De novo assembly and identification of antennal transcriptome reveals an olfactory system in *Heortia vitessoides* (Lepidoptera: Crambidae)

Jie Cheng [1], **Chun-Yan Wang** [1], **Zi-Hao Lyu** [1], **Jing-Xiang Chen** [1], **Tong Lin** Corresp. [1]

[1] College of Forestry and Landscape Architecture, South China Agricultural University, GuangZhou, GuangDong, China

Corresponding Author: Tong Lin
Email address: lintong@scau.edu.cn

Olfactory systems in insects are used to detect volatile chemical odors, and play crucial roles in survival, reproduction, and mediating key behaviors. Here, RNA sequencing technology was used to sequence and assemble the antennal transcriptome of *Heortia vitessoides*, a defoliating pest in *Aquilaria sinensis* (Loureiro) Sprenger forests and a non-model species with no genomic resources. Analysis of the transcriptome of female and male antennae generated 22.16 gigabases of genomic data, from which 52,383 unigenes were assembled. We identified 80 candidate olfactory genes: eight for odorant binding proteins (OBPs), 14 for chemosensory proteins (CSPs), 35 odorant receptors (ORs), 18 ionotropic receptors (IRs), three gustatory receptors (GRs), and two for sensory neuron membrane proteins (SNMPs). Furthermore, phylogenetic trees and fragments per kilobase of transcript per million fragments mapped (FPKM) were used to analyze these olfactory genes. This study is the first comprehensive antennal transcriptome analysis for *H. vitessoides,* and these novel olfactory genes will increase understanding of the molecular mechanism of chemoreception and further contribute to exploring strategies to manage this insect.

De novo assembly and identification of antennal transcriptome reveals an olfactory system in *Heortia vitessoides* (Lepidoptera: Crambidae)

Jie Cheng[1], Chun-Yan Wang[1], Zi-Hao Lyu[1], Jing-Xiang Chen[1], Tong Lin[1]

[1]College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, China

Corresponding Author:

Tong Lin[1]

Wushan street, Guangzhou, Guangdong 510642, China,

Email address: lintong@scau.edu.cn

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3433v1 | CC BY 4.0 Open Access | rec: 26 Nov 2017, publ: 26 Nov 2017

Abstract

Olfactory systems in insects are used to detect volatile chemical odors, and play crucial roles in survival, reproduction, and mediating key behaviors. Here, RNA sequencing technology was used to sequence and assemble the antennal transcriptome of *Heortia vitessoides*, a defoliating pest in *Aquilaria sinensis* (Loureiro) Sprenger forests and a non-model species with no genomic resources. Analysis of the transcriptome of female and male antennae generated 22.16 gigabases of genomic data, from which 52,383 unigenes were assembled. We identified 80 candidate olfactory genes: eight for odorant binding proteins (OBPs), 14 for chemosensory proteins (CSPs), 35 odorant receptors (ORs), 18 ionotropic receptors (IRs), three gustatory receptors (GRs), and two for sensory neuron membrane proteins (SNMPs). Furthermore, phylogenetic trees and fragments per kilobase of transcript per million fragments mapped (FPKM) were used to analyze these olfactory genes. This study is the first comprehensive antennal transcriptome analysis for *H. vitessoides,* and these novel olfactory genes will increase understanding of the molecular mechanism of chemoreception and further contribute to exploring strategies to manage this insect.

Keywords

Heortia vitessoides Antennae Transcriptome analysis Olfactory genes

Introduction

Chemical signals are detected by insects using olfactory systems (Field, Pickett & Wadhams, 2000). The olfactory could communicate signals with the environment and make the behaviors, including habits choosing, couple searching, food hunting, gathering, tropism, and signal communication (Fatouros et al. ,2008; Leal, 2013). Olfactory-related proteins, mainly located in the sensillum hemolymph on the antennae, and to a lesser extent on other sensory appendages (Vosshall & Stocker, 2007; Zhang et al., 2016a; Sheng et al., 2017). These proteins include many types, namely, odorant-binding proteins (OBPs), chemosensory proteins (CSPs), olfactory receptors (ORs), ionotropic receptors (IRs), gustatory receptors (GRs), and sensory neuron membrane proteins (SNMPs) (Fan et al., 2011; Chen et al., 2016; Wang, Liu & Wang, 2017). Among them, the non-receptor proteins are encoded by three gene families, the OBPs, CSPs and SNMPs. ORs, IRs and GRs belonging to receptors proteins (Zhao et al., 2016).

OBPs of insects are soluble acidic proteins, are highly concentrated in the lymph of chemosensory sensilla of insect antennae, and belong to four major classes: PBP (pheromone-binding protein), general OBP1, general OBP2, and OBP-like (Vogt, Rybcynski & Lemer, 1991; Wang, Guo & Wu, 2002). OBPs are widely involved in olfactory perception, and play a key role in transporting hydrophobic odorants to the ORs (Sanchez-Gracia, Vieira & Rozas, 2009; Ji et al., 2013). CSPs are small soluble acidic proteins, formed during long evolution, and are abundantly distributed in the antennae, tarsi and other sensory appendages of insects (Maleszka & Stange, 1997). CSPs among insects are well-conserved, and the identity is generally 30%–90%. They are believed to be involved in chemical communication, including the perception, identification, transport, and transduction of semiochemicals from the environment (including olfaction, taste, and others), and may be associated with the regulation of circadian rhythms and maturation of tissue or appendages (Marchese et al., 2000; Anholt & Mackay, 2001; Briand et al., 2002). SNMPs, being the only identified neuronal members of the CD36 family, belong to a larger gene family of receptor proteins characterized by the human protein CD36. The insect SNMP/CD36 genes are divided into three major groups, and SNMPs belong to group three, which are divided into two subgroups SNMP1 and SNMP2 (Nichols & Vogt, 2008). The SNMP1 subgroup is specifically located in the antennae, and are related to pheromone-specific olfactory sensory neurons (OSNs). The SNMP2 subgroup is expressed in neurons, and to a greater extent in sensilla support cells. It is also associated with pheromone detection (Benton, Vannice & Vosshall, 2007; Forstner et al., 2008; Vogt et al., 2009).

Insect OSNs express two types of ORs: conventional ORs, a highly divergent family of receptors that are expressed in small subpopulations of OSNs; and a member of the olfactory co-receptor family, *Orco* (formerly called *OR83b*), a receptor without odor sensitivity, which is expressed in the majority of OSNs and is remarkably conserved across insect species (Elmore & Smith, 2001; Larsson et al., 2004; Pitts, Fox & Zwiebel, 2004; Smith, 2007).

Membrane topology of ORs shows an inverted orientation with G-protein-coupled receptors (Benton et al., 2006). Insect ORs function via *Orco*, which transduces odorant signals by both heteromeric ligand-gated ion channels and cyclic nucleotide-activated cation channels, and therefore can respond to them rapidly, transiently, sensitively, and efficiently (Jones et al., 2005; Sato et al., 2008; Wicher et al., 2008). IRs are a new chemosensory receptor family, and is another variant subfamily of ionotropic glutamate receptors (iGluRs). The function of IRs is associated with the detection of chemical signals (Benton et al., 2006; Benton et al., 2009). Insect IRs can be divided into two distinct subfamilies based on the expression location: the conserved "antennal IRs" and the species-specific "divergent IRs" (Croset et al., 2010; Abuin et al., 2011). In antennal IRs, similarly to *Orco*, *IR8a* and *IR25a* act as co-receptors, are broadly expressed, and play essential roles in tuning IR sensory cilia targeting and IR-based sensory channels (Abuin et al., 2011; Zhang et al., 2016b). GR sequences are conserved to a much greater extent than OR sequences (McBride & Arguello, 2007; Gardiner et al., 2008). GRs are dominantly expressed in the gustatory organs, such as the mouthparts (Dong et al., 2016). When the taste organs are exposed to environmental stimuli, GRs in the taste neurons can encode these chemical signals, convert them into electrical signals, and then transmit these signals to the central nervous system through nerve axons in the form of pulses (Schoonhoven, Loon & Dicke, 2005).

Studies on the identification of olfactory genes in insects are making significant progress. Next-generation sequencing techniques such as RNA sequencing (RNA-seq), aided by decreasing costs along with technical advancements, have become valuable tools for researchers to obtain vast amounts of genetic information from non-model organisms without any prior sequence knowledge (Liu et al., 2016). Over the past several years, RNA-seq has been used to identify olfactory genes by antennal transcriptome sequencing in many Lepidoptera species, such as *Ostrinia furnacalis* (Zhang et al., 2015a), *Chilo suppressalis* (Cao et al., 2014), *Helicoverpa armigera* (Zhang et al., 2015b), and *Agrotis ipsilon* (Gu et al., 2014).

Heortia vitessoides Moore (Lepidoptera: Crambidae) is a serious defoliating pest of *Aquilaria sinensis* (Loureiro) Sprenger, an endemic species to China that produces valuable agarwood, a fragrant wood widely used in traditional medicine and the incense industry (Qiao et al., 2013; Jin et al., 2016). Field investigations have indicated that a large-scale outbreak of *H. vitessoides* can cause serious defoliation of *A. sinensis,* and therefore lead to severe damage to tree growth and agarwood production (Lu et al., 2014; Wen et al., 2016; Zhou et al., 2016). The biological characteristics, pesticide control methods, and behavior of *H. vitessoides* have been extensively studied for applications in insect control (Cheng, Chen & Lin, 2017). However, there are no reports detailing an effective means to control this pest. Genetic information about the olfactory genes in *H. vitessoides* is scarce. In this study, the female and male antennal transcriptome of *H. vitessoides* were sequenced on a BGISEQ-500 platform, and identified 80 candidate olfactory genes (eight OBPs, 14 CSPs, 35 ORs, 18 IRs, three GRs and two SNMPs). We conducted a comprehensive phylogenetic analysis of these olfactory genes. Furthermore, using fragments per kilobase of transcript per million fragments mapped (FPKM), the gene-level expressions of these candidate olfactory genes were measured and compared between the female and male antennae of *H. vitessoides*. Our results will aid genetic and genomic research related to the chemosensory system in this species.

Materials and methods

Insect rearing and antennae collection

H. vitessoides larvae were collected in May, 2017, from an *A. sinensis* plantation (22° 01′ N, 110° 25′E) in Huazhou, Guangdong, China. No chemical treatment was applied before or during collection. The insects were fed with fresh *A. sinensis* leaves in the laboratory under conditions of 26 ± 2°C with 70 ± 2% relative humidity, and a consistent 14 h:10 h light/dark cycle. After pupation, pupae were sexed according to the morphology of the eighth and ninth abdominal segments (Cao et al., 2013). Female and male pupae were stored in separate plastic containers. After eclosion, adults were fed with 7% honey solution. The antennae were collected (100 pairs from each sex) two to three days after eclosion and stored in liquid nitrogen immediately.

RNA sample preparation

Total RNA was extracted using the E.Z.N.A.™ Total RNA Kit II (OMEGA Biotec, Norcross, GA, USA) following the supplier's instructions, and was treated with DNase I (Invitrogen, Life Technologies, Carlsbad, CA, USA). A Nanodrop 2000 spectrophotometer (NanoDrop Products, Wilmington, DE, USA) was used to check sample purity, while a Qubit1 2.0 fluorometer (Life Technologies, Gaithersburg, MD, USA) and Quantifluor-ST fluorometer with Agilent 2100 Bioanalyzer (Promega, Madison, WI, USA) were used to measure for concentration and integrity, respectively.

cDNA library construction

The qualified RNA samples were used for transcriptome sequencing. The first step involved purifying the poly-(A)-containing mRNA molecules using poly-T oligo-attached magnetic beads. Following purification, the mRNA was fragmented using divalent cations under elevated temperature. The cleaved RNA fragments were copied into first-strand cDNA using reverse transcriptase and random primers. This was followed by second-strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments were then subjected to the addition of a single 'A' base, and subsequently the ligation of the DNA adapters. The products were then purified and enriched with PCR amplification. We then quantified the PCR yield using Qubit, and pooled samples together to make a single strand DNA (ssDNA) circle, which gave the final library. The cDNA library was subjected to sequencing on the BGISEQ-500 platform (BGI, Shenzhen, China).

De novo assembly and function annotation

Clean reads were obtained by removing reads containing adaptor sequences, more than 5% unknown nucleotides, more than 50% bases with Q-value ≤ 20, and empty reads. Then, the Q30 and GC-content were used to assess the sequencing quality. Next, *de novo* assembly was performed using Trinity software (Grabherr et al., 2011) to generate unigenes. The TGICL software (Pertea et al., 2003) was used to assemble the unigenes to form a single set of non-redundant unigenes. To acquire comprehensive information on gene functions, assembled unigenes were searched against the Nr (NCBI non-redundant protein sequences), Nt (NCBI nucleotide), KOG (eukaryotic ortholog groups), SwissProt, KEGG (Kyoto Encyclopedia of Genes and Genomes) database using BLASTx and BLASTn with an E-value < 10^{-5}. Blast2GO (Conesa et al., 2005) was used for GO (Gene Ontology) annotation with an E-value < 10^{-5} based on the protein annotation results of the Nr database. InterPro functional annotation was performed using InterProScan5 (Quevillon et al., 2005).

Identification of *H. vitessoides* olfactory genes

The candidate unigenes encoding putative OBPs, CSPs, ORs, IRs, GRs, and SNMPs were retrieved from the assembled sequences. A BLASTn analysis was also performed by searching these putative genes using the sequences of homologous genes from other lepidopteran species. The sequences representing each family of genes that showed highest homology with each other according to the alignment results were chosen for further analysis.

