

# SCIENTIFIC REPORTS



OPEN

## Identification and tissue distribution of chemosensory protein and odorant binding protein genes in *Tropidothorax elegans* Distant (Hemiptera: Lygaeidae)

Yue-Qin Song<sup>1</sup>, Hui-Zhong Sun<sup>1</sup> & Jun Du<sup>2</sup>

*Tropidothorax elegans* Distant (Hemiptera: Lygaeidae) is an insect pest that inflicts damage to vegetables and flowering plants across China. The olfactory system regulates insect behavior, such as feeding, mating, oviposition and predator avoidance. Odorant-binding proteins (OBPs) and the chemosensory proteins (CSPs) are two groups of small soluble proteins that initiate olfactory signal transduction in insects. In this study, we generated antennal transcriptomes of male and female *T. elegans*, and identified 19 putative OBP (14 classic OBPs and five plus-C OBPs) and seven CSP genes. Through real-time quantitative PCR analysis, we found that 14 of the 19 OBP genes were highly expressed in the antennae of both adult females and males, and 3 OBP genes were expressed higher in the antennae of males than females. Some OBP genes were also highly expressed in the legs or wings. Three CSP genes were highly expressed in the antennae of both sexes, and *TeleCSP7* showed higher expression in male antennae compare to females. Interestingly, one CSP gene, *TeleCSP2*, was expressed in all olfactory tissues. Our results provide molecular insights into further investigating of the olfactory system of an important plant pest, *T. elegans*.

The olfactory system plays an important role in regulating behaviors in insects, such as finding food, mating, oviposition and avoiding predators<sup>1,2</sup>. The process of olfactory perception is mediated by a series of proteins that include odorant-binding proteins (OBPs), chemosensory proteins (CSPs), olfactory receptors (ORs), sensory neuron membrane proteins (SNMPs) and olfactory degrading enzymes (ODEs). OBPs and CSPs are produced in the lymph of the chemosensilla and bind to chemical signals from the external environment, then transfer odorants through the sensillar lymph to the ORs, activating signal transduction<sup>2,3</sup>. Therefore OBPs and CSPs are two families of small water-soluble polypeptides in the lymph of chemosensilla and initiate the biochemical recognition of olfactory signals in insect<sup>4–6</sup>.

All insect OBPs have six highly conserved cysteine residues, paired in three interlocked disulphide bridges<sup>7</sup>. OBPs are classified into five groups: 'Classic' OBPs with six Cys residues, 'dimer' OBPs with two classical Cys signature motifs, 'Minus-C' OBPs that lacks two of the six conserved Cys, 'Plus-C' OBPs that has a Pro residue and two additional Cys residues and 'Atypical' OBPs with 9–10 Cys residues<sup>5</sup>. OBPs are highly expressed in the antennae, and the primary function of OBPs are to transport odorant molecules to ORs in the antenna<sup>8,9</sup>. However, recent studies have found expression of OBPs in nonsensory organs, such as pheromone glands and reproductive organs. This suggests that OBPs serve an additional function as chemical compound carrier, which may be analogous to the vertebrates urinary and salivary proteins as well as some insect chemosensory proteins<sup>10,11</sup>.

<sup>1</sup>Forestry College, Henan University of Science and Technology, Luoyang, 471000, China. <sup>2</sup>Institute of Plant Nutrition and Resource Environment, Henan Academy of Agricultural Sciences, Zhengzhou, 450002, China. Correspondence and requests for materials should be addressed to H.-Z.S. (email: [huizhong66@163.com](mailto:huizhong66@163.com))

