

RESEARCH ARTICLE

Open Access



Chemosensory genes in the antennal transcriptome of two syrphid species, *Episyrphus balteatus* and *Eupeodes corollae* (Diptera: Syrphidae)

Bing Wang, Yang Liu* and Gui-Rong Wang*

Abstract

Background: Predatory syrphid larvae are an important natural enemy of aphids in cotton agro-ecosystems in China. Their behaviors in prey foraging, localization and oviposition greatly rely on the perception of chemical cues. As a first step to better understand syrphid olfaction at the molecular level, we have performed a systematic identification of their major chemosensory genes.

Results: Male and female antennal transcriptomes of *Episyrphus balteatus* and *Eupeodes corollae* were sequenced and assembled using Illumina HiSeq2000 technology. A total of 154 chemosensory genes in *E. balteatus* transcriptome, including candidate 51 odorant receptors (ORs), 32 ionotropic receptors (IRs), 14 gustatory receptors (GRs), 49 odorant-binding proteins (OBPs), 6 chemosensory proteins (CSPs) and 2 sensory neuron membrane proteins (SNMPs) were identified. In *E. corollae* transcriptome, we identified 134 genes including 42 ORs, 23 IRs, 16 GRs, 44 OBPs, 7 CSPs and 2 SNMPs. We have provided full-length sequences of the highly conserved co-receptor Orco, IR8a/25a family and carbon dioxide gustatory receptor in both syrphid species. The expression of candidate OR genes in the two syrphid species was evaluated by semi-quantitative reverse transcription PCR. There were no significant differences of transcript abundances in the respective male and female antenna, which is consistent with differentially expressed genes (DEGs) analysis using the FPKM value. The sequences of candidate chemosensory genes were confirmed and phylogenetic analysis was performed.

Conclusions: This research comprehensively analyzed and identified many novel candidate chemosensory genes regarding syrphid olfaction. It provides an opportunity for understanding how syrphid insects use chemical cues to conduct their behaviors among tritrophic interactions of plants, herbivorous insects, and natural enemies in agricultural ecosystems.

Keywords: *Episyrphus balteatus*, *Eupeodes corollae*, Transcriptome, Chemosensory genes, Odorant receptors, Syrphid olfaction

Background

Natural enemies are important in controlling the population of insect pests in agricultural ecosystems [1–6]. The presence of natural enemies is detected through chemical cues emitted by prey and host plants, alone or in association, and received by the antennal sensory

system in the multi-trophic environment to result in various behavioral choices including prey forage, localization, oviposition and escape [7–17]. Recently the study of insect peripheral sensory system and chemosignal transduction has experienced considerable progress due to the development of bioinformatics-based approaches and protein function prediction methods [18]. In particular, a large amount of information is provided by antennal transcriptome projects from various insect taxa [19–28].

* Correspondence: yangliu@ippcaas.cn; wangguirong@caas.cn
State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China



In general, the process of chemoreception, including olfaction and taste, involves several families of genes, including odorant receptors (ORs), ionotropic receptors (IRs), and gustatory receptors (GRs) [20, 29–31]. In addition, odorant binding proteins (OBPs), chemosensory proteins (CSPs) and sensory neuron membrane proteins (SNMPs) also play crucial roles in chemoreception [32–36]. The insect chemoreceptor superfamily including ORs and GRs was first identified in the *Drosophila melanogaster* genome [18, 21]. Insect odorant and gustatory receptors were once thought to be G-protein-coupled receptors just like ORs in worms and vertebrates, but subsequent studies have shown a lack of homology to vertebrate ORs [37]. One such superfamily encoding ORs is highly divergent across insect taxa with sequences and frequencies varying to a large extent [18, 38, 39]. ORs are broadly tuned to alcohols, ketones, and esters generally present in the environment [40, 41]. Another family encoding GRs, or receptors for taste or contact stimuli, is also very divergent across insect taxa [31]. On the contrary, one example of exceptionally conserved GRs are GR21a and GR63a, which work together as a CO₂ receptor in *Drosophila* [42]. Such chemoreceptors play an important role in host seeking behaviors in many insects but, especially seen in mosquitoes [43, 44]. A new insect chemosensory family was identified recently and given the name ionotropic receptors (IRs). These IRs belong to the ionotropic glutamate receptor superfamily (iGluRs) and were identified in both the olfactory and gustatory systems [30, 45]. IRs are more greatly conserved than ORs and GRs but considerable variations can be observed in ligand-binding domains. They are mainly tuned to acids, amines and other odorants that are not detected by ORs [30, 45, 46].

Since chemosensory gene families were characterized in two important model species, *D. melanogaster* and *Anopheles gambiae* [47, 48], a growing number of chemosensory genes have been identified from many Dipteran species, such as *Musca domestica* [49], *Bactrocera dorsalis* (genome: assembly ASM78921v2), *Calliphora stygia* [50], *Glossina morsitans morsitans* [51], and *Mayetiola destructor* Say [20]. Protein prediction methods have been the first step for functional identification of chemosensory genes. All information regarding insect chemosensory was obtained bioinformatically and has been beneficial in understanding insect processing of diverse volatile compounds and cross-species differences in chemical communication.

The syrphids belong within the Diptera order and their larvae are aphid-specific natural enemies [1, 2, 4, 5, 8]. Due to the larvae's agricultural importance via potential applications, several reports have been published on the chemical ecology of these insects. In *Episyrphus balteatus* DeGeer, larvae may use a sesquiterpene as a kairomone [17, 52] and other potential semiochemicals to locate their prey [8, 53]. In *Sphaerophoria rueppellii*, adult females are

strongly attracted to odors from aphid colonies showing that specific volatile compounds are important to detect their prey [1]. Some studies on the relationships between aphid or host plant volatile emissions and aphid localization and foraging behavior have shown strong associations with syrphid recognition. One striking finding has shown that volatiles from plants attacked by aphids produce strong electrophysiological responses from the antennae of syrphids [7, 8]. These studies indicate that detecting prey-derived volatiles ((E)- β -farnesene), herbivore-induced plant volatiles (monoterpenes and sesquiterpenes), or naturally occurring general leaf volatiles (GLVs; alcohols, aldehydes and esters) help natural enemies to select oviposition sites and locate their prey [8, 12, 13, 17, 52–55].

Despite these reports on chemosensory behavior, little is known on the molecular basis of syrphid olfaction. Therefore, the identification of predatory syrphid chemosensory gene families will help reveal how syrphids forage on their prey and choose oviposition sites. In this study we selected two syrphid species, *E. balteatus* and *Eupeodes corollae* Fabricius, active in northern China cotton fields to perform antennal transcriptome sequencing in order to explore and compare chemosensory genes in the two species. A total of 154 and 134 chemosensory candidate genes were identified in *E. balteatus* and *E. corollae* transcriptomes, respectively, including ORs, IRs, GRs, OBPs, CSPs and SNMPs. Furthermore, we report the expression profile of the OR families found in each insect transcriptome. A comparison between these two syrphids and other insect species revealed candidate chemosensory genes that could be involved in prey selection and plant volatile recognition. The discovery of putative chemosensory genes gives way for further exploration into functional assessments regarding chemoreception association.

Results

Antennal transcriptome sequencing and sequence assembly

E. balteatus and *E. corollae* antennal transcriptomes were sequenced using the Illumina HiSeq 2000 platform combined with Trinity assembly. Approximately 68.71 million and 77.28 million raw-reads were obtained from *E. balteatus* male and female antenna, respectively, reduced after filtering, to 65.69 and 74.25 million clean-reads. These were assembled into 57,950 unigenes for male and 68,165 for female. A final transcript dataset with 53,575 unigenes was obtained, consisting of 17,407 distinct clusters and 36,168 distinct singletons. The dataset was 47.61 megabase (Mb) in size with a mean length of 889 bp and N50 of 1724 bp (Additional file 1: Table S1). Parallel experiments generated 80.15 million and 77.38 million raw-reads in *E. corollae* male and female, respectively, and 65.69 and 74.25 million clean-reads. From these data 54,116 and 61,220 unigenes were obtained for male and female, respectively. The final

transcript dataset of *E. corollae* contained 50,942 unigenes with a mean length of 1039 bp and N50 length of 2104 bp, consisting of 18,054 distinct clusters and 32,888 distinct singletons (Additional file 1: Table S1). In addition, unigenes with a sequence length > 500 bp accounted for 47.64% and 43.59% of the *E. corollae* and *E. balteatus* transcriptome assembly, respectively.

Homology analysis and gene ontology (GO) annotation

A BLASTX homology search against the NCBI non-redundant protein database indicated 23,680 (44.2%) and 25,606 (50.3%) unigenes from *E. balteatus* and *E. corollae*, respectively, with sequence similarities to known proteins using a cut-off E-value of 10^{-5} . For *E. balteatus*, the larger number of similar genes (27.4%) belonged to *Ceratitis capitata* followed by *M. domestica* (20.4%), *D. melanogaster* (7.4%), *Drosophila willistoni* (3.1%), *Drosophila virilis* (3.0%) and *Drosophila mojavensis* (2.5%). For *E. corollae*, again *C. capitata* was best represented (25.1%), followed by *M. domestica* (18.8%), *D. melanogaster* (7.1%), *D. willistoni* (2.9%), *Acyrtosiphon pisum* (2.8%), and *D. virilis* (2.5%) (Additional file 2: Fig. S1A).

Gene ontology (GO) annotations were used to classify the transcripts into functional groups in accordance with specific GO categories. A total of 12,441 (23.22%) of all predicted proteins from *E. balteatus* and 12,425 (24.39%) predicted proteins from *E. corollae* were assigned to at least one GO term (Additional file 2: Fig. S1B). The GO terms distribution in the three categories were similar in the two species. In the “molecular function” category, the most abundant GO terms were “binding” and “catalytic activity”. In the “biological process” category, “cellular process”, “single-organism process” and “metabolic process” were the most represented. Finally, “cell”, “cell part”, and “organelle” were the most abundant GO terms in “cellular component” category (Additional file 2: Fig. S1B). GO terms associated with chemosensory genes were distributed in the “biological process” category (e.g. “cellular process”, “developmental process”, “response to stimulus”, “establishment of localization”, and “biological regulation”, etc.), “molecular function” category (e.g. “molecular transducer activity”, etc.) and “cellular component” category (e.g. “extracellular region”, “membrane part”, “membrane”, etc.).

Candidate ORs in *E. balteatus* and *E. corollae*

Based on our analysis of the antennal transcriptomes in the two species, 51 and 42 transcripts for candidate ORs were identified in the combined male and female data sets from *E. balteatus* and *E. corollae*, respectively (Additional file 3: Table S2). A total of 21 *E. balteatus* ORs (EbalORs) and 29 *E. corollae* ORs (EcorORs) contained full-length open reading frames (ORFs), whose translation products are predicted to possess 2–8 transmembrane domains (TMDs). Other partial length transcripts encoded proteins

exhibiting overlapping regions with low identity and were classified as unique genes. After a more exhaustive comparison with OR genes from other insect species, we found that all putative EbalORs shared between 22% and 86% amino acid identity with other ORs, with almost identical values (22% to 87%) for EcorORs. Detailed information is reported in Additional file 3 and Table S2.