Phylogenetic analyses

The phylogenetic trees were constructed for phylogenetic analyses of *H. vitessoides* OBPs (*HvitOBPs*), *HvitCSPs*, *HvitORs*, *HvitIRs*, *HvitGRs*, and *HvitSNMPs*, based on these genes and the sequences of other insects. The OBP dataset contained eight sequences from *H. vitessoides* and 24 from lepidopteran insects (eight from *Cnaphalocrocis medinalis*, eight from *Ostrinia furnacalis*, and eight from *Conogethes punctiferalis*). The CSP dataset contained 14 sequences from *H. vitessoides* and 42 from lepidopteran insects (14 from *O. furnacalis*, 14 from *C. medinalis*, and 14 from *Bombyx mori*). The OR dataset contained 35 sequences from *H. vitessoides* and 84 from lepidopteran insects (32 from *C. punctiferalis*, 25 from *C. medinalis*, 16 from *Helicoverpa assulta*, and 11 from *Helicoverpa armigera*). The IR dataset contained 18 sequences from *H. vitessoides* and 48 from lepidopteran and dipteran insects (nine from *C. punctiferalis*, 20 from *Drosophila melanogaster*, eight from *Dendrolimus houi*, and 11 from *H. assulta*). The GR dataset contained three sequences from *H. vitessoides* and nine from lepidopteran insects (three from *C. punctiferalis*, three from *H. assulta*, and three from *H. armigera*). The SNMP data set contained two sequences from *H. vitessoides* and six from lepidopteran insects (two from *O. furnacalis*, two from *Chilo suppressalis*, and two from *C. medinalis*). Amino acid sequences were aligned with ClustalX 1.83 (Thompson et al., 1997). Phylogenetic trees were constructed by the neighbor-joining method, with a *p*-distance model and a pairwise deletion of gaps, as implemented in MEGA5.0 software (Tamura et al., 2011). Node support was assessed using a bootstrap procedure based on 1,000 replicates, and node support values < 50% are not shown.

Expression abundance analyses

Gene-level expression measurement of OBPs, CSPs, ORs, IRs, GRs, and SNMPs between female and male antennae of *H. vitessoides* were reported in FPKM. Clean sequencing reads of each sample were compared with the unigene database using the Bowtie2 program (Langmead & Salzberg, 2012). RSEM software (Li & Dewey, 2011) was used to measure the gene expression values, and the FPKM value was used to represent expression abundance of the genes. The FPKM of a gene is calculated with the formula:

$$\mathrm{FPKM} = \frac{cDNA\ Fragments}{Mapped\ Fragments(Millions) \times Transcript\ \ Length(kb)},$$

where cDNA fragments represent the number of aligned fragments within certain transcripts and indicate the number of double reads. Mapped fragments (Millions) depict the total number of aligned fragments within transcripts with a unit of 10^6. Transcript length (kb) is the length of transcripts with a unit of 10^3 bases.

Results

Sequencing and *de novo* assembly

To identify the olfactory genes from *H. vitessoides*, cDNA from male and female antennae were sequenced using the BGISEQ-500 platform. Sequencing yielded a total of 117.04 Mb and 117.04 Mb of raw reads for the male and female antennae samples, respectively. After removing adaptor sequences, low quality sequences, and N-containing sequences, 110.52 Mb and 111.07 Mb of clean reads were generated from the male and female antennae raw data, respectively. All clean reads from male and female antennae were assembled to generate 46,797 (mean length 1,146 bp) and 46,347 (mean length 1,105 bp) unigenes. We assembled all of the clean reads from male and female antennae and generated 52,383 unigenes (mean length 1,238 bp) (Table 1). The size distribution of the assembled unigenes is shown in Fig 1.

Sequence annotation

The number of matched unigenes and the match percentages at different values are shown in Table 2. In summary, 24,805 (47.35%), 13,981 (26.69%), 17,772 (33.93%), 18,917 (36.11%), 17,333 (33.09%), 17,384 (33.19%), and 2,346 (4.48%) unigenes had homologous sequences in the Nr, Nt, Swiss-Prot, KEGG, KOG, Interpro, and GO databases, respectively. The number of unigenes that were annotated by any of the seven functional databases was 51.59% of the total. Only 2.43% of the unigenes were annotated in all databases. For species distribution, the highest match percentage was to *Amyelois transitella* (36.15%) sequences, followed by the sequences of *Bombyx mori* (13.03%), *Papilio xuthus* (10.86%), and *Papilio machaon* (8.99%) (Fig 2).

GO annotation was used to classify the function of transcripts according to the GO terms (Fig S1). In the biological process terms, cellular process (1,247), metabolic processes (1,174), and single-organism processes (729) were the most abundant. In the cellular component terms, cell (931), cell part (924), and membrane (618) were the highest classified. In the molecular function terms, the numbers of genes involved in binding (1,045), catalytic activity (936), and structural molecule activity (254) were the greatest.

In total, 17,333 unigenes were assigned to 25 KOG functional categories (Fig S2). Of these categories, 'General function prediction only' represented the largest group, containing 4,405 unigenes, followed by 'Signal transduction mechanisms' (3,714) and 'Function unknown' (2,004). The 'Cell motility' (58), 'Coenzyme transport and metabolism' (145), and 'Nucleotide transport and metabolism' (238) categories were the smallest clusters represented.

To understand the biological pathways involved in the antennae of *H. vitessoides*, the sequences were mapped to reference canonical pathways in KEGG (Fig S3). In summary, 18,917 unigenes were classified into six groups, Cellular Processes, Environmental Information Processing, Genetic Information Processing, Human Diseases, and Metabolism and Organismal Systems. Transport and catabolism (1,389), Signal transduction (2,736), Translation (1,212), Cancers: Overview (1,680), Global and overview maps (2,347), and Immune system (1,465) were the dominant pathways in each group, respectively.

Candidate OBPs in *H. vitessoides* antennae

In total, we identified eight transcripts belonging to the OBP family (Table 3). A phylogenetic analysis was conducted to evaluate the relationships between *HvitOBPs* and OBPs for other lepidopteran species. The neighbor-joining tree showed that the eight *HvitOBPs* were spread across several branches, and that six *HvitOBPs* (OBP1, OBP2, OBP3, OBP4, OBP5, and OBP8) were clustered with orthologs of other lepidopteran species (Fig 3). *HvitOBP1* was clustered with OBPs from *C. medinalis* and *O. furnacalis* (Fig 3). Of these OBPs, *HvitOBP3*, *HvitOBP1*, and *HvitOBP2* showed the highest expression level in females (FPKM = 7,752.16, 3,955.53, 1,324.36, respectively) and males (FPKM = 6,798.6, 1,368.05, 903.07, respectively) (Fig 9).

Candidate CSPs in *H. vitessoides* antennae

In the current study, 14 candidate CSPs from the *H. vitessoides* antennal transcriptome were identified (Table 4). The phylogenetic tree constructed using CSP sequences from *H. vitessoides* and three other lepidopteran species revealed that four *HvitCSPs* (*HvitCSP3*, *HvitCSP5*, *HvitCSP8*, and *HvitCSP10*) were clustered together (Fig 4). The FPKM results showed that *HvitCSP2* (FPKM = 1,402.46), *HvitCSP1* (FPKM = 1,368.66), and *HvitCSP5* (FPKM = 1,122.45) displayed the highest expression levels in females (Fig 10). *HvitCSP1* (FPKM = 1,099.85), *HvitCSP2* (FPKM = 988.35), and *HvitCSP5* (FPKM = 850.83) displayed the highest expression levels in males (Fig 10).

Candidate ORs in *H. vitessoides* antennae

In total, 35 different sequences that encode candidate OR genes were identified by bioinformatic analysis (Table 5). We next performed a phylogenetic analysis using our candidate ORs and the ORs from four other lepidopteran species. The results showed that 35 *HvitORs* were spread across several branches, and *HvitOrco* was clustered with other lepidopteran *Orco* sequences (*CpunOR2*, *CmedOrco*, and *HassOR83b*), which formed one small branch (Fig 5). Amino acid alignment of the candidate *HvitOrco* with other insects revealed that the similarity of *Orco* was higher than 91% in *O. furnacalis*, *D. indica*, *C. punctiferalis*, *C. medinalis* and *C. suppressalis* (Fig 6). Of the co-receptors, *Orco/OR2* was most abundant in females (FPKM = 68.75) and males (FPKM =79.46) (Fig 11).

Candidate IRs in *H. vitessoides* antennae

In total, 18 different genes were annotated as candidate IRs from the *H. vitessoides* antennal transcriptome (Table 6). In the neighbor-joining tree of IRs, *HvitIR25a* grouped with the highly conserved *IR25a/IR8a* (Fig 7). *HvitIR25a* showed the highest homology with *C. punctiferalis* (97%), and had amino acid sequence homology with the *IR25a* gene from five species of insects, such as *C. pomonella* (90%–93%) (Fig 8). *HvitIR25a*, *HvitIR1*, and *HvitIR4* had high expression levels in females (FPKM values of 91.86, 39.13, and 18.96, respectively) and males (FPKM values of 116.76, 30.77, and 19.62, respectively) (Fig 12). Similar to *Orco*, *IR8a* and *IR25a* act as co-receptors, and *HvitIR25a* had the highest expression level in female and male antennae (Fig 12).

Other candidate olfactory genes in *H. vitessoides* antennae

Antennal transcriptome of *H. vitessoides* assembly led to the identification of three GR candidates (Table S2). Using the phylogenetic tree of the GRs from *H. vitessoides* and other lepidopteran species, we found that *HvitGR3* was clustered with *CpunGR3* and *HarmGR1* (Fig S4). The FPKM results showed that *HvitGR3* displayed the highest expression levels in both female (FPKM =2.04) and male (FPKM =1.63) antennae (Fig S6).

We also obtained two SNMPs from the antennal transcriptome of *H. vitessoides* (Table S2). Based on the phylogenetic analysis, we found that *HvitSNMP1* was clustered with *OfurSNMP1*, *CmedSNMP1*, and *CsupSNMP1*, which was expected since they belong to the same family (Fig S5). *HvitSNMP2* showed the same phylogeny as *HvitSNMP1* (Fig S5). Between the two SNMPs, *HvitSNMP2* (FPKM = 313.84) had the highest expression level in female antennae and *HvitSNMP1* (FPKM = 660.36) was dominant in male antennae (Fig S7).

Discussion

H. vitessoides, an insect belonging to Lepidoptera, Crambidae, is a defoliator that is widely distributed throughout the southern coastal areas of China (Chen et al., 2011). To date, many studies on *H. vitessoides* have focused on the elucidation of the ecological and biological aspects, while few have been aimed at molecular biology, and no studies on olfactory genes have been reported in this pest. With the rapid development of next-generation sequencing technology, a large number of functional genes in non-model organisms have been identified (Yan et al., 2015). In the present study, using next-generation sequencing technology, we generated the first transcriptome databases from the female and male antennae of *H. vitessoides*. In total, 80 transcripts were identified, including eight OBPs, 14 CSPs, 35 ORs, 18 IRs, three GRs, and two SNMPs. All 80 transcripts were novel for *H. vitessoides*. These results supply the basis for elucidating molecular mechanisms of olfactory behaviors in *H. vitessoides*, and have determined a large number of target genes to develop specific and environmentally-friendly pesticides.

By using homology alignment, we identified eight OBPs based on the female and male antennal transcriptome of *H. vitessoides*. Compared with the insects in which the OBPs were identified by analyzing antenna transcriptomes, the number of the candidate OBPs identified by the current study in *H. vitessoides* was similar to the numbers found in *Atta vollenweideri* (eight OBPs) (Koch et al., 2013), *Sclerodermus sp.* (10 OBPs) (Zhou et al., 2015), and *Chrysoperla sinica* (12 OBPs) (Li et al., 2015), but was lower than that in *Agrotis ipsilon* (33 OBPs) (Newcomb et al., 2014) and *Bombyx mori* (46 OBPs) (Ou et al., 2014). The gene expression analysis of OBPs in *H. vitessoides* might help to characterize the function of these proteins in future research. *HvitOBP3*, *HvitOBP1*, and *HvitOBP2* showed the highest expression level in female and male antennal, which suggests that they may have important functions during olfactory related behaviors.

Insect CSPs represent a novel group of olfactory proteins that probably function in a manner similar to OBPs in insect chemoreception (Zhao et al., 2016). In total, 14 CSPs were identified in the antennal transcriptome of *H. vitessoides*. The number of CSP genes is highly variable among lepidopteran species, ranging from 10 in *O. furnacalis* (Zhang et al., 2015a) to 24 in *Sesamia inferens* (Zhang et al., 2013), according to the transcriptome data. The FPKM results showed that *HvitCSP3*, *HvitCSP9*, and *HvitCSP14* had the lowest expression levels in females (FPKM = 0.07, 6.33, and 1.75, respectively) and males (FPKM = 0.07, 3.55, and 2.14, respectively). There is a possible reason for this result; CSPs have broad expression profiles in olfactory organs such as the antennae, maxillary palps, and labial palps, and non-olfactory tissues such as the legs, wings, and pheromone glands (Sheng et al., 2017). We constructed the antenna transcriptome, and therefore may not have detected CSP expression from other tissues.