Gene name	Acc. no.	ORF(aa)	Molecular weight (kDa)	Isoelectric point	Signal peptide	Full length	Best Blastx hit					
							Gene annotation	Species	Protein ID	Score(bits)	E-value	Identify
TeleOBP1	MF593898	205	22.66	5.18	1-21	yes	OBP7	Adelphocoris suturalis	ANA10233.1	191	3e-58	53%
TeleOBP2	MF593899	102	11.41	5.16	no	No	OBP5	Adelphocoris lineolatus	ACZ58031.1	164	4e-50	75%
TeleOBP3	MF593900	152	17.04	4.27	1-29	yes	OBP3	Euschistus heros	AIU64820.1	85.8	3e-18	36%
TeleOBP4	MF593901	140	15.85	5.43	1-17	yes	OBP4	Chinavia ubica	AIU64827.1	100	3e-24	39%
TeleOBP5	MF593902	383	45.43	7.96	1-23	yes	OBP5	Halyomorpha halys	AOV87022.1	209	2e-61	40%
TeleOBP6	MF593903	156	17.09	8.86	1-13	no	OBP25	Chilo suppressalis	ANC68513.1	56.6	7e-07	29%
TeleOBP7	MF593904	200	22.96	5.4	1-20	yes	OBP3	Sogatella furcifera	AGZ04903.1	64.3	3e-09	32%
TeleOBP8	MF593905	203	22.83	6.34	1-20	yes	OBP31	Apolygus lucorum	AMQ76484.1	104	2e-24	34%
TeleOBP9	MF593906	172	19.11	7.56	no	yes	OBP1	Pachypeltis micranthus	ARO46430.1	82	9e-17	33%
TeleOBP10	MF593907	143	15.64	5.33	1-18	yes	OBP2	Chinavia ubica	AIU64825.1	110	2e-28	42%
TeleOBP11	MF593908	155	17.69	5.26	1-19	yes	OBP4	Chinavia ubica	AIU64827.1	63.2	1e-09	32%
TeleOBP12	MF593909	111	12.56	9.18	no	yes	OBP33	Apolygus lucorum	AMQ76486.1	93.6	4e-22	45%
TeleOBP13	MF593910	175	19.69	4.84	1-17	yes	OBP13	Halyomorpha halys	AOV87030.1	63.5	2e-09	33%
TeleOBP14	MF593911	152	17.07	5.57	1-20	yes	OBP2	Telenomus podisi	AIU64830.1	166	4e-50	53%
TeleOBP15	MF593912	145	16.04	6.31	1-18	yes	OBP15	Halyomorpha halys	AOV87032.1	98.2	2e-23	31%
TeleOBP16	MF593913	152	17.04	5.16	1-19	yes	OBP27	Halyomorpha halys	AOV87044.1	103	2e-25	35%
TeleOBP17	MF593914	140	15.78	4.97	1-22	yes	OBP4	Lygus lineolaris	AHF71031.1	162	4e-49	53%
TeleOBP18	MF593915	262	30.92	7.05	1-20	yes	OBP24	Anopheles funestus	ADQ01710.1	53.1	4e-05	31%
TeleOBP19	MF593916	210	23.38	6.32	1-25	yes	OBP25	Apolygus lucorum	AMQ76478.1	122	2e-31	34%
TeleCSP1	MF598723	134	15.29	8.93	1-21	yes	CSP8	Nilaparvata lugens	ACJ64054.1	149	9e-44	57%
TeleCSP2	MF598724	115	12.88	9.27	1-26	no	CSP13	Lygus hesperus	APB88051.1	96.3	1e-23	42%
TeleCSP3	MF598725	125	14.06	5.59	1-19	yes	CSP3	Apolygus lucorum	AEP95757.1	152	3e-45	58%
TeleCSP4	MF598726	141	16.56	6.41	1-23	yes	CSP4	Adelphocoris suturalis	ANA10246.1	124	3e-34	47%
TeleCSP5	MF598727	127	14.47	4.70	1-17	yes	CSP8	Lygus hesperus	APB88044.1	130	1e-36	48%
TeleCSP6	MF598728	123	13.98	9.18	1-17	yes	CSP3	Apolygus lucorum	AGD80083.1	209	8e-68	86%
TeleCSP7	MF598729	135	15.66	5.74	1-18	yes	CSP8	Lygus hesperus	APB88044.1	119	2e-32	46%

**Table 1.** Sequence information of TeleOBPs and TeleCSPs.

In contrast to OBPs, CSPs have four conserved Cys residues<sup>5,12</sup>. Initially, these proteins were named as olfactory specific protein D (OS-D)<sup>13</sup> or A10<sup>14</sup>, but were then named chemosensory proteins (CSPs). CSPs are expressed in chemosensory tissues, such as antennae<sup>15</sup>, maxillary palps<sup>16</sup>, proboscis<sup>17</sup>, labial palps<sup>16,18</sup>, wings<sup>19</sup>, legs<sup>20</sup>, and non-chemosensory tissues, including pheromone gland<sup>21,22</sup> and ejaculatory bulb<sup>23</sup>. CSPs may play important role as carriers for odorant molecules through the sensillar lymph to transmembrane chemoreceptors but can also be involved in other physiological and behavioral functions.

The number of OBPs and CSPs is highly variable even between closely related species. For example, 38 OBPs were identified in *Apolygus lucorum* Meyer-Dür<sup>9</sup>, 33 OBPs in *Lygus lineolaris* Palisot de Beauvois<sup>8</sup>, 28 OBPs and 16 CSPs in *Nysius ericae* Schilling<sup>24</sup>, 16 OBPs and 8 CSPs in *Adelphocoris suturalis* Jakovlev<sup>25</sup>, 14 OBPs and 3 CSPs in *Adelphocoris lineolatus* Goeze<sup>26,27</sup>, 10 OBPs and 5 CSPs in *Cyrtorhinus lividipennis* Reuter<sup>28</sup>, 40 OBPs and 11 CSPs in *Empoasca onukii* Matsuda<sup>29</sup>, 8 OBPs and 13 CSPs in *Bemisia tabaci* Gennadius<sup>30</sup>, 30 OBPs in *Halyomorpha halys* Stål<sup>31</sup>, 9 OBPs and 3 CSPs in *Pachypeltis micranthus* Mu et Liu<sup>32</sup>.