We next performed a phylogenetic analysis using our candidate ORs and the ORs from four other Diptera species including *B. dorsalis*, *C. stygia*, *D. melanogaster* and *M. domestica* (Fig. 1). Clustered with DmOR83b, the highly conserved co-receptor Orco, orthologous genes were identified in the antennal transcriptomes of both syrphid species, and named EbalOrco and EcorOrco. As expected, sequence identity between EbalOrco and EcorOrco is very high (97.27%). Among the other ORs, five EbalORs (EbalOR9, 16, 18, 22 and 37) and three EcorORs (EcorOR8, 24 and 29) clustered with DmelOR67d, the pheromone receptor from *D. melanogaster*. This OR67d specific clade also included the OR67d orthologues from *M. domestica* and *B. dorsalis*. Two of these genes, EbalOR16 and EcorOR24, are full-length transcripts with 71.65% amino acid identity. The remaining ORs in this group were highly divergent among different species. Within the Dipteran OR sequences, we found a species-specific clade including eight members from *E. balteatus* (EbalOR7, 10, 30, 32, 41, 46, 47 and 48) and seven from *E. corollae* (EcorOR9, 10, 11, 15, 16, 34 and 38) that shared low identities with other Dipteran ORs (Fig. 1).

Amino acid identity between gene products correlates with similarity between genes. An amino acid sequence comparison between EbalORs and EcorORs revealed 33 pairs of orthologous ORs (including Orco) with 73.86% identity and the amino acid identity of complete OR ORFs was 84.00%. Sequence similarity percentages of the 28 pairs of homologous ORs are greater than 60% (Additional file 4: Table S3). In addition to the Orco family, orthologous groups with identities higher than 90% include the EbalOR29/EcorOR36, EbalOR3/EcorOR5, EbalOR13/EcorOR19 EcorOR22 with/EcorOR22 with 94.88%, 92.20%, 91.46% and 91.30% sequence identities, respectively. In addition, 10 EbalOR sequences (EbalOR2, 5, 7, 8, 11, 14, 21, 23, 25 and 33) are closely related to EcorOR homologues (EcorOR3, 4, 7, 9, 13, 14, 23, 28, 31 and 35) with identities greater than 80% (Additional file 4: Table S3; Fig. 1). All of these highly homologous proteins may play important roles in olfactory recognition.

Candidate GRs in *E. balteatus* and *E. corollae*

We have identified 14 and 16 candidate GR genes from *E. balteatus* and *E. corollae* transcriptomes, respectively (Additional file 3: Table S2). The majority of candidate EbalGRs and EcorGRs were partial fragments, with only three from *E. balteatus* and six from *E. corollae* encoding full-length proteins. These complete sequences all show

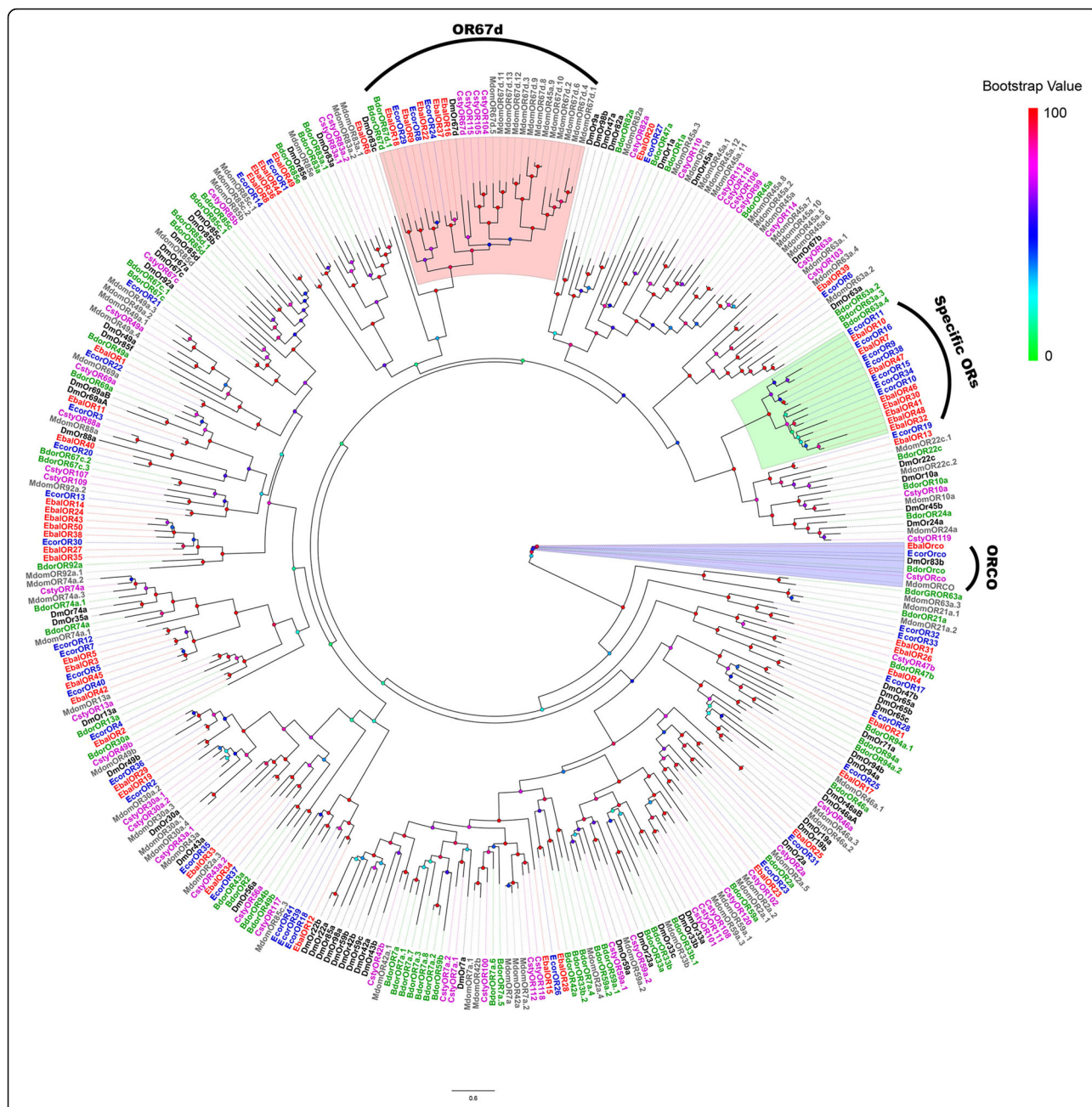


Fig. 1 Phylogenetic tree of candidate *E. balteatus* and *E. corollae* ORs and other Dipteran ORs. The distance tree was rooted by the conservative Orco gene orthologous. Bootstrap values are shown. The Orco clade, OR67d clade and specific EbalORs and EcorORs clade are shown. Species in this phylogeny include *E. balteatus* (Ebal, red), *E. corollae* (Ecor, blue), *Drosophila melanogaster* (Dm, dark), *Bactrocera dorsalis* (Bdor, green), *Calliphora stygia* (Csty, magenta), and *Musca domestica* (Mdom, gray)

six or seven TMDs with an intracellular N-terminus and extracellular C-terminus. Phylogenetic analysis with GRs from six Dipteran species suggest that *Drosophila* GR21a and GR63a, reported as carbon dioxide sensors [42, 56], clustered first with EbalGR2 and EcorGR2 and second with EbalGR1 and EcorGR1. In addition, EbalGR4, EbalGR13 and EcorGR4 showed high identities to thermoreceptor DmelGR28b responsible for rapid warmth avoidance [57].

Several other GRs clustered with members of candidate sugar detection GR (GR5a, GR61a and GR64a-f) sub-family (Fig. 2) [58–62].

Candidate IRs in *E. balteatus* and *E. corollae*

We identified 32 transcripts for putative ionotropic receptors in *E. balteatus* and 23 in *E. corollae*. Of these, seven EbalIRs and 14 EcorIRs contained full-length ORFs, with

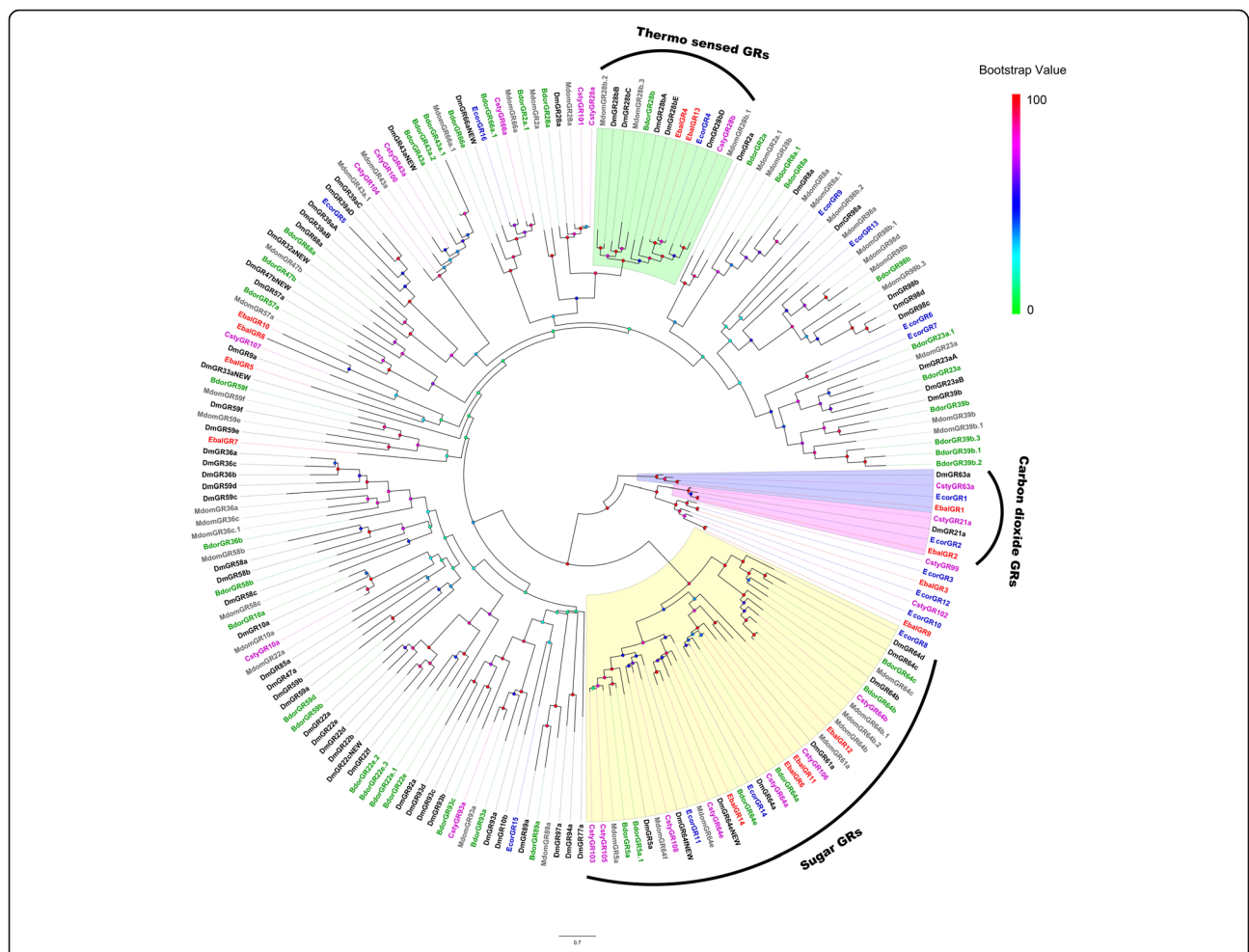


Fig. 2 Phylogenetic tree of candidate *E. balteatus* and *E. corollae* GRs and other Dipteran GRs. The distance tree was rooted by the conservative carbon dioxide GRs gene orthologous. Bootstrap values are shown. The carbon dioxide GRs clade, thermos-sensed GRs clade and sugar GRs clade are shown. This tree was constructed using the species *E. balteatus* (Ebal, red), *E. corollae* (Ecor, blue), *D. melanogaster* (Dm, dark), *B. dorsalis* (Bdor, green), *C. stygia* (Csty, magenta), and *M. domestica* (Mdom, gray)

two to five TMDs (Additional file 3: Table S2). Among these we found the common conserved co-receptors IR8a (EbalIR8a and EcorIR8a) and IR25a (EbalIR25a and EcorIR25a) in both species. Other candidate IRs were found as partial sequences (Fig. 3).