In chemosensory signal transduction processes, ORs play a central role as a bio-transducer, facilitating the conversion of the chemical signal to an electrical signal (Cao et al., 2014). Insects of different orders usually have a large number of antennal-expressed ORs. For example, in Lepidoptera: *Helicoverpa assulta* (64 ORs) (Zhang et al., 2015b), *Helicoverpa armigera* (60 ORs) (Zhang et al., 2015b), and *Bombyx mori* (66 ORs) (Wanner & Robertson, 2008; Tanaka et al., 2009); in Coleoptera: *Rhynchophorus ferrugineus* (77 ORs) (Antony et al., 2016) and *Anomala corpulenta* (43 ORs) (Li et al., 2015a); in Hymenoptera: *Chouioia cunea* (80 ORs) (Zhao et al., 2016) and *Atta vollenweideri* (70 ORs) (Koch et al., 2013); in Diptera: *Bactrocera dorsalis* (43 ORs) (Jin et al., 2017); and in Neuroptera: *Chrysoperla sinica* (37 ORs) (Li et al., 2015b). As a receptor that is strikingly different from all other insect olfactory receptors, *Orco* is highly conserved in insect species and widely expressed. *Orco* plays a critical role in insect olfaction. (Touhara & Vosshall, 2009). From the antennal transcriptome of *H. vitessoides*, we found a co-receptor *HvitOrco* that showed a high degree of similarity with *Orco* genes of other insects, indicating homologous functions of *HvitOrco* in *H. vitessoides*. *HvitOrco* had the highest expression level in both female and male antennae, as with other insects, such as *Nysius ericae* (Zhang et al., 2016) and *Conogethes punctiferalis* (Ge et al., 2016).

In total, 18 IRs were identified in the antennal transcriptome of *H. vitessoides*. The number of *H. vitessoides* IRs identified was more than that identified in *Sesamia inferens* (three IRs) (Zhang et al., 2013) and *Ips typographus* (seven IRs) (Andersson et al., 2013), and was similar to the numbers found in *Helicoverpa armigera* (19 IRs) (Zhang et al., 2015b) and *Helicoverpa assulta* (19 IRs) (Zhang et al., 2015b). Like *Orco*, *IR8a* and *IR25a* play the role of co-receptors in the insect IR system, and could be co-expressed together with other IRs to ensure precise and efficient responses to odorants (Benton et al., 2009). In the present study, *HvitIR25a* was found clustered in the *IR25a*/*IR8a* group with other insect species, and was similar to *IR25a* in other insects, with a higher antennal expression (Olivier et al., 2011; Feng et al., 2015). Although we identified *IR25a* genes in *H. vitessoides*, *IR8a* was not found.

Additionally, we identified two SNMPs and only three GRs in our transcriptomic data. The number of GR genes identified in this study was clearly less than in other insects, such as *Cydia pomonella* (20 GRs) (Walker et al., 2016). There are two possible reasons to explain this problem. Firstly, adult *H. vitessoides* antennae are not primary gustatory organs (Clyne, Warr & Carlson, 2000). Secondly, the current sequencing technology may not be powerful enough to screen all candidate genes, especially those transcripts with extremely low abundance in the antennae (Liu et al., 2015; Sheng et al., 2017). SNMP is one of the important proteins involved in insect odor recognition, and two SNMP subfamilies (SNMP1 and SNMP2) have been discovered in insects (Vogt et al., 2009). For most lepidopteran species, two SNMPs were detected from an antennal transcriptome (Xie, Nian & Su, 2016). Here, *HvitSNMP1* and *HvitSNMP2* were identified in *H. vitessoides*. The phylogenetic tree clearly showed *HvitSNMP1* and *HvitSNMP2* belonging to separate phylogenetic groups, indicating a functional conservatism among lepidopteran insects (Chang et al., 2017).

Conclusions

In conclusion, we successfully constructed a female and male antennal transcriptome dataset for *H. vitessoides*. Furthermore, we have identified 80 novel olfactory genes including eight OBPs, 14 CSPs, 35 ORs, 18 IRs, three GRs, and two SNMPs. Our findings make it possible for future research of the olfactory system of *H. vitessoides* at the molecular level, and these olfactory genes are promising targets for further research into controlling this pest in an eco-friendly manner.

Acknowledgments

We thank lecturer Xiao-Sheng Chen (South China Agricultural University) and Masters student Wei-Dong Huang (South China Agricultural University) for their assistance with the phylogenetic analyses.

References

Abuin L,Bargeton B,Ulbrich MH,Isacoff EY,Kellenberger S,Benton R.2011. Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* 69:44-60 DOI: 10.1016/j.neuron.2010.11.042

Anholt RR, Mackay TF.2001. The genetic architecture of odor-guided behavior in *Drosophila melanogaster*. *Behavior Genetics* 31:17-27 DOI:10.1023/A:1010201723966

Antony B, Soffan A, Jakse J, Abdelazim MM, Aldosari SA, Aldawood AS, Pain A.2016. Identification of the genes involved in odorant reception and detection in the palm weevil *Rhynchophorus ferrugineus*, an important quarantine pest, by antennal transcriptome analysis. *BMC Genomics* 17(1):69 DOI: 10.1186/s12864-016-2362-6

Andersson MN, Grosse-Wilde E, Keeling CI, Bengtsson JM, Yuen MMS, Li M, Hillbur Y, Bohlmann J, Hansson BS, Schlyter F. 2013. 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(1):1 DOI: 10.1186/1471-2164-14-198

Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB.2009. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*.*Cell* 136:149-162 DOI: 10.1016/j.cell.2008.12.001

Benton R, Vannice KS, Vosshall LB.2007. An essential role for a CD36-related in pheromone detection in *Drosophila*.*Nature* 450:289-293 DOI: 10.1038/nature06328

Benton R, Sachse S, Michnick SW, Vosshall LB.2006. Atypical membrane topology and heteromeric function of Drosophila odorant receptors in vivo.*PLOS Biology* 4:e20

DOI: 10.1371/journal.pbio.0040020

Briand L, Swasdipan N, Nespoulous C, Bezirard V, Blon F, Huet JC, Ebert P, Pernollet JC.2002. Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera* L.) as a brood pheromone carrier. *European Journal of Biochemistry* 269(18):4586-4596 DOI: 10.1046/j.1432-1033.2002.03156.x

Cao CL, Zheng LF, Li YZ, Wen XJ, Yang XC. 2013. A method used for distinguishing the pupae sexuality of *Heortia vitessoide*.*China Forestry Science and Technology* 27(2):121-122

DOI:10.3969/j.issn.1000-8101.2013.02.033

Cao D, Liu Y, Wei J, Liao X, Walker WB, Li J, Wang G.2014. Identification of candidate olfactory genes in Chilo suppressalis by antennal transcriptome analysis.*International Journal of Biological Sciences* 10(8):846-860 DOI: 10.7150/ijbs.9297

Chen Y, He M,Li ZQ, Zhang YN, He P.2016. Identification and tissue expression profile of genes from three chemoreceptorfamilies in an urban pest, *Periplaneta americana*.*Scientific Reports* 6:27495 DOI: 10.1038/srep27495

Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson, TJ, Benton, R.2010. Ancient protostome origin of chemosensory ionotropic glutamate receptoers and the evolution of insect taste and olfaction.*PLOS Genetics* 6:e1001064 DOI: 10.1371/journal.pgen.1001064

Cheng J,Chen JX, Lin T. 2017.*De novo* assembly and analysis of the *Heortia vitessoides* transcriptome *via* high-throughput Illumina sequencing. *Journal of Asia-Pacific Entomology* 20:1241-1248 DOI:10.1016/j.aspen.2017.08.014

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.*Bioinformatics* 21:3674-3676 DOI: 10.1093/bioinformatics/bti610

Chen ZY, Li DW, Wang L, Li YZ, Huang XR, Qin CS.2011.Studies on biological characteristics of *Heortia vitessoides* Moore on *Aquilaris sinensis*.*China Plant Protection* 31(11):10-14 DOI:10.3969/j.issn.1672-6820.2011.11.002

Clyne P.J, Warr CG, Carlson JR.2000. Candidate taste receptors in *Drosophila*. *Science* 287:1830-1834 DOI:10.1126/science.287.5459.1830

Chang XQ, Nie XP, Zhang Z, Zeng FF, Lv L, Zhang S, Wang MQ.2017.De novo analysis of the oriental armyworm *Mythimna separata* antennal transcriptome and expression patterns of odorant-binding proteins.*Comparative Biochemistry and Physiology Part D* 22:120-130

DOI: 10.1016/j.cbd.2017.03.001

Dong JF, Song YQ, Li WL, Shi J, Wang ZY. 2016. Identification of Putative Chemosensory Receptor Genes from the *Athetis dissimilis* Antennal Transcriptome. *PLoS ONE* 11(1): e0147768 DOI: 10.1371/journal.pone.0147768

Elmore T,Smith DP.2001. Putative *Drosophila* odor receptor OR43b bcalizes to dendrites of olfactory neurons. *Insect Biochemistry and Molecular Biology* 31(8):791-798 DOI:10.1016/S0965-1748(00)00184-3

Field LM,Pickett JA,Wadhams LJ.2000. Molecular studies in insect olfaction.*Insect Molecular Biology* 9:545-551 DOI:10.1046/j.1365-2583.2000.00221.x

Fatouros NE,Dicke M,Mumm R,Meiners T,Hilker M.2008. Foraging behavior of egg parasitoids exploiting chemical information.*Behavioral Ecology* 19:677-689 DOI: 10.1093/beheco/arn011

Fan J, Francis F, Liu Y, Chen JL, Cheng DF.2011. An overview of odorant-binding protein functions in insect peripheral olfactory reception. *Genetics and Molecular Research* 10:3056-3069 DOI: 10.4238/2011.December.8.2

Forstner M,Gohl T,Gondesen I,Raming K,Breer H,Krieger J.2008. Differential expression of SNMP-1 and SNMP-2 proteins in pheromone-sensitive hairs of moths.*Chemical Senses* 33(3):291-299 DOI: 10.1093/chemse/bjm087

Feng B, Lin XD, Zheng KD, Qian K, Chang YC, Du YJ. 2015. Transcriptome and expression

profiling analysis link patterns of gene expression to antennal responses in *Spodoptera litura*. *BMC Genomics* 16:269 DOI: 10.1186/s12864-015-1375-x

Gardiner A, Barker D, Butlin RK, Jordan WC, Richie MG.2008.*Drosophila* chemoreceptor gene evolution:selection,specialization and genome size. *Molecular Ecology* 17:1648-1657 DOI: 10.1111/j.1365-294X.2008.03713.x

Grabherr MG,Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma F, Birren BW,

Nusbaum C, Lindblad-Toh K, Friedma N, Regev A.2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7):644-U130 DOI: 10.1038/nbt.1883

Gu SH, Sun L, Yang RN, Wu KM, Guo YY, Li XC, Zhou JJ, Zhang YJ.2014. Molecular characterization and differential expression of olfactory genes in the antennae of the black cutworm moth *Agrotis ipsilon*.*PLoS ONE* 9(8):e103420 DOI: 10.1371/journal.pone.0103420

Ge X, Zhang TT, Wang ZY, He KL, Bai SX.2016. Identification of putative chemosensory receptor genes from yellow peach moth *Conogethes punctiferalis* (Guenée) antennae Transcriptome.*Scientific Reports* 6:32636 DOI: 10.1038/srep32636

Ji P, Liu JT, Gu SH, Zhu XQ, Zhang YJ, Guo YY.2013. Expression and binding specificity analysis of odorant binding protein AlucOBP7 from *Apolygus lucorum* (Hemiptera: Miridae).*Acta Entomologica Sinica* 56(6): 575-583 DOI:10.16380/j.kcxb.2013.06.0

Jones WD, Nguyen TAT, Kloss B, Lee KJ, Vosshall LB.2005. Functional conservation of an insect odorant receptor gene across 250 million years of evolution.*Current Biology* 15:R119-R121

DOI: 10.1016/j.cub.2005.02.007

Jin XF, Ma T, Chang MS, Wu YJ, Liu ZT, Sun ZH, Shan TJ, Chen XY, Wen XJ, Wang C.2016. Aggregation and Feeding Preference of Gregarious *Heortia vitessoides* (Lepidoptera: Crambidae) Larvae to *Aquilaria sinensis* (Thymelaeaceae). *Journal of Entomological Science* 51(3):209-218 DOI: http://dx.doi.org/10.18474/JES15-39

Jin S, Zhou X, Gu F, Zhong G,Yi X.2017. Olfactory Plasticity: Variation in the Expression of Chemosensory Receptors in *Bactrocera dorsalis* in Different Physiological States. *Frontiers in Physiology* 8:672 DOI: 10.3389/fphys.2017.00672

Koch SI, Groh K, Vogel H, Hansson BS, Kleineidam CJ, Grosse-Wilde E.2013. Caste-specific expression patterns of immune response and chemosensory related genes in the leaf-cutting ant,*Atta vollenweideri*.*PLoS ONE* 8(11):e81518 DOI: 10.1371/journal.pone.0081518

Leal WS.2013. Odorant reception in insects:roles of receptors, binding proteins, and degrading enzymes.*Annual Review of Entomology* 58:373-391 DOI: 10.1146/annurev-ento-120811-153635

Larsson MC,Domingos AI,Jones WD,Chiappe ME,Amrein H,Vosshall LB.2004.*Or83b* encodes a broadly expressed odorant receptor essential for *Drosophila* olfantion.*Neuron* 43(5):703-714 DOI: 10.1016/j.neuron.2004.08.019

Liu YJ, Qi MG, Chi YC, Wuriyanghan HD. 2016. De Nove Assembly of the Transcriptome for Oriental Armyworm *Mythimna separata* (Lepidoptera:Noctuidae) and Analysis on Insecticide Resistance-Related Genes. .*Journal of Insect Science* 16(1):92;1-8 DOI: 10.1093/jisesa/iew079

Lu J, Liu Y, Li LL, Zhang M, Wen XJ, Li YZ.2014.A Experiment on 7 kinds of Insecticides Against *Heortia vitessoides* Moore.*Journal of Shandong Forestry Science and Technology* 6:37-39.