*Tropidothorax elegans* Distant (Hemiptera: Lygaeidae) is a polyphagous insect pest that can severely damage commercially important plants, such as locust, pepper, cucurbit, red sage, oilseed rape, Chinese cabbage, and wheat<sup>33</sup>. Insects detect the chemical information of a plant to determine whether it is a suitable host. However, the molecular mechanisms underlying the detection of host plants remain elusive. In this study, we identified OBP and CSP genes from *T. elegans* by screening the antennal transcriptome and subsequently examined their expression in different tissues by real-time quantitative PCR (RT-qPCR). Our results provided resources for further functional characterization of *T. elegans* OBPs/CSPs.

## Results

**Identification of CSP and OBP genes in *T. elegans*.** After removing adaptors and low quality reads, a total of 36,445,539 clean reads were obtained with a Q30 percentage of 91.90%. The assembly of all clean reads together led to the generation of 118,998 (mean length 852.84 bp) transcripts and 93,940 (mean length 667.34 bp) unigenes with N50 length of 1,545 and 1,046 bp. We obtained 33,079 unigenes that were longer than 500 bp, which accounted for 35.21% of all unigenes (Table S1).

The tBLASTn results showed 19 unigenes encoding putative OBP genes (GenBank Accession Nos. MF593898–MF593916). Of the 19 TeleOBPs that were identified, 17 TeleOBPs had intact ORFs that were 140–383 aa in length, and 15 TeleOBPs had a signal peptide at the N-terminus. Two TeleOBPs (TeleOBP2 and TeleOBP6) were partial sequences with the N-terminus missing. The full-length protein sequences of TeleOBP1–19 had



**Figure 1.** Alignment of *Tropicodthorax elegans* ‘classic’ odorant-binding proteins (OBPs). The six conserved cysteine residues are indicated with the letter “c” under the sequence.



**Figure 2.** Alignment of *Tropicodthorax elegans* ‘Plus-C’ odorant-binding proteins (OBPs). The conserved residues in the ‘Plus-C’ OBPs motif are indicated with the letter “c” under the sequence.



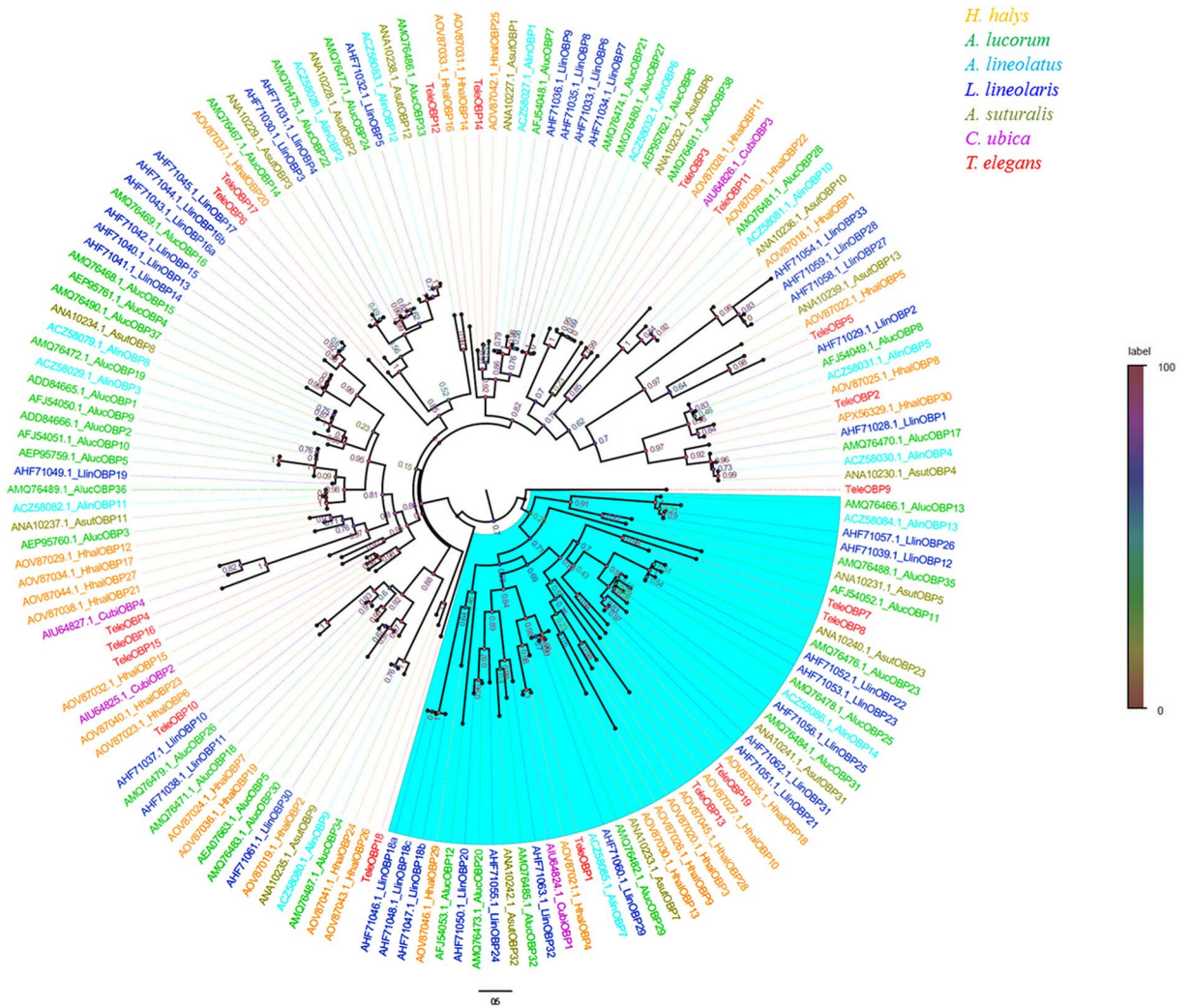
**Figure 3.** Alignment of *Tropicodthorax elegans* chemosensory proteins (CSPs). Four conserved cysteine residues are indicated with the letter “c” under the sequence.

