

OPEN

Identification of candidate chemosensory genes of *Ophraella communa* LeSage (Coleoptera: Chrysomelidae) based on antennal transcriptome analysis

Chao Ma¹, Chenchen Zhao¹, Shaowei Cui^{1,2}, Yan Zhang¹, Guangmei Chen¹, Hongsong Chen^{1,3}, Fanghao Wan¹ & Zhongshi Zhou^{1*}

Antennal olfaction plays a key role in insect survival, which mediates important behaviors like host search, mate choice, and oviposition site selection. As an oligophagous insect, olfaction is extremely important for *Ophraella communa* to locate host plants. However, information on the olfactory genes has been lacking in *O. communa*. Using next generation sequencing, we assembled the antennal transcriptome of *O. communa* and first reported the major chemosensory genes necessary for olfaction in this species. In this study, a total 105 candidate chemosensory genes were identified in *O. communa* antennae, including 25 odorant-binding proteins (OBPs), 11 chemosensory proteins (CSPs), four sensory neuron membrane proteins (SNMPs), 30 odorant receptors (ORs), 18 ionotropic receptors (IRs), and 17 gustatory receptors (GRs). We also identified full-length sequences of the highly conserved ORco and IR8a/25a family in *O. communa*. In addition, the expression profile of 15 ORs and four OBPs were validated by quantitative real-time polymerase chain reaction (qPCR). We found that *OcomOR2/4/19* and *OcomOBP19/20* had a biased expression in male antennae, and *OcomOR8* had a biased expression in the female antennae. This large number of chemosensory genes handled by homology analysis and qPCR results will provide the first insights into molecular basis for the olfactory systems of *O. communa* as well as advance our understanding of olfactory mechanisms in Coleoptera.

The olfactory system is very important to insects because it is involved in various insect behaviors, such as locating suitable hosts, avoiding predators, identifying oviposition sites, and finding sexual partners¹. The antenna is the major organ for insect olfactory sensing, especially for olfaction. Mounting evidence suggests that diverse olfactory genes are involved in the signal recognition process, including odorant receptors (ORs), ionotropic receptors (IRs), odorant binding proteins (OBPs), gustatory receptors (GRs), chemosensory proteins (CSPs), and sensory neuron membrane proteins (SNMPs)^{2–4}. Olfactory signal transduction can be summarized as follows: Firstly, the hydrophobic chemical compounds penetrate the sensillar lymph through pores, wherein they are recognized and bound by OBPs^{5,6} or CSPs⁷. Secondly, it was speculated that the OBPs or CSPs were the transporters that transferred odorants through the sensillar lymph to ORs, a family of integral membrane proteins, located on the dendrites of olfactory receptor neurons (ORNs)^{8–10}. Additionally, SNMPs¹¹, IRs^{12,13} and GRs¹⁴ have also been proposed to play a role in insect olfaction.

Although insect ORs are seven-transmembrane domain (TMD) proteins with a reversed membrane topology (intracellular N-terminus)^{15,16}, they do not belong to the G protein-coupled receptors. In the transduction process, ORs appear to be the primary mechanism by which insects detect volatile chemicals, facilitating the conversion of the chemical message to an electrical signal, such as a biological transducer^{17,18}. It is generally thought that each ORN expresses a highly conserved OR co-receptor (Orco protein) and a divergent, conventional ORx, such

¹State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, China. ²School of Plant Protection, Jilin Agricultural University, Changchun, 130118, China. ³Guangxi Key Laboratory for Biology of Crop Diseases and Insect Pests, Institute of Plant Protection, Guangxi Academy of Agricultural Sciences, Nanning, 530007, China. *email: zhongshizhou@yahoo.com

| Transcripts | Total Number |
|---------------------------------|----------------|
| Female antenna-Raw Reads (nt) | 34,825,000 |
| Male antenna-Raw Reads (nt) | 31,106,518 |
| Female antenna-Clean reads (nt) | 32,944,204 |
| Male antenna-Clean reads (nt) | 29,369,754 |
| Female antenna-Clean bases (Gb) | 4.9 |
| Male antenna-Clean bases (Gb) | 4.4 |
| Total transcripts | 153,276 |
| Total Unigene | 92,259 |
| N50 of unigenes (nt) | 2,068 |
| Mean length (nt) | 1,229 |
| Unigenes larger than 1,000 nt | 35,508 (38.5%) |
| Unigenes with homolog in nr | 43,779 (47.5%) |
| Unigenes annotated to GO term | 33,258 (36.0%) |

Table 1. Assembly summary of *O. communa* antennal transcriptome. nt = nucleotides, Gb = gigabases.

that the heterodimer of the Orco-OR forms an ion channel and mediates odorant-binding specificity^{19–21}. ORs are broadly tuned to a variety of volatile chemicals, including pheromones, plant volatiles, and odor molecules present in the environment^{17,22,23}.

Coleoptera species account for approximately 25% of all known species of animal life-forms²⁴. Almost 40% of all previously described insect species are beetles²⁵. However, compared with Lepidoptera, olfaction genes identified in Coleoptera are poorly known. To date, there are only about 20 species of Coleoptera for which olfactory genes have been identified, such as *Tribolium castaneum*^{25,26}, *Megacyllene caryae*²⁷, *Leptinotarsa decemlineata*²⁸, *Phyllotreta striolata*²⁹, *Colaphellus bowringi*³⁰, *Pyrrhalta maculicollis*³¹, *P. aenescens*³¹, *Ambrostoma quadriimpressum*³² and *Galeruca daurica*³³. Thus, much work is needed to investigate and better understand olfaction and its associated molecular biology in other species of Coleoptera. *Ophraella communa* LeSage (Coleoptera: Chrysomelidae) originated in North America, and it is considered a potential biological control agent of common ragweed, *Ambrosia artemisiifolia* L. (Asteraceae)³⁴. Both adults and larvae feed on the leaves of common ragweed, resulting in severe defoliation³⁵. Since the beetle was first discovered in Nanjing, Jiangsu province (China) in 2001³⁶, it has been reported widely in eastern and central China, where it has significantly suppressed the population of common ragweed³⁷. Zhou *et al.*³⁸ reported that when the olfaction of male *O. communa* was hindered by covering their antennae with paint, the males spent significantly more time seeking mates in the arena, indicating that olfaction is important to the mating process. However, the molecular mechanism of olfaction recognition in this insect is still unknown. In this study, we performed a transcriptome analysis of the adult antennae of *O. communa* and identified 105 candidate chemosensory genes, including 30 ORs, 25 OBPs, 11 CSPs, 18 IRs, 17 GRs, and four SNMPs. Furthermore, we conducted a comprehensive and comparative phylogenetic analysis and examined 19 genes expression profiles using quantitative real-time polymerase chain reaction (qPCR). These results could help us better understand the olfactory signal transduction mechanisms in this insect.

Results

Transcriptome overview. Using an Illumina HiSeq 2000TM platform, a total of 31.1 million and 34.8 million raw reads were yielded, respectively, from the libraries of male and female antennae. After removing low-quality and adaptor reads, 29.4 million and 32.9 million clean-reads were generated. The total bases of sequence data were approximately 4.4 and 4.9 gigabases from male and female samples, respectively. Overall, 153,276 transcripts were generated, and we identified 92,259 unigenes by clustering and redundancy filtering. The mean length of unigenes was 1,229 nt and the N50 length reached 2,068 nt. In total, 35,508 unigenes were larger than 1,000 nt in length, which comprised 38.5% of all unigenes (Table 1). Homology searches of all unigenes with respect to other insect species showed that the highest percentage of unigenes matched *T. castaneum* (47.5%), followed by *Dendroctonus ponderosae* (12.8%), *Lasius niger* (3.7%), *Acyrtosiphon pisum* (3.2%), and *Plutella xylostella* (2.3%). The remaining 30.7% of the sequences showed similarity with the sequences of other insects (Fig. 1).

OBPs. We identified 25 different sequences encoding odorant binding proteins in *O. communa* antennal transcriptomes. Sequence analysis results showed that 20 unigenes had a putative full-length open reading frame (ORF) and 19 unigenes had predicted signal peptide sequences. All of the candidate OBPs sequences Blastx best hits were similar to known Coleoptera OBPs. The length of all putative full-length *OcomOBPs* ranged from 119 to 198 amino acids. Compared to ORs, insect OBPs were highly conserved. Twenty-one of 25 putative OBPs had more than 50% similarity with OBPs from *G. daurica*, *P. maculicollis*, and *P. aenescens*. Based on phylogenetic analysis, *OcomOBPs* were split in various branches and they formed small subgroups together with OBPs from other beetles. These groups were strongly supported by high bootstrap values. Remarkably, we found *OcomOBP19*, a pheromone binding protein (PBP), which clustered with other Coleoptera PBPs in a clade (Fig. 2). Information, including unigene reference, length, and best Blastx hit for all 25 OBPs are listed in Table 2.

CSPs. We identified 11 unigenes encoding candidate chemosensory proteins in *O. communa* antennal transcriptome. Notably, all putative chemosensory proteins were predicted with a putative full-length ORF and signal

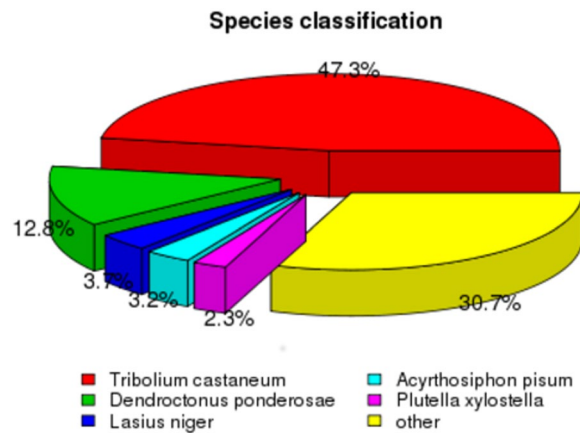


Figure 1. All unigenes sequences (92,259) that had blast annotations against the nr database with a cut-off E-value 10^{-5} were analyzed for species distribution.