DOI:10.3969/j.issn.1002-2724.2014.06.009

Langmead B, Salzberg SL.2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357-359 DOI: 10.1038/NMETH.1923

Li B, Dewey CN.2011.RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.*BMC Bioinformatics* 12:323 DOI: 10.1186/1471-2105-12-323

Li X, Ju Q, Jie WC, Li F, Jiang XJ, Hu JJ, Qu MJ, 2015a. Chemosensory gene families in adult antennae of *Anomala corpulenta* Motschulsky (Coleoptera: Scarabaeidae: Rutelinae). *PLoS ONE* 10(4):e0121504 DOI: 10.1371/journal.pone.0121504

Li ZQ, Zhang S, Luo JY, Wang SB, Wang CY, Lv LM, Dong SL, Cui JJ.2015b. Identification and expression pattern of candidate olfactory genes in *Chrysoperla sinica* by antennal transcriptome analysis. *Comparative Biochemistry and Physiology Part D* 15:28-38

DOI: 10.1016/j.cbd.2015.05.002

Liu S, Rao XJ, Li MT, Feng MF, He MZ, Li SG.2015. Identification of candidate chemosensory genes in the antennal transcriptome of *Tenebrio molitor* (Coleoptera: Tenebrionidae).*Comparative Biochemistry and Physiology Part D* 13:44-51 DOI: 10.1016/j.cbd.2015.01.004

Maleszka R, Stange G.1997. Molecular clonging, by a noval approach,of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Catoblastis cactorum*.*Gene* 202:39-43 DOI:10.1016/S0378-1119(97)00448-4

Marchese S, Angeli S, Andolfo A, Scaoni A, Brandazza A, Mazza M, Picimbon JF, Leal WS, Pelosi P.2000. Soluble proteins from chemosensory organs of *Eurycantha calcarata* (Insects, Phasmatodea). *Insect Biochemistry and Molecular Biology* 30:1091-1098 DOI:10.1016/S0965-1748(00)00084-9

McBride CS, Arguello JR.2007.Five Drosophila genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics* 177:1395-1657 DOI: 10.1534/genetics.109.107946

Nichols Z, Vogt RG.2008. The SNMP/CD36 gene family in Diptera,Hymenoptera and Coleoptera: *Drosophila melanogaster*, *D.pseudoobscura*,*Anopheles gambiae*,*Aedes aegypti*, *Apis mellifera*, and *Tribolium castaneum*.*Insect Biochemistry and Molecular Biology* 38(4):398-415 DOI: 10.1016/j.ibmb.2007.11.003

Newcomb RD, Gu SH, Sun L, Yang RN, Wu KM, Guo YY, Li XC, Zhou JJ, Zhang YJ. 2014. Molecular characterization and differential expression of olfactory genes in the antennae of the black cutworm moth *Agrotis ipsilon*.*PLoS ONE* 9(8):e103420 DOI: 10.1371/journal.pone.0103420

Ou J, Deng HM, Zheng SC, Huang LH, Feng QL, Liu L.2014. Transcriptomic analysis of developmental features of Bombyx mori wing disc during metamorphosis.*BMC Genomics* 15:820 DOI: 10.1186/1471-2164-15-820

Olivier V, Monsempes C, Francois MC, Poivet E, Jacquin-Joly E. 2011. Candidate chemosensory ionotropic receptors in a Lepidoptera.*Insect Molecular Biology* 20:189-199 DOI: 10.1111/j.1365-2583.2010.01057.x

Pitts RJ, Fox AN, Zwiebel LJ.2004.A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector *Anopheles gambiae*.*Proceedings of the National Academy of Sciences of the United States of America* 101(14):5058-5063 DOI: 10.1073/pnas.0308146101

Pertea G, Huang XQ, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B.2003.TIGR Gene Indices clustering tools(TGICL):a software system for fast clustering of large EST datasets.*Bioinformatics* 19:651-652 DOI: 10.1093/bioinformatics/btg034

Qiao HL, Lu PF, Chen J, Xu CQ, Ma WS, Qin RM, Li XM, Cheng HZ. 2013. Biological characteristics and occurrence patterns of *Heortia vitessoides*. *Chinese Journal of Applied Entomology* 50(5):1244-1252 DOI: 10.7679 /j.issn.2095 -1353.2013.170

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R.2005. InterProScan: protein domains identifier.*Nucleic Acids Research* 33:W116-20 DOI: 10.1093/nar/gki442

Sheng S, Liao CW, Zheng Y, Zhou Y, Xu Y, Song WM,He P, Zhang J, Wu FA.2017. Candidate chemosensory genes identified in the endoparasitoid *Meteorus pulchricornis* (Hymenoptera: Braconidae) by antennal transcriptome analysis. *Comparative Biochemistry and Physiology Part D* 22:20-31 DOI: 10.1016/j.cbd.2017.01.002

Sanchez-Gracia A, Vieira FG, Rozas J.2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity(Edinb)* 103:208-216 DOI: 10.1038/hdy.2009.55

Smith DP.2007. Odor and pheromone detection in *Drosophila melanogaster*. *Pflugers Archiv-european Journal of Physiology* 454(5):749-758 DOI: 10.1007/s00424-006-0190-2

Sato K, Pellegrino M, Nakagawa T, Nakagawa T, Vosshall LB, Touhara K.2008. Insect olfactory receptors are heteromeric ligand-gated ion channels.*Nature* 452:1002-1006 DOI: 10.1038/nature06850

Schoonhoven LM, Loon JJAV, Dicke M.2005. Insect-Plant Biology. Oxford: Oxford University Press 183-189.

Thompson JD,Gibson TJ,Plewniak F,Jeanmouin F,Higgins DG.1997. The CLUSTAL_X windows inferface: flexible strategies for multiple sequence alignment aided by quality analysis tools.*Nucleic Acids Research* 25:4876-4882 DOI:10.1093/nar/25.24.4876

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011.MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28:2731-2739 DOI:10.1093/molbev/msr121

Tanaka K, Uda Y, Ono Y, Nakagawa T, Suwa M, Yamaoka R, Touhara k.2009. Highly Selective Tuning of a Silkworm Olfactory Receptor to a Key Mulberry Leaf Volatile.*Current Biology* 19:881-890

DOI: 10.1016/j.cub.2009.04.035

Touhara K, Vosshall LB. 2009. Sensing odorants and pheromones with chemosensory

receptors.*Annual Review of Physiology* 71:307-332

DOI: 10.1146/annurev.physiol.010908.163209

Vosshall LB,Stocker RF.2007. Molecular architecture of small and taste in *Drosophila*. *Annual Review of Neuroscience* 30:505-533 DOI: 10.1146/annurev.neuro.30.051606.094306

Vogt RG, Rybcynski R, Lemer M R.1991. Molecular cloning and sequencing of general-odorant binding proteins GOBP1 and GOBP2 from the tobacco hawk moth *Manduca sexta*:Comparisons with other insect OBPs and their signal peptides.*The Journal of neuroscience* 11:2972-2984.

Vogt RG, Miller NE, Litvack R, Fandino RA, Sparks J, Staples J, Friedman R, Dickens JC.2009. The insect SNMP gene family. *Insect Biochemistry and Molecular Biology* 39:448-456 DOI: 10.1016/j.ibmb.2009.03.007

Wang B, Liu Y, Wang GR.2017. Chemosensory genes in the antennal transcriptome of two syrphid species, *Episyrphus balteatus* and *Eupeodes corollae*(Diptera: Syrphidae).*BMC Genomics* 18:586 DOI: 10.1186/s12864-017-3939-4

Wang GR, Guo YY, Wu KM.2002. Progress in the studies of antenna odorant binding proteins of insects. *Acta Entomologica Sinica* 45 (1):131-137 DOI:10.16380/j.kcxb.2002.01.023

Wicher D, Schfer R, Bauernfeind R, Stensmyr MC, Heller R, Heinemann SH, Hansson BS.2008.*Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels.*Nature* 452(7190):1007-1011 DOI: 10.1038/nature06861

Wen YZ, Jin XF, Zhu CQ, Chen X, Ma T, Zhang SN, Zhang Y, Zeng SC, Chen XY, Sun ZH, Wen XJ, Wang C.2016. Effect of Substrate Type and Moisture on Pupation and Emergence of *Heortia vitessoides* (Lepidoptera:Crambidae): Choice and No-Choice Studies.*Journal of Insect Behavior* 29:473–489

DOI: 10.1007/s10905-016-9572-2

Wanner KW, Robertson HM. 2008. The gustatory receptor family in the silkworm moth *Bombyx mori* is characterized by a large expansion of a single lineage of putative bitter receptors. *Insect Molecular Biology* 17:621-629 DOI: 10.1111/j.1365-2583.2008.00836.x

Walker WB, Gonzalez F, Garczynski SF, Witzgall P.2016. The chemosensory receptors of codling moth *Cydia pomonella*-expression in larvae and adults. *Scientific Reports* 6:23518

DOI: 10.1038/srep23518

Xie CQ, Nian HY, Su SK.2016. Advances in research on antennal transcriptome in insects. *Chinese Journal of Applied Entomology* 53(6):1157-1164 DOI: 10.7679/j.issn.2095-1353.2016.143

Yan W, Liu L, Qin WQ,Li CX,Peng ZQ.2015. Transcriptomic identification of chemoreceptors genes the red palm weevil *Phynchophorus ferrugineus*.*Genetics and Molecular Research* 14:7469-7480 DOI: 10.4238/2015.July.3.23

Zhang YN,Zhu XY,Zhang Q,Yin CY,Dong ZP,Zuo LH,Deng DG,Sun L,Li XM.2016a. De novo assmbly and characterization of antennal transcriptome reveal chemosensory system in *Nysius ericae*.*Journal of Asia-Pacific Entomology* 19:1077-1087 DOI: 10.1016/j.aspen.2016.09.013

Zhang YN, Ma JF, Dong ZP, Li ZQ, Zhu XY, Wang Y, Wang L, Deng DG, Li JB.2016b. Molecular identification and sex distribution of two chemosensory receptor families in *Athetis lepigone* by antennal transcriptome analysis. *Journal of Asia-Pacific Entomology* 19:571-580

DOI: 10.1016/j.aspen.2016.05.009

Zhang TT, Coates BS, Ge X, Bai SX, He KL, Wang ZY. 2015a. Male- and female-biased gene expression of olfactory-related genes in the antennae of asian corn borer,*Ostrinia furnacalis* (Guence)(Lepidoptera: Crambidae). *PLoS ONE* 10(6):e0128550

DOI: 10.1371/journal.pone.0128550

Zhang J, Wang B, Dong SL, Cao DP, Dong JF, Walker WB, Liu Y, Wang GR.2015b. Antennal transcriptome analysis and comparison of chemosensory gene families in two closely related noctuidae moths, *Helicoverpa armigera* and *H.assulta*. *PLoS ONE* 10(2):e0117054 DOI: 10.1371/journal.pone.0117054

Zhao Y, Wang F, Zhang X, Zhang S, Guo S,Zhu G, Liu Q, Li M. 2016. Transcriptome and Expression Patterns of Chemosensory Genes in Antennae of the Parasitoid Wasp *Chouioia cunea*. *PLoS ONE* 11(2):e0148159 DOI:10.1371/journal.pone.014815911

Zhou KY, Zhan QQ, Lu LL, Yang Y, Liu YY, Gan BC. 2016. Toxicity and efficacy evaluation of six biopesticides against *Heortia vitessoides* larvae.*Forest Pest and Disease* 1:31-33 DOI:10.3969/j.issn.1671-0886.2016.01.009