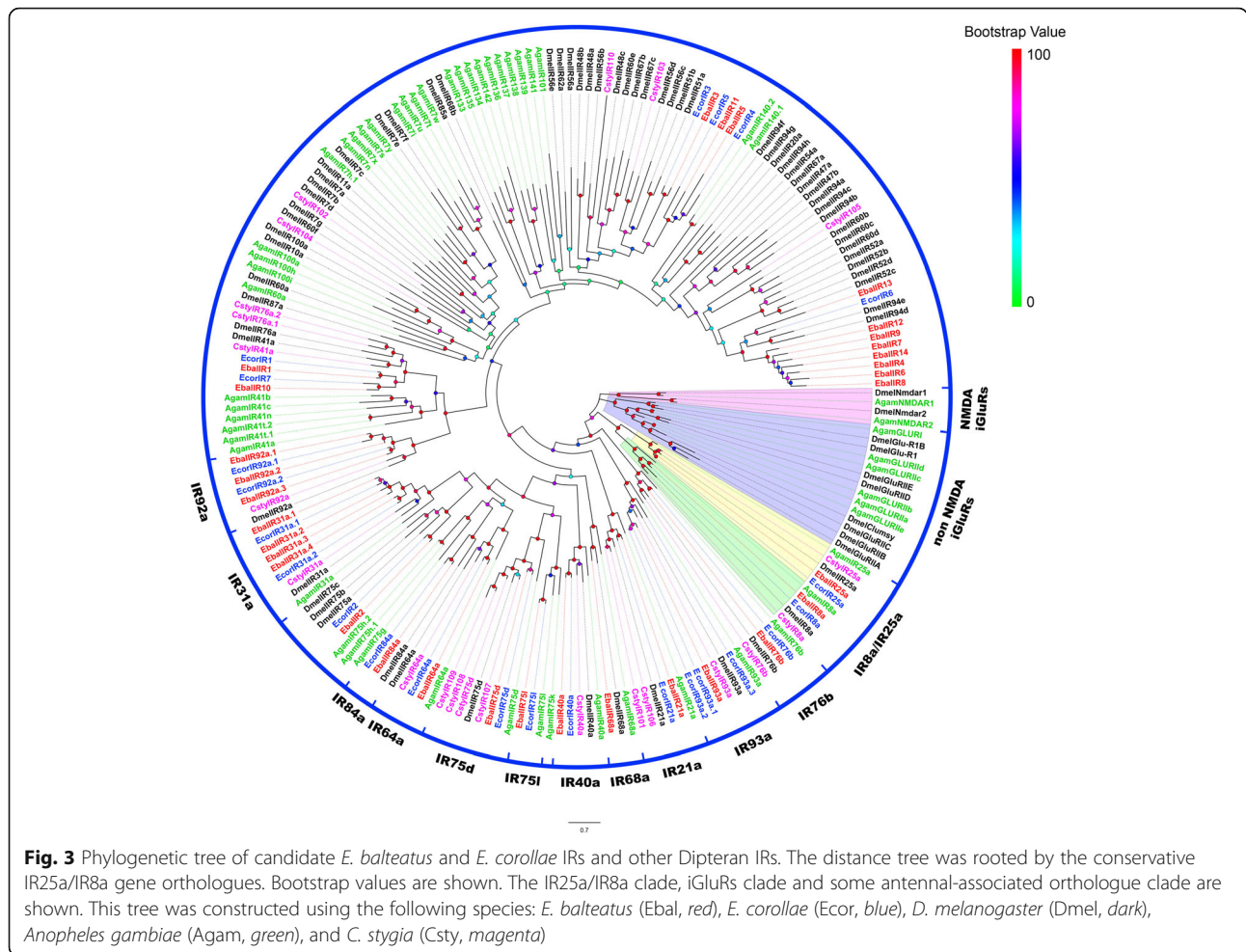
In order to further distinguish putative IRs from iGluRs, all EbalIRs and EcorIRs were aligned with IRs from *A. gambiae*, *C. stygia* and *D. melanogaster*, as well as some AgamiGluRs and DmeliGluRs for phylogenetic analysis. The results showed that the candidate EbalIRs and EcorIRs clustered with presumed “antennal” orthologues IR76b, IR93a, IR21a, IR68a, IR40a, IR75l, IR75d, IR64a, IR84a, IR31a and IR92a, and were well separated from the AgamiGluRs and DmeliGluRs clade (Fig. 3) [63]. Interestingly, the conserved “antennal” orthologues, IR60a, was lacking from *E. balteatus* and *E. corollae* transcriptome assemblies, while IR68a was only absent from *E. corollae*. The sequences of *E. balteatus* clustering with DmelIR94d and DmelIR94e

were quite divergent (Fig. 3; Additional file 5: Fig. S2). When compared to the orthologues within other species, these IRs may play different roles in olfaction.

Candidate OBPs in *E. balteatus* and *E. corollae*

We identified 49 different transcripts encoding candidate OBPs in *E. balteatus* and 44 in *E. corollae*, numbers similar to the 52 OBPs of *D. melanogaster* [64]. Of these, 38 transcripts of EbalOBPs and 31 EcorOBPs contained full-length ORFs with predicted signal peptide sequences (with EbalOBP31 as the only exception) (Additional file 3: Table S2).

A phylogenetic tree was built with these sequences and those of orthologous from *B. dorsalis*, *C. stygia*, *D. melanogaster* and *M. domestica*. Among EbalOBPs, thirty-one showed the classic motif of six conserved cysteines, three were Plus-C (EbalOBP2, 3, 4) and fifteen were Minus-C (EbalOBP5, 7, 8, 9, 11, 12, 14, 15, 16, 18, 21, 25, 26, 44 and 45) (Fig. 4). For EcorOBPs, we found 30 classic, 5 Plus-C



(EcorOBP1, 3, 4, 5, 41) and 9 Minus-C (EcorOBP6, 7, 8, 10, 11, 13, 18, 36 and 37) (Fig. 4) [64, 65]. One large group of classic OBPs, including 15 EbalOBPs and 14 EcorOBPs showed large differences compared to sequences of other species and could represent OBPs specific of syrphids (Fig. 4; Additional file 6: Fig. S3). We found the orthologue of DmelOBP-lush in both species, EbalOBP17 and EcorOBP14. These two proteins are 98.64% identical at the amino acid level between each other and 39.87% and 40.52% identity to DmelOBP-lush.

Candidate CSPs in *E. balteatus* and *E. corollae*

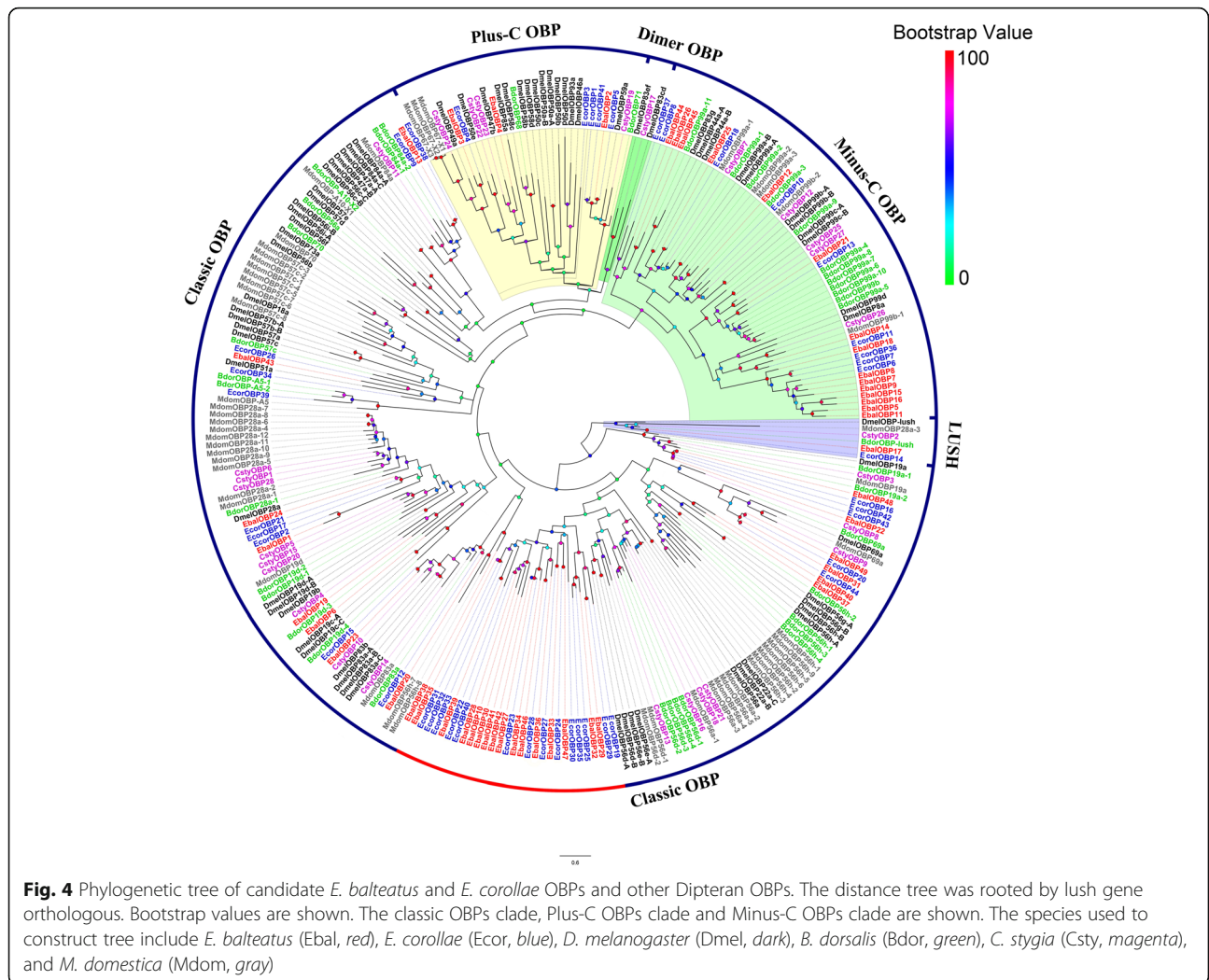
Through bioinformatic analysis, six and seven different transcripts encoding candidate CSPs were identified from *E. balteatus* and *E. corollae* transcriptomes, respectively. Five EbalCSPs and six EcorCSPs represented full-length proteins and only EbalCSP6 lacked a signal peptide (Additional file 3: Table S2). All of the identified amino acid sequences possessed the highly conserved four-cysteine profile. A phylogenetic tree was built with all the syrphid CSPs and those of *A. gambiae*, *C. stygia*, *D. melanogaster* (Fig. 5).