molecular weights of 15.64–23.38 kDa and isoelectric points of 4.27–8.86. BLAST analyses revealed that the TeleOBP sequences were similar to putative OBP sequences identified from other hemipteran species. The highest degree of sequence identity (75%) was found in the putative OBP from the alfalfa plant bug species, *A. lineolaris*, followed by *A. suturalis* (53%), *L. lineolaris* (53%) and *Telenomus podisi* (53%) (Table 1).

Except for TeleOBP2 where the 5' information was missing, the other 18 TeleOBPs contained OBP cysteine patterns that were ‘classic’ OBP or ‘Plus-C’ OBP. Based on the Hemiptera ‘classic’ OBP Cys motif (C1-X<sub>15-39</sub>-C2-X<sub>3</sub>-C3-X<sub>21-44</sub>-C4-X<sub>7-12</sub>-C5-X<sub>8</sub>-C6), we classified 13 TeleOBPs (TeleOBP3–6, TeleOBP9–12 and TeleOBP14–18) as ‘classic’ OBPs (Fig. 1). The remaining five TeleOBPs (TeleOBP1, TeleOBP7–8, TeleOBP13 and TeleOBP19) were categorized as ‘Plus-C’ OBPs, because they contained two additional conserved cysteins and a conserved proline immediately after the sixth cysteine (Fig. 2).

We next searched for CSP genes and identified seven putative TeleCSP genes from the *T. elegans* antennal transcriptome. Aside from TeleCSP2, six TeleCSPs contained complete ORF ranging 115–141 aa. The sequences were named TeleCSP1 to TeleCSP7 (GenBank accession numbers MF585723–MF585729). The full-length protein sequences of TeleCSP1 to TeleCSP7 had molecular weights of 13.96–16.56 kDa and isoelectric points of 4.70–9.18. All TeleCSPs shared high sequence similarities to known hemipteran CSPs, and in particular, TeleCSP6 showed highest identity (86%) with a CSP from *A. lucorum* (Table 1).