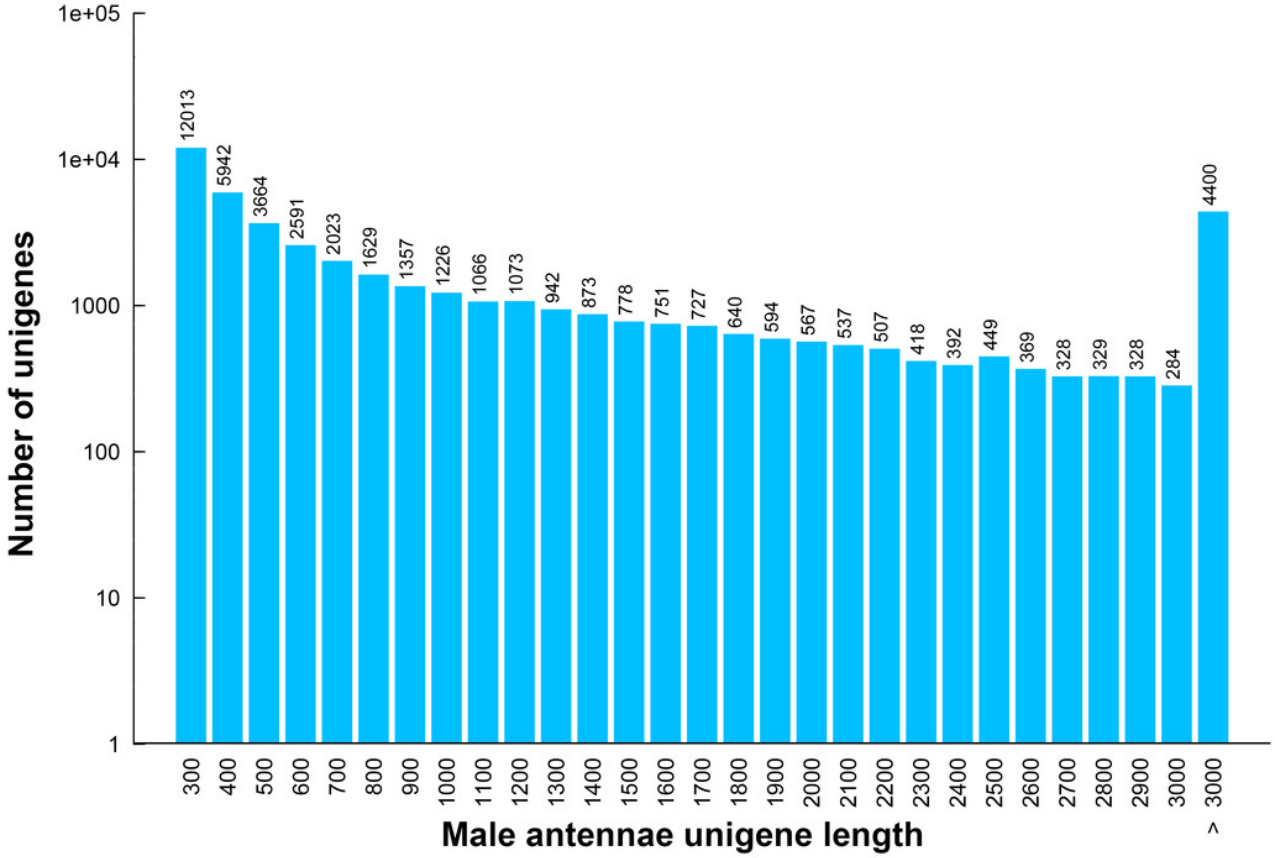
Zhou CX, Min SF, Yan-Long T, Wang MQ.2015. Analysis of antennal transcriptome and odorant binding protein expression profiles of the recently identified parasitoid wasp,*Sclerodermus sp*.*Comparative Biochemistry and Physiology Part D* 16:10-19 DOI: 10.1016/j.cbd.2015.06.003

Zhang YN, Jin JY, Jin R, Xia YH, Zhou JJ, Deng JY, Dong SL. 2013. Differential Expression Patterns in Chemosensory and Non-Chemosensory Tissues of Putative Chemosensory Genes Identified by Transcriptome Analysis of Insect Pest the Purple Stem Borer *Sesamia inferens* (Walker). *PLoS ONE* 8:e69715 DOI: 10.1371/journal.pone.0069715

Figure 1

The size distribution of the assembled unigenes from *H. vitessoides* male and female antennal transcriptomes.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3433v1 | CC BY 4.0 Open Access | rec: 26 Nov 2017, publ: 26 Nov 2017


Number of unigenes
1e+05
1e+04
1000
100
10
1
12013
5942
3664
2591
2023
1629
1357
1226
1066
1073
942
873
778
751
727
640
594
567
537
507
418
392
449
369
328
329
328
284
4400
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
2800
2900
3000
> 3000
Male antennae unigene length

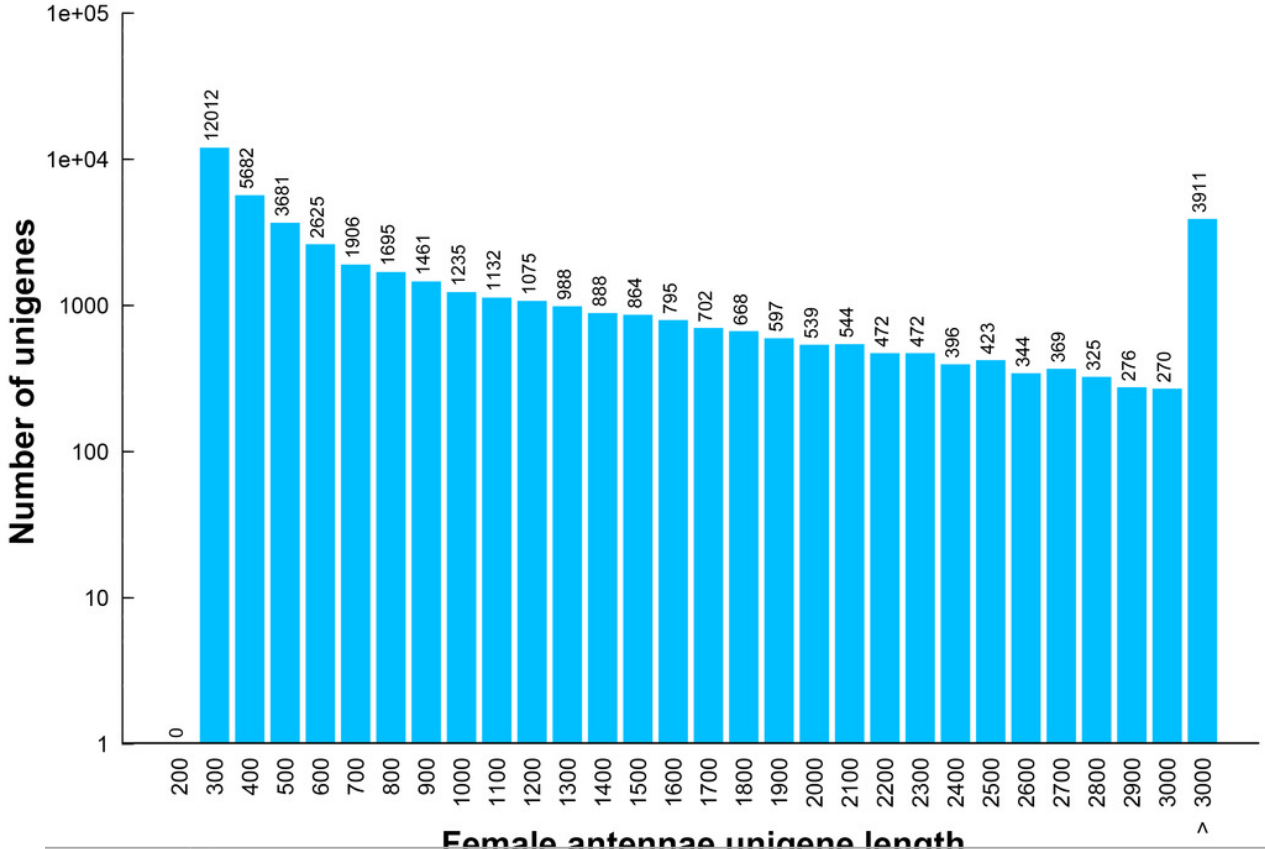

Number of unigenes
1e+05
1e+04
1000
100
10
1
0
12012
5682
3681
2625
1906
1695
1461
1235
1132
1075
988
888
864
795
702
668
597
539
544
472
472
396
423
344
369
325
276
270
3911
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
2800
2900
3000
> 3000
Female antennae unigene length

Figure 2

Species distribution.

Species distribution of the unigene BLASTx matches against the NR protein database with a cut-off E-value of E^{-5}. The first hit of each unigene was used for analysis, and the proportion of homologous sequences in each species is shown.

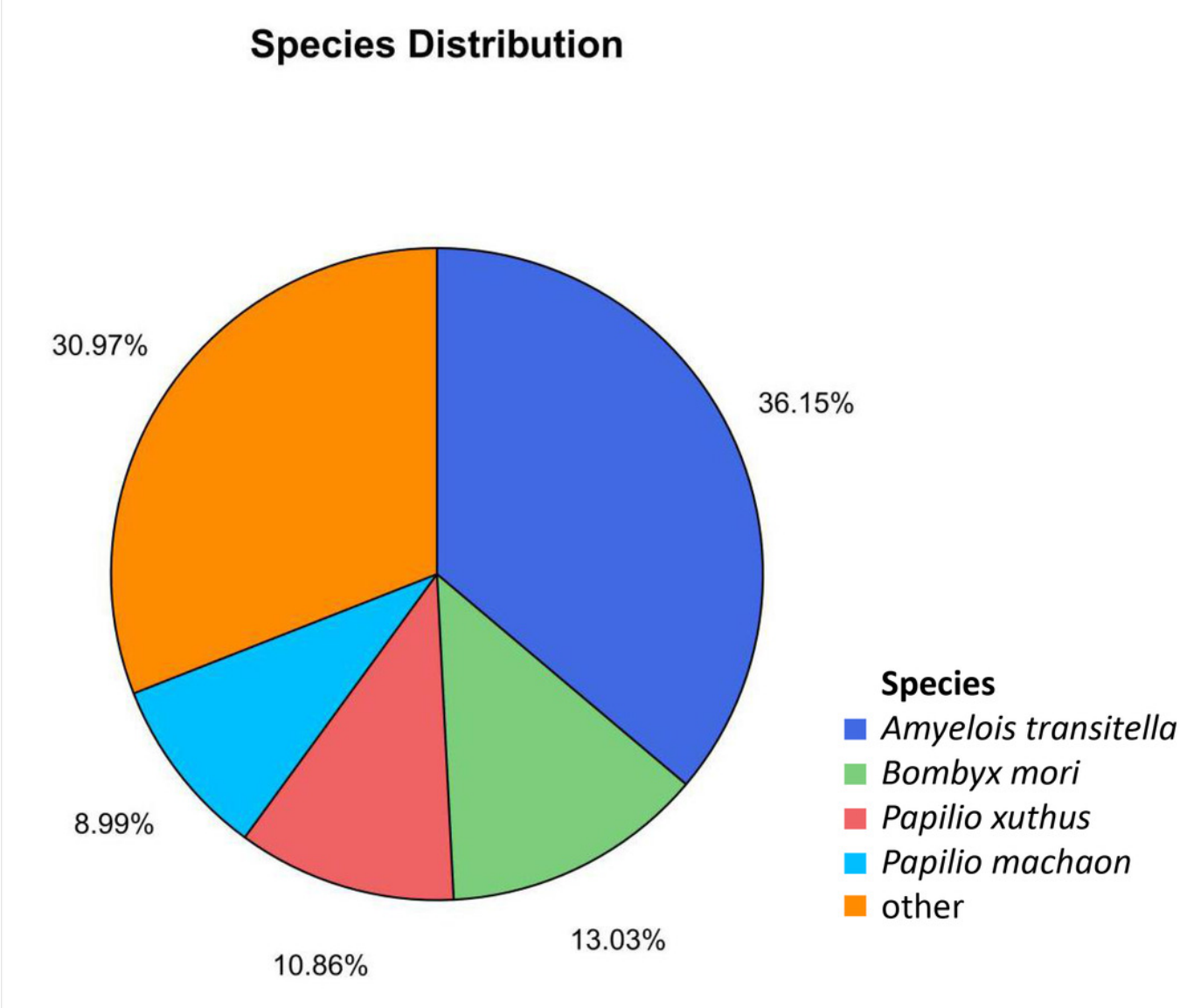


Figure 3

Phylogenetic tree of odorant-binding proteins (OBPs) from lepidopteran insects.

Hvit, *H. vitessoides*; Cmed, *C. medinalis*; Ofur, *O. furnacalis*; Cpun, *C. punctiferalis*. GenBank accession numbers and amino acid sequences used for the tree are given in Table S1.