Candidate SNMPs in *E. balteatus* and *E. corollae*

In both species, two SNMPs with full-length ORFs were identified possessing two TMDs (with EbalSNMP1 having a single TMD as an exception) (Additional file 3: Table S2). EbalSNMP1 and EcorSNMP1 are very similar to DmelSNMP1, a protein shown to be required for correct pheromone detection [50, 66–68]. EbalSNMP2 and EcorSNMP2 are similar to DmelSNMP2, reported to be expressed in supporting cells (Fig. 6) [27, 69, 70].

Differentially expressed genes (DEGs) analysis

Gene expression levels of all male and female antennae-associated chemosensory genes in both *E. balteatus* and *E. corollae* were assessed using fragments per kilobase per million fragments (FPKM) values, represented in a heatmap (Fig. 7). Normalised antennal expression levels of candidate *E. balteatus* and *E. corollae* ORs are shown in Additional file 7. Of all ORs, Orco had the highest expression level of transcripts in both sexes of each species. There were no significant differences of OR transcript abundances (FPKM value) in the respective male and female antenna, except for EcorOR14 (Additional file 7). A combined analysis of



false discovery rate (FDR) ≤ 0.001 and $|\log_2 \text{Ratio}| \geq 1$ showed that EcorOBP shared highest number of differentially expressed genes (DEGs), including eleven high-expression in male and six high-expression in female syrphids (Additional file 7). In addition, candidate carbon dioxide receptor GR1 and GR2, and SNMP1 in both sexes showed a high expression level (Fig. 7).

Tissue- and sex- specific expression of candidate *E. balteatus* and *E. corollae* OR genes

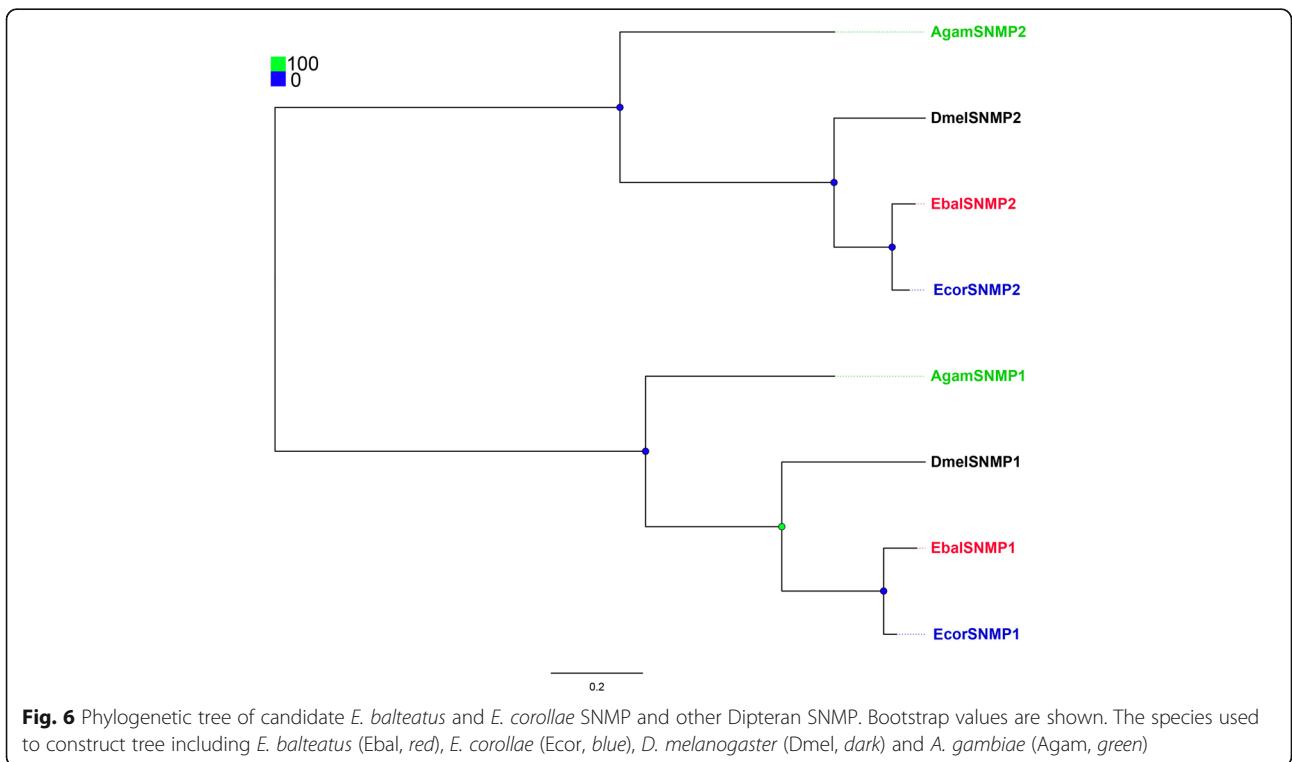
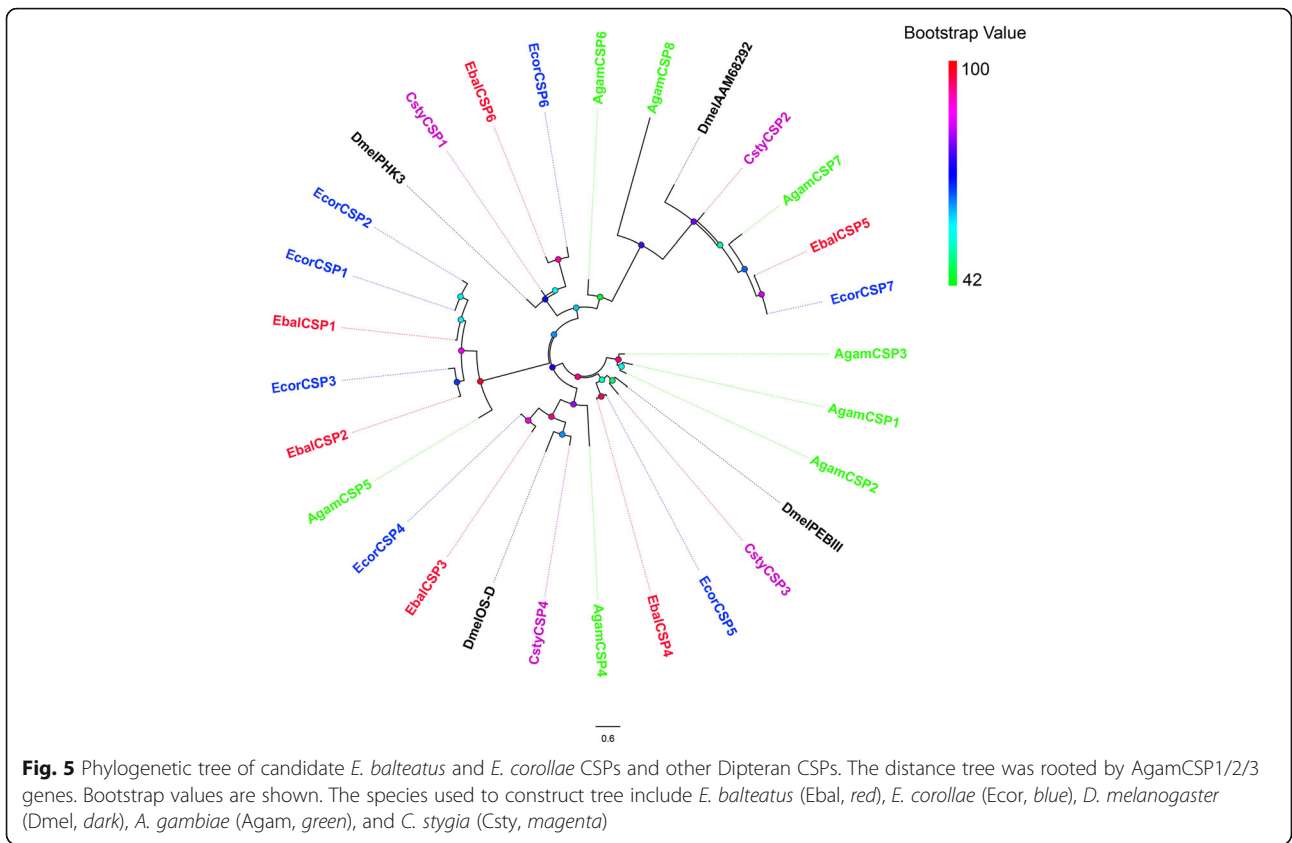
The expression of the candidate ORs in *E. balteatus* and *E. corollae* male and female antennae and legs (control sample) was analyzed using semi-quantitative reverse transcription PCR (RT-PCR). All 51 EbalORs and 42 EcorORs were detected in the antennae at high expressing level. Only EbalOR49 was found to be mainly expressed in legs. There were no significant differences of transcript abundances in the respective male and female antenna (Fig. 8). The Orco co-receptor gene also showed a high expression level in both syrphid species. This is consistent

