (GR) proteins<sup>7</sup>.

Beyond *Drosophila*, however, only limited information is available for species such as those in the family Tephritidae that include many major agricultural pests<sup>8–10</sup>.

One of these is the melon fly, a highly polyphagous pest that attacks a wide range of economically important cucurbit<sup>11</sup>.

<sup>1</sup>Commonwealth Science and Industry Organization (CSIRO), Biosecurity flagship, P.O. BOX 1700, Canberra, ACT 2601, Australia. <sup>2</sup>National Taiwan University, Department of Bio-industrial Mechatronics and Engineering, Taipei, Taiwan. <sup>3</sup>National Taiwan University, Department of Entomology, Taipei, Taiwan. <sup>4</sup>Information and Computer Sciences, University of Hawaii at Manoa, Honolulu, Hawaii, USA. <sup>5</sup>Department of Cell and Molecular Biology, University of Hawaii at Manoa, Honolulu, Hawaii, USA. Correspondence and requests for materials should be addressed to S.E. (email: samia.elfekih@csiro.au)

	<i>B. cucurbitae</i> Assembly (K25)
# of reads	44,400,736*2
Min length	201
Max length	27,802
2 <sup>nd</sup> long contig	27,771
3 <sup>rd</sup> long contig	27,766
Total length	81,038,823
Avg. length	1,469
N50	3,117
>0 bp	55,141
(>100 bp)	55,141
(>500 bp)	30,810
(>1000 bp)	21,708

**Table 1. Statistics of the Trinity Assembly of the Melon fruit fly transcriptome.**

Here, using a newly generated transcriptome dataset, we have identified putative chemosensory proteins in this species and compared them to counterparts in congeneric species such as the oriental fruit fly, *Bactrocera dorsalis* and more distantly related species within the family Tephritidae such as the Mediterranean fruit fly *Ceratitis capitata*, the apple maggot *Rhagoletis pomonella* and the walnut husk fly *R. suavis*. These species, although also classified as economic pests<sup>12–14</sup>, are known to show both qualitative and quantitative differences in critical aspects of chemosensory perception. In addition, where appropriate, alignments to protein and gene sequences from *Drosophila*, a non-economically important species, were also made.

## Results

**Transcriptome assembly.** The newly-generated melon fruit fly transcriptome was derived from RNA obtained from a total of 30 four-day old melon fly adults. The assembly was performed using Trinity, with an optimal K-mer length set to 25. The assembly generated 55,141 isotigs with a minimum sequence length of 100 bp. The N50 was 3,117 bp and the average contig length was 1,469 bp (Table 1). The transcriptome assembly was deposited in the SRA database and is available under accession identifier SRP058791.

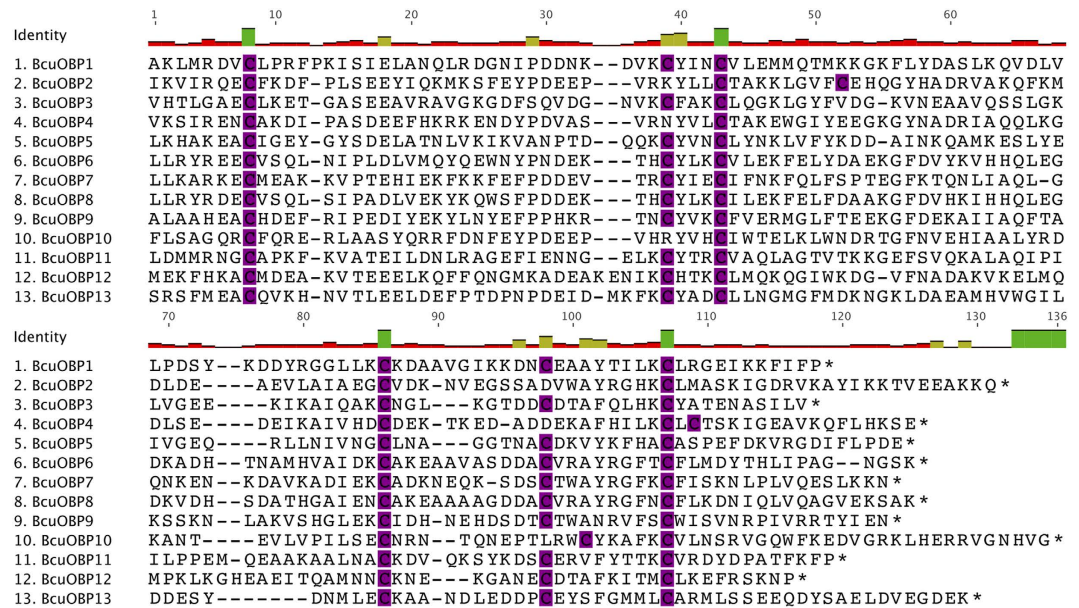
**Comparison of *B. cucurbitae* transcriptome to published datasets.** In order to assess the quality of our assembled transcriptome, we compared our results to those obtained from the combined assembly of four *B. cucurbitae* datasets, produced from egg, larvae, pupae and adult stages (SRAS: SRS691534, SRS691533, SRS691532 and SRS691531 respectively). While both our assembly and the one previously published<sup>15</sup> yielded similar GC rates (~39%), our results contained 14,863 and 21,062 fewer unigenes and isoforms respectively. Approximately 21.4% of the difference in isoforms was attributed to transcripts shorter than 300 bp, and 37.7% of the difference consisted of transcripts ranging in length between 300 and 800 bp. Additionally, our assembly identified 227 more transcripts of length greater than 10,000 bp and more proteins with a complete ORFs (25,943 in our assembly versus 12,936 in the published dataset<sup>15</sup>). The presence of fewer short transcripts in favor of longer ones could be an indication of a less fragmented assembly, which can be supported by our higher N50 (3,117 bp), versus 2,626 bp in the assembly retrieved from Genbank<sup>15</sup>. Based on this comparative analysis, the present melon fly transcriptome is of high quality and can, therefore, be used to investigate aspects related to the biology of this important species, such as the chemosensory perception genes and proteins.