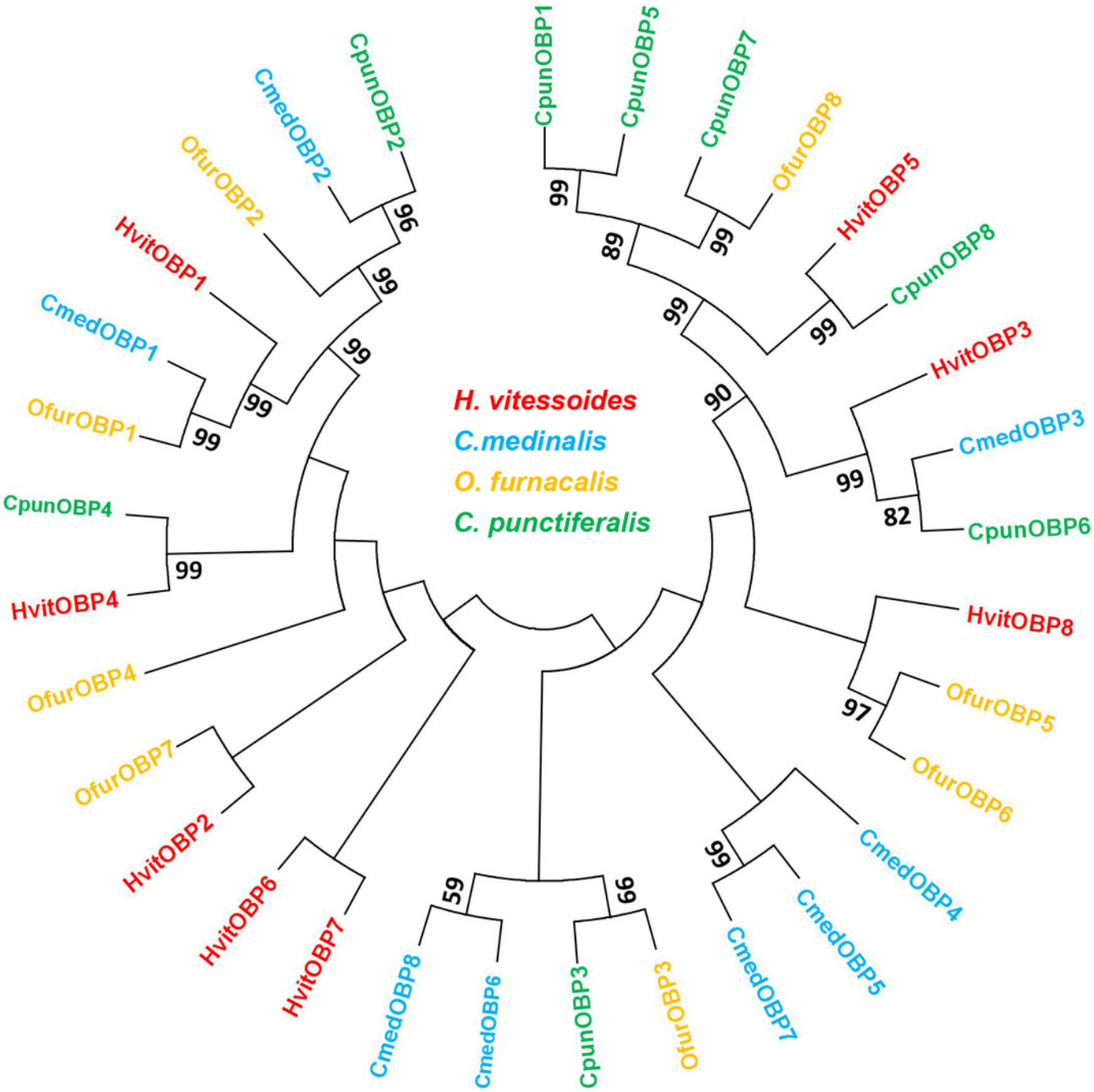


Figure 4

Phylogenetic tree of chemosensory proteins (CSPs) from lepidopteran insects.

Hvit, *H. vitessoides*; Ofur, *O. furnacalis*; Cmed, *C. medinalis*; Bmor, *B. mori*. GenBank accession numbers and amino acid sequences used for the tree are given in Table S1.

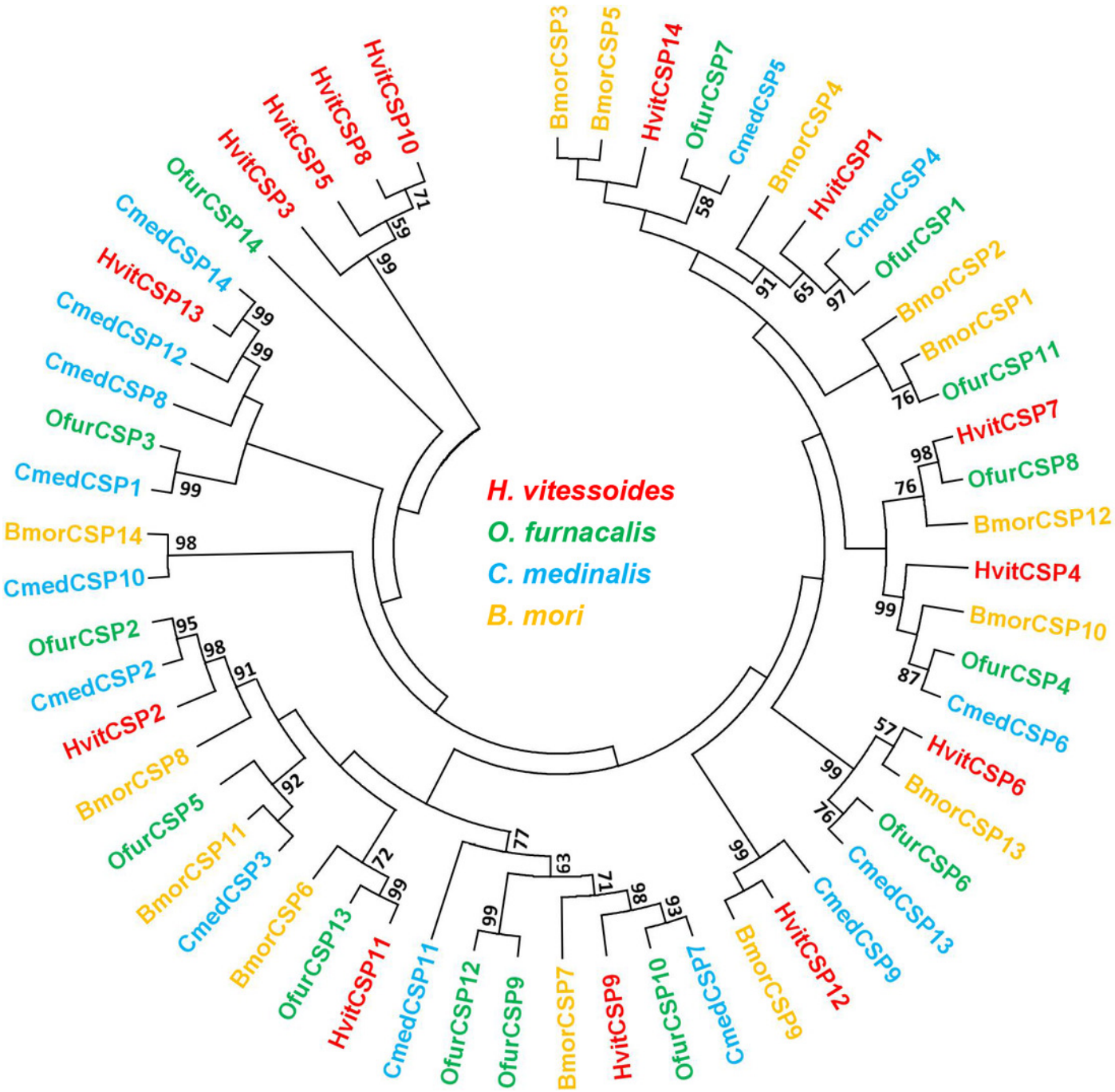


Figure 5

Phylogenetic tree of odorant receptors (ORs) from lepidopteran insects.

The black triangle indicates *Orco/OR2/OR83b.* Hvit, *H. vitessoides*; Cpun, *C. punctiferalis*; Cmed, *C. medinalis*; Hass, *H. assulta*; Harm, *H. armigera*. GenBank accession numbers and amino acid sequences used for the tree are given in Table S1.

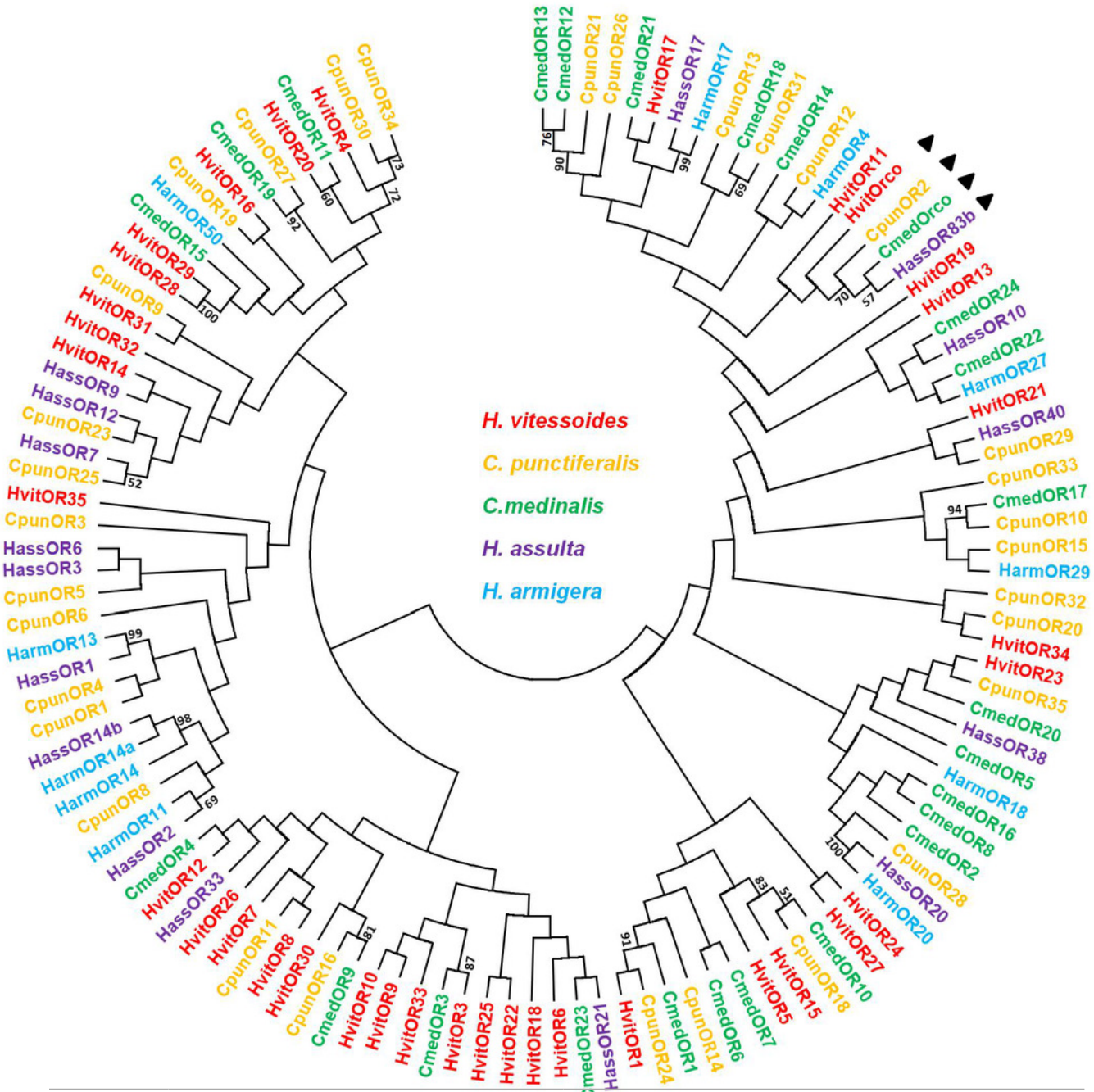


Figure 6

Amino acid sequence alignment of *Orco* from *H. vitessoides* and other insects.

Ofur, *O. furnacalis*; Dind, *D. indica*; Cpun, *C. punctiferalis*; Cmed, *C. medinalis*; Csup, *C. suppressalis*. Blue shading indicates the same sequence among insects. Pink shading indicates amino acids that show 75% identity between sequences. Green shading indicates amino acids that show 50% identity between sequences.

Figure 7

Phylogenetic tree of ionotropic receptors (IRs) from lepidopteran and dipteran insects.

The *IR25a/IR8a* clade is shown. Hvit, *H. vitessoides*; Hass, *H. assulta*; Dhou, *D. houi*; Dmel, *D. melanogaster*; Cpun, *C. punctiferalis*. GenBank accession numbers and amino acid sequences used for the tree are given in Table S1.

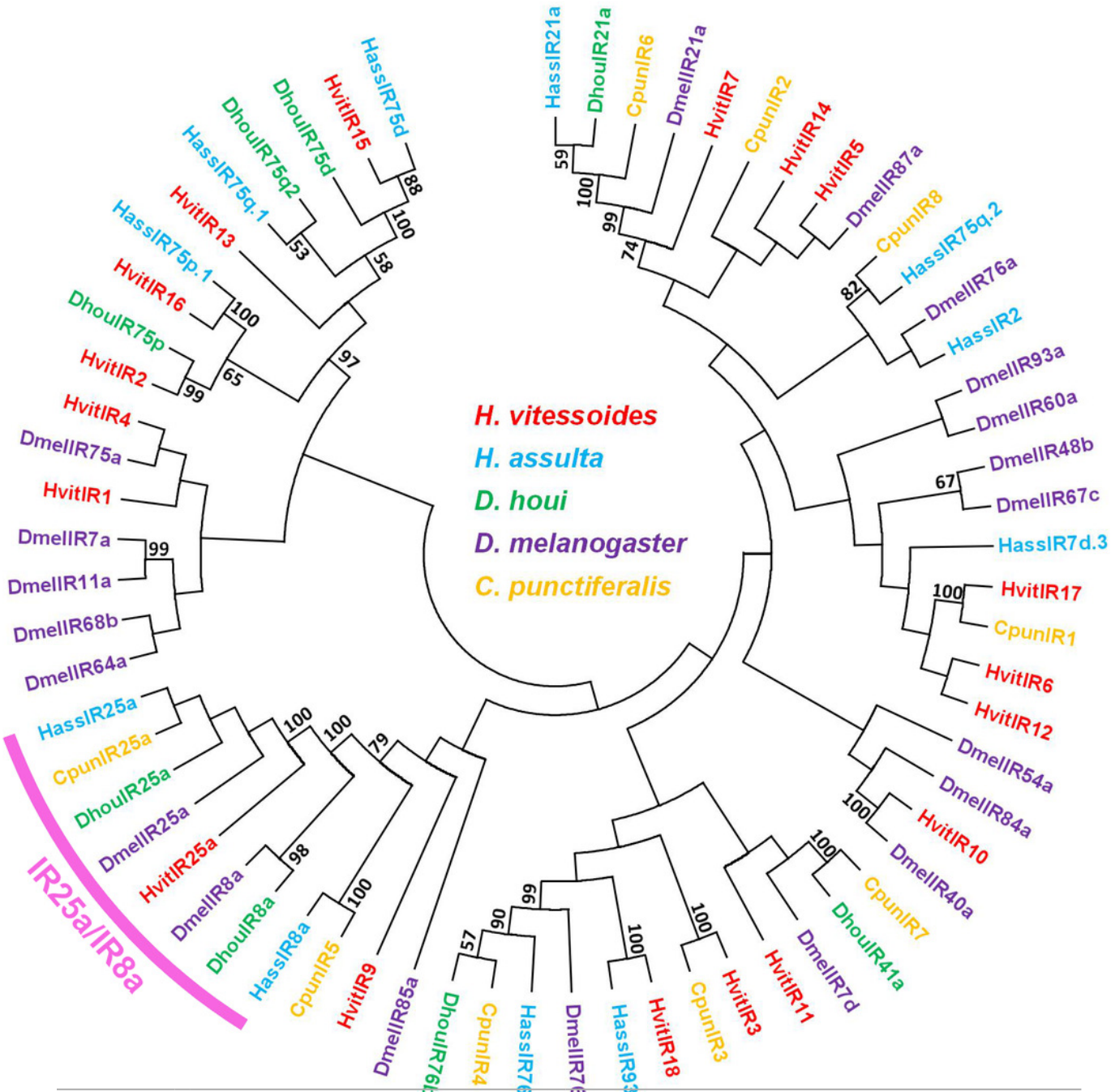


Figure 8

Amino acid sequence alignment of *IR25a* from *H. vitessoides* and other insects.