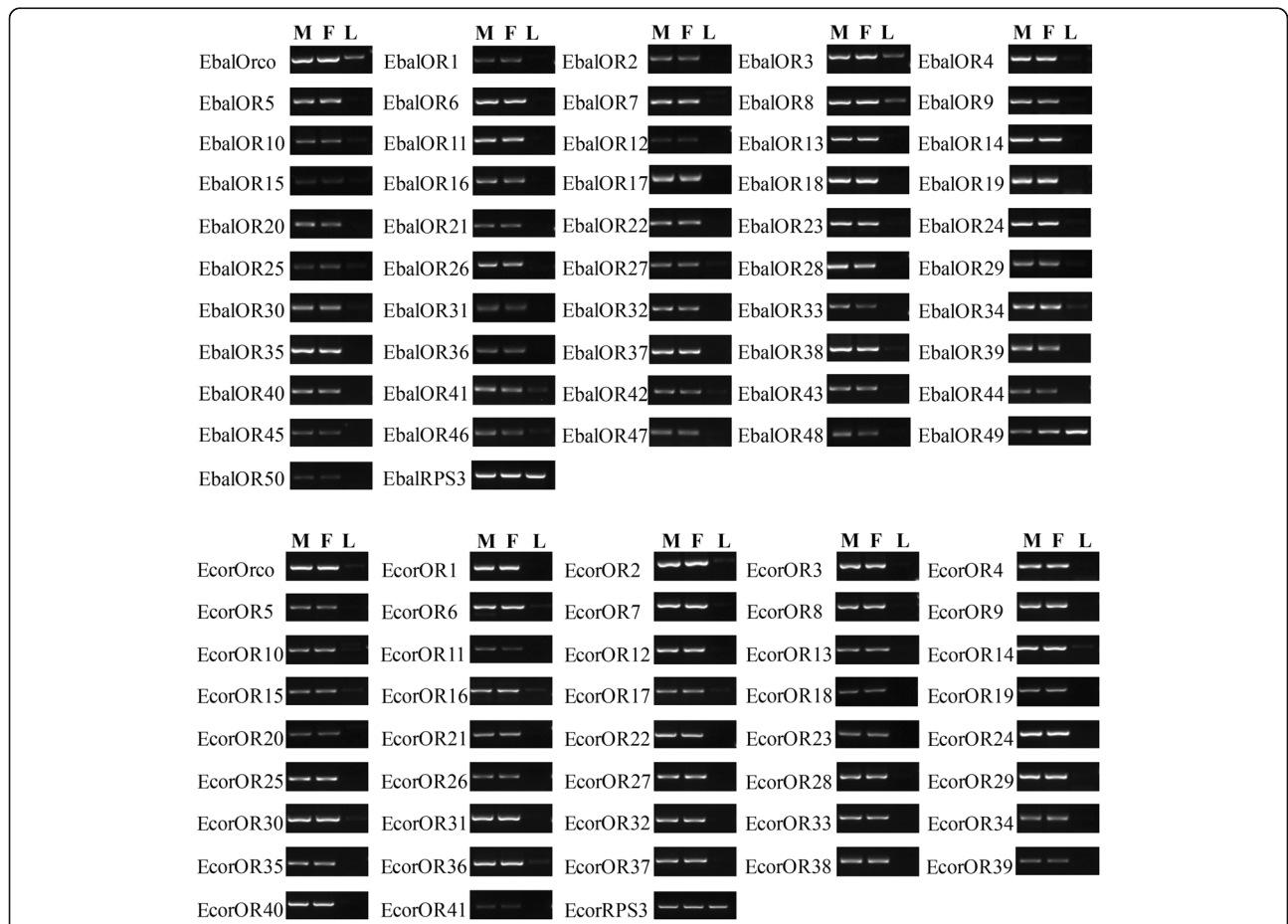
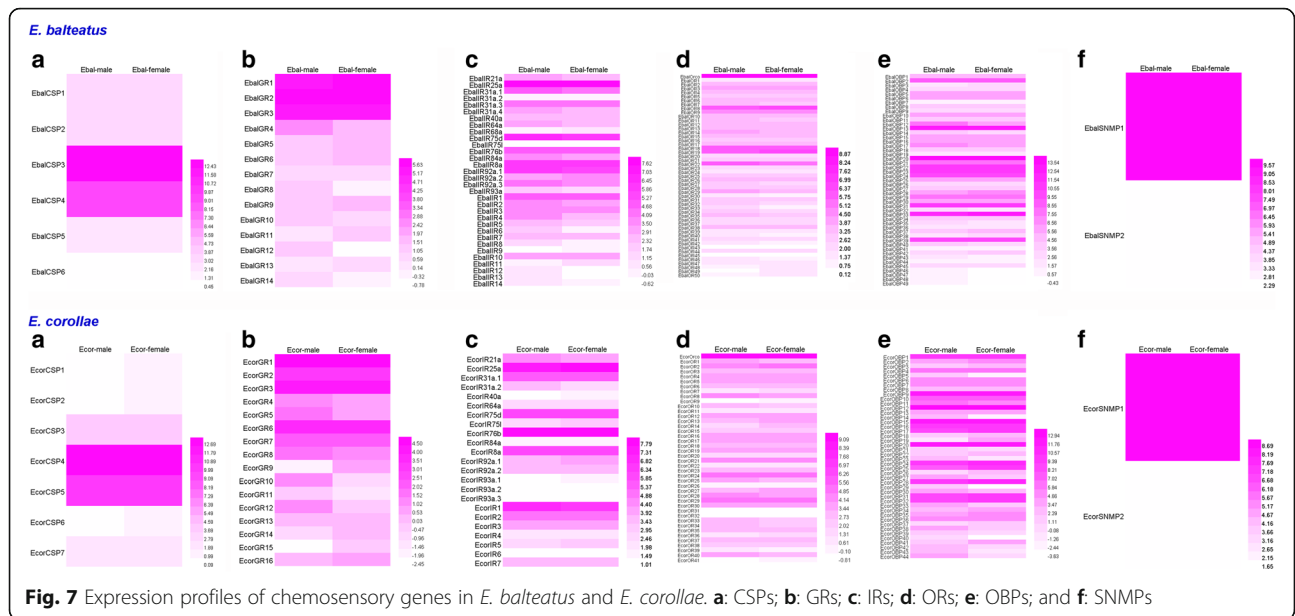
with DEGs analysis of OR transcript abundances using the FPKM value.

Discussion

The syrphids *E. balteatus* and *E. corollae* are aphid-specific predators and predominately inhabit northern China wheat and cotton fields. Typical of most insects, chemical cues drive several aspects of their behavior, such as foraging on prey and choosing oviposition sites [7, 8, 10]. Chemosensory proteins play an important role in this process. We analyzed antennal transcriptomes of *E. balteatus* and *E. corollae* and searched for chemosensory genes with the purpose of understanding chemical communication of tritrophic interactions among plants, herbivorous insects, and natural enemies.

In our study, we sequenced *E. balteatus* and *E. corollae* antennal transcriptomes using next generation sequencing technology on the Illumina HiSeq 2000 platform. The total RNA was converted into a template library for high throughput DNA sequencing, allowing





us to obtain all expressed transcripts. De novo assembly of transcripts using the Trinity method gives high-efficiency and reliable full-length transcripts across extensive expression levels, even without genome information [71]. Our sequence assembly yielded a final transcript dataset of 50,942 unigenes from *E. corollae* and 53,575 *E. balteatus* unigenes. Total unigenes counts resulted in 44.2% unigenes from *E. balteatus* and 50.3% unigenes from *E. corollae* shared sequence similarities to known proteins using the BLASTX homology search from the NCBI non-redundant protein database. These percentages are very similar to other Dipteran species [50, 72]. Remaining transcripts without associated GO terms may represent species-specific genes. The antennal transcriptome analysis proved to be a powerful tool to identify chemosensory genes in insects without genome information. It has been successfully employed in many insect orders including Lepidoptera, Coleoptera, Hymenoptera and Hemiptera. In Diptera, the chemosensory genes were successfully identified in *C. stygia*, *B. dorsalis* and *Scaeva pyrastris* antennal transcriptomes [50, 72, 73]. Here, we identified 154 and 134 candidate chemosensory genes in *E. balteatus* and *E. corollae*, respectively, a number similar to other Diptera antennal transcriptomes (e.g. 128 in *C. stygia*) [50] but less than the chemosensory genes identified in *D. melanogaster* (254), *M. domestica* (386) and *A. gambiae* (292) genome [30, 33, 47, 49, 63, 65, 74]. This could be the result of differential expression based on developmental stages of the insect larva or adult olfactory organ development such as maxillary palp and proboscis. All data shows that the chemosensory genes identified by antennal transcriptome sequencing are accurate and reliable.

We identified 154 candidate chemosensory genes (51 ORs, 32 IRs, 14 GRs, 49 OBPs, 6 CSPs and 2 SNMPs) in *E. balteatus* and 134 (42 ORs, 23 IRs, 16 GRs, 44 OBPs, 7 CSPs and 2 SNMPs) were identified in *E. corollae*, numbers slightly different compared with those of other Dipteran species [50, 73, 75, 76]. Such differences could be due to sequencing methods, coverage and/or depth. The number of chemosensory genes is higher in *E. balteatus* than in *E. corollae*. However, assembling and splicing quality (unigene number and N50 length) in *E. corollae* is better than in *E. balteatus*. The differences in the number and quality of transcripts identified could arise from variations in sample preparation or could be due to evolution [77] and adaptation to the environment (tritrophic interactions).

A total of 49 and 44 OBPs were identified in *E. balteatus* and *E. corollae* transcriptomes, respectively. The number of OBPs is variable across species, with 52 members in *D. melanogaster*, 66 in *A. gambiae* (Diptera), 21 in *Apis mellifera* (Hymenoptera), 34 in *Helicoverpa armigera*, 29 in *Helicoverpa assulta* (Lepidoptera), 26 in *Colaphellus bowringi*, 46 in *Tribolium castaneum* (Coleoptera) and 15 in

A. pisum (Hemiptera) [27, 74, 77–83]. Meanwhile, the OBPs of these two syrphid species are highly divergent with those of other insects. These evolutionary differences may result from different physiological functions or ecological niches. Compared with OBPs, only a small amount of CSPs were detected in Diptera. They are only 4 CSPs in *D. melanogaster*, 4 in *C. stygia* and 7 in *A. gambiae* (Diptera). These numbers are much lower than in other insect orders, such as 18 CSPs found in *H. armigera* and 17 in *H. assulta* (Lepidoptera) [27, 50, 80, 81]. In our study, six EbalCSPs and seven EcorCSPs are identified in transcriptome sequencing, revealing that the numbers of CSP gene family differ among species. CSPs show a high evolutionary diversity in insecta, probably related to different physiological functions.

We identified 51 ORs from *E. balteatus* and 42 ORs from *E. corollae*, respectively. Compared with other Dipteran species, these numbers are similar to those identified in *C. stygia* (50) [50] and *G. morsitans morsitans* (46) [51] but lower than those of *D. melanogaster* (62), *M. domestica* (86), *A. gambiae* (79) [47, 49, 84–86], suggesting that sequencing method/depth may be different between studies yielding less genes that may be difficult to detect because of low expression [77]. Here, we were able to detect species-specific OR transcripts in *E. balteatus* and *E. corollae*. This clade of ORs may have a greater impact on recognizing specific odors, particularly perception of aphids-derived volatiles and herbivore-induced plant volatiles granting syrphid localization access of its prey.

The tissue- and sex-specific expression analysis showed no differences between male and female, which is consistent with DEGs analysis of OR transcript abundances using FPKM values. Lepidopteran ORs have shown male-specific expression that is usually involved in the detection of the sex pheromone [19, 26, 27], but this does not seem to be the case in syrphids. Additional real time quantitative PCR, in situ hybridization and single-sensilla recordings would be required to validate OR expressions and functions.