**Gene Ontology analysis.** The Blast2GO annotation was used, and the results were visualized in the protein classification system PANTHER (<http://www.pantherdb.org>)<sup>16,17</sup>. A total of 25,943 transcripts were predicted, which represents 47% of the total number of generated contigs (55,141). Of these predicted transcripts, 19,071 (73%) were associated with GO terms and 4,661 transcripts were assigned to three main GO classes, specifically: Biological process (1,918), molecular function (1,086) and cellular components (98) (Supplementary material Fig. S1). Within the molecular function class, the most abundant GO associations were linked to catalytic (GO:0003824) [32.9%] and binding (GO:0005488) activities [32.1%]. At the biological process level, the metabolic process (GO: 0008152) was the most abundant association [27%] followed by cellular process (GO:0009987) [15.10%]. Other GO terms, such as hydrolase, transferase, transcription factor and nucleic acid binding proteins were also represented but at relatively lower levels of abundance within the dataset.

Further annotation of the transcriptome of the melon fruit fly permitted the extraction of transcripts corresponding to putative chemosensory genes and gene families and their corresponding proteins as described in the next sections.

**Candidate chemosensory genes.** A total number of 52 sequences were extracted and were putatively classified as follows: 13 Odorant-Binding Proteins (OBPs) and 1 Odorant receptor co-receptor (*BcuOrco*), 31 Ionotropic Glutamate Receptors (iGluRs), and 7 Gustatory receptors (GRs).

**OBP proteins.** The 13 candidate OBP proteins were further characterized by initially aligning them to each other (Fig. 1) and to similar sequences from *D. melanogaster*. This allowed us to classify them using already established nomenclature<sup>18</sup> and to organize them into different classes based on key features, such as the number of cysteine motifs present in each transcript.



**Figure 1.** Alignment of the Odorant-Binding protein putative sequences of *B. cucurbitae*. The conserved cysteine motifs are highlighted in purple.

The predicted OBPs were grouped into two main classes based on the following descriptors: First, “Classic” OBPs that contain six cysteine motifs at conserved positions. This class also includes the antennal binding proteins (ABPs). Second, the “Minus-C” class that include sequences lacking 2 cysteine residues, usually C2 and C5<sup>18</sup>. Overall, we detected three OBPs belonging to the Minus-C class and 10 falling into the “Classic” category. Members of an additional class designated “Plus-C” that possess 4 to 6 more cysteine residues were not identified in our dataset.

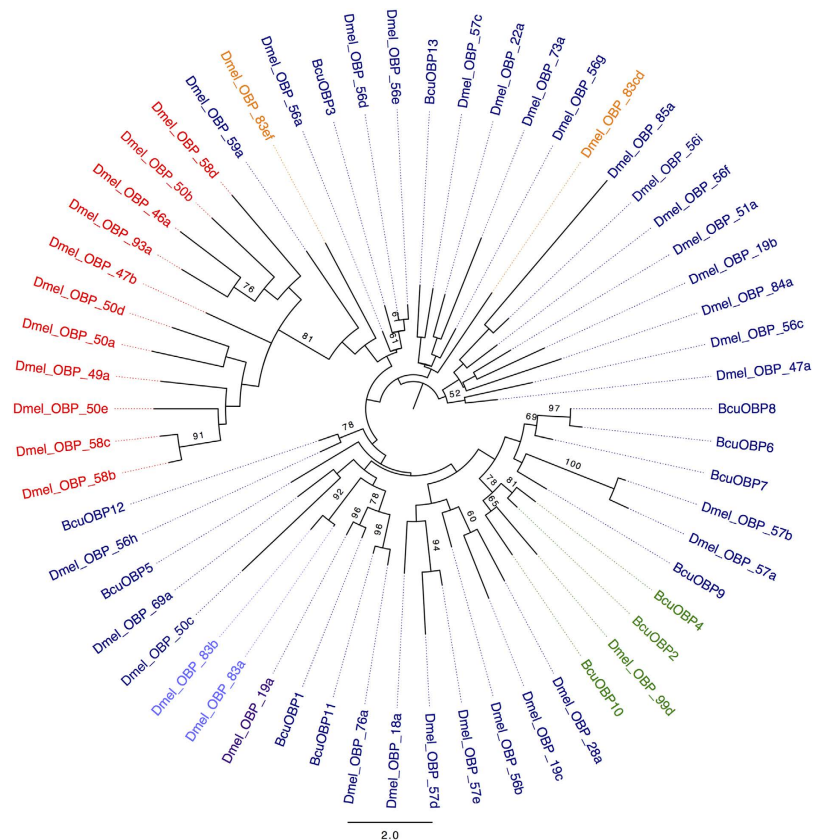
We used all the putative OBPs from *B. cucurbitae* and representative homologous sequences from *D. melanogaster* to build a Maximum Likelihood phylogenetic tree (Fig. 2). The tree shows a clear cluster representing the Minus-C OBP class described previously consisting of three of the BcuOBPs (BcuOBP2, BcuOBP4, BcuOBP10) with specific counterparts from *D. melanogaster* (Dmel\_OBP\_99d). This result also shows a large clade containing the classic OBPs, and there appears to be an expansion of sequences from *B. cucurbitae* including BcuOBP6, BcuOBP7, BcuOBP8, BcuOBP9, clustering with the *Drosophila* sequence Dmel\_OBP57a and Dmel\_OBP57b.

The Bcu\_OBP13 protein clusters with Dmel\_OBP57c, while the BcuOBP1 sequence clusters with Dmel\_OBP19a, which is known to be an antennal binding protein. Another of the classical OBP protein identified here (Bcu\_OBP11) appears to be closely related to the sequence corresponding to Dmel\_OBP76a. This is also known as LUSH, a protein involved in pheromone binding activities.