Cpun, *C. punctiferalis*; Cpom, *C. pomonella*; Hnub, *H. nubiferana*; Bmor, *B. mori*; Hass, *H. assulta*. Blue shading indicates the same sequence among insects. Pink shading indicates amino acids that show 75% identity between sequences. Green shading indicates amino acids that show 50% identity between sequences.

Figure 9

Expression abundance of eight odorant-binding proteins (OBPs) in the *H. vitessoides* antennal transcriptome dataset.

The gene expression abundance is expressed as the fragments per kilobase per million mapped fragments (FPKM) values.

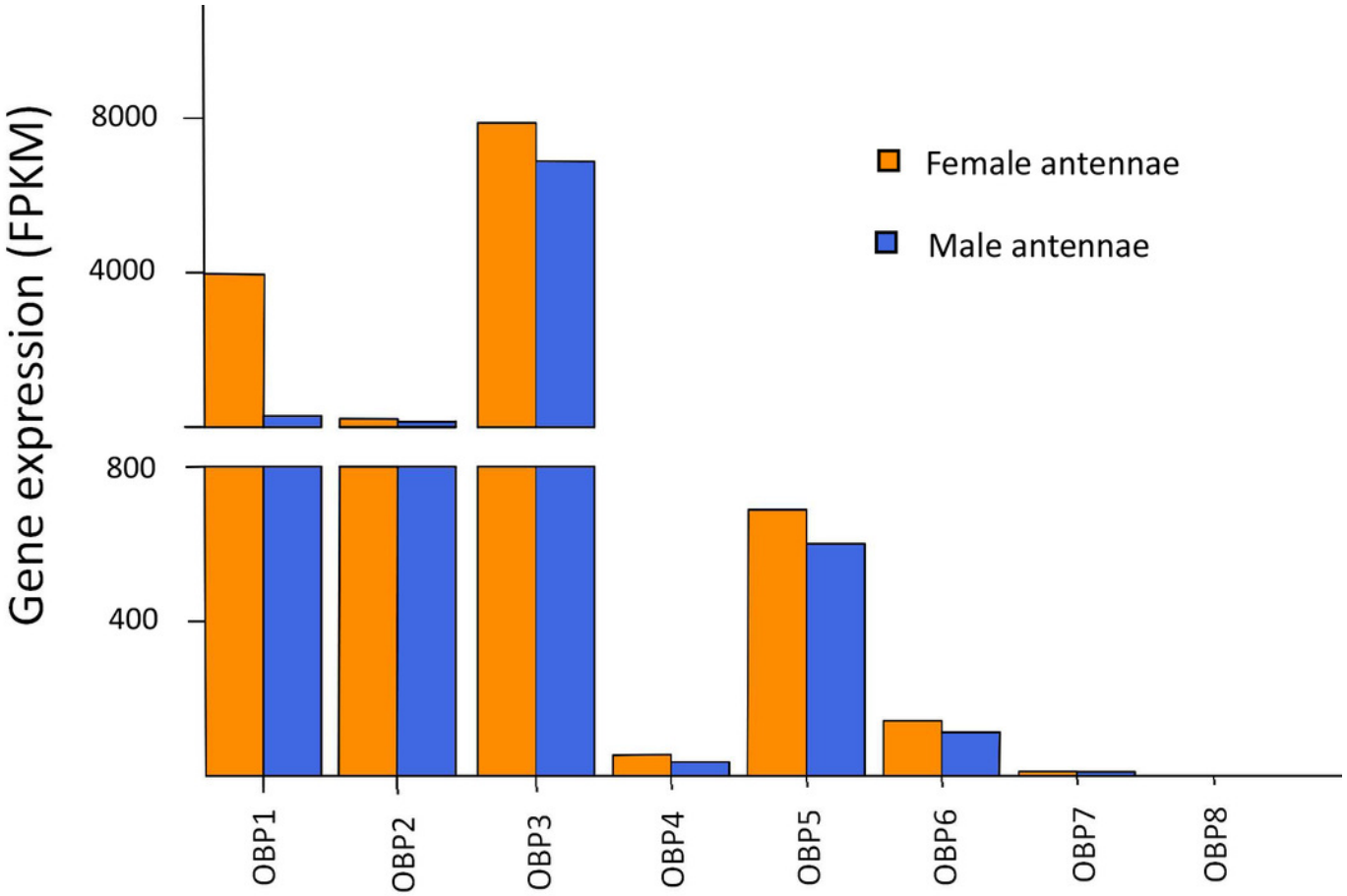


Figure 10

Expression abundance of 14 chemosensory protein (CSPs) in the *H. vitessoides* antennal transcriptome dataset.

The gene expression abundance is expressed as the fragments per kilobase per million mapped fragments (FPKM) values.

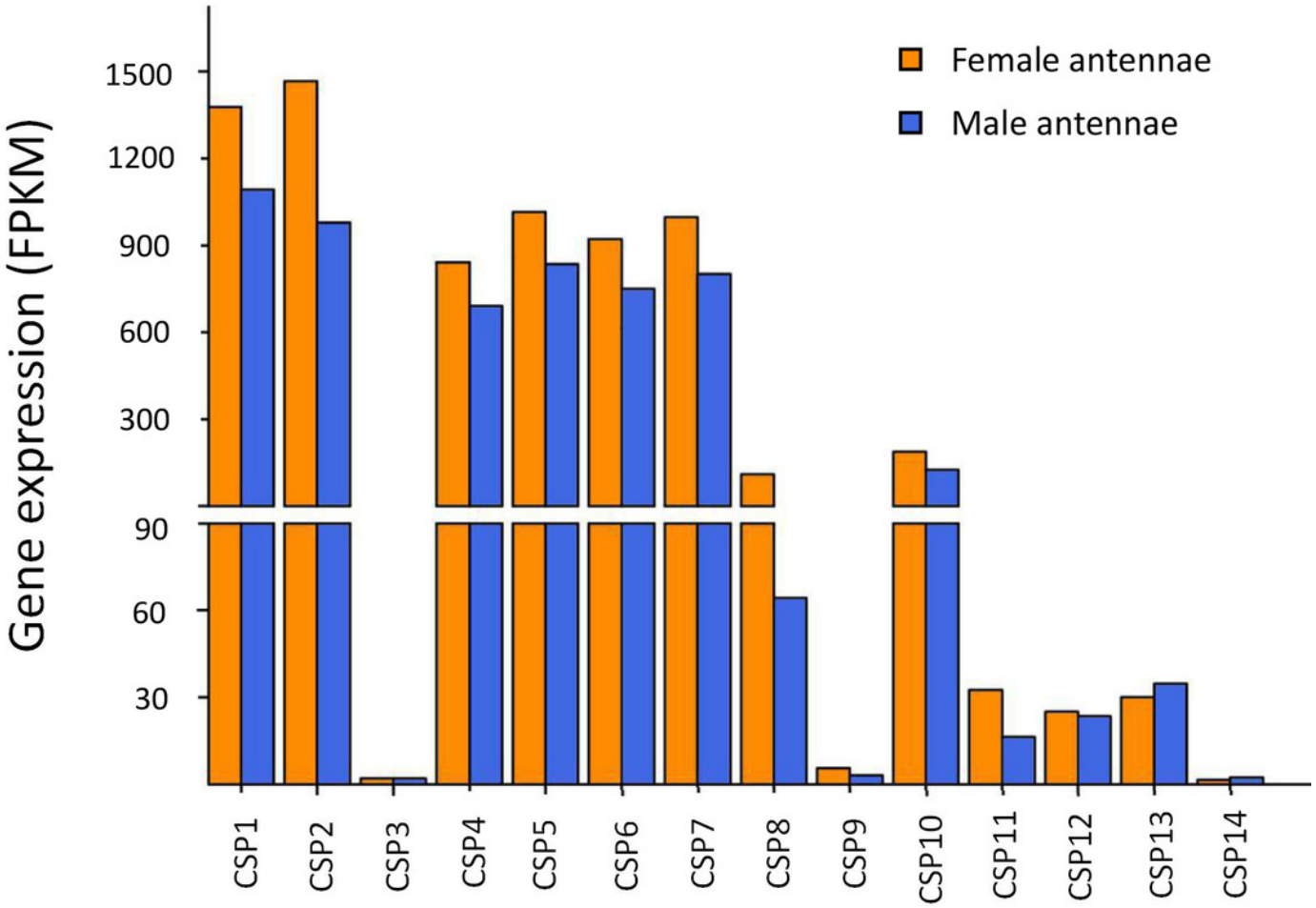


Figure 11

Expression abundance of 35 odorant receptors (ORs) in the *H. vitessoides* antennal transcriptome dataset.

The gene expression abundance is indicated as the fragments per kilobase per million mapped fragments (FPKM) values.

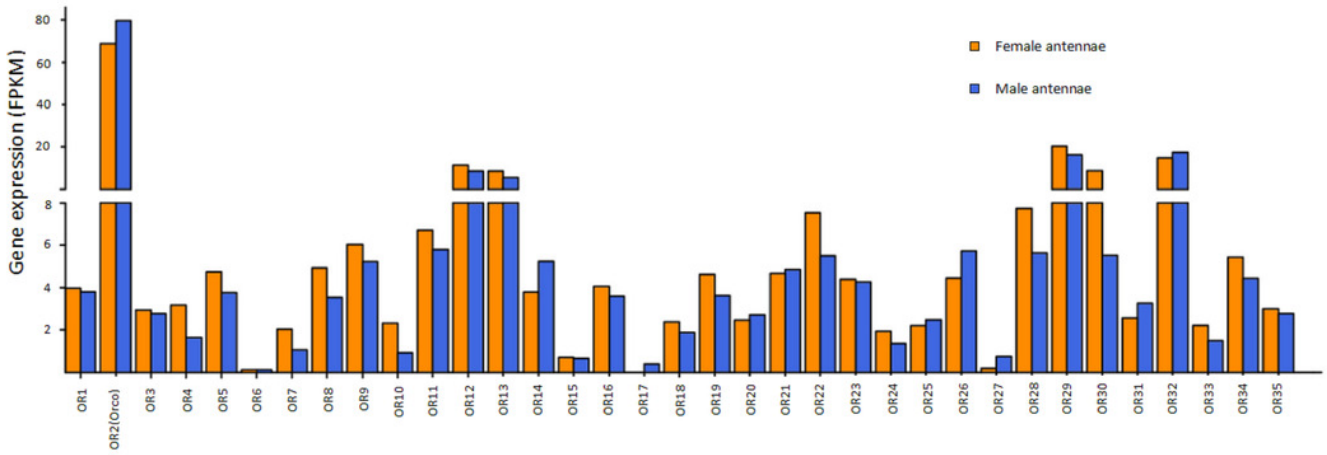


Figure 12

Expression abundance of 18 ionotropic receptors (IRs) in the *H. vitessoides* antennal transcriptome dataset.

The gene expression abundance is expressed as the fragments per kilobase per million mapped fragments (FPKM) values.