In *D. melanogaster*, males release the volatile sex pheromone *cis*-vaccenyl acetate (cVA) [87–89]. The perception of sex pheromone cVA is mediated by OR67d [88], OR65a [90, 91], LUSH [92], and SNMP1 [34]. In our two syrphid species, EbalOBP17 and EcorOBP14 are the orthologues of the DmelOBP-lush gene, while EbalOR16 and EcorOR24 are the orthologues of DmelOR67d, and EbalSNMP1 and EcorSNMP1 are very similar to DmelSNMP1, suggesting that these proteins may be involved in detection of their yet unidentified pheromones. Therefore, further functional characterization of these candidate proteins will help reveal any mechanism associated with pheromone reception in *E. balteatus* and *E. corollae*.

The *E. balteatus* and *E. corollae* IR family is relatively conserved, especially with respect to common receptors IR8a and IR25a, which are expressed in both olfactory and

gustatory systems [30, 45]. The numbers of IRs identified in *E. balteatus* (32) and *E. corollae* (23) are similar to that of *C. stygia* (22) [50], but lower than those of *D. melanogaster* (66) and *A. gambiae* (46) [63]. It is possible that some IRs do not express in antennae tissues or perhaps the number of IRs varies between species and is dependent on natural habitats. A large number of EballRs and EcorIRs are clustered with “antennal” orthologues in *Drosophila*, indicating that IRs are highly conserved in Diptera. Furthermore, the IRs identified in these two species may be activated by acids, amines and other odorants that are not sensed by ORs [30, 45, 46].

In the antennae of *E. balteatus* and *E. corollae*, we identified 14 and 16 candidate GRs, respectively. The total number of GRs in these two species may be much larger, because some members could be exclusively expressed in other gustatory organs, such as maxillary palps, proboscises and legs. However, the numbers are still lower than those reported in other Dipteran antennal transcriptomes [50]. The conserved receptors identified in the two syrphid species may be involved in CO₂ perception. However, we infer that the mechanism of CO₂ perception is different from mosquitoes which concerns host-seeking [43, 44, 48, 93]. Some GRs may function as taste or contact receptors [31], particularly with reference to their specific pollination behavior [94, 95]. Some GRs from these two species are clustered with thermos-sensing GRs and sugar-detecting GRs from *Drosophila*, indicating that they may perform similar functions. Functional analysis of the candidate *E. balteatus* and *E. corollae* chemosensory proteins is required to identify their physiological roles.

Conclusions

We have identified and annotated 154 transcripts encoding putative chemosensory proteins in antennal transcriptome of *E. balteatus* and 134 in *E. corollae*. Comparisons between the two syrphid species and among other Dipteran species were deduced using sequence information. This work gives a foundation for future studies aimed at understanding chemical communication in syrphids and tritrophic interactions between plants, herbivorous insects, and natural enemies in agricultural ecosystems.

Methods

Insect rearing and tissue collection

E. balteatus and *E. corollae* larvae were fed with aphids (*Aphis gossypii* Glover) and maintained at 22 ± 1 °C with a 12 h light: 12 h dark photo-period at the Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China. Following eclosion, adult males and females were separated and provided with pollen and 10% honey solution.

Antennae were excised from 2- to 5-day-old adult males and females respectively, and legs were collected

together, then immediately frozen and stored in liquid nitrogen.

cDNA library construction and Illumina sequencing

Total RNA of male and female antennae was extracted from *E. balteatus* and *E. corollae* using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The method for RNA extraction followed in the manufacturer's instruction. Total RNA was dissolved in RNase-free water and RNA integrity was verified by gel electrophoresis. RNA concentration and purity were measured on a Nanodrop ND-2000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). Ten micrograms total RNA of each sample was used to construct the cDNA library. The cDNA library construction and Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) sequencing of the samples was performed at Beijing Genomics Institute (BGI, Shenzhen, China). The insert sequence length was around 200 bp and these libraries were pair-end sequenced using PE100 strategy [22, 27].

Assembly and function annotation

Raw reads were pre-processed by filtering low quality reads, trimming low quality nucleotides at each ends and removing 3' adaptors and poly-A/T tails. Each clean-read dataset of male and female antenna was fed to Trinity [71]. The Trinity assembly procedure, including Inchworm, Chrysalis and Butterfly were followed using Grabherr *et al.*, 2011 as a reference [71]. In the first step of Trinity, Inchworm assembles reads into the unique sequences of transcripts using the default parameters (default *k*-mers = 25). Next, Chrysalis clusters related contigs that correspond to portions of alternatively spliced transcripts or otherwise unique portions of paralogous genes. Finally, Butterfly uses read sequences, read-pairings and Chrysalis' read mappings to select the paths that are best supported by read sequences [71].

The Trinity outputs were clustered by TGICL [96]. The consensus cluster sequences and singletons make up the unigenes dataset [22]. The unigenes annotation was performed by NCBI BLASTX against a pooled database of non-redundant and SwissProt protein sequences with *e*-value <1e-5. The BLASTX results were then imported into Blast2GO pipeline for GO annotation [97].

Identification of chemosensory genes

Candidate unigenes encoding putative ORs, IRs, OBPs, CSPs, SNMPs and GRs were found by running Perl scripts against transcriptome assembly and annotation in the remote server. Perl scripts were written to extract sequence from functional annotation results using olfaction keywords. Subsequently, all candidate chemosensory genes were manually checked by BLASTX against local non-redundant database with *e*-value <1e-5. Using the BLASTX NCBI database, we manually performed alignments comparing transcripts against all known

proteins to examine full-length coverage. The full-length transcripts contain start and termination codons. The ORFs of all putative chemosensory genes were predicted by using ExPASy (Expert Protein Analysis System) server version (<http://web.expasy.org/translate/>) according to the BLASTX best hit result [98]. Putative N-terminal signal peptide of OBPs and CSPs were predicted by SignalP 4.0 server version with default parameters [99]. The TMDs of ORs, IRs and GRs were predicted using TMHMM server version 2.0 [100].

Sequence and phylogenetic analysis

After removing redundancy, alignments of amino acid sequences were performed by MAFFT (<https://www.ebi.ac.uk/Tools/msa/mafft/>). The phylogenetic trees of *E. balteatus* and *E. corollae* chemosensory genes were constructed by RaxML version 8 with Jones-Taylor-Thornton amino acid substitution model (JTT) [101] with the putative chemosensory genes in other Dipteran species (Additional file 8: Table S4). Node support was assessed using a bootstrap method based on 1000 replicates. The OR data set contained OR sequences identified in Dipteran (51 from *E. balteatus*, 42 from *E. corollae*, 62 from *D. melanogaster* [85, 86], 61 from *B. dorsalis* [genome: assembly ASM78921v2], 50 from *C. stygia* [50] and 81 from *M. domestica* [49]). The GR data set contained GR sequences identified in Dipteran (14 from *E. balteatus*, 16 from *E. corollae*, 68 from *D. melanogaster* [47], 40 from *B. dorsalis* [genome: assembly ASM78921v2], 21 from *C. stygia* [50] and 43 from *M. domestica* [49]). The IR data set contained IR sequences identified in Dipteran (32 from *E. balteatus*, 23 from *E. corollae*, 76 from *D. melanogaster* [30, 63], 22 from *C. stygia* [50] and 54 from *A. gambiae* [63, 84]). The OBP data set contained OBP sequences identified in Dipteran (49 from *E. balteatus*, 44 from *E. corollae*, 71 from *D. melanogaster* [64], 40 from *B. dorsalis* [genome: assembly ASM78921v2], 28 from *C. stygia* [50] and 52 from *M. domestica* [49]). The CSP data set contained CSP sequences identified in Dipteran (7 from *E. balteatus*, 9 from *E. corollae*, 4 from *D. melanogaster* [74], 4 from *C. stygia* [50] and 8 from *A. gambiae* [74]). The SNMP data set contained SNMP sequences identified in Dipteran (2 from *E. balteatus*, 2 from *E. corollae*, 2 from *D. melanogaster* [34, 35] and 2 from *A. gambiae* [35]).

DEGs analysis

A mapping-based expression profiling analysis of the chemosensory genes was conducted to compare gene expression between male and female antennae. All of the clean reads were remapped onto the transcripts using SOAPaligner (<http://soap.genomics.org.cn/soapaligner.html>), allowing up to three base mismatches and a minimum length of 40 bp. The FPKM method was used for calculating unigene expression levels [20, 50, 102, 103]. The suitable *P*-values were calculated to identify differentially expressed genes according to the hypergeometric test [103]. The FDR was a statistical

method used in multiple hypothesis testing to correct for *P*-value. Criteria for estimating significant differential expression was set at $FDR \leq 0.001$ and $|\log_2 \text{Ratio}| \geq 1$. Heatmaps of differential gene expression between male antennae and female antennae in both species were generated by Heml 1.0 software [104].

Expression analysis by semi-quantitative RT-PCR

Semi-quantitative RT-PCR was performed to verify the expression of candidate chemosensory genes. Male and female antennae and legs were collected from adult *E. balteatus* and *E. corollae* after eclosion. The extraction of total RNA followed the manufacturer's instruction [27]. The cDNA was synthesized from total RNA using RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA). Gene specific primers were designed using PrimerQuest Tool (<http://sg.idtdna.com/Primerquest/Home/Index>) (Additional file 9: Table S5) and synthesized by Sangon Biotech Co., Ltd. (Shanghai, China). A Taq MasterMix (CW BIO, Beijing, China) was used for PCR reactions under the general three-step amplification of 94 °C for 30s, 55 °C for 30s, 72 °C for 30s. RT-PCR products were separated on 2% agarose gels, stained by ethidium bromide (EB), and photographed under UV light in Gel Doc XR+ Gel Documentation System with Image Lab Software (Bio-Rad, Hercules, CA, USA).