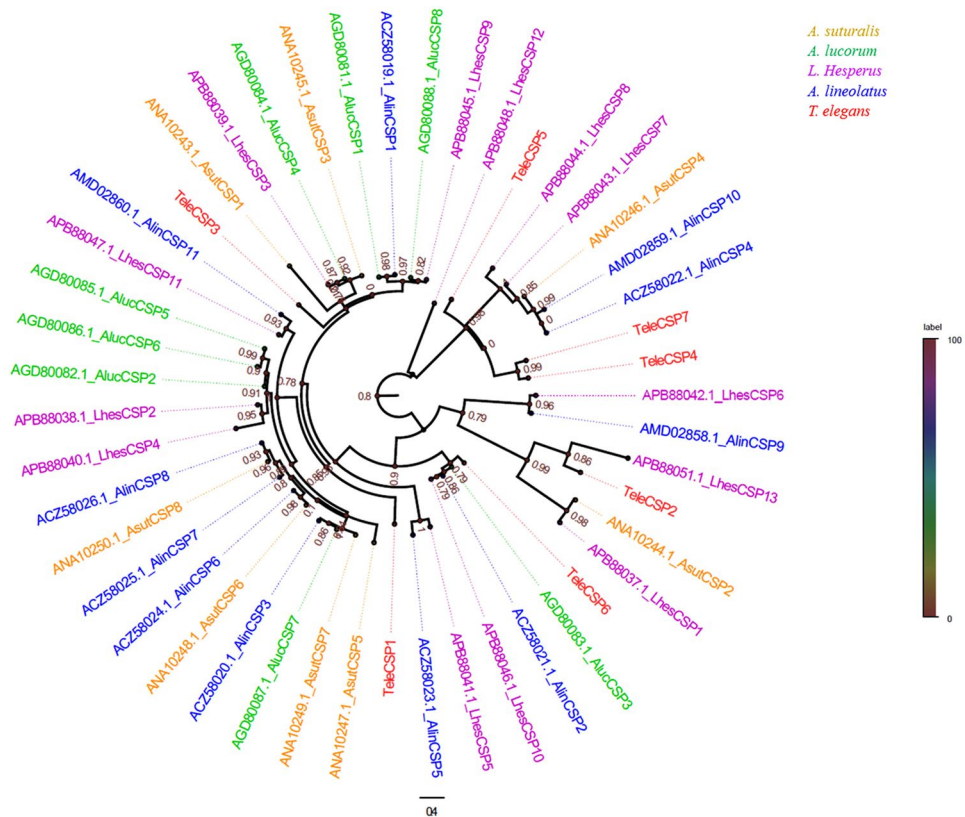
**Figure 4.** Phylogenetic relationships of 160 odorant-binding proteins (OBPs) from seven Hemiptera species: *Halyomorpha halys*, *Apolygus lucorum*, *Adelphocoris lineolatus*, *Lygus lineolaris*, *Adelphocoris suturalis*, *Chinavia ubica*, and *Tropidothorax elegans*. The branch leading to ‘Plus-C’ odorant-binding proteins (OBPs) is marked in blue.

All TeleCSP genes contained a 17–26 aa N-terminal signal peptide sequence and a highly conserved four cysteine residues with the spacing pattern C<sub>1</sub>-X<sub>6</sub>-C<sub>2</sub>-X<sub>18</sub>-C<sub>3</sub>-X<sub>2</sub>-C<sub>4</sub> (Fig. 3).

**Phylogenetic analysis.** To conduct a phylogenetic analysis of OBPs, we used a total of 160 OBP protein sequences from seven different Hemiptera insects, including the 19 predicted OBPs we identified from *T. elegans*, 30 OBPs from *H. halys*, 39 OBPs from *A. lucorum*, 14 OBPs from *A. lineolatus*, 39 OBPs from *L. lineolaris*, 15 OBPs from *A. suturalis*, and four OBPs from *Chinavia ubica*. The resulting phylogenetic tree showed that OBPs were segregated into two clades: ‘classic’ OBP and ‘Plus-C’ OBPs. OBPs from the same species but in different OBP family (classic vs Plus-C) were equally divergent from each other. Most *T. elegans* OBPs were located in the same branch along with the orthologous sequences. However TeleOBP9 and TeleOBP18 had no orthologous sequences and segregated into unique clades (Fig. 4).

We next constructed a phylogenetic tree of CSP protein sequences using 47 CSP sequences from five hemipteran species, including seven CSPs from *T. elegans*, eight CSPs from *A. suturalis*, eight CSPs from *A. lucorum*, 13 CSPs from *Lygus hesperus*, and 11 CSPs from *A. lineolatus*. CSPs from the same species were also equally divergent from each other (Fig. 5).

**Tissue-specific expression of *T. elegans* OBP and CSP genes.** To understand the function of TeleOBPs, we measured relative expression levels of OBP genes in different tissues of *T. elegans* by RT-qPCR. The antennal transcript levels of 14 TeleOBP genes (*TeleOBP1*, *TeleOBP2*, *TeleOBP3*, *TeleOBP4*, *TeleOBP5*, *TeleOBP6*, *TeleOBP9*, *TeleOBP10*, *TeleOBP11*, *TeleOBP12*, *TeleOBP13*, *TeleOBP14*, *TeleOBP15* and *TeleOBP19*) were significantly higher in both female and male antennae compared to other tissues. Furthermore, the expression levels of *TeleOBP2*, *TeleOBP5* and *TeleOBP6* were significantly higher in the antennae of males than the antennae of females. The expression levels of *TeleOBP1*, *TeleOBP4*, *TeleOBP9*, *TeleOBP13*, *TeleOBP14*, *TeleOBP15* and



**Figure 5.** Phylogenetic relationships of 47 chemosensory proteins (CSPs) from five hemipteran species: *Adelphocoris suturalis*, *Apolygus lucorum*, *Lygus Hesperus*, *Adelphocoris lineolatus*, *Tropidothorax elegans*.