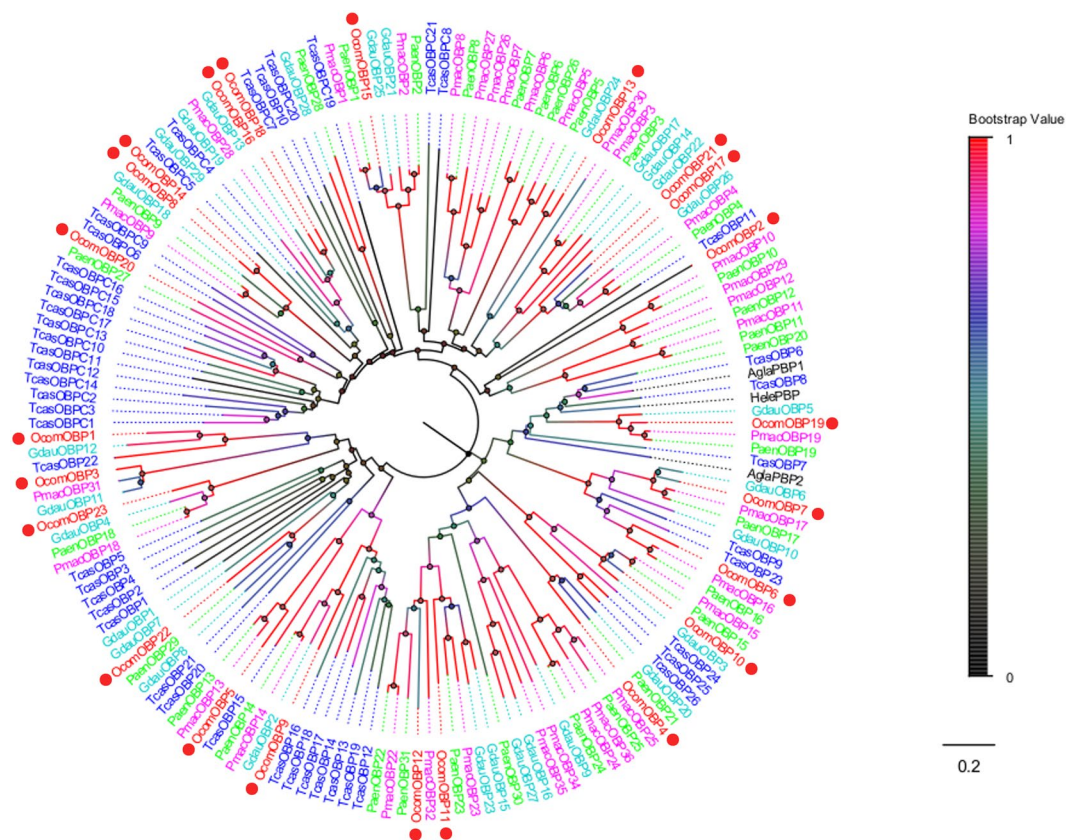


Figure 2. Neighbor joining phylogenetic tree of candidate *OcomOBPs* with known Coleopteran OBP sequences. *Tcas*, *Tribolium castaneum* (N = 47); *Pmac*, *Pyrrhalta maculicollis* (N = 33); *Paen*, *Pyrrhalta aenescens* (N = 31); *Gdua*, *Galeruca daurica* (N = 29); *Agla*, *Anoplophora glabripennis* (N = 2); *Hele*, *Hylamorphia elegans* (N = 1). Candidate *OcomOBPs* were indicated by red circles.

peptide through sequence analysis. The length of all putative full-length *OcomCSPs* ranged from 118 to 261 amino acids. In addition, all of the *OcomCSPs* followed the highly conserved pattern with four cysteines arranged with an exact spacing of $C_1X_6C_2X_{18}C_3X_2C_4$ (Fig. 3). Insect CSPs are more conserved than ORs or OBPs, and all *OcomCSPs* amino acid sequences have more than 65% similarity with CSPs from *P. maculicollis*, *P. aenescens*, *G. daurica*, and *C. bowringi*. Homology analysis showed that the *OcomCSPs* were present on different branches throughout the dendrogram and supported by high bootstrap values (Fig. 4). Information, including unigene reference, length, and the best Blastx hit of all 11 CSPs are listed in Supplementary Material S3.

| Unigene ID | NAME | Length | ORF (aa) | Status | Signal peptide | Blastx best-hit | Species | Score | E-value | Ident | Accession |
|--------------------|-----------|--------|----------|--------------|----------------|-------------------------------------|-------------------------------|-------|-----------|-------|------------|
| Cluster-9750.32846 | OcomOBP1 | 1109 | 198 | Complete ORF | Y | odorant-binding protein | <i>Galeruca daurica</i> | 188 | 5.00E-55 | 54% | AQY18976.1 |
| Cluster-9750.2277 | OcomOBP2 | 604 | 192 | 3' missing | Y | odorant binding protein 25 | <i>Colaphellus bowringi</i> | 154 | 3.00E-43 | 45% | ALR72513.1 |
| Cluster-9750.23752 | OcomOBP3 | 1228 | 182 | Complete ORF | Y | odorant-binding protein 31 | <i>Pyrrhalta maculicollis</i> | 312 | 3.00E-103 | 82% | APC94203.1 |
| Cluster-9750.32204 | OcomOBP4 | 593 | 169 | Complete ORF | N | odorant-binding protein 21 | <i>Pyrrhalta maculicollis</i> | 267 | 2.00E-61 | 61% | APC94176.1 |
| Cluster-3339.0 | OcomOBP5 | 551 | 153 | Complete ORF | Y | odorant-binding protein 13, partial | <i>Pyrrhalta aenescens</i> | 232 | 1.00E-75 | 71% | APC94286.1 |
| Cluster-9750.21729 | OcomOBP6 | 738 | 153 | Complete ORF | N | odorant-binding protein 16 | <i>Pyrrhalta aenescens</i> | 219 | 5.00E-69 | 71% | APC94289.1 |
| Cluster-9750.32217 | OcomOBP7 | 644 | 152 | Complete ORF | Y | odorant-binding protein 17 | <i>Pyrrhalta aenescens</i> | 240 | 2.00E-78 | 79% | APC94290.1 |
| Cluster-17495.0 | OcomOBP8 | 551 | 151 | Complete ORF | N | odorant-binding protein 9 | <i>Pyrrhalta aenescens</i> | 177 | 5.00E-54 | 64% | APC94272.1 |
| Cluster-9750.33001 | OcomOBP9 | 657 | 148 | Complete ORF | Y | odorant-binding protein 14 | <i>Pyrrhalta aenescens</i> | 193 | 9.00E-60 | 60% | APC94287.1 |
| Cluster-9750.32206 | OcomOBP10 | 738 | 146 | Complete ORF | N | odorant-binding protein 15 | <i>Pyrrhalta aenescens</i> | 254 | 4.00E-83 | 88% | APC94288.1 |
| Cluster-9853.0 | OcomOBP11 | 628 | 144 | Complete ORF | Y | odorant-binding protein 23 | <i>Pyrrhalta maculicollis</i> | 139 | 1.00E-38 | 48% | APC94180.1 |
| Cluster-1804.1 | OcomOBP12 | 693 | 144 | Complete ORF | Y | odorant-binding protein 22 | <i>Pyrrhalta maculicollis</i> | 122 | 2.00E-31 | 44% | APC94177.1 |
| Cluster-8649.0 | OcomOBP13 | 748 | 141 | Complete ORF | Y | odorant-binding protein | <i>Galeruca daurica</i> | 238 | 6.00E-77 | 79% | AQY18988.1 |
| Cluster-15202.0 | OcomOBP14 | 493 | 141 | Complete ORF | Y | odorant-binding protein 9 | <i>Pyrrhalta maculicollis</i> | 197 | 3.00E-62 | 62% | APC94188.1 |
| Cluster-9750.28158 | OcomOBP15 | 603 | 137 | Complete ORF | Y | odorant-binding protein | <i>Galeruca daurica</i> | 222 | 1.00E-71 | 82% | AQY18989.1 |
| Cluster-9750.32896 | OcomOBP16 | 778 | 135 | Complete ORF | Y | odorant-binding protein 28 | <i>Pyrrhalta maculicollis</i> | 177 | 5.00E-53 | 69% | APC94187.1 |
| Cluster-16599.0 | OcomOBP17 | 418 | 135 | 3' missing | Y | odorant-binding protein | <i>Galeruca daurica</i> | 160 | 6.00E-48 | 62% | AQY18990.1 |
| Cluster-9750.26603 | OcomOBP18 | 412 | 131 | 3' missing | Y | odorant-binding protein 28 | <i>Pyrrhalta maculicollis</i> | 177 | 8.00E-55 | 62% | APC94187.1 |
| Cluster-9750.32590 | OcomOBP19 | 819 | 131 | Complete ORF | Y | odorant-binding protein 19 | <i>Pyrrhalta aenescens</i> | 216 | 5.00E-68 | 87% | APC94292.1 |
| Cluster-9750.34868 | OcomOBP20 | 574 | 130 | Complete ORF | Y | odorant-binding protein 7 | <i>Monochamus alternatus</i> | 142 | 2.00E-19 | 37% | AIX97022.1 |
| Cluster-11640.0 | OcomOBP21 | 401 | 127 | Complete ORF | Y | odorant-binding protein 4 | <i>Pyrrhalta aenescens</i> | 138 | 4.00E-39 | 61% | APC94279.1 |
| Cluster-9750.24950 | OcomOBP22 | 540 | 120 | Complete ORF | Y | odorant-binding protein | <i>Galeruca daurica</i> | 118 | 4.00E-31 | 50% | AQY18972.1 |
| Cluster-9750.33322 | OcomOBP23 | 618 | 119 | Complete ORF | Y | odorant-binding protein 18 | <i>Pyrrhalta aenescens</i> | 166 | 2.00E-49 | 74% | APC94291.1 |
| Cluster-9750.56133 | OcomOBP24 | 513 | 99 | 5' missing | N | odorant-binding protein 28 | <i>Pyrrhalta maculicollis</i> | 132 | 2.00E-36 | 55% | APC94187.1 |
| Cluster-9750.32950 | OcomOBP25 | 562 | 95 | 5' missing | N | odorant-binding protein 21 | <i>Pyrrhalta aenescens</i> | 254 | 2.00E-57 | 66% | APC94263.1 |

Table 2. The Blastx match of *O. communis* candidate candidate odorant binding proteins.