A protein BLAST analysis comparing the putative OBPs from the melon fly to their counterparts in *B. dorsalis*, a congeneric species, was also performed. Table 2 shows the results. Here, the similarity values range from 20 to 93%. In order to compare the melon fruit fly OBPs to other species from the Tephritidae family, we built a Maximum Likelihood phylogenetic tree using the 13 BcuOBPs identified and OBP sequences of other tephritid fruit flies already published and were retrieved from Genbank (Supplementary Material Table S1). These included 11 OBPs from the congeneric species *B. dorsalis*, designated as BdorOBPs<sup>19</sup>, 17 OBP from the Mediterranean fruit fly *C. capitata*, designated as Ccap OBPs<sup>20</sup>, 15 OBP from the apple maggot fly *Rhagoletis pomonella* designated as RpOBPs<sup>21</sup> and 9 OBP from *R. suavis* designated as RsOBPs<sup>22</sup>. The mid-point rooted ML tree (Fig. 3) shows clustering by OBP class. The melon fly Minus-C OBPs (BcuOBP2, BcuOBP4) cluster with their homologues in the oriental fruit fly (BdorOBP10) and in the medfly (CcapOBP99c). Within the same Minus-C clade, another cluster harbors BcuOBP10 and CcapOBP8a. The majority of the remaining OBPs belong to the classic clade and are grouped according to their percentage of similarity among the tephritid species. The Dimer and Plus-C OBPs are scattered among the classic OBPs clade, since we did not detect any OBPs belonging to those two classes in the melon fly.

**Orco gene.** The odorant receptor co-receptor in *B. cucurbitae* (BcuOrco) was identified as a protein sequence with 473 amino acids. When aligned with homologs from other fruit fly species (Fig. 4) i.e. *B. dorsalis* (ACC86853) and *C. capitata* (XP\_012156143) and to *D. melanogaster* (NP\_524235), the highest percentage of identity is scored between *B. cucurbitae* and *B. dorsalis*, (98%), followed by *C. capitata* (96%), then *D. melanogaster* (87%).

**Glutamate receptor proteins.** A total number of 31 putative members of the iGluR gene family were also identified from this analysis of the melon fly transcriptome (Supplementary Table S2). Based on similarities to sequences from *D. melanogaster*, we found one protein from the melon fly corresponding to the  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) subfamily and one corresponding to the NMDA subfamily<sup>23</sup>. The remaining transcripts were classified as representing members of the Kainate subfamily. Searches against



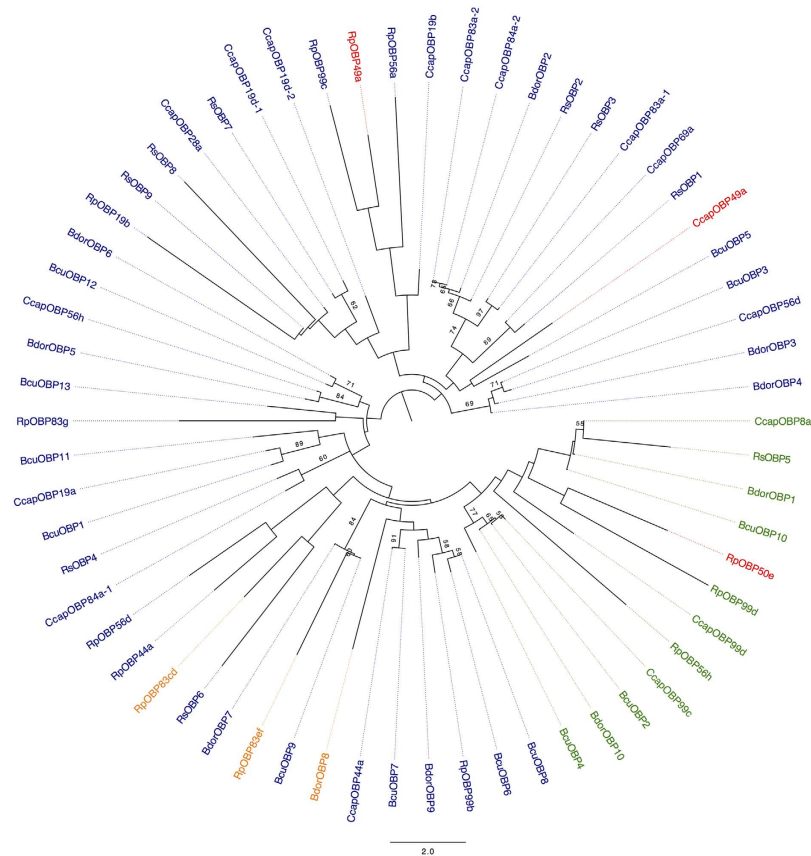
**Figure 2.** Maximum Likelihood Phylogenetic tree of *B. cucurbitae* and *D. melanogaster* OBPs. Bootstrap values greater than 50% are displayed (1000 replications). The colors refer to the OBP class (blue: Classical, red: Plus-C, green: Minus-C, orange: Dimer, Grape: ABPII).

ID	OBP Class	Transcript	NCBI accession	Length (AA)	Best hit	NCBI Accession	E-value	Identity%
BcuOBP1	ABPII	m.3589	KR998336	148	BdorOBP3	AGS08185	4e-06	24
BcuOBP2	Minus-C	m.5870	KR998337	150	BdorOBP10	AGS08192	3e-102	89
BcuOBP3	Classic	m.6160	KR998338	139	BdorOBP3	AGS08185	7e-65	70
BcuOBP4	Minus-C	m.6807	KR998339	142	BdorOBP10	AGS08192	2e-44	49
BcuOBP5	ABPII	m.7193	KR998340	166	BdorOBP4	AGS08186	5e-13	23
BcuOBP6	Classic	m.7840	KR998341	179	BdorOBP4	AGS08186	0.003	23
BcuOBP7	Classic	m.11279	KR998342	144	BdorOBP9	AGS08191	2e-41	45
BcuOBP8	Classic	m.11490	KR998343	155	BdorOBP9	AGS08191	1e-41	42
BcuOBP9	Classic	m.11812	KR998344	143	BdorOBP7	AGS08189	6e-104	93
BcuOBP10	Minus-C	m.16988	KR998345	160	BdorOBP1	AGS08192	2e-14	27
BcuOBP11	ABPII	m.51638	KR998346	146	BdorOBP5	GS08187	2e-06	20
BcuOBP12	Classic	m.1076	KR998347	136	BdorOBP5	GS08187	2e-17	32
BcuOBP13	Classic	m.2892	KR998348	173	BdorOBP5	GS08187	2e-11	28