Additional files

Additional file 1: Table S1. Assembly summary of *E. balteatus* and *E. corollae* antennal transcriptome. (DOCX 16 kb)

Additional file 2: Fig. S1. (A) Species distribution and annotation summaries in the *E. balteatus* (Ebal) and *E. corollae* (Ecor) antennal transcriptome assembly. (B) Gene ontology classifications of the *E. balteatus* and *E. corollae* unigenes with Blast2GO program, including categories with biological process, molecular function and cellular component. (TIFF 3397 kb)

Additional file 3: Table S2. Candidate *E. balteatus* and *E. corollae* antennal chemosensory genes. Unigenes of candidate odorant receptors (2–1), gustatory receptors (2–2), ionotropic receptors (2–3), odorant binding proteins (2–4), chemosensory proteins (2–5) and sensory neuron membrane proteins (2–6) with gene name, length, ORF, best BLASTX hit and identity. (DOCX 112 kb)

Additional file 4: Table S3. Comparison of homologous ORs in *E. balteatus* and *E. corollae*. (DOCX 61 kb)

Additional file 5: Fig. S2. Protein domain analysis of the species-specific IR clade with *Drosophila* iGluRs and DmellR94d /e. Amino acid alignments shows the ligand binding domains (S1 and S2), the ion channel pore (P), and TMD (M1, M2 and M3) of ionotropic receptors. The key ligand binding residues are marked in red box. (JPEG 4154 kb)

Additional file 6: Fig. S3. Amino acid alignments of the species-specific OBPs clade in the *E. balteatus* and *E. corollae*. The motif of six conserved cysteines are marked with asterisks at the top. (JPEG 3067 kb)

Additional file 7: Antennal expression levels of candidate *E. balteatus* and *E. corollae* odorant receptors. (XLSX 99 kb)

Additional file 8: Table S4. GenBank accession numbers of chemosensory genes used in phylogenetic analyses. (XLSX 34 kb)

Additional file 9: Table S5. Primers of candidate ORs in *E. balteatus* and *E. corollae* used for RT-PCR. (DOCX 20 kb)

Abbreviations

cDNA: Complementary DNA; CSP: Chemosensory protein; DEGs: Differentially expressed genes; FDR: False discovery rate; FPKM: Fragments per kilobase per million fragments; GLV: Green leaf volatiles; GO: Gene ontology; GR: Gustatory receptor; IR: Ionotropic receptor; OBP: Odorant-binding protein; OR: Odorant receptor; ORF: Open reading frame; PCR: Polymerase chain reaction; RT-PCR: Semi-quantitative reverse transcription PCR; SNMP: Sensory neuron membrane protein; TMD: Transmembrane domain

Acknowledgements

We thank Ph.D. students Kang He, Song Cao, Rui-Bin Zhang and M.S. student Yan-Wei Wang (Institute of Plant Protection, Chinese Academy of Agricultural Sciences) for help collecting insects. We thank Ms. Chun-Yan Wang and Ms. Li-Yan Yang (Institute of Plant Protection, Chinese Academy of Agricultural Sciences) for rearing insects. We are grateful to thank M.S. Xiang-Zhi Liang, Dr. Yong-Lei Liu and Dr. Ya-Nan Zhang for technical assistance. Thanks to Prof. Paolo Pelosi for editorial assistance and comments on the manuscript.

Funding

This project is supported by National Natural Science Foundation of China (31,621,064; 31,230,062&31,321,004 to GW, 31,402,023 to BW, 31,471,833 to YL), China Postdoctoral Science Foundation (2014 M550905 to BW), China Postdoctoral Science Foundation 2017T100121 to BW.

Availability of data and materials

The clean reads of the four antennal transcriptomes in this study have been stored in the NCBI SRA database, under the accession number of SRX1812754 (*E. balteatus* male), SRX1812787 (*E. balteatus* female), SRX1829150 (*E. corollae* male), and SRX1829282 (*E. corollae* female).

Authors' contributions

All authors contributed to research design and manuscript preparation. Conceived and designed the experiments: BW, YL, GW. Performed the experiments: BW. Analyzed the data: BW, YL. Contributed reagents/materials/analysis tools: BW, YL, GW. Wrote the paper: BW, YL, GW. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 January 2017 Accepted: 12 July 2017

Published online: 07 August 2017

References

- Amoros-Jimenez R, Robert CA, Marcos-Garcia MA, Fereres A, Turlings TC. A differential role of volatiles from conspecific and heterospecific competitors in the selection of oviposition sites by the Aphidophagous hoverfly *Sphaerophoria rueppellii*. *J Chem Ecol*. 2015;41(5):493–500.
- Brewer MJ, Elliott NC. Biological control of cereal aphids in north america and mediating effects of host plant and habitat manipulations. *Annu Rev Entomol*. 2004;49:219–42.
- Amoros-Jimenez R, Pineda A, Fereres A, Marcos-Garcia MA. Prey availability and abiotic requirements of immature stages of the aphid predator *Sphaerophoria rueppellii*. *Biol Control*. 2012;63(1):17–24.
- Freier B, Triltsch H, Mowes M, Moll E. The potential of predators in natural control of aphids in wheat: results of a ten-year field study in two German landscapes. *BioControl*. 2007;52(6):775–88.
- Haenke S, Scheid B, Schaefer M, Tschamtker T, Thies C. Increasing syrphid fly diversity and density in sown flower strips within simple vs. complex landscapes. *J Appl Ecol*. 2009;46(5):1106–14.
- Latham DR, Mills NJ. Quantifying insect predation: a comparison of three methods for estimating daily per capita consumption of two Aphidophagous predators. *Environ Entomol*. 2009;38(4):1117–25.
- Francis F, Lognay G, Haubruge E. Olfactory responses to aphid and host plant volatile releases: (E)-beta-farnesene an effective kairomone for the predator *Adalia bipunctata*. *J Chem Ecol*. 2004;30(4):741–55.
- Verheggen FJ, Arnaud L, Bartram S, Gohy M, Haubruge E. Aphid and plant volatiles induce oviposition in an aphidophagous hoverfly. *J Chem Ecol*. 2008;34(3):301–7.
- Poppy GM. Tritrophic interactions: improving ecological understanding and biological control? *Endeavour*. 1997;21(2):61–5.
- Vet LEM, Dicke M. Ecology of infochemical use by natural enemies in a tritrophic context. *Annu Rev Entomol*. 1992;37:141–72.
- Bargen H, Sauthof K, Poehling HM. Prey finding by larvae and adult females of *Episyrphus balteatus*. *Entomol Exp Appl*. 1998;87(3):245–54.
- Sadeghi H, Gilbert F. Oviposition preferences of aphidophagous hoverflies. *Ecol Entomol*. 2000;25(1):91–100.
- Sadeghi H, Gilbert F. Aphid suitability and its relationship to oviposition preference in predatory hoverflies. *J Anim Ecol*. 2000;69(5):771–84.
- Zhu J, Obrycki JJ, Ochieng SA, Baker TC, Pickett JA, Smiley D. Attraction of two lacewing species to volatiles produced by host plants and aphid prey. *Die Naturwissenschaften*. 2005;92(6):277–81.
- Almohamad R, Verheggen FJ, Francis F, Haubruge E. Predatory hoverflies select their oviposition site according to aphid host plant and aphid species. *Entomol Exp Appl*. 2007;125(1):13–21.
- Verheggen F, Ryne C, Olsson POC, Arnaud L, Lognay G, Hogberg HE, et al. Electrophysiological and behavioral activity of secondary metabolites in the confused flour beetle. *Tribolium confusum* *J Chem Ecol*. 2007;33(3):525–39.
- Verheggen FJ, Fagel Q, Heuskin S, Lognay G, Francis F, Haubruge E. Electrophysiological and behavioral responses of the multicolored asian lady beetle, *Harmonia axyridis* Pallas, to sesquiterpene semiochemicals. *J Chem Ecol*. 2007;33(11):2148–55.
- Suh E, Bohbot J, Zwiebel LJ. Peripheral olfactory signaling in insects. *Curr opin insect sci*. 2014;6:86–92.
- Liu Y, Gu SH, Zhang YJ, Guo YY, Wang GR. Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS One*. 2012;7(10):e48260.
- Andersson MN, Videvall E, Walden KKO, Harris MO, Robertson HM, Lofstedt C. Sex- and tissue-specific profiles of chemosensory gene expression in a herbivorous gall-inducing fly (Diptera: Cecidomyiidae). *BMC Genomics*. 2014; 15:501.
- Cao DP, Liu Y, Walker WB, Li JH, Wang GR. Molecular characterization of the *Aphis gossypii* olfactory receptor gene families. *PLoS One*. 2014;9(6):e101187.
- Cao DP, Liu Y, Wei JJ, Liao XY, Walker WB, Li JH, et al. Identification of candidate olfactory genes in *Chilo suppressalis* by antennal transcriptome analysis. *Int J Biol Sci*. 2014;10(8):846–60.
- Dippel S, Oberhofer G, Kahnt J, Gerischer L, Opitz L, Schachtner J, et al. Tissue-specific transcriptomics, chromosomal localization, and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics*. 2014;15:1141.
- Gu SH, Sun L, Yang RN, Wu KM, Guo YY, Li XC, et al. Molecular characterization and differential expression of olfactory genes in the antennae of the black cutworm moth *Agrotis ipsilon*. *PLoS One*. 2014;9(8): e103420.
- Hodges TK, Cosme LV, Athrey G, Pathikonda S, Takken W, Slotman MA. Species-specific chemosensory gene expression in the olfactory organs of the malaria vector *Anopheles gambiae* (Retracted article. See vol. 16, 572, 2015). *BMC genomics*. 2014; 15:1089.
- Zhang SF, Zhang Z, Wang HB, Kong XB. Antennal transcriptome analysis and comparison of olfactory genes in two sympatric defoliators, *Dendrolimus houi* and *Dendrolimus kikuchii* (Lepidoptera: Lasiocampidae). *Insect Biochem Mol Biol*. 2014;52:69–81.
- Zhang J, Wang B, Dong SL, Cao DP, Dong JF, Walker WB, et al. Antennal transcriptome analysis and comparison of chemosensory gene families in two closely related Noctuidae moths, *Helicoverpa armigera* and *H. assulta*. *PLoS one*. 2015;10(2):e0117054.
- Yang B, Ozaki K, Ishikawa Y, Matsuo T. Identification of candidate odorant receptors in Asian corn borer *Ostrinia furnacalis*. *PLoS One*. 2015;10(3): e0121261.
- Buck L, Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*. 1991;65:175–87.
- Benton R, Vannice KS, Gomez-Diaz C, Vossball LB. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell*. 2009; 136(1):149–62.
- Clyne PJ, Warr CG, Carlson JR. Candidate taste receptors in *Drosophila*. *Sci*. 2000;287(5459):1830–4.