SNMPs. We identified four SNMP genes in the antennal transcriptome (Fig. 5). Lengths of all candidate *OocmSNMPs* were over 500 amino acids and three of them were predicted to have a putative full-length ORF. Furthermore, all *OocmSNMPs* had more than 50% identity with SNMPs of *P. aenescens*, *P. striolata*, and *C. bowringi*. Information, including unigene reference, length, and the best Blastx hit of all four SNMPs are listed in Supplementary Material S3.

ORs. Thirty putative OR transcripts were identified in the *O. communis* antennal transcriptome. *OcomOR1* (*OcomORco*) gene was easily identified because it had an intact ORF and seven transmembrane domains, which were characteristic of typical insects ORs. The amino acid sequences of *OcomOR1* shared 91% identity with the co-receptor of *C. bowringi*. Except for *OcomORco*, 13 putative ORs were predicted to have a full-length ORF, encoding proteins with more than 335 amino acids. The putative *OcomORs* transcripts encoded complete proteins that were predicted to have three to seven transmembrane domains. Ten *OcomORs* were highly divergent and they had low levels of identity (<50%) with other reported insect ORs. Following the phylogenetic analysis, the OR sequences were clustered into several subgroups (Fig. 6). The *OcomOR1* was clustered with other insects ORco containing *PstrORco*, *CbowORco*, and *TcasORco*. In addition, *OcomOR5*, *OcomOR9*, *OcomOR24*, and *OcomOR26* were grouped into the same clade. Interestingly, *OcomOR2* and *OcomOR4* clustered with *McarOR3* and *McarOR5* in the same clade, *OcomOR12* and *OcomOR28* grouped together with *McarOR20*, meaning *OcomOR2*, *OcomOR4*, *OcomOR12*, and *OcomOR28* may play a role in pheromone identity function, because *McarOR3*, *McarOR5*, and *McarOR20* have been demonstrated to be tuned to the male-produced pheromone chemicals of *M. caryae*. Information, including unigene reference, length, and the best Blastx hit of all 30 ORs are listed in Table 3.

GRs. We found 17 candidate GRs transcripts in the *O. communis* antennal transcriptome (Fig. 7). The majority of candidate *OcomGRs* were partial fragments (only four were predicted to have a putative full-length ORF), encoding overlapping but distinct sequences. Eleven *OcomGRs* had more than 50% identity with GRs of *P. aenescens*, *P. striolata*, *C. bowringi*, *Monochamus alternatus*³⁹, *Anomala corpulenta*⁴⁰, and *Anoplophora glabripennis*⁴¹. Information, including unigene reference, length, and the best Blastx hit of all 17 GRs are listed in Supplementary Material S3.

IRs. We identified 18 transcripts encoding candidate IRs in the *O. communis* antennal transcriptome. Of these, eight *OcomIRs* contained a putative full-length ORF, with three to four TMDs. Based on the Blastx results, all *OcomIRs* had high levels of identity (>58%) with other reported insect IRs, indicating IRs were relatively

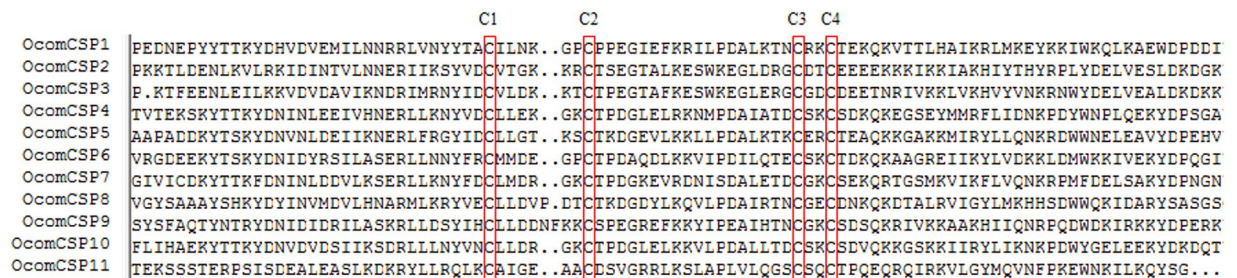


Figure 3. Sequences alignment of candidate *OcomCSPs* amino acid sequences. The conserved cysteine residues were marked with red box.

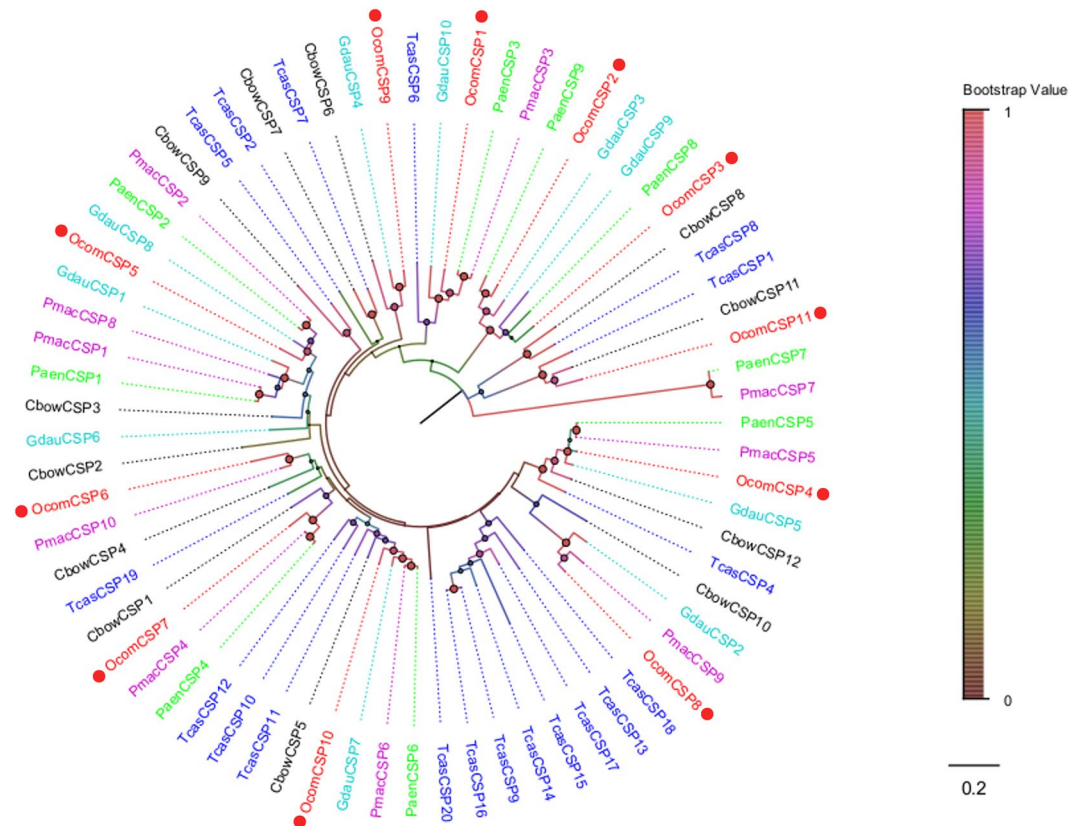


Figure 4. Neighbor joining phylogenetic tree of candidate *OcomCSPs* with known Coleopteran CSP sequences. *Tcas*, *Tribolium castaneum* (N = 19); *Pmac*, *Pyrrhalta maculicollis* (N = 10); *Paen*, *Pyrrhalta aenescens* (N = 9); *Gdau*, *Galeruca daurica* (N = 10); *Cbow*, *Colaphellus bowringi* (N = 12). Candidate *OcomCSPs* were indicated by red circles.

conserved in Coleoptera insects. In the phylogenetic analysis, *OcomIRs* were grouped into different clades with high-level bootstrap values. *OcomIR3* and *OcomIR7* clustered with the IR8a/IR25a clades (including *TcasIR8a*, *TcasIR25a*, *CbowIR6*, *CbowIR8a*, *PstrIR19*, and *PstrIR49*), indicating they may be the co-receptor of *OcomIRs* (Fig. 8). Information, including unigene reference, length, and the best Blastx hit of all 18 IRs are listed in Supplementary Material S3.

Fluorescence quantitative real-time PCR. To verify the expression of olfactory genes in male or female antennae and characterize the expression profiles of chemosensory genes in different parts (including male antennae, female antennae, heads, legs, and the remainder of the body), 15 ORs and four OBPs were selected for qPCR. The qPCR results showed that all 15 *OcomORs* were predominately expressed in the antennae, indicating their function related to insect olfaction. Although we did not find apparent sex-specific OR genes in *O. communa*, we found *OcomOR4*, *OcomOR19*, and *OcomOR2* had significantly higher expression levels in the male antennae, whereas *OcomOR8* had a significantly higher expression level in the female antennae. Furthermore, *OcomOBP19*, *OcomOBP10*, and *OcomOBP20* were specifically expressed in the antennae, whereas *OcomOBP2*