**Table 2.** Classification of odorant binding proteins in *B. cucurbitae* according to the best-hit matches to OBPs in *B. dorsalis*.

the NCBI non redundant protein (nr) database using BLASTP (protein- protein BLAST), returned hits corresponding to putatively homologous genes in *B. dorsalis* and *C. capitata*<sup>19,20</sup> (Supplementary Table S3). These were added to our dataset and used to build a Maximum Likelihood phylogenetic tree (Fig. 5). The tree reveals three major clades: the first represents the N-Methyl-D-Aspartate (NMDA) iGluR sub-family and includes receptor sequences from *B. cucurbitae*, *B. dorsalis*, *C. capitata* and *D. melanogaster*. The second clade includes ionotropic co-receptor sequences. The third, and largest, clade has the members of the iGluR from the Kainate subfamily.

**Gustatory receptor proteins.** We also identified six candidate gustatory receptor genes in the melon fly transcriptome dataset. Similar to earlier descriptions, BLAST searches of the nr sequences database returned



**Figure 3. Maximum Likelihood Phylogenetic tree of Tephritid fruit flies Odorant-Binding Proteins (OBPs).** Bootstrap values greater than 50% are displayed (1000 replications).

putative homologous sequences from *C. capitata* and *D. melanogaster* (Supplementary Table S4). Figure 6 shows the Maximum Likelihood phylogenetic tree for these gustatory receptor proteins. It reveals that the gustatory receptor sequences from *B. cucurbitae* (Bcu\_GR) are distributed among three different clades. The receptors Bcu\_GR3 and Bcu\_GR4 are present in the clade of taste sensing receptors, which includes Dmel\_GR\_98.

The BcuGR1 sequence clusters with the Ccap\_GR\_39b and DmelGR\_39b sequences, which are receptors involved in mediating the acceptance or avoidance of specific substances. Finally, the Bcu\_GR5 sequence is found in the clade containing the Dmel\_GR\_22 gustatory receptors (Dmel\_GR\_22a, Dmel\_GR\_22b, Dmel\_GR\_22c, Dmel\_GR\_22d, Dmel\_GR\_22e, Dmel\_GR\_22f). The three other melon fly gustatory receptors (Bcu\_GR2, Bcu\_GR6 and Bcu\_GR7) cluster with Dmel\_GR\_66a and Ccap\_GR\_66a.

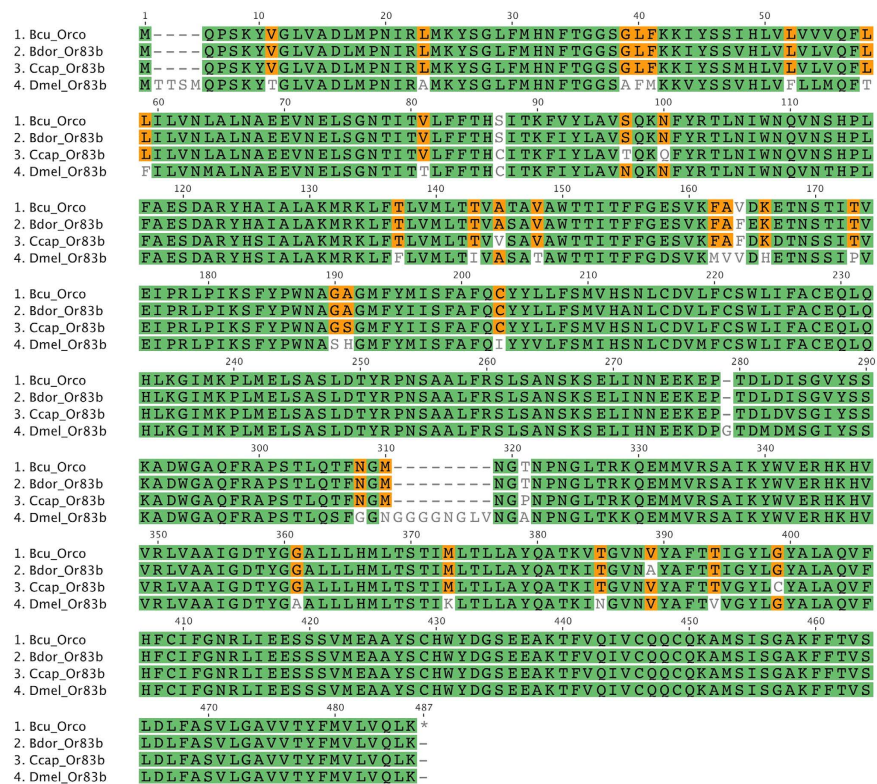
## Discussion

Transcriptome analysis, along with genome annotation and next-generation sequencing methods, permit the discovery and characterization of multiple genes in insects representing a wide range of functional categories<sup>24–26</sup>. Starting from the model organism *D. melanogaster*<sup>27</sup> and extending to other insect species such as beetle<sup>28</sup> and mosquito species<sup>29</sup> among others, this approach has also been used to identify a wide range of genes involved in the specific area of chemosensory perception. Recently, this approach has also been used to identify and characterize genes in some species of the Tephritidae or true fruit flies that are major agricultural pests, including the Mediterranean fruit fly (*C. capitata*)<sup>20,30</sup> and the oriental fruit fly (*B. dorsalis*)<sup>19,31</sup>.

Using bioinformatic tools for the analysis of transcriptome level data, and taking advantage of the existing databases<sup>6,19,20</sup>, we identified and characterized here a wide range of putative proteins involved in chemosensory perception in the melon fly, *B. cucurbitae*.

This species is a major pest of many cucurbit crops. Despite being a member of the genus *Bactrocera*, it infests a distinctly different set of host plants compared to the oriental fruit fly *B. dorsalis*, and has also long been known to exhibit a number of other biological differences<sup>11</sup>. For instance, extensive quantifications of the differential responses of these tephritid fruit flies were tested for a wide range of chemically defined kairomones such as Methyl Eugenol, a sex attractant molecule, occurring naturally in several tropical plants and widely used in male annihilation control strategies<sup>32</sup>. In many cases, these species showed response values that differed by up to 1,000 times or more<sup>33</sup>.