*TeleOBP17* were significantly higher in the female antennae than the male antennae. Interestingly, *TeleOBP7*, *TeleOBP16* and *TeleOBP18* were highly expressed in the legs. In addition, the Plus-C OBP *TeleOBP8* had highest expression levels in the wings of female and male compared to other tissues (Fig. 6).

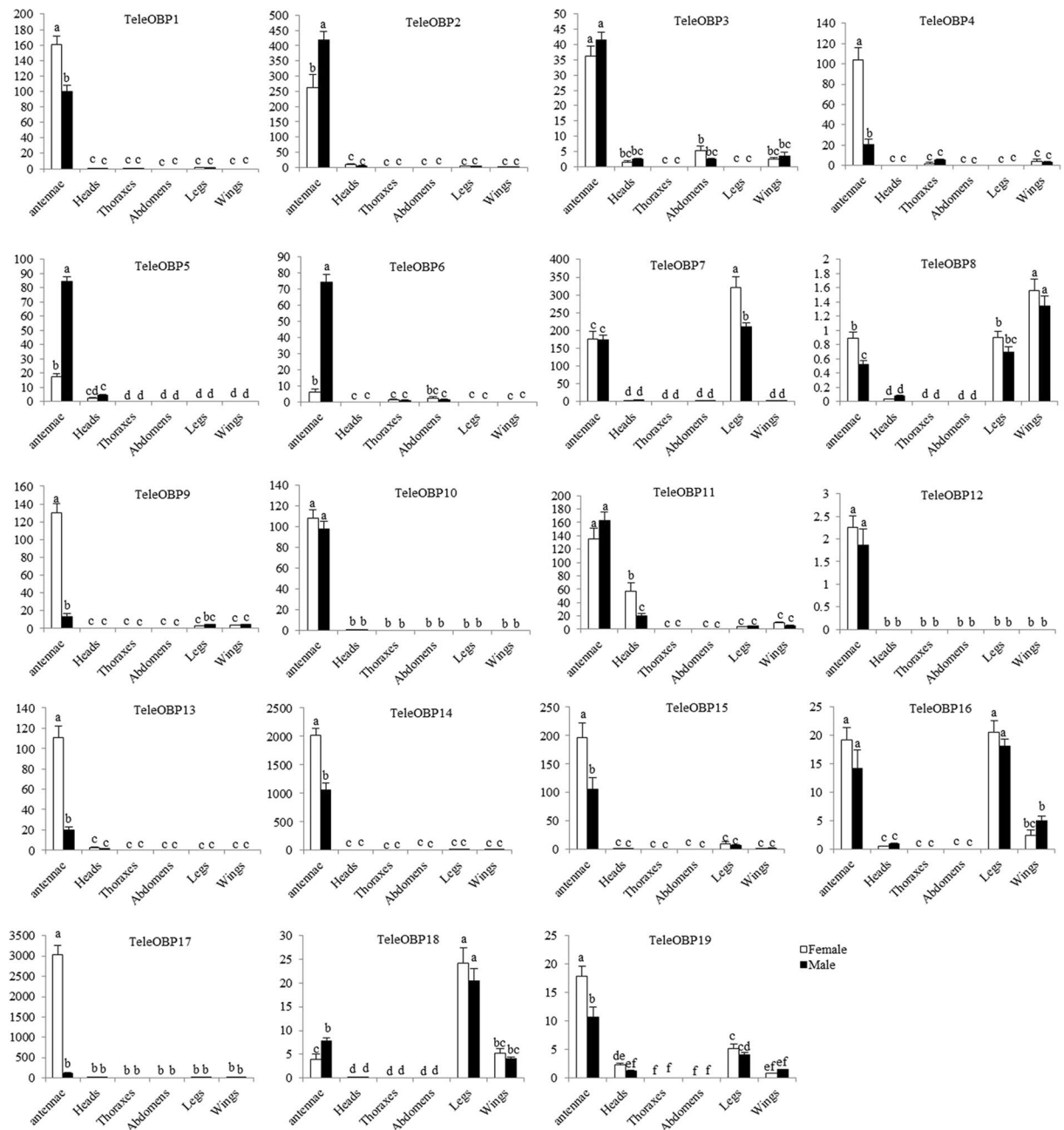
Of the seven *TeleCSP* genes, we found that *TeleCSP1*, *TeleCSP4* and *TeleCSP6* expression was highest in the antennae of females and males. *TeleCSP3*, *TeleCSP4* and *TeleCSP5* expression levels were significantly higher in the female antennae than the male antennae, while *TeleCSP1* and *TeleCSP6* expression levels were similar in females and in males. Interestingly, *TeleCSP2* was expressed in all olfactory tissues that were analyzed, which were antennae, head, legs and wings (Fig. 7).