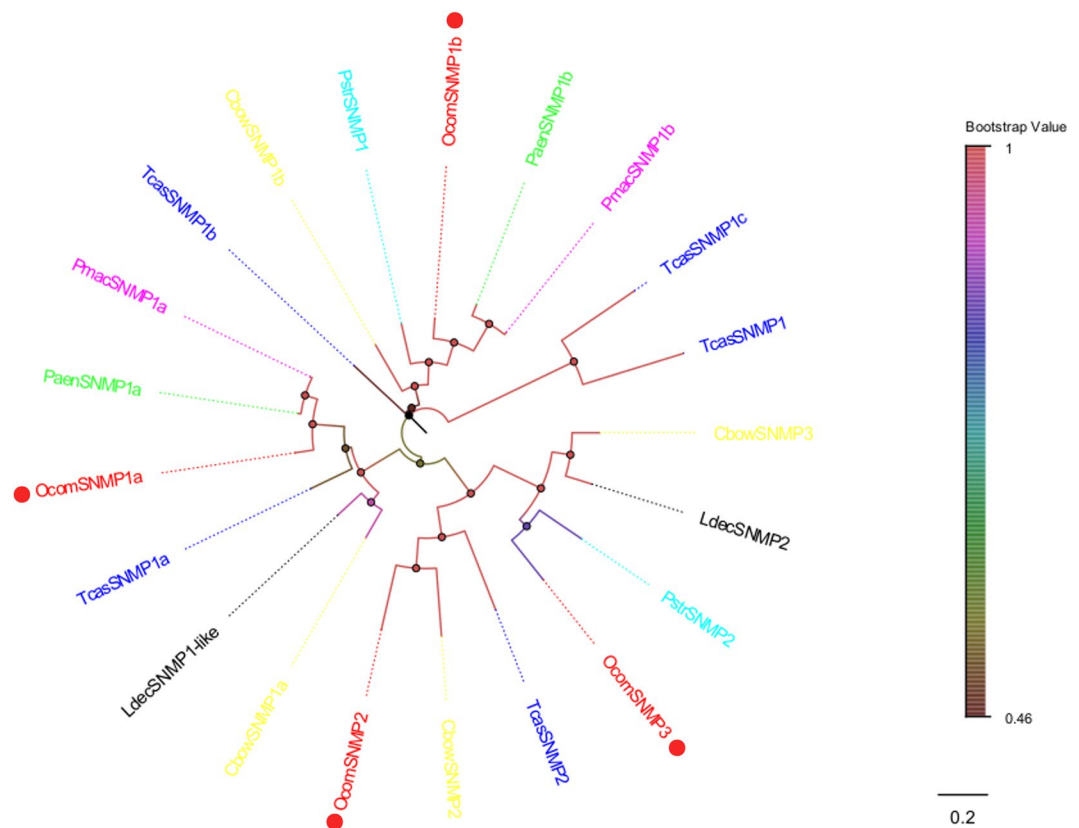


Figure 5. Neighbor joining phylogenetic tree of candidate *OcomSNMPs* with known Coleopteran SNMP sequences. *Tcas*, *Tribolium castaneum* (N = 5); *Pmac*, *Pyrrhalta maculicollis* (N = 2); *Paen*, *Pyrrhalta aenescens* (N = 2); *Cbow*, *Colaphellus bowringi* (N = 4); *Ldec*, *Leptinotarsa decemlineata* (N = 2); *Pstr*, *Phyllotreta striolata* (N = 2). Candidate *OcomSNMPs* were indicated by red circles.

was expressed not only in the antennae, but also slightly expressed in the head, body, and leg. Importantly, we found *OcomOBP19* and *OcomOBP20* were expressed significantly higher in male antennae than in female antennae (Fig. 9).

Discussion

Compared with Dipterans and Lepidopterans, the molecular underpinnings of the olfactory system of Coleoptera are poorly understood. Based on the deep RNA sequencing, we analyzed the transcriptome antennae of *O. communis*. Among the 92,259 unigenes identified, only 47% gene translations shared significant similarity with entries in the NCBI non-redundant (nr) protein database, and only 36% unigenes could be annotated with one or more gene ontology (GO) term; this is similar to that reported in other Coleoptera species^{42,43}. Thus, this amount of *O. communis* genes did not have any GO term because they were specific or fast-evolution genes of *O. communis*. However, the N50 of *O. communis* antennal transcriptome reached 2,068 bp, longer than those in *P. aenescens*³¹, *P. maculicollis*³¹, *C. bowringi*³⁰, *L. decemlineata*²⁸, and *A. quadriimpressum*³². The high quality of our transcriptome sequencing laid a foundation for olfactory annotation and further exploration of the molecular chemosensory mechanism of *O. communis*. On the basis of *O. communis* transcriptome results, we identified 105 candidate chemosensory genes, including 30 ORs, 25 OBPs, 11 CSPs, 18 IRs, 17 GRs, and four SNMPs, and this analysis substantially extended our knowledge of olfactory-related genes in Coleoptera insects. Moreover, we validated the expression profile of 15 ORs and four OBPs in different tissues of *O. communis* by qPCR, which facilitated the exploration of the function of these olfactory genes.

OBPs play an important role in odor processing by insects, facilitating the transport of odorant molecules through the sensillar lymph, and serving as the liaison between the external environment and ORs^{1,44}. In our study, we identified 25 transcripts encoding OBP genes in the *O. communis* antennal transcriptome. The numbers of OBPs are clearly lower than *T. castaneum* (49 OBPs)²⁴, *G. daurica* (29 OBPs)³², *P. aenescens* (31 OBPs)³⁰, and *P. maculicollis* (36 OBPs)³⁰, similar to the number of OBP genes in *C. bowringi* (26 OBPs)³⁰, *L. decemlineata* (26 OBPs)²⁸, but higher than that of *A. quadriimpressum* (16 OBPs)³². This is because the genome data of *T. castaneum* provide more a comprehensive list of olfactory genes than antennal transcriptome of other Coleoptera species. On the other hand, some genes may have been missed in our transcriptome results because some genes were expressed in other tissues than antennae⁴⁵ or at different life history stages^{46,47}. Relatively low coverage of the RNA-seq may also results in missing low transcribed olfactory genes. Based on the phylogenetic analysis,

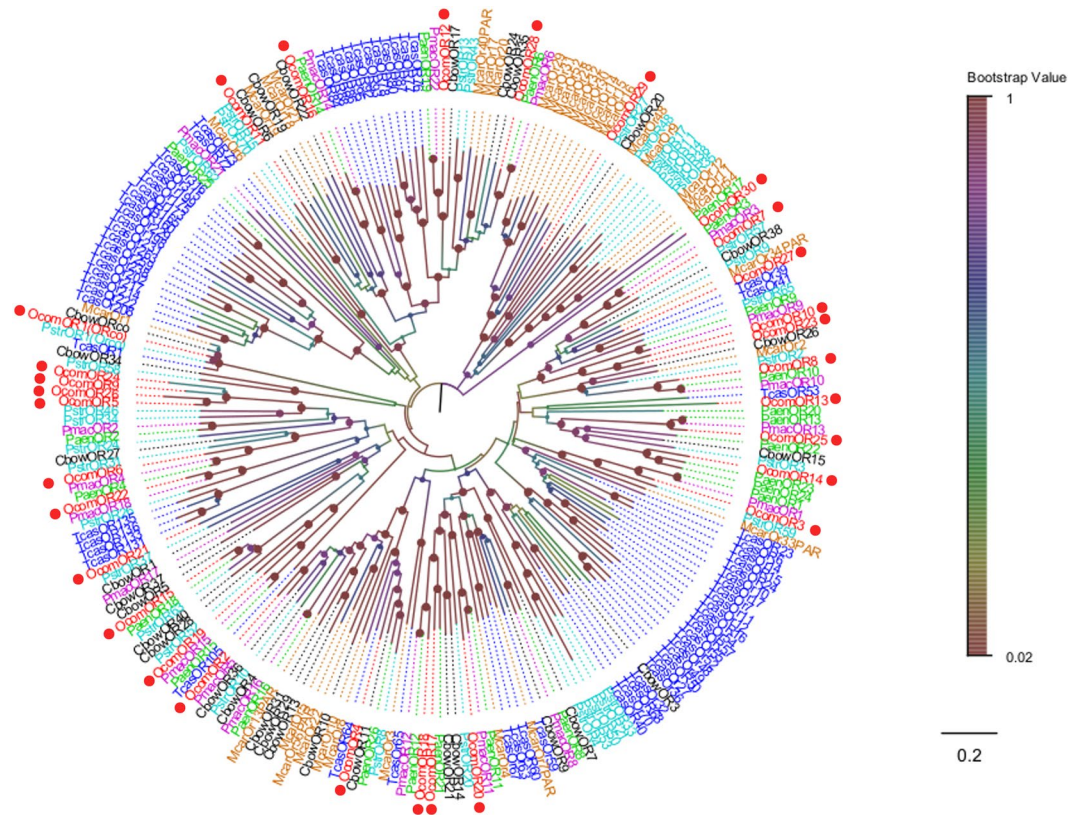


Figure 6. Neighbor joining phylogenetic tree of candidate *OcomORs* with known Coleopteran OR sequences. *Tcas*, *Tribolium castaneum* (N = 70); *Pmac*, *Pyrrhalta maculicollis* (N = 18); *Paen*, *Pyrrhalta aenescens* (N = 23); *Cbow*, *Colaphellus bowringi* (N = 30); *Pstr*, *Phyllotreta striolata* (N = 36); *Mcar*, *Megacyllene caryae* (N = 34). Candidate *OcomORs* were indicated by red circles.

OcomOBP19 grouped together with *HelePBP* and *AglaPBP1*, indicating that *OcomOBP19* may be involved in the pheromone recognition process. Further, the expression level of *OcomOBP19* in the male antennae was significantly higher than that in the female antennae, which confirms that the function of *OcomOBP19* may be related to pheromone identification in *O. communis*. Similarly, *OcomOBP20* expression in male antennae was significantly higher than that in the female antennae, indicating the function of *OcomOBP20* may be related to sex pheromone recognition or male-specific behaviors like *OcomOBP19*.

CSPs were another class of soluble proteins in the sensillum lymph with abundant expression⁴⁸. A total of 11 candidate CSP genes were found in our transcriptome data. All of them were predicted to have a putative full-length ORF and amino acid sequences range from 118 to 261. In addition, the high level similarities found in Blastx best-hit results demonstrated that CSPs were highly conserved proteins between insects. Comparing CSP gene numbers in *O. communis* with that in other Coleoptera species, there was less than 20 CSPs in *T. castaneum*²⁴, 15 CSPs in *L. decemlineata*²⁷, and similar to CSP genes in *P. aenescens* (nine CSPs)³⁰, *P. maculicollis* (ten CSPs)³⁰ and *A. quadriimpressum* (ten CSPs)³¹. Thus, the number of CSP genes identified in our study was comparable with that of previous reports on these latter three beetles. SNMPs were first identified in pheromone-sensitive neurons of Lepidopterans⁴⁹ and its function was thought to be related to pheromone detection⁵⁰. Generally, SNMPs were classified into two families, SNMP1 and SNMP2. Two SNMPs were identified in *P. aenescens*³⁰, *P. maculicollis*³⁰ and *A. quadriimpressum*³¹, whereas there were three and four SNMPs identified in the *C. bowringi*³⁰ and *L. decemlineata*²⁸ transcriptome, respectively. In this study, four SNMPs were identified in the *O. communis* antennal transcriptome as well.