Our goal here was to use RNAseq to identify putative protein sequences related to olfaction and chemosensory perception in the melon fly and compare them to closely related species from the Tephritidae family.



**Figure 4.** Alignment of the Odorant receptor co-receptor (*Orco*) protein sequences corresponding to *B. cucurbitae*, *B. dorsalis* and *D. melanogaster*.

The BLAST analysis of the OBP proteins identified in the melon fly allowed first for characterization of sequences based on the presence of specific structural features in counterparts previously described in other tephritid species along with *D. melanogaster*. Overall, when the sequences recovered here from the melon fly were compared to those of the oriental fruit fly, which is the most closely related species considered here, the percentage of identity values varied from 24 to 93%. This is consistent with levels of divergence of OBP proteins seen among other Tephritid species<sup>19</sup>.

Regarding their classification, the odorant binding proteins identified in the melon fly fell into either the “Classic” or the “Minus-C” sub-family. For instance, we identified a classic OBP protein in the melon fly (BcuOBP9), very similar to the OBP protein in the oriental fruit fly *B. dorsalis* (BdorOBP7), which is highly expressed in the taste sensilla on the leg, and may have a role in perceiving non-volatile chemical compounds<sup>34</sup>. We did not, however, identify any OBPs from the Plus-C or the dimer subfamilies in the melon fly transcriptome data. The presence of these OBPs was reported previously in the case of the oriental fruit fly *B. dorsalis*<sup>15</sup> and the Mediterranean fruit fly *C. capitata*<sup>20,35,36</sup>.

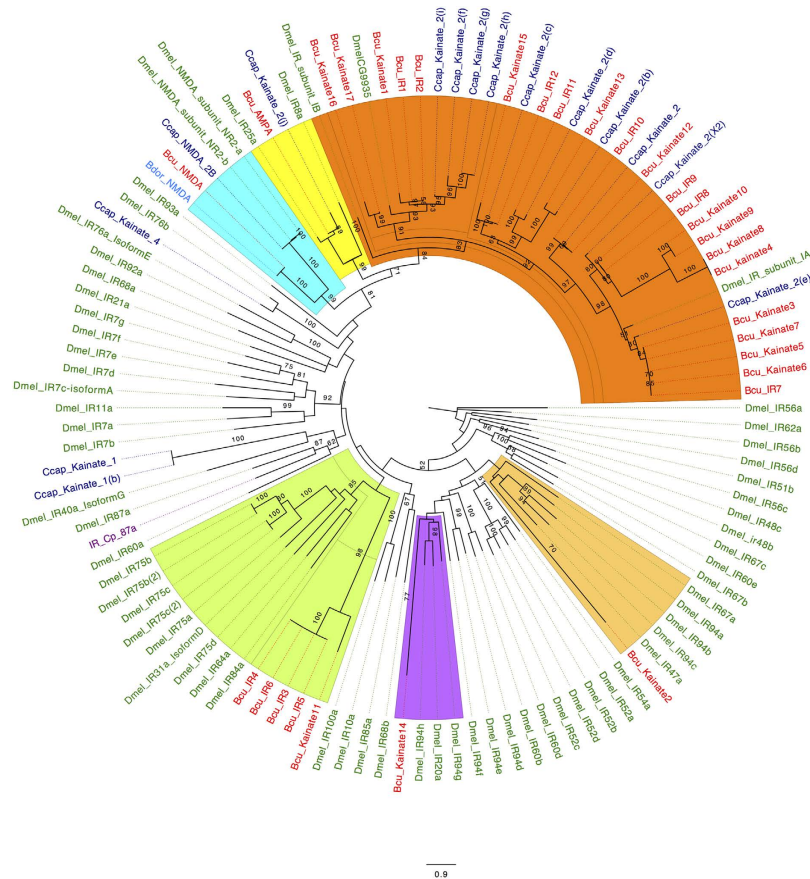
The absence of this category in the melon fly transcriptome reflects either (1) key biological differences between these species or (2) the necessity of generating a more informative dataset, from the antennal part instead of the whole body of the insect<sup>37</sup>.

Within the Odorant receptors family, the *BcuOrco* putative protein showed high levels of similarity with counterparts in closely related species from the tephritidae family, which is consistent with the fact that this gene is well conserved specifically among tephritids<sup>38</sup> and among insects in general<sup>39,40</sup>. The high conservation of this gene is a strong indication of its crucial role in odor detection<sup>41</sup>.

For the sequences classified by the GO analysis as representing ionotropic receptors (IRs), in *Drosophila* the IRs have been shown to play an important role in the detection of biological decomposed material (acids, ammonia etc.)<sup>42</sup>. For the melon fly, the number of IR sequences we identified is higher than those previously reported for other species<sup>43,44</sup>. This larger number may reflect the ecological specificities of the melon fly species, which feeds only on decaying or damaged fruits<sup>45</sup>.

The phylogenetic analysis of the Ionotropic Glutamate Receptors (iGluRs) of the melon fly and those from *D. melanogaster*<sup>46</sup> along with two other tephritid species (*C. capitata* and *R. pomonella*), revealed three iGluR subfamilies. Specifically, the NMDA receptor, the Kainate receptors and the AMPA receptors, appear to be highly conserved among species that are both closely and distantly related to the melon fly.

Moreover, gustatory receptor sequences were identified in the melon fly using homology to known sequences from *D. melanogaster*, *C. capitata* and *B. dorsalis*. The melon fly sequences clustered with their *D. melanogaster* counterparts, including those previously reported to have a key role in mediating avoidance or acceptance of substrates and compounds<sup>47</sup>.



**Figure 5. Maximum Likelihood Phylogenetic tree of Tephritid fruit flies Ionotropic Receptors Bootstrap values greater than 50% are displayed (1000 replications).** (Ionotropic receptors of *B. cucurbitae*, *C. capitata*; *D. melanogaster* was used as an outgroup).