## Discussion

Here, we identified 19 candidate OBP genes and seven CSP genes by searching the *T. elegans* antennal transcriptome. Although the number of OBP genes we identified in *T. elegans* is lower than other species, such as 38 OBPs in *A. lucorum*<sup>9</sup>, 33 OBPs in *L. lineolaris*<sup>8</sup> and 28 OBPs in *N. ericae*<sup>24</sup>, but was similar to some Hemiptera species, such as 16 OBPs in *A. suturalis*<sup>25</sup> and 14 OBPs in *A. lineolatus*<sup>26</sup>. The host range of *T. elegans* is broad, and there may be more OBP and CSP genes to detect odor molecules. Our analysis would not detect OBPs and CSPs primarily expressed in other tissues, or expressed at very low levels in the antennae, therefore the repertoire of *TeleOBPs* and *TeleCSPs* we identified may be incomplete.

The number of 'Plus-C' OBPs is variable even within the same order. For example, in five Hemiptera insects *L. lineolaris*, *N. ericae*, *A. lucorum*, *A. suturalis* and *A. lineolatus*, the number of 'Plus-C' OBPs is 10, 7, 6, 4 and 2, respectively<sup>8,9,24–26</sup>. Here, we identified five 'Plus-C' OBPs in *T. elegans*. The phylogenetic tree of Hemiptera OBPs revealed that except for *TeleOBP9*, all OBPs of *T. elegans* were assigned to orthologous Hemiptera OBPs clades. The 'classic' and 'Plus-C' OBP genes were clearly clustered in two clades, which may be due to the functional divergence of 'classic' OBP and 'Plus-C' OBP genes during evolution.

Expression patterns of OBPs can better help us to understand the function of these proteins. In Hemipterans, OBP genes are primarily expressed in the antenna. For example, 13 of 16 OBPs in *A. suturalis*<sup>[25]</sup>, 12 of 14 OBPs in *A. lineolatus*<sup>[26]</sup>, at least 6 of 10 OBPs in *Nilaparvata lugens*<sup>34</sup>, and 21 of 33 OBPs in *L. lineolaris*<sup>[8]</sup> are expressed in the antennae. We found that 15 of 19 OBPs in *T. elegans* showed highest expression levels in the antennae, indicating a vital olfactory role of OBPs. Furthermore, eight *TeleOBPs* showed female-biased expression and three *TeleOBPs* showed male-biased expression. The occurrence of sex-biased expression suggests that these OBP genes may be involved in recognition of plant volatiles for oviposition behavior or pheromones<sup>35–37</sup>. OBPs expressed in taste sensilla on legs regulate behavioral adaptation to the host plant in *Drosophila sechellia*<sup>38,39</sup>. *TeleOBP7*, *TeleOBP16* and *TeleOBP18* were highly expressed in the legs, which may be related to host plant adaptation of *T. elegans*. Interestingly, the Plus-C OBP *TeleOBP8* was expressed higher in the wings of females and male than