ORs are important to insects olfactory system, which determine the sensitivity and specificity of odorant reception, being the centerpiece of peripheral olfactory reception in insects¹. The numbers of putative OR-encoding transcripts identified in *O. communis* are close to the number reported in the antennal transcriptome of *P. aenescens* (26 ORs)³¹, *P. maculicollis* (22 ORs)³¹, *C. bowringi* (43 ORs)³⁰, *L. decemlineata* (37 ORs)²⁸, and *A. quadriimpressum* (34 ORs)³², but much lower than the number in the *T. castaneum* genome (341 OR-encoding genes, including 79 pseudogenes)²⁴, suggesting the antennal transcriptome data may have missed some OR genes. Obviously, *OcomOR1* was grouped with *PstrORco*, *CbowORco*, and *TcasOR1*, and formed a specific co-receptor lineage, indicating that *OcomOR1* could be the ORco of *O. communis*. Similar to the reported OR genes of *T. castaneum*, *M. caryae*, and *A. corpulenta*, a species-specific expansion of ORs (*OcomOR5/9/24/26*) was also found in *O. communis*, which may suggest that these distinct species inhabit different ecological niches. The OR gene function in beetles was first characterized in *M. caryae*²⁷. *McarOR3*, *McarOR5*, and *McarOR20* were sensitive to three compounds of male-produced pheromones in *M. caryae*, indicating the function of these three ORs may

| Unigene ID | NAME | Length | ORF (aa) | Status | TMD | Blastx best-hit | Species | Score | E-value | Ident | Accession |
|--------------------|---------------|--------|----------|---------------|-----|---------------------------|----------------------------------|-------|-----------|-------|----------------|
| Cluster-9750.33622 | OcomOR1(ORco) | 4800 | 479 | Complete ORF | 7 | odorant receptor ORco | <i>Colaphellus bowringi</i> | 902 | 0.0E+00 | 91% | ALR72547.1 |
| Cluster-9750.66916 | OcomOR2 | 1594 | 446 | Complete ORF | 7 | odorant receptor 5 | <i>Pyrrhalta maculicollis</i> | 591 | 0.0E+00 | 74% | APC94229.1 |
| Cluster-9750.12389 | OcomOR3 | 1353 | 423 | Complete ORF | 3 | odorant receptor 1 | <i>Pyrrhalta aenescens</i> | 537 | 0.0E+00 | 75% | APC94305.1 |
| Cluster-8368.1 | OcomOR4 | 1665 | 415 | 5' missing | 6 | odorant receptor 26 | <i>Pyrrhalta aenescens</i> | 585 | 0.0E+00 | 65% | APC94330.1 |
| Cluster-9750.24474 | OcomOR5 | 1478 | 410 | Complete ORF | 6 | odorant receptor OR27 | <i>Colaphellus bowringi</i> | 185 | 6.00E-50 | 28% | ALR72570.1 |
| Cluster-9750.589 | OcomOR6 | 1330 | 407 | 5' missing | 5 | odorant receptor 4 | <i>Pyrrhalta aenescens</i> | 505 | 7.00E-175 | 64% | APC94309.1 |
| Cluster-9254.0 | OcomOR7 | 1399 | 406 | 5' missing | 7 | odorant receptor 3 | <i>Pyrrhalta maculicollis</i> | 622 | 0.0E+00 | 79% | APC94226.1 |
| Cluster-9750.2139 | OcomOR8 | 1392 | 400 | Complete ORF | 6 | odorant receptor 10 | <i>Pyrrhalta aenescens</i> | 669 | 0.0E+00 | 79% | APC94317.1 |
| Cluster-18330.0 | OcomOR9 | 1212 | 399 | 5' 3' missing | 6 | odorant receptor 2 | <i>Pyrrhalta aenescens</i> | 171 | 1.00E-45 | 30% | APC94306.1 |
| Cluster-13422.1 | OcomOR10 | 1305 | 393 | 5' missing | 6 | odorant receptor 9 | <i>Pyrrhalta aenescens</i> | 368 | 2.00E-121 | 49% | APC94316.1 |
| Cluster-11030.0 | OcomOR11 | 1331 | 387 | Complete ORF | 7 | odorant receptor 4-like | <i>Leptinotarsa decemlineata</i> | 272 | 4.00E-84 | 38% | XP_023017287.1 |
| Cluster-9750.5088 | OcomOR12 | 1336 | 387 | Complete ORF | 7 | odorant receptor 19 | <i>Pyrrhalta aenescens</i> | 538 | 0.0E+00 | 72% | APC94313.1 |
| Cluster-9750.1861 | OcomOR13 | 1352 | 385 | Complete ORF | 7 | odorant receptor OR15 | <i>Colaphellus bowringi</i> | 178 | 7.00E-48 | 29% | ALR72560.1 |
| Cluster-9750.9485 | OcomOR14 | 1245 | 384 | 5' missing | 6 | odorant receptor 23 | <i>Pyrrhalta aenescens</i> | 556 | 0.0E+00 | 68% | APC94324.1 |
| Cluster-9750.31554 | OcomOR15 | 1677 | 383 | Complete ORF | 5 | odorant receptor 18 | <i>Pyrrhalta aenescens</i> | 543 | 0.0E+00 | 85% | APC94311.1 |
| Cluster-18259.1 | OcomOR16 | 1384 | 382 | Complete ORF | 7 | odorant receptor 14 | <i>Pyrrhalta aenescens</i> | 463 | 1.00E-158 | 64% | APC94327.1 |
| Cluster-17525.0 | OcomOR17 | 1303 | 380 | 5' missing | 4 | odorant receptor 21 | <i>Pyrrhalta aenescens</i> | 501 | 8.00E-174 | 67% | APC94321.1 |
| Cluster-17306.1 | OcomOR18 | 1357 | 376 | Complete ORF | 6 | odorant receptor 12 | <i>Pyrrhalta maculicollis</i> | 583 | 0.0E+00 | 74% | APC94239.1 |
| Cluster-7195.0 | OcomOR19 | 1264 | 375 | Complete ORF | 6 | odorant receptor 15 | <i>Pyrrhalta aenescens</i> | 431 | 5.00E-147 | 61% | APC94328.1 |
| Cluster-13662.1 | OcomOR20 | 1250 | 372 | 5' missing | 7 | odorant receptor 11 | <i>Pyrrhalta maculicollis</i> | 469 | 1.00E-161 | 67% | APC94238.1 |
| Cluster-18700.0 | OcomOR21 | 1200 | 357 | Complete ORF | 6 | odorant receptor OR1 | <i>Colaphellus bowringi</i> | 134 | 3.00E-32 | 29% | ALR72546.1 |
| Cluster-8036.1 | OcomOR22 | 1088 | 349 | 3' missing | 5 | odorant receptor 18 | <i>Pyrrhalta maculicollis</i> | 657 | 0.0E+00 | 87% | APC94230.1 |
| Cluster-5881.0 | OcomOR23 | 1277 | 339 | 5' missing | 4 | odorant receptor 9 | <i>Pyrrhalta maculicollis</i> | 275 | 1.00E-85 | 46% | APC94236.1 |
| Cluster-9750.52233 | OcomOR24 | 1076 | 337 | 3' missing | 5 | odorant receptor 2 | <i>Pyrrhalta aenescens</i> | 94.4 | 1.00E-17 | 26% | APC94306.1 |
| Cluster-6485.1 | OcomOR25 | 1132 | 335 | Complete ORF | 5 | odorant receptor 13 | <i>Pyrrhalta maculicollis</i> | 589 | 0.0E+00 | 79% | APC94240.1 |
| Cluster-9750.24472 | OcomOR26 | 1219 | 326 | 5' 3' missing | 5 | odorant receptor 4 | <i>Pyrrhalta aenescens</i> | 104 | 5.00E-21 | 27% | APC94309.1 |
| Cluster-11404.0 | OcomOR27 | 1396 | 315 | 5' missing | 5 | odorant receptor Or1-like | <i>Anoplophora glabripennis</i> | 376 | 1.00E-122 | 55% | XP_023310030.1 |
| Cluster-1406.0 | OcomOR28 | 956 | 268 | 5' missing | 4 | odorant receptor 6 | <i>Pyrrhalta maculicollis</i> | 378 | 2.00E-77 | 71% | APC94231.1 |
| Cluster-17313.0 | OcomOR29 | 1032 | 212 | 5' missing | 3 | odorant receptor | <i>Anoplophora chinensis</i> | 184 | 2.00E-51 | 35% | AUF73043.1 |
| Cluster-9750.12020 | OcomOR30 | 484 | 120 | 5' missing | 1 | odorant receptor 24 | <i>Pyrrhalta aenescens</i> | 93.6 | 2.00E-19 | 79% | APC94325.1 |

Table 3. The Blastx match of *O. communis* candidate odorant receptors.

be related to pheromone recognition²⁷. In the phylogenetic analysis, *OcomOR2* and *OcomOR4* clustered with *McarOR3* and *McarOR5* in the same clade, and *OcomOR12* and *OcomOR28* grouped together with *McarOR20*, indicating that the function of these four ORs in *O. communis* may be pheromone identification similar to that of other lepidopteran pheromone receptors (PRs)⁵¹. In addition, qPCR results revealed that the expression level of *OcomOR2*, *OcomOR4*, and *OcomOR12* in male antennae was higher than in the female antennae, and the difference between *OcomOR2* and *OcomOR4* reached statistical significance. This evidence further demonstrates that *OcomOR2*, *OcomOR4*, and *OcomOR12* may play a role in pheromone identification in *O. communis*. *OcomOR8* expression levels in female antennae were significantly higher than in male antennae; therefore, *OcomOR8* may be related to female critical behaviors, such as oviposition cues or male-produced courtship pheromones. The sex-specific functions of these *OcomORs* need to be further investigated in the future.