32. Pelosi P, Zhou JJ, Ban LP, Calvello M. Soluble proteins in insect chemical communication. *Cell Mol Life Sci.* 2006;63(14):1658–76.
33. Pelosi P, Iovinella I, Felicioli A, Dani FR. Soluble proteins of chemical communication: an overview across arthropods. *Front Physiol.* 2014;5:320.
34. Benton R, Vannice KS, Vosshall LB. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nat.* 2007;450(7167):289–93.
35. Vogt RG, Miller NE, Litvack R, Fandino RA, Sparks J, Staples J, et al. The insect SNMP gene family. *Insect Biochem Mol Biol.* 2009;39(7):448–56.
36. Jin X, Ha TS, Smith DP. SNMP is a signaling component required for pheromone sensitivity in *Drosophila*. *Proc Natl Acad Sci U S A.* 2008;105(31):10996–1001.
37. Song HG, Kwon JY, Han HS, Bae YC, Moon C. First contact to odors: our current knowledge about odorant receptors. *Sensors.* 2008;8(10):6303–20.
38. Kirkness EF, Haas BJ, Sun WL, Braig HR, Perotti MA, Clark JM, et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci U S A.* 2010;107(27):12168–73.
39. Zhou XF, Slone JD, Rokas A, Berger SL, Liebig J, Ray A, et al. Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoS Genet.* 2012;8(8):e1002930.
40. Hallem EA, Ho MG, Carlson JR. The molecular basis of odor coding in the *Drosophila* antenna. *Cell.* 2004;117(7):965–79.
41. Hallem EA, Carlson JR. Coding of odors by a receptor repertoire. *Cell.* 2006;125(1):143–60.
42. Kwon JY, Dahanukar A, Weiss LA, Carlson JR. The molecular basis of CO2 reception in *Drosophila*. *Proc Natl Acad Sci U S A.* 2007;104(9):3574–8.
43. Gillies MT. The role of carbon dioxide in host-finding by mosquitoes (Diptera: Culicidae): a review. *Bull Entomol Res.* 1980;70:525–32.
44. Erdelyan CNG, Mahood TH, Bader TSY, Whyard S. Functional validation of the carbon dioxide receptor genes in *Aedes aegypti* mosquitoes using RNA interference. *Insect Mol Biol.* 2012;21(1):119–27.
45. Ai MR, Blais S, Park JY, Min S, Neubert TA, Suh GSB. Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex *in vivo* in *Drosophila*. *J Neurosci.* 2013;33(26):10741–9.
46. Ai M, Min S, Grosjean Y, Leblanc C, Bell R, Benton R, et al. Acid sensing by the *Drosophila* olfactory system. *Nat.* 2010;468(7324):691–5.
47. Robertson HM, Warr CG, Carlson JR. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A.* 2003;100:14537–42.
48. Rinker DC, Zhou XF, Pitts RJ, Rokas A, Zwiebel LJ, Consortium A. Antennal transcriptome profiles of anopheline mosquitoes reveal human host olfactory specialization in *Anopheles gambiae*. *BMC Genomics.* 2013;14:749.
49. Scott JG, Warren WC, Beukeboom LW, Bopp D, Clark AG, Giers SD, et al. Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. *Genome Biol.* 2014;15(10):466.
50. Leitch O, Papanicolaou A, Lennard C, Kirkbride KP, Anderson A. Chemosensory genes identified in the antennal transcriptome of the blowfly *Calliphora stygia*. *BMC Genomics.* 2015;16:255.
51. Obiero GF, Mireji PO, Nyanjom SR, Christoffels A, Robertson HM, Masiga DK. Odorant and gustatory receptors in the tsetse fly *Glossina morsitans morsitans*. *PLoS Negl Trop Dis.* 2014;8(4):e2663.
52. Francis FD, Martin T, Lognay G, Haubruge E. Role of (E)-beta-farnesene in systematic aphid prey location by *Episyrphus balteatus* larvae (Diptera: Syrphidae). *Eur J Entomol.* 2005;102(3):431–6.
53. Harmel N, Almohamad R, Fauconnier ML, Du Jardin P, Verheggen F, Marlier M, et al. Role of terpenes from aphid-infested potato on searching and oviposition behavior of *Episyrphus balteatus*. *Insect Sci.* 2007;14(1):57–63.
54. Turlings TCJ, Tumlinson JH. Systemic release of chemical signals by herbivore-injured corn. *Proc Natl Acad Sci U S A.* 1992;89:8399–402.
55. Pare PW, Tumlinson JH. De novo biosynthesis of volatiles induced by insect herbivory in cotton plants. *Plant Physiol.* 1997;114(4):1161–7.
56. Jones WD, Cayirlioglu P, Kadow IG, Vosshall LB. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nat.* 2007;445(7123):86–90.
57. Ni L, Bronk P, Chang EC, Lowell AM, Flam JO, Panzano VC, et al. A gustatory receptor paralogue controls rapid warmth avoidance in *Drosophila*. *Nat.* 2013;500(7464):580–4.
58. Dahanukar A, Lei YT, Kwon JY, Carlson JR. Two gr genes underlie sugar reception in *Drosophila*. *Neuron.* 2007;56(3):503–16.
59. Slone J, Daniels J, Amrein H. Sugar receptors in *Drosophila*. *Curr Biol.* 2007;17(20):1809–16.
60. Jiao Y, Moon SJ, Montell C. A *Drosophila* gustatory receptor required for the responses to sucrose, glucose, and maltose identified by mRNA tagging. *Proc Natl Acad Sci U S A.* 2007;104(35):14110–5.
61. Jiao Y, Moon SJ, Wang X, Ren Q, Montell C. Gr64f is required in combination with other gustatory receptors for sugar detection in *Drosophila*. *Curr Biol.* 2008;18(22):1797–801.
62. Fujii S, Yavuz A, Slone J, Jagge C, Song X, Amrein H. *Drosophila* sugar receptors in sweet taste perception, olfaction, and internal nutrient sensing. *Curr Biol.* 2015;25(5):621–7.
63. Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, et al. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* 2010;6(8):e1001064.
64. Hekmat-Scafe D, Scafe CR, McKinney AJ, Tanouye MA. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res.* 2002;12:1357–69.
65. Sanchez-Gracia A, Vieira FG, Rozas J. Molecular evolution of the major chemosensory gene families in insects. *Heredity.* 2009;103(3):208–16.
66. Krieger J, Raming K, Dewey YME, Bette S, Conzelmann S, Breer H. A divergent gene family encoding candidate olfactory receptors of the moth *Heliothis virescens*. *Eur J Neurosci.* 2002;16(4):619–28.
67. Rogers ME, Sun M, Lerner MR, Vogt RG. Snmp-1, a novel membrane protein of olfactory neurons of the silk moth *Antheraea polyphemus* with homology to the CD36 family of membrane proteins. *J Biol Chem.* 1997;272(23):14792–9.
68. Rogers ME, Krieger J, Vogt RG. Antennal SNMPs (sensor neuron membrane proteins) of lepidoptera define a unique family of invertebrate CD36-like proteins. *J Neurobiol.* 2001;49(1):47–61.
69. Forstner M, Gohl T, Gondesens I, Raming K, Breer H, Krieger J. Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths. *Chem Senses.* 2008;33(3):291–9.
70. Gu SH, Yang RN, Guo MB, Wang GR, Wu KM, Guo YY, et al. Molecular identification and differential expression of sensory neuron membrane proteins in the antennae of the black cutworm moth *Agrotis ipsilon*. *J Insect Physiol.* 2013;59(4):430–43.
71. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29(7):644–52.
72. Li XM, Zhu XY, He P, Xu L, Sun L, Chen L, et al. Molecular characterization and sex distribution of chemosensory receptor gene family based on transcriptome analysis of *Scaeva pyrastris*. *PLoS One.* 2016;11(5):e0155323.
73. Liu Z, Smagghe G, Lei Z, Wang JJ. Identification of male- and female-specific olfaction genes in antennae of the oriental fruit fly (*Bactrocera dorsalis*). *PLoS One.* 2016;11(2):e0147783.
74. Vieira FG, 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.* 2011;3:476–90.
75. Pitts RJ, Rinker DC, Jones PL, Rokas A, Zwiebel LJ. Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. *BMC Genomics.* 2011;12:271.
76. Riveron J, Boto T, Alcorta E. Transcriptomic basis of the acclimation to high environmental temperature at the olfactory receptor organs of *Drosophila melanogaster*. *BMC Genomics.* 2013;14:259.
77. Li XM, Zhu XY, Wang ZQ, Wang Y, He P, Chen G, et al. Candidate chemosensory genes identified in *Colaphellus bowringi* by antennal transcriptome analysis. *BMC Genomics.* 2015;16:1028.
78. Foret S, Maleszka R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* 2006;16(11):1404–13.
79. Foret S, Wanner KW, Maleszka R. Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expression profiling. *Insect Biochem Mol Biol.* 2007;37(1):19–28.
80. Wanner KW, Willis LG, Theilmann DA, Isman MB, Feng QL, Plettner E. Analysis of the insect os-d-like gene family. *J Chem Ecol.* 2004;30(5):889–911.
81. Zhou JJ, Vieira FG, He XL, Smadja C, Liu R, Rozas J, Field LM. Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrtosiphon pisum*. *Insect Mol Biol.* 2010;19:113–22.
82. Wang LJ, Wang SZ, Li YH, Paradesi MSR, Brown SJ. BeetleBase: the model organism database for *Tribolium castaneum*. *Nucleic Acids Res.* 2007;35:D476–9.
83. Xu PX, Zwiebel LJ, Smith DP. Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. *Insect Mol Biol.* 2003;12(6):549–60.

84. Fox AN, Pitts RJ, Robertson HM, Carlson JR, Zwiebel LJ. Candidate odorant receptors from the malaria vector mosquito *Anopheles gambiae* and evidence of down-regulation in response to blood feeding. *Proc Natl Acad Sci U S A*. 2001;98(25):14693–7.
85. Clyne PJ, Warr CG, Freeman MR, Lessing D, Kim JH, Carlson JR. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron*. 1999;22(2):327–38.
86. Gao Q, Chess A. Identification of candidate *Drosophila* olfactory receptors from genomic DNA sequence. *Genomics*. 1999;60(1):31–9.
87. Bartelt RJ, Schaner AM, Jackson LL. Cis-Vaccenyl acetate as an aggregation pheromone in *Drosophila melanogaster*. *J Chem Ecol*. 1985;11:1747–56.
88. Kurtovic A, Widmer A, Dickson BJ. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nat*. 2007;446(7135):542–6.
89. Lebreton S, Trona F, Borrero-Echeverry F, Bilz F, Grabe V, Becher PG, et al. Feeding regulates sex pheromone attraction and courtship in *Drosophila* females. *Sci Rep*. 2015;5:13132.
90. Liu W, Liang X, Gong J, Yang Z, Zhang YH, Zhang JX, et al. Social regulation of aggression by pheromonal activation of Or65a olfactory neurons in *Drosophila*. *Nat Neurosci*. 2011;14(7):896–902.
91. Lebreton S, Grabe V, Omondi AB, Ignell R, Becher PG, Hansson BS, et al. Love makes smell blind: mating suppresses pheromone attraction in *Drosophila* females via Or65a olfactory neurons. *Sci Rep*. 2014;4:7119.
92. Kim MS, Repp A, Smith DP. LUSH odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. *Genetics*. 1998;150(2):711–21.
93. Dekker T, Geier M, Carde RT. Carbon dioxide instantly sensitizes female yellow fever mosquitoes to human skin odours. *J Exp Biol*. 2005;208(15):2963–72.
94. Jauker F, Diekötter T, Schwarzbach F, Wolters V. Pollinator dispersal in an agricultural matrix: opposing responses of wild bees and hoverflies to landscape structure and distance from main habitat. *Landsc Ecol*. 2009;24(4):547–55.
95. Raymond L, Plantegenest M, Vialatte A. Migration and dispersal may drive to high genetic variation and significant genetic mixing: the case of two agriculturally important, continental hoverflies (*Episyrphus balteatus* and *Sphaerophoria scripta*). *Mol Ecol*. 2013;22(21):5329–39.
96. Perte G, Huang XQ, Liang F, Antonescu V, Sultana R, Karamycheva S, et al. TIGR gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*. 2003;19(5):651–2.
97. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674–6.
98. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 2003;31(13):3784–8.
99. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785–6.
100. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001;305(3):567–80.
101. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312–3.
102. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5(7):621–8.
103. Audic S, Claverie JM. The significance of digital gene expression profiles. *Genome Res*. 1997;7(10):986–95.
104. Deng WK, Wang YB, Liu ZX, Cheng H, Xue Y. Heml: a toolkit for illustrating Heatmaps. *PLoS One*. 2014;9(11):e111988.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