In addition to taste sensing, in *Drosophila*, some of these receptors also have sensory functions in organs like the abdominal ganglia<sup>48</sup>. To which extent these genes exhibit a similar spatial pattern of expression in the melon fly is yet to be determined. However, it is clear that four of the gustatory receptors identified from the melon fly dataset cluster with DmelGR\_22a and DmelGR\_66a, which are gustatory receptors specific to bitter tasting substrates in *D. melanogaster*<sup>49</sup>. This might have important implications in pest management efforts to control the melon fly<sup>50,51</sup>, since some bait sprays use plant-derived semiochemicals such as cucurbitacin, a toxin found in cucurbit plants. The melon fly may be able to recognize these toxic compounds through its gustatory receptors.

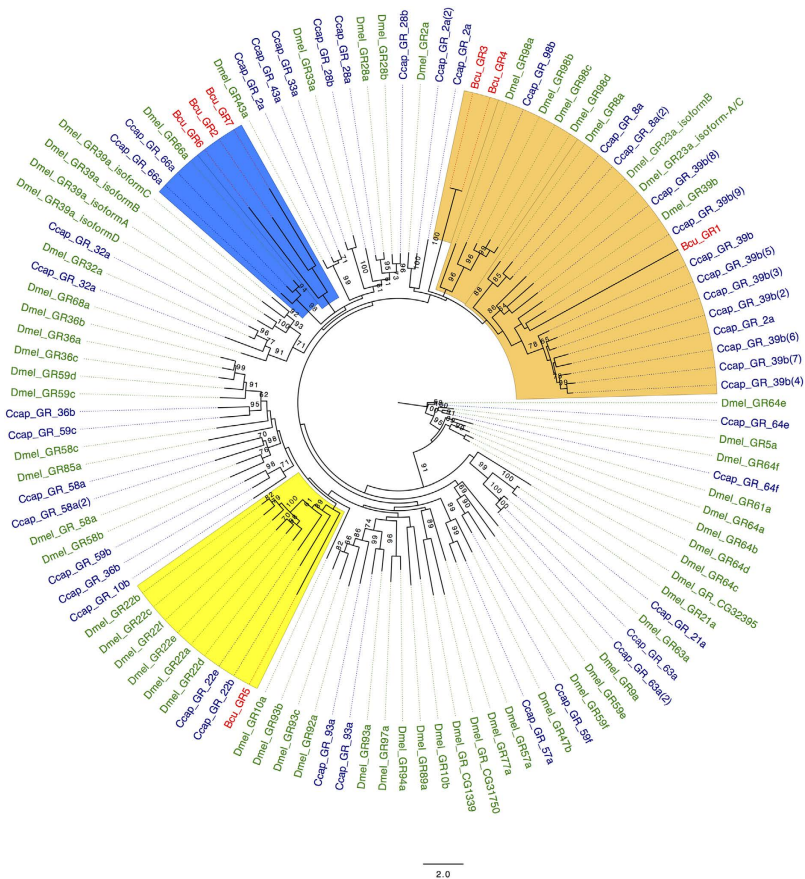
Overall, the work presented here brings a significant contribution to the study of chemosensory receptors in tephritid fruit flies. Using a transcriptome-based approach, we were able to identify and partially characterize several important genes. These include several genes that likely play a key role in the chemosensory perception activities of the melon fruit fly. The identification and characterization of these candidate chemosensory proteins in the melon fly could be of great help in the development of novel and species specific semiochemicals used in pest management strategies<sup>52</sup>.

## Methods

**Insect materials, RNA purification and cDNA synthesis.** The melon fly *B. cucurbitae* specimens used in this study, were from samples collected in Taiwan and were reared in the lab for several generations since 1997. Four-day melon fly adults were used for total RNA extraction with TRIzol reagent (Invitrogen, Carlsbad, CA, USA).

The purified RNA was quantified using a Nanodrop ND-2000 (Thermo Scientific, Waltham, MA, USA) and the average concentration was approximately 1.98 ng/μl ( $OD_{260/280} = 2.06$ ).

**Sequencing, assembly and annotation.** The RNA samples were sequenced at the Hawaii Institute for Marine Biology (HIMB) genomics core using the Illumina Genome Analyzer Iix platform. The reads were assembled using the Trinity platform and the contigs were annotated using Trinotate: an automatic functional annotation pipeline of *de novo* assembled transcriptomes (<http://trinotate.sourceforge.net>)<sup>53,54</sup>. Gene Ontology (GO) terms were assigned to each contig using Blast2GO<sup>55</sup>. The melon fruit fly transcripts were next searched for sequence homologies using BLASTX analysis in Genbank to identify candidate chemosensory protein receptors. The putative protein sequences were then compared to orthologous genes from other insect species using BLASTP<sup>56</sup>.



**Figure 6. Maximum Likelihood Phylogenetic tree of Tephritid fruit flies gustatory receptors.** Bootstrap values greater than 50% are displayed (1000 replications). (Gustatory receptors of: *B. cucurbitae* (BcuGR), *C. capitata* (CcapGR), *D. melanogaster* (DmelGR) was used as an outgroup).

**Protein naming scheme.** The putative proteins identified in the melon fly transcriptome, were labeled following the already established nomenclature<sup>57</sup>. The label consists of the abbreviation of the species Latin name followed by the candidate protein name and a number from 1 upwards.