Furthermore, 18 putative transcripts encoding IRs were identified in *O. communis*. The IRs number of *O. communis* was greater than in most Coleoptera species, such as nine IRs in *C. bowringi*³⁰, ten IRs in *L. decemlineata*²⁸, eight IRs in *P. aenescens*³¹, seven IRs in *P. maculicollis*³¹, three IRs in *Dendroctonus valens*⁴³, and four IRs in *A. glabripennis*⁴¹. In addition, the IR number of *O. communis* was similar to that of 20 IRs in *A. quadriimpressum*³¹. Similar to the ORco, both IR8 and IR25 were considered to act as co-receptors because they were co-expressed along with other IRs. From the phylogenetic tree of IRs, IR8 and IR25 formed a conserved IR clade, which agreed with the analysis results of *C. bowringi*³⁰. *OcomIR3* and *OcomIR7* were clustered into conserved IR25a/IR8a clades, indicating they belong to this co-expression group. Furthermore, IRs in insects were more conserved than ORs, we can predict that the function of IRs is probably conserved among Coleoptera. We identified 17 candidate GRs in *O. communis*. The GR numbers of *O. communis* was greater than most of those previously reported in beetles, and we also believe that there are more GRs expressed in other tissues, such as maxillary palps, proboscises, and legs. In the previous study of *P. aenescens* antennal transcriptome, *PaenGR12* was predicted to be the CO₂ receptor³¹. In the phylogenetic tree of GRs, *OcomGR16* grouped together with *PaenGR12*, so we predicted that *OcomGR16* acted as the CO₂ receptor in *O. communis*.

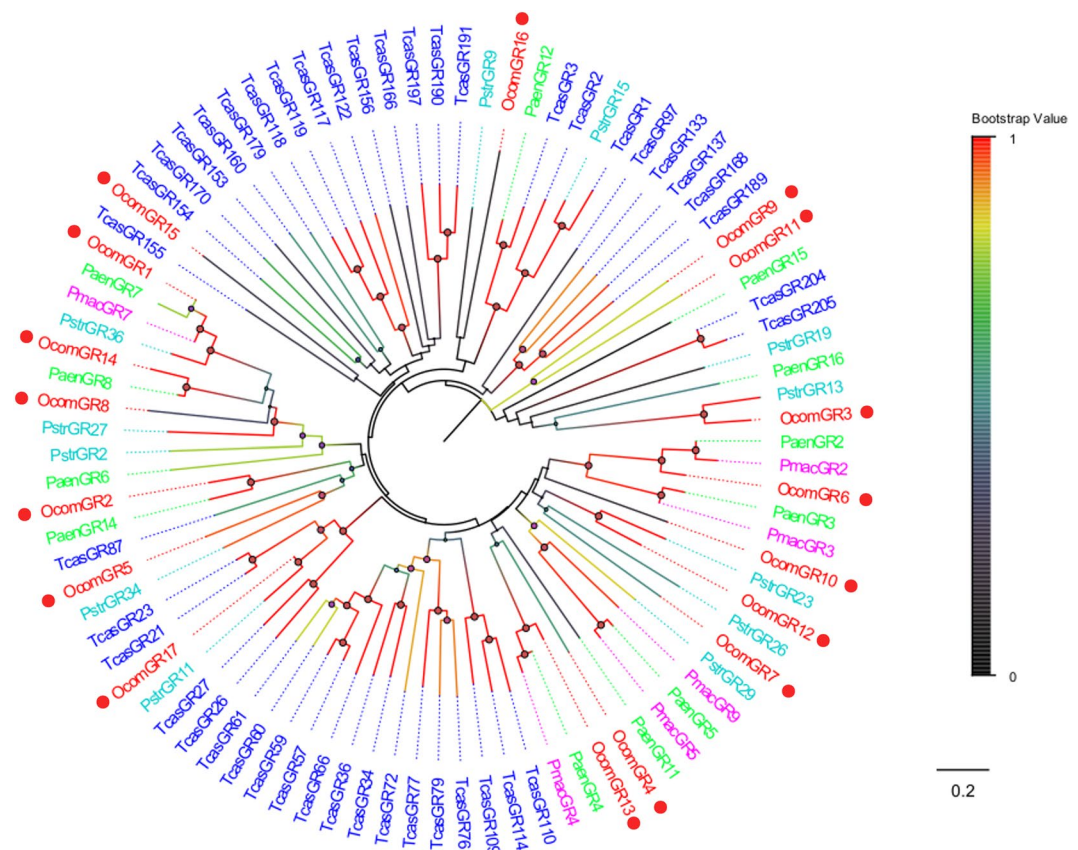


Figure 7. Neighbor joining phylogenetic tree of candidate *OcomGRs* with known Coleopteran GR sequences. *Tcas*, *Tribolium castaneum* (N = 44); *Pmac*, *Pyrrhalta maculicollis* (N = 6); *Paen*, *Pyrrhalta aenescens* (N = 12); *Pstr*, *Phyllotreta striolata* (N = 12). Candidate *OcomGRs* were indicated by red circles.

Conclusion

Using next-generation sequencing technology, we first reported large-scale olfactory gene information for *O. communa* and identified 30 ORs, 25 OBPs, 11 CSPs, 18 IRs, 17 GRs, and four SNMPs. This large number of insect chemosensory genes will provide the molecular basis for the olfactory systems of *O. communa* and will advance our understanding of olfactory mechanisms in Coleoptera. In addition, homology analysis and qPCR were performed to confirm the tissue- and sex-specific patterns of these chemosensory genes, which can help us to predict their function. Further analysis is needed to explore the function of these genes using integrated functional studies.

Materials and Methods

Insects. *O. communa* adults were collected from Laibin City, Guangxi province, southern China in June 2017, mixed, and reared together with common ragweed plants in cages in an insect breeding room at 26 °C, under 14 h light: 10 h dark cycle, and 70%–80% humidity. After the beetles laid eggs, the adults were removed from the cages and the next generation reared on common ragweed plant in the same breeding room. After eclosion, the male and female adults were separated under microscope and kept in separate cages. The antennae of the unmated male and female individuals were collected two days after eclosion. The antennae were pulled off with tweezers by grasping at the very root of the antennae, and subsequently transferred to Eppendorf tubes. For the study of gene expression profiles in different tissues, male antennae (M-T), female antennae (F-T), heads, legs, and the rest of the body were collected. All samples were immediately frozen in liquid nitrogen and stored at –80 °C until RNA extraction.

RNA extracting, cDNA library construction and Illumina sequencing. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions, in which a DNaseI digestion step was included to avoid contamination of genomic DNA. RNA quality was checked with a spectrophotometer (NanoDrop™ 1000, Thermo Fisher Scientific, USA) and 1% agarose gels, and its concentration was measured using Qubit® RNA Assay Kit with a Qubit® 2.0 Fluorometer (Life Technologies, CA, USA). The complementary DNA (cDNA) library construction and Illumina sequencing methods followed Li *et al.*³⁰. Briefly, The mRNA samples were purified and fragmented using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. Random hexamer primers were used to synthesize the first-strand cDNA, followed by synthesis of the second-strand cDNA using buffer, dNTPs, RNase H, and DNA polymerase I, and then end repair and the ligation of adaptors were handled. The cDNA library created by amplifying the products using

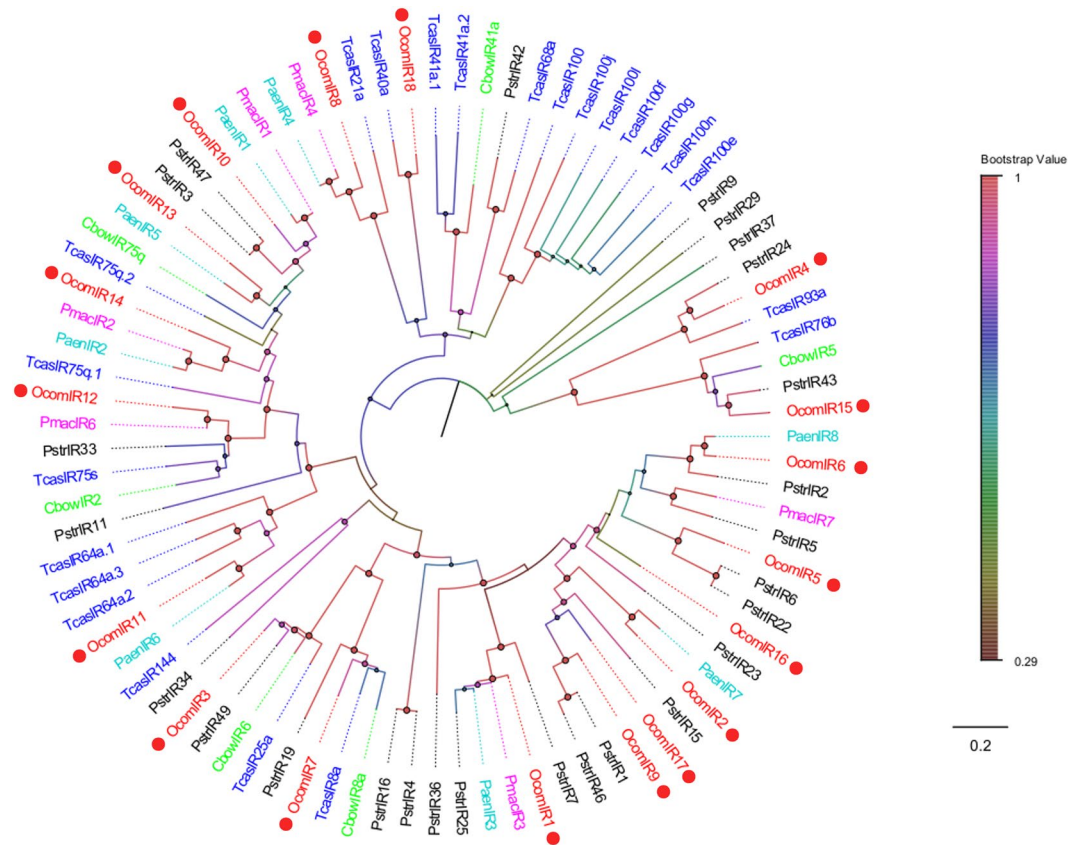


Figure 8. Neighbor joining phylogenetic tree of candidate *OcomIRs* with known Coleopteran IR sequences. *Tcas*, *Tribolium castaneum* (N = 23); *Pmac*, *Pyrrhalta maculicollis* (N = 6); *Paen*, *Pyrrhalta aenescens* (N = 8); *Pstr*, *Phyllotreta striolata* (N = 26); *Cbow*, *Colaphellus bowringi* (N = 6). Candidate *OcomIRs* were indicated by red circles.

polymerase chain reaction (PCR) and quantifying precisely using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA). The cDNA library was sequenced on the HiSeq 2000™ platform.

de novo assembly and gene annotation. The *de novo* assembly and gene annotation methods followed Li *et al.*³⁰. All raw reads were processed to remove low-quality and adaptor sequences. And then the clean reads were assembled by Trinity v2.3.1^{52,53} using the default parameters to generate unigenes. The annotation of unigenes was performed by Blastx searches (<http://www.ncbi.nlm.nih.gov>) against nr, Swiss-Prot, KEGG, and COG protein databases (E-value < 10⁻⁵). Blast2GO program⁵⁴ was used to obtain the GO annotation and WEGO software⁵⁵ was used to get GO functional classification of these unigenes.