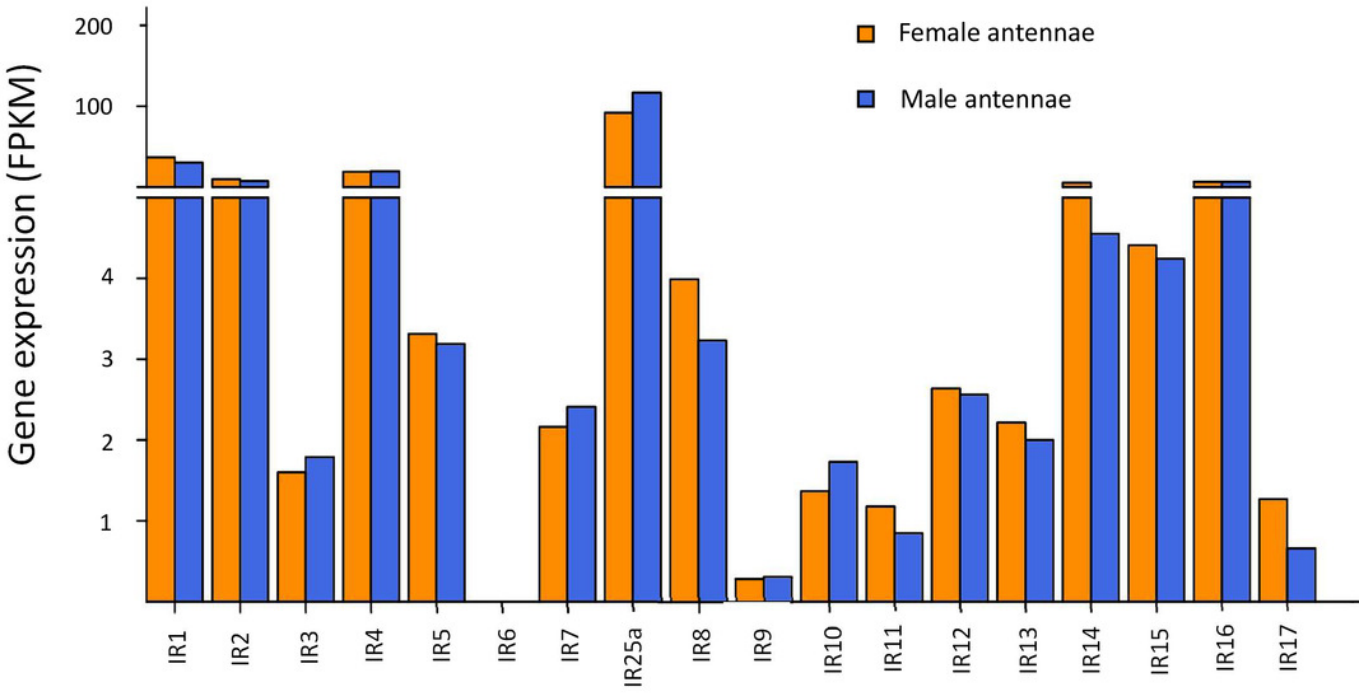


Table 1(on next page)

Summary of assembled unigenes.

Samples	FA-Unigene	MA-Unigene	ALL-Unigene
Total number	46,347	46,797	52,383
Total length	51,246,120	53,639,720	64,883,116
Mean length	1,105	1,146	1,238
N50	2,084	2,230	2,454
GC%	41.83	41.57	41.64

1 FA: Female antennae

2 MA: Male antennae

3 ALL: Merged result

4

Table 2(on next page)

Summary of annotations of unigenes.

Values	Number of unigenes	Percentage (%)
Total	52,383	100%
Nr	24,805	47.35%
Nt	13,981	26.69%
Swiss-Prot	17,772	33.93%
KEGG	18,917	36.11%
KOG	17,333	33.09%
Interpro	17,384	33.19%
GO	2,346	4.48%
Intersection	1,275	2.43%
Overall	27,023	51.59%

1 Intersection: The number of unigenes that were annotated by all seven functional databases.

2 Overall: The number of unigenes that were annotated by any of the functional databases.

3

Table 3(on next page)

Candidate OBPs identified in *H. vitessoides* antennae.

Gene ID	Gene name	Top BLASTx hit				
		Description	E-value	%ID	Species	Acc.no
CL5240.Contig1	OBP1	General odorant-binding protein 1	2e-76	76%	*Chilo suppressalis*	EU825766.1
CL257.Contig1	OBP2	General odorant-binding protein 2	6e-120	82%	*Conogethes punctiferalis*	KT983812.1
CL4797.Contig1	OBP3	General odorant-binding protein 3	2e-103	83%	*Cnaphalocrocis medinalis*	KC507179.1
Unigene4790	OBP4	Odorant-binding protein 4	2e-69	74%	*Conogethes punctiferalis*	KP985222.1
CL5501.Contig1	OBP5	Odorant-binding protein 8	9e-96	79%	*Conogethes punctiferalis*	KP985226.1
CL1934.Contig1	OBP6	Odorant-binding protein 10	4e-84	71%	*Ostrinia furnacalis*	LC027694.1
Unigene10057	OBP7	Odorant-binding protein 13	4e-83	77%	*Ostrinia furnacalis*	LC027697.1
Unigene34601	OBP8	Putative odorant binding protein 18	4e-56	73%	*Nasonia vitripennis*	HE578203.1

1

Table 4(on next page)

Candidate CSPs identified in *H. vitessoides* antennae.

Gene ID	Gene name	Top BLASTx hit				
		Description	E-value	%ID	Species	Acc.no
CL3548.Contig1	CSP1	Chemosensory protein 1	1e-80	77%	*Ostrinia furnacalis*	LC027702.1
Unigene626	CSP2	Chemosensory protein 2	2e-73	77%	*Ostrinia furnacalis*	LC027703.1
Unigene34610	CSP3	Chemosensory protein 3	4e-49	73%	*Athetis dissimilis*	KT357397.1
CL1524.Contig2	CSP4	Chemosensory protein 4	7e-70	76%	*Ostrinia furnacalis*	LC027705.1
Unigene5725	CSP5	Chemosensory protein 5	3e-47	74%	*Ostrinia furnacalis*	LC027706.1
Unigene17488	CSP6	Chemosensory protein 6	1e-49	76%	*Ostrinia furnacalis*	LC027707.1
CL5098.Contig3	CSP7	Chemosensory protein 8	8e-55	76%	*Ostrinia furnacalis*	LC027709.1
Unigene13728	CSP8	Chemosensory protein 9	2e-48	75%	*Helicoverpa armigera*	JX305304.1
CL6043.Contig1	CSP9	Chemosensory protein 10	5e-48	74%	*Ostrinia furnacalis*	LC027711.1
Unigene7854	CSP10	Chemosensory protein 11	1e-63	78%	*Ostrinia furnacalis*	LC027712.1
CL4345.Contig1	CSP11	Chemosensory protein 13	1e-90	80%	*Ostrinia furnacalis*	LC027714.1
Unigene8885	CSP12	Chemosensory protein 15	6e-93	81%	*Ostrinia furnacalis*	LC027716.1
Unigene2777	CSP13	Chemosensory protein 17	4e-61	80%	*Ostrinia furnacalis*	LC027718.1
Unigene14958	CSP14	Chemosensory protein 19	6e-47	74%	*Ostrinia furnacalis*	LC027720.1

1

Table 5(on next page)

Candidate ORs identified in *H. vitessoides* antennae.

Gene ID	Gene name	Top BLASTx hit				
		Description	E-value	%ID	species	Acc.no
CL495.Contig1	OR1	Odorant receptor	0.0	74%	*Cnaphalocrocis medinalis*	KP975136.1
CL1654.Contig2	OR2(Orco)	Odorant receptor co-receptor (Orco)	0.0	80%	*Cnaphalocrocis medinalis*	KP975160.1
CL1046.Contig8	OR3	Odorant receptor	3e-50	70%	*Cnaphalocrocis medinalis*	KP975138.1
CL6305.Contig1	OR4	Odorant receptor	0.0	75%	*Helicoverpa armigera*	XM_021330911.1
Unigene19900	OR5	Odorant receptor	5e-79	67%	*Eogystia hippophaecolus*	KX655960.1
CL6415.Contig2	OR6	Odorant receptor	1e-101	85%	*Ostrinia nubilalis*	AB597006.1
Unigene6842	OR7	Putative olfactory receptor 10	1e-70	70%	*Ostrinia furnacalis*	LC002704.1
CL2882.Contig2	OR8	Putative olfactory receptor 11	1e-97	68%	*Ostrinia furnacalis*	LC002705.1
Unigene20257	OR9	Putative olfactory receptor 12	5e-141	70%	*Ostrinia furnacalis*	LC002706.1
CL6863.Contig1	OR10	Odorant receptor 13	4e-124	69%	*Conogethes punctiferalis*	KX084464.1
Unigene15674	OR11	Odorant receptor 14	5e-163	71%	*Conogethes punctiferalis*	KX084465.1
Unigene20835	OR12	Odorant receptor 15	5e-134	70%	*Conogethes punctiferalis*	KX084466.1
CL207.Contig2	OR13	Odorant receptor 16	1e-69	67%	*Athetis dissimilis*	KR935713.1
CL215.Contig2	OR14	Olfactory receptor 17	1e-48	65%	*Manduca sexta*	LN885111.1
CL783.Contig1	OR15	Odorant receptor 18	2e-74	76%	*Conogethes punctiferalis*	KX084469.1
Unigene18369	OR16	Odorant receptor	2e-50	68%	*Eogystia hippophaecolus*	KX655975.1
CL82.Contig1	OR17	Putative olfactory receptor 21	2e-144	78%	*Ostrinia furnacalis*	LC002715.1
CL4622.Contig2	OR18	Putative olfactory receptor 22	5e-149	71%	*Ostrinia furnacalis*	LC002716.1
Unigene18334	OR19	Odorant receptor 25	1e-92	70%	*Conogethes punctiferalis*	KX084475.1
Unigene4128	OR20	Putative olfactory receptor 26	2e-149	71%	*Ostrinia furnacalis*	LC002720.1
Unigene1688	OR21	Putative olfactory receptor 27	0.0	74%	*Ostrinia furnacalis*	LC002721.1

Unigene18315	OR22	Putative olfactory receptor 33	3e-158	71%	*Ostrinia furnacalis*	LC002727.1
Unigene10517	OR23	Odorant receptor 35	0.0	77%	*Conogethes punctiferalis*	KX084485.1
CL3232.Contig1	OR24	Putative olfactory receptor 36	3e-138	69%	*Ostrinia furnacalis*	LC002730.1
Unigene2995	OR25	Putative olfactory receptor 37	1e-157	72%	*Ostrinia furnacalis*	LC002731.1
Unigene16302	OR26	Putative olfactory receptor 42	8e-107	69%	*Ostrinia furnacalis*	LC002736.1
CL2190.Contig1	OR27	Putative olfactory receptor 44	4e-78	74%	*Ostrinia furnacalis*	LC002738.1
CL3089.Contig1	OR28	Putative olfactory receptor 46	6e-134	75%	*Ostrinia furnacalis*	LC002740.1
CL1046.Contig7	OR29	Odorant receptor 49	6e-79	67%	*Conogethes punctiferalis*	KX084499.1
Unigene4829	OR30	Odorant receptor 50	5e-73	66%	*Conogethes punctiferalis*	KX084500.1
Unigene8713	OR31	Odorant receptor 51	1e-59	67%	*Conogethes punctiferalis*	KX084501.1
CL4619.Contig2	OR32	Putative olfactory receptor 53	6e-53	67%	*Ostrinia furnacalis*	LC002747.1
Unigene7544	OR33	Odorant receptor 54	1e-47	67%	*Conogethes punctiferalis*	KX084504.1
Unigene13831	OR34	Odorant receptor 55	3e-79	66%	*Conogethes punctiferalis*	KX084505.1
Unigene11077	OR35	Odorant receptor 60	4e-70	72%	*Athetis dissimilis*	KR935751.1

1

Table 6(on next page)

Candidate IRs identified in *H. vitessoides* antennae.

Gene ID	Gene name	Top BLASTx hit				
		Description	E-value	%ID	Species	Acc.no
CL869.Contig2	IR1	Ionotropic receptor 1	0.0	76%	*Cnaphalocrocis medinalis*	KP975100.1
CL245.Contig1	IR2	Ionotropic receptor 2	2e-97	70%	*Conogethes punctiferalis*	KP975101.1
Unigene3557	IR3	Ionotropic receptor 3	0.0	74%	*Conogethes punctiferalis*	KX084511.1
Unigene14578	IR4	Ionotropic receptor 4	0.0	71%	*Conogethes punctiferalis*	KX084512.1
CL144.Contig2	IR5	Ionotropic receptor 7	9e-37	67%	*Conogethes punctiferalis*	KX084515.1
Unigene36533	IR6	Ionotropic receptor 9	2e-90	80%	*Conogethes punctiferalis*	KX096209.1
CL5500.Contig1	IR7	Ionotropic receptor	1e-96	76%	*Ostrinia furnacalis*	LC017783.1
Unigene9996	IR25a	Ionotropic receptor 25a	0.0	81%	*Conogethes punctiferalis*	KX084508.1
CL6851.Contig1	IR8	Ionotropic receptor	9e-80	68%	*Eogystia hippophaecolus*	KX655900.1
Unigene12388	IR9	Ionotropic receptor	3e-88	77%	*Ostrinia furnacalis*	LC017785.1
Unigene20350	IR10	Ionotropic receptor	3e-49	71%	*Eogystia hippophaecolus*	KX655897.1
Unigene13651	IR11	Ionotropic receptor	6e-51	67%	*Helicoverpa armigera*	KF768725.1
CL5259.Contig1	IR12	Ionotropic receptor	1e-171	73%	*Ostrinia furnacalis*	LC017787.1
CL1638.Contig1	IR13	Ionotropic receptor	9e-170	78%	*Ostrinia furnacalis*	LC017789.1
CL647.Contig1	IR14	Ionotropic receptor 75d	0.0	74%	*Athetis dissimilis*	KR912017.1
Unigene13775	IR15	Ionotropic receptor	0.0	77%	*Ostrinia furnacalis*	LC017793.1
Unigene6992	IR16	Ionotropic receptor	0.0	80%	*Ostrinia furnacalis*	KX084509.1
CL2367.Contig1	IR17	Ionotropic receptor	0.0	74%	*Ostrinia furnacalis*	LC017797.1

1