**Phylogenetic analysis.** Protein Sequence alignments corresponding to the melon fruit fly *B. cucurbitae* and sequences related to other tephritid fruit flies, were generated using ClustalW<sup>58</sup> as implemented in Geneious V. 8.0.5 (<http://www.geneious.com>)<sup>59</sup>. These alignments served as input for the program RAXML<sup>60</sup> used here to construct phylogenetic trees based on a Maximum Likelihood approach with JTT substitution model and PROTGAMMA as the GAMMA model of rate heterogeneity, with 1000 bootstrap replicates. The trees were visualized and formatted in FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

## References

- Vieira, F. G. & Rozas, J. Comparative Genomics of the Odorant-Binding and Chemosensory Protein Gene Families across the Arthropoda: Origin and Evolutionary History of the Chemosensory System. *Genome Biol Evol.* **3**, 476–490 (2011).
- Leal, W. S. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annu Rev Entomol.* **58**, 373–391 (2013).
- Scott, K. A Chemosensory Gene Family Encoding Candidate Gustatory and Olfactory Receptors in *Drosophila*. *Cell.* **104**, 661–673 (2001).
- Zhou, J.-J., He, X.-L., Pickett, J. A. & Field, L. M. Identification of odorant-binding proteins of the yellow fever mosquito *Aedes aegypti*: genome annotation and comparative analyses. *Insect Mol Biol.* **17**, 147–163 (2008).
- Gong, D. P., Zhang, H.-J., Zhao, P., Xia, Q.-Y. & Xiang, Z.-H. The Odorant Binding Protein Gene Family from the Genome of Silkworm, *Bombyx mori*. *BMC Genomics.* **10**, doi: Artn 33210.1186/1471-2164-10-332 (2009).
- Hsu, J.-C. *et al.* Discovery of Genes Related to Insecticide Resistance in *Bactrocera dorsalis* by Functional Genomic Analysis of a *De Novo* Assembled Transcriptome. *PLoS One.* **7**, e40950 (2012).
- Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in *Drosophila*. *Cell.* **136**, 149–162 (2009).
- Elfékih, S. & Haymer, D. S. Molecular Taxonomic identification of *Dacus* and *Ceratitis* species from Sub-Saharan Africa using Mitochondrial Haplotypes. *J Econ Entomol.* **103**, 1842–1847 (2010).
- Elfékih, S. & Haymer, D. S. Characterization of *Dacus punctatifrons* (KARSCH) and *D. bivittatus* (BIGOT) populations using mitochondrial DNA markers. *Mitteilungen der Deutschen Gesellschaft für allgemeine und angewandte Entomologie* **17**, 289–291 (2009).
- Aluja, M. & Norrbom, A. eds. *Fruit flies (Tephritidae): phylogeny and evolution of behavior*. CRC Press (1999).