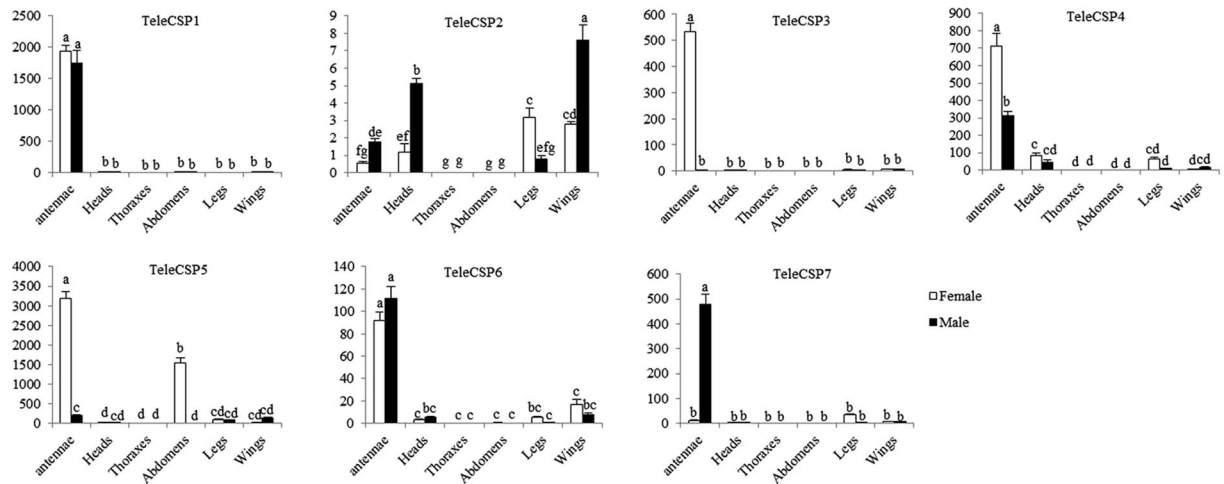


**Figure 6.** Relative expression of TeleOBP genes in female and male adult tissues of *Tropidothorax elegans* determined by RT-qPCR. TeleActin gene was used as an internal control. Data was normalized to actin expression levels and shown relative to the expression of female thoraxes each gene. Error bars represent standard errors of the means for three biological replicates and different lower cases letters indicate significant differences ( $p < 0.05$ ).

other tissues. OBP expression in non-olfactory tissues, which have been observed in a variety of species, suggests that they also may function as carriers of chemicals during developmental and physiological processes<sup>9</sup>.

In general, there are fewer CSP genes in Hemiptera insects than OBP genes. For example, there are 28 OBPs and 16 CSPs in *N. ericae*<sup>24</sup>, 16 OBPs and 8 CSPs in *A. suturalis*<sup>25</sup>, 14 OBPs and 3 CSPs in *A. lineolatus*<sup>26</sup>, 10 OBPs and 5 CSPs in *C. lividipennis*<sup>28</sup>, 40 OBPs and 11 CSPs in *E. onukii*<sup>29</sup>, and 9 OBPs and 3 CSPs in *P. micranthus*<sup>32</sup>. We identified seven CSPs in the antennal transcriptome of *T. elegans*. In our phylogenetic analysis, we found that there is low amino acid sequence conservation of CSPs among true bugs. Only one of seven TeleCSPs, TeleCSP6, clustered with AlucCSP3 with over 80% amino acid sequence identity. The distribution of CSP orthologs in other hemipterans suggests that CSP genes originating from a common ancestor may have similar functions or that they acquired novel functions via subfunctionalization<sup>9</sup>. These results suggest that CSP proteins in hemipteran





**Figure 7.** Relative expression of TeleCSP genes in female and male adult tissues of *Tropidothorax elegans* determined by RT-qPCR. TeleActin gene was used as an internal control. Data was normalized to actin expression levels and shown relative to the expression of female thoraxes each gene. Error bars represent standard errors of the means for three biological replicates and different lower cases letters indicate significant differences ( $p < 0.05$ ).

insects undergo extensive gene duplication and divergence by natural selection, strongly indicating that they may have diverse functions<sup>40</sup>.

CSPs play an important role in a variety of biological process, including chemosensation<sup>41</sup>, leg regeneration<sup>42,43</sup>, and embryonic development<sup>44</sup>. *TeleCSP1*, *TeleCSP4* and *TeleCSP6* were expressed highly in the antennae, and maybe involved in recognizing sex pheromones and plant volatiles<sup>27,45,46</sup>. *TeleCSP2* was widely expressed in chemosensory and non-chemosensory tissues (antennae, head, legs and wings) and may have different physiological functions in adult tissues.

Overall, we generated the transcriptome of the female and male antenna of *T. elegans* by next-generation high throughput sequencing, and identified 19 OBP and seven CSP genes. Furthermore, we identified the gene expression patterns of CSPs and OBPs in different adult tissues. These findings provide important insights into the function of OBP and CSPs, and their role in odorant reception.

## Materials and Methods

**Insects samples and RNA extraction.** A laboratory strain of *T. elegans* was generated from a population collected on a vegetable field in Luoyang, Henan, China (112°26'E, 34°43'N) in 2014. The population was reared on *Metaplexis japonica* (Thunb.) in a greenhouse maintained at  $25 \pm 2$  °C, 14 h: 10 h light/dark cycle with 60–80% relative humidity. For the transcriptome, about 500 pairs of 3–4 days old female and male adult antennae were dissected, immediately frozen in liquid nitrogen, and stored at  $-80$  °C until RNA isolation.

Total RNA was extracted using the RNAiso Plus kit (TaKaRa, Dalian, China) and treated with RNase-free DNase I (TaKaRa, Dalian, China) to remove residual DNA. The quantity and integrity of RNA was tested using 1.0% agarose gel electrophoresis and NanoDrop 2000c spectrophotometer (Thermo Scientific, USA), Qubit 2.0 (Life Technologies, USA) and Agilent 2100 (Agilent, USA).

**Antennal cDNA library construction, sequencing and analysis.** Following the TruSeq RNA Sample Preparation Guide v2