Sequence analysis and phylogenetic analysis. The ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) and the NCBI-BLAST network server (<http://blast.ncbi.nlm.nih.gov/>) were respectively used to identify the ORFs and perform the similarity searches of the candidate chemosensory genes. TMHMM Server Version 2.0 (<http://www.cbs.dtu.dk/services/TMHMM>) was used to predict the TMDs of ORs, GRs, and IRs. The signal peptides of OBPs, CSPs, and SNMPs protein sequences were predicted by Signal IP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>)⁵⁶ with default parameters.

The amino acid sequence alignment of the candidate OBPs, CSPs, ORs, GRs, and SNMPs from *O. communa* and other insect species were performed using the ClustalW method⁵⁷ implemented in the Mega v6.0 software package⁵⁸. The OBP dataset contained 33 sequences from *P. maculicollis*, 31 from *P. aenescens*, 29 from *G. daurica*, and 47 from *T. castaneum*. The CSP dataset contained 10 sequences from *P. maculicollis*, nine from *P. aenescens*, 12 from *C. bowringi*, and 19 sequences from *T. castaneum*. The SNMP dataset contained two sequences from *P. maculicollis*, two from *P. aenescens*, four from *C. bowringi*, two from *L. decemlineata*, two from *P. striolata*, and five sequences from *T. castaneum*. The OR dataset contained 18 sequences from *P. maculicollis*, 23 from *P. aenescens*, 30 from *C. bowringi*, 36 from *P. striolata*, 34 from *M. caryae*, and 70 sequences from *T. castaneum*. The GR dataset contained six sequences from *P. maculicollis*, 12 from *P. aenescens*, 12 from *P. striolata*, and 44 sequences from *T. castaneum*. The IR dataset contained six sequences from *P. maculicollis*, eight from *P. aenescens*, 26 from *P. striolata*, six from *C. bowringi*, and 23 sequences from *T. castaneum*. All amino acid sequences of *O. communa* and other insects used in the phylogenetic analyses are listed in Supplementary Material S1. The phylogenetic tree was constructed using the neighbor-joining (NJ) method⁵⁹ with P-distance modeling and pairwise deletion of gaps performed in the Mega v6.0 software package⁵⁸ and the dendrograms were colored in Fig Tree v1.4.3

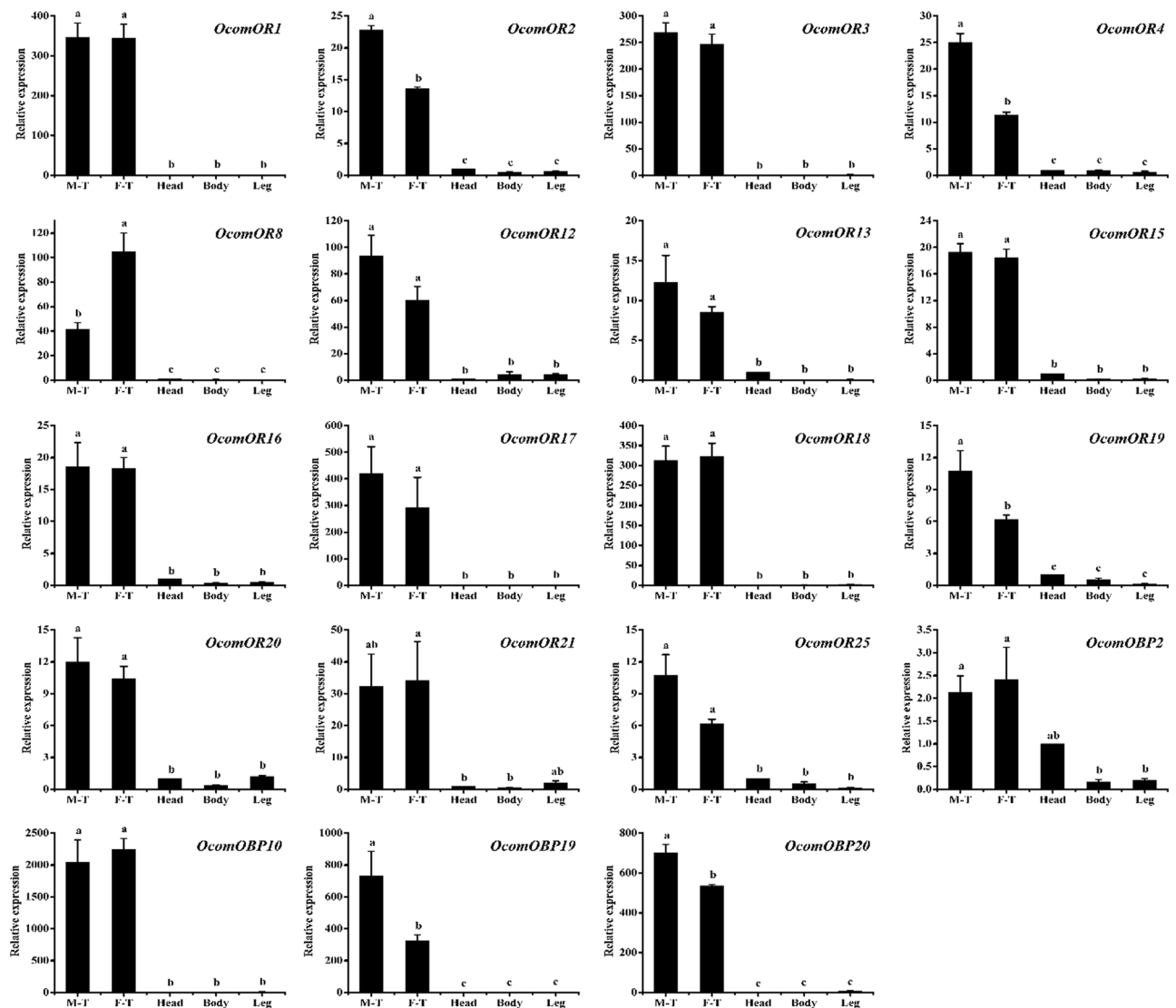


Figure 9. Relative expression levels of 15 ORs and four OBPs in adult antennae, head, leg and the rest of body using qPCR. M-T, male antennae; F-T, female antennae. The relative expression level is indicated as mean \pm SE (N = 3). Different capital letters mean significant difference between tissues ($P < 0.05$, ANOVA, LSD).

software package. The reliability of the tree structure and node support was assessed using a bootstrap procedure based on 1,000 replicates. To ensure greater accuracy in the analyses and make sure that the analyzed transcripts corresponded to individual genes, incomplete transcripts without sufficient overlap in alignments and protein sequence length less than 100 amino acids in length were excluded from the phylogenetic analyses. Six group candidate chemosensory genes were named “*OcomOBP*,” “*OcomCSP*,” “*OcomSNMP*,” “*OcomOR*,” “*OcomGR*,” and “*OcomIR*,” and were followed by a numeral in descending order of their coding region lengths.

Quantitative real-time PCR validation. We selected 15 ORs and four OBPs to verify their expression profiles because their relative high abundance from fragments per kilobase of exon per million reads mapped (FPKM) data in antennal transcriptome. The expression profiles of 15 ORs and four OBPs were analyzed using qPCR experiments. Total RNA was isolated from the five tissues as described above. The concentration of each RNA sample was standardized to one $\mu\text{g}/\mu\text{l}$ and the cDNA was synthesized using a first-strand cDNA synthesis kit (Transgen Biotech, Beijing, China) according to the manufacturer’s protocol. Ribosomal protein (RL4) was used as an internal control and its specific primer sequences were RL4-F: “TGTGGTAATGCTGTGGTAT” and RL4-R: “TCTAGCACTGCATGAACA”. The qPCR was performed on an ABI 7500 (Thermo Scientific, Waltham, MA, USA) with TransStar Tip Top Green qPCR Supermix (Transgen Biotech, Beijing, China). The PCR reaction programs were 30 s at 94 °C, 40 cycles of 94 °C for 5 s, and 60 °C for 34 s. All qPCR primers were designed using Primer Premier 5.0 (PREMIER Biosoft International) and the efficiency of these primers was validated before gene expression analysis. All primer sequences were listed in Supplementary Material S2. Each qPCR reaction was performed using three technical replicates and three biological replicates.

Statistical analysis. Data analysis was performed using the $2^{-\Delta\Delta CT}$ method and data were analyzed using SAS 9.0 (SAS Institute Inc., Cary, NC, USA). Statistical significance was assessed by an analysis of variance (ANOVA) followed by a Tukey multiple comparison tests. A value of $P < 0.05$ was considered statistically significant. Figures were made using OriginPro 9.1 (Northampton, Massachusetts, USA).

Data deposition. All the Illumina sequencing data of the antennal transcriptome in this study have been stored in the NCBI SRA database, under the accession number of SRR8372148 (*O. communa* male antennae) and SRR8372149 (*O. communa* female antennae).