11. White, I. M. & Elson-Harris, M. *Fruit Flies of Economic Significance: Their Identification and Bionomics*. CAB International in association with ACIAR (1992).
12. Elfékih, S., Makni, M. & Haymer, D. S. Detection of novel mitochondrial haplotype variants in populations of the Mediterranean fruit fly, *Ceratitis capitata*, from Tunisia, Israel and Morocco. *J Appl Entomol.* **134**, 647–651 (2010).
13. Yee, W. L., Hernández-Ortiz, V., Rull, J., Sinclair, B. J. & Neven, L. G. Status of *Rhagoletis* (Diptera: Tephritidae) Pests in the NAPPO Countries. *J Econ Entomol.* **107**, 11–28 (2014).
14. Shelly, T. *et al.* *Trapping and the Detection, Control, and Regulation of Tephritid Fruit Flies*. Springer Netherlands (2014).
15. Sim, S. B., Calla, B., Hall, B., DeRego, T. & Geib, S. M. Reconstructing a comprehensive transcriptome assembly of a white-pupal translocated strain of the pest fruit fly *Bactrocera cucurbitae*. *GigaScience.* **4**, 1–5 (2015).
16. Mi, H., Muruganujan, A. & Thomas, P. D. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **41**, D377–D386 (2012).
17. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc.* **8**, 1551–1566 (2013).
18. Hekmat-Scafe, D. S. Genome-Wide Analysis of the Odorant-Binding Protein Gene Family in *Drosophila melanogaster*. *Genome Res* **12**, 1357–1369 (2002).
19. Zheng, W. *et al.* Identification and Expression Profile Analysis of Odorant Binding Proteins in the Oriental Fruit Fly *Bactrocera dorsalis*. *Int J Mol Sci.* **14**, 14936–14949 (2013).
20. Siciliano, P. *et al.* Sniffing Out Chemosensory Genes from the Mediterranean fruit fly, *Ceratitis capitata*. *PLoS One.* **9**, e85523 (2014).
21. Schwarz, D. *et al.* Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella*. *BMC Genomics.* **10**, doi: Artn 63310.1186/1471-2164-10-633 (2009).
22. Ramsdell, K. M. M. *et al.* Expressed Sequence Tags from Cephalic Chemosensory Organs of the Northern Walnut Husk Fly, *Rhagoletis suavis*, Including a Putative Canonical Odorant Receptor. *J Insect Sci.* **10**, 1–11 (2010).
23. Van Dongen A. M. *Biology of the NMDA Receptor*. Boca Raton (FL) CRC Press (2009).
24. Siepel, A. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
25. Pelosi, P. & Maida, R. Odorant-binding proteins in insects. *Comp Biochem Phys B.* **111**, 503–514 (1995).
26. Misof, B. *et al.* Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).
27. Anholt, R. R. H. *et al.* The genetic architecture of odor-guided behavior in *Drosophila*: epistasis and the transcriptome. *Nat Genet.* **35**, 180–184 (2003).
28. Andersson, M. N. *et al.* Antennal transcriptome analysis of the chemosensory gene families in the tree killing bark beetles, *Ips typographus* and *Dendroctonus ponderosae* (Coleoptera: Curculionidae: Scolytinae). *BMC Genomics.* **14**, doi: Artn 19810.1186/1471-2164-14-198 (2013).
29. Pelletier, J. & Leal, W. S. Genome Analysis and Expression Patterns of Odorant-Binding Proteins from the Southern House Mosquito *Culex pipiens quinquefasciatus*. *PLoS One.* **4**, e6237 (2009).
30. 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**, doi: Artn S1110.1186/1471-2156-15-S2-S11 (2014).
31. Wu, Z. *et al.* Discovery of Chemosensory Genes in the oriental fruit fly *Bactrocera Dorsalis*. *PLoS One.* **10**, e0129794 (2015).
32. Tan, K. H. & Nishida, R. Methyl Eugenol: Its Occurrence, Distribution, and Role in Nature, Especially in Relation to Insect Behavior and Pollination. *J Insect Sci.* **12**, 1–60 (2012).
33. Metcalf, R. L., Mitchell, W. C. & Metcalf, E. R. Olfactory receptors in the melon fly *Dacus cucurbitae* and the oriental fruit fly *Dacus dorsalis*. *Proc. Natl Acad Sci. USA* **80**, 3143–3147 (1983).
34. Zheng, W. W., Peng, T., He, W. & Zhang, H. Y. High-Throughput Sequencing to Reveal Genes Involved in Reproduction and Development in *Bactrocera dorsalis* (Diptera: Tephritidae). *PLoS One.* **7**, e36463 (2012).
35. Gomulski, L. M. *et al.* Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly, *Ceratitis capitata*. *BMC Genomics.* **9**, doi: Artn 24310.1186/1471-2164-9-243 (2008).
36. Siciliano. *et al.* Identification of pheromone components and their binding affinity to the odorant binding protein CcapOBP83a-2 of the Mediterranean fruit fly, *Ceratitis capitata*. *Insect Biochem Molec Biol.* **48**, 51–62 (2014).
37. Schultze, A., Schymura, D., Forstner, M. & Krieger, J. Expression pattern of a 'Plus-C' class odorant binding protein in the antenna of the malaria vector *Anopheles gambiae*. *Insect Mol Biol.* **21**, 187–195 (2011).
38. Zheng, W., Zhu, C., Peng, T. & Zhang, H. Odorant receptor co-receptor *Orco* is upregulated by methyl eugenol in male *Bactrocera dorsalis* (Diptera: Tephritidae). *J Insect Physiol.* **58**, 1122–1127 (2012).
39. Benton, R., Sachse, S., Michnick, S. W. & Vosshall, L. B. Atypical Membrane Topology and Heteromeric Function of *Drosophila* Odorant Receptors *In Vivo*. *PLoS Biol.* **4**, 240 (2006).
40. Missbach C. *et al.* Evolution of insect olfactory receptors. *eLife* **3**, e02115 (2014).
41. Stengl, M. & Funk, N. W. The role of the coreceptor *Orco* in insect olfactory transduction. *J Comp Physiol A.* **199**, 897–909 (2013).
42. Rytz, R., Croset, V. & Benton, R. Ionotropic Receptors (IRs): Chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem Molec Biol.* **43**, 888–897 (2013).
43. Leitch, O., Papanicolaou, A., Lennard, C., Kirkbride, P. K. & Anderson, A. Chemosensory genes identified in the antennal transcriptome of the blowfly *Calliphora stygia*. *BMC Genomics.* **16**, doi: 10.1186/s12864-015-1466-8 (2015).
44. Riveron, J., Boto, T. & Alcorta, E. Transcriptional basis of the acclimation to high environmental temperature at the olfactory receptor organs of *Drosophila melanogaster*. *BMC Genomics.* **14**, doi: Artn 25910.1186/1471-2164-14-259 (2013).
45. Hill, D. S. *Pests of crops in warmer climates and their control*. Springer Science & Business Media (2008).
46. Croset, V. *et al.* Ancient Protostome Origin of Chemosensory Ionotropic Glutamate Receptors and the Evolution of Insect Taste and Olfaction. *PLoS Genet.* **6**, e1001064 (2010).
47. Ai, undefined M. *et al.* Ionotropic Glutamate Receptors IR64a and IR8a Form a Functional Odorant Receptor Complex *In Vivo* in *Drosophila*. *J Neurosci.* **33**, 10741–10749 (2013).
48. Park, J. H. & Kwon, J. Y. A systematic analysis of *Drosophila* gustatory receptor gene expression in abdominal neurons, which project to the central nervous system. *Mol Cells.* **32**, 375–381 (2011).
49. Weiss, L. A., Dahanukar, A., Kwon, J. Y., Banerjee, D. & Carlson, J. R. The Molecular and Cellular Basis of Bitter Taste in *Drosophila*. *Neuron.* **69**, 258–272 (2011).
50. Piñero, J. C., Jácome, I., Vargas, R. & Prokopy, R. J. Response of female melon fly, *Bactrocera cucurbitae*, to host-associated visual and olfactory stimuli. *Entomol Exp Appl.* **121**, 261–269 (2006).
51. Miller, N. W., Vargas, R. L., Prokopy, R. J. & Mackey, B. E. State-Dependent Attractiveness of Protein Bait and Host Fruit Odor to *Bactrocera cucurbitae* (Diptera: Tephritidae) Females. *Ann Entomol Soc Am.* **97**, 1063–1068 (2004).
52. Carlson, J. R. & Wynand van der Goes van Naters. Insects as chemosensors of humans and crops. *Nature.* **444**, 302–307 (2006).
53. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnol.* **29**, 644–652 (2011).
54. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* **8**, 1494–1512 (2013).
55. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

56. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **17**, 3389–3402 (1997).
57. Guo, S. & Kim, J. Molecular Evolution of *Drosophila* odorant receptor genes. *Mol Biol Evol.* **24**, 1198–1207 (2007).
58. Thompson *et al.* Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
59. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* **28**, 1647–1649 (2012).
60. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* **22**, 2688–2690 (2006).

## Acknowledgements

The authors gratefully acknowledge the support of the bioinformatics team in the Chen lab: Dr. Ting-Ying Chien, Ms. Mei-Ju May Chen, Dr. Tony Chien-Yen Kuo and Ms. Ching Chang. S.E. was supported by an EMBO short-term grant (ASTF-7-2013) and a UNESCO-L'Oréal for women in Science fellowship (2011-2013).

## Author Contributions

C.Y.C., J.H. and D.H. conceived and designed the experiments. S.E., C.Y.C., M.B. and J.H. performed the experiments and analyzed the data. S.E. and D.H. wrote the paper. All authors read and approved the final 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:** Elfekih, S. *et al.* Identification and preliminary characterization of chemosensory perception-associated proteins in the melon fly *Bactrocera cucurbitae* using RNA-seq. *Sci. Rep.* **6**, 19112; doi: 10.1038/srep19112 (2016).



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/>