Received: 20 June 2019; Accepted: 4 October 2019;

Published online: 29 October 2019

References

- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391 (2013).
- Rützler, M. & Zwiebel, L. J. Molecular biology of insect olfaction: recent progress and conceptual models. *J. Comp. Physiol.* **191**, 777–790 (2005).
- Bruyne, M. D. & Baker, T. C. Odor detection in insects: volatile codes. *J. Chem. Ecol.* **34**, 882–897 (2008).
- Sato, K. & Touhara, K. Insect olfaction: receptors, signal transduction, and behavior. *Results Probl. Cell Differ.* **47**, 121–138 (2009).
- Vogt, R. G. Biochemical diversity of odor detection: OBPs, ODEs and SNMPs (ed. Blomquist, G. J. & Vogt, R. G.) 397–451 (Elsevier, 2003).
- Zhou, J. J. Odorant-binding proteins in insects. *Vitam. Horm.* **83**, 241–72 (2010).
- Pelosi, P., Calvello, M. & Ban, L. Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chem. Senses* **30**, 291–292 (2005).
- Robertson, H. M., Warr, C. G. & Carlson, J. R. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **100**, 14537–422003 (2003).
- Sanchez, G. A., Vieira, F. G. & Rozas, J. Molecular evolution of the major chemosensory gene families in insects. *Heredity* **103**, 208–216 (2009).
- Zhang, J., Walker, W. B. & Wang, G. R. Pheromone reception in moths: from molecules to behaviors. *Prog. Mol. Biol. Transl. Sci.* **130**, 109–128 (2015).
- Vogt, R. G. *et al.* The insect SNMP gene family. *Insect Biochem. Mol. Biol.* **39**, 448–456 (2009).
- Benton, R., Vannice, K. S., Gomez, D. C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136**, 149–162 (2009).
- Tunstall, N. E. & Warr, C. G. Chemical communication in insects: the peripheral odour coding system of *Drosophila melanogaster*. *Adv. Exp. Med. Biol.* **739**, 59–77 (2012).
- Montell, C. A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol.* **19**, 345–353 (2009).
- Clyne, P. J. *et al.* A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**, 327–338 (1999).
- Vosshall, L., Amrein, H., Morozov, P., Rzhetsky, A. & Axel, R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**, 725–736 (1999).
- Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
- Wang, G. R., Carey, A. F., Carlson, J. R. & Zwiebel, L. J. Molecular basis of odor coding in the malaria vector mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* **107**, 4418–4423 (2010).
- Sato, K., Pellegrino, M., Nakagawa, T., Vosshall, L. B. & Touhara, K. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–1006 (2008).
- Wicher, D. *et al.* *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**, 1007–1011 (2008).
- Jones, P. L., Pask, G. M., Rinker, D. C. & Zwiebel, L. J. Functional agonism of insect odorant receptor ion channels. *Proc. Natl. Acad. Sci. USA* **108**, 8821–8825 (2011).
- Sakurai, T. *et al.* Identification and functional characterization of a sex pheromone receptor in the silkworm *Bombyx mori*. *Proc. Natl. Acad. Sci. USA* **101**, 16653–16658 (2004).
- Carey, A. F., Wang, G. R., Su, C. Y., Zwiebel, L. J. & Carlson, J. R. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464**, 66–71 (2010).
- Hunt, T. *et al.* A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* **318**, 1913–1916 (2007).
- Liu, Y. *et al.* Identification of candidate olfactory genes in *Leptinotarsa decemlineata* by antennal transcriptome analysis. *Front. Ecol. Evol.* **3**, 60 (2015).
- Richards, S. *et al.* The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**, 949–955 (2008).
- Kim, H. S. *et al.* BeetleBase in 2010: revisions to provide comprehensive genomic information for *Tribolium castaneum*. *Nucleic Acids Res.* **38**, 437–442 (2010).
- Mitchell, R. F. *et al.* Sequencing and characterizing odorant receptors of the cerambycid beetle *Megacyllene caryae*. *Insect Biochem. Mol. Biol.* **42**, 499–505 (2012).
- Wu, Z. Z. *et al.* Differential expression analysis of chemoreception genes in the striped flea beetle *Phyllotreta striolata* using a transcriptomic approach. *PLoS One* **11**, e0153067 (2016).
- Li, X. M. *et al.* Candidate chemosensory genes identified in *Colaphellus bowringi* by antennal transcriptome analysis. *BMC Genomics* **16**, 1028 (2015).
- Zhang, B. *et al.* Comparative transcriptome analysis of chemosensory genes in two sister leaf beetles provides insights into chemosensory speciation. *Insect Biochem. Molec.* **79**, 108–118 (2016).
- Wang, Y. L., Chen, Q., Zhao, H. B. & Ren, B. Z. Identification and comparison of candidate olfactory genes in the olfactory and non-olfactory organs of elm pest *Ambrostoma quadriimpressum* (Coleoptera: Chrysomelidae) based on transcriptome analysis. *PLoS One* **11**, e0147144 (2016).
- Li, L., Zhou, Y. T., Tan, Y., Zhou, X. R. & Pang, B. P. Identification of odorant-binding protein genes in *Galeruca daurica* (Coleoptera: Chrysomelidae) and analysis of their expression profiles. *B. Entomol. Res.* **107**, 550–561 (2017).
- Palmer, W. A. & Goeden, R. D. The host range of *Ophraella communa* Lesage (Coleoptera: Chrysomelidae). *Coleopt. Bull.* **45**, 115–120 (1991).
- Yamazaki, K., Imai, C. & Natuhara, Y. Rapid population growth and food-plant exploitation pattern in an exotic leaf beetle, *Ophraella communa* LeSage (Coleoptera: Chrysomelidae), in western Japan. *Appl. Entomol. Zool.* **35**, 215–223 (2000).

36. Meng, L. & Li, B. P. Advances on biology and host specificity of the newly introduced beetle, *Ophraella communa* Lesage (Coleoptera: Chrysomelidae), attacking *Ambrosia artemisiifolia* (Compositae) in continent of China. *Chinese Journal of Biological Control* **21**, 65–69 (2005).
37. Zhou, Z. S., Guo, J. Y., Chen, H. S. & Wan, F. H. Effects of temperature on survival, development, longevity and fecundity of *Ophraella communa* (Coleoptera: Chrysomelidae), a biological control agent against invasive ragweed, *Ambrosia artemisiifolia* L. (Asterales: Asteraceae). *Environ. Entomol.* **39**, 1021–1027 (2010).
38. Zhou, Z. S., Guo, J. Y., Guo, W. & Wan, F. H. Synergistic effects of olfactory and tactile cues in short-range mate finding of *Ophraella communa*. *Entomol. Exp. Appl.* **138**, 48–54 (2010).
39. Wang, J. *et al.* Analysis of chemosensory gene families in the beetle *Monochamus alternatus* and its parasitoid *Dastarcus helophoroides*. *Comp. Biochem. Physiol.* **11**, 1–8 (2014).
40. Chen, H. L. *et al.* De novo sequencing, assembly and characterization of antennal transcriptome of *Anomala corpulenta* Motschulsky (Coleoptera: Rutelidae). *PLoS One* **9**, e114238 (2014).
41. Hu, P., Wang, J., Cui, M., Tao, J. & Luo, Y. Antennal transcriptome analysis of the Asian longhorned beetle *Anoplophora glabripennis*. *Sci Rep-UK* **6**, 26652 (2016).
42. Andersson, M. N. *et al.* Antennal transcriptome analysis of the chemosensory gene families in the tree killing dark beetles, *Ips typographus* and *Dendroctonus ponderosae* (Coleoptera: Curculionidae: Scolytinae). *BMC Genomics* **14**, 1–16 (2013).
43. Gu, X. C., Zhang, Y. N., Kang, K., Dong, S. L. & Zhang, L. W. Antennal transcriptome analysis of odorant reception genes in the Red Turpentine Beetle (RTB), *Dendroctonus valens*. *PLoS One* **10**, e0125159 (2015).
44. Liu, Y., Gu, S. H., Zhang, Y. J., Guo, Y. Y. & Wang, G. R. Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS One* **7**, e48260 (2012).
45. Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell Mol. Life Sci.* **63**, 1658–1676 (2006).
46. Engson, P. *et al.* The red flour beetle's large nose. An expanded odorant receptor gene family in *Tribolium castaneum*. *Insect Biochem. Mol. Biol.* **38**, 387–397 (2008).
47. Xia, Y. *et al.* The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl. Acad. Sci. USA* **105**, 6433–6438 (2008).
48. Foret, S., Wanner, K. W. & Maleszka, R. Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expression profiling. *Insect Biochem. Mol. Biol.* **37**, 19–28 (2007).
49. Rogers, M. E., Krieger, J. & Vogt, R. G. Antennal SNMPs (sensor neuron membrane proteins) of lepidoptera define a unique family of invertebrate CD36-like proteins. *J. Neurobiol.* **49**, 47–61 (2001).
50. Benton, R., Vannice, K. S. & Vosshall, L. B. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature* **450**, 289–293 (2007).
51. Zhang, J. *et al.* Identification and functional characterization of sex pheromone receptors in the common cutworm (*Spodoptera litura*). *Chem. Senses* **40**, 7–16 (2015).
52. Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20**, 265–272 (2010).
53. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
54. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
55. Ye, J. *et al.* WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* **34**, 293–297 (2006).
56. Petersen, T. N., Brunak, S., Heijne, G. V. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
57. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
58. Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
59. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).

Acknowledgements

We thank Dr. Yang Liu and Ph.D. students Li-xiao Du (Institute of Plant Protection, Chinese Academy of Agricultural Sciences) for help analysing data. We thank Dr. Lin Qiu (College of Plant Protection, Hunan Agricultural University) for editorial assistance and comments on the manuscript. This work was supported by the National Natural Science Foundation of China (No. 31672089) and National Natural Science Foundation of China for Excellent Young Scholars (No. 31322046). The funding bodies themselves did not participate in the design of the study or the collection, analysis, and interpretation of data, or in writing the manuscript.

Author contributions

All authors contributed to research design and manuscript preparation. C.M. designed research, conducted experiments, analyzed data, and wrote the manuscript. C.Z., S.C., Y.Z. and G.C. participated in data reduction. H.C., F.W. and Z.Z. revised the manuscript. All the authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-52149-x>.

Correspondence and requests for materials should be addressed to Z.Z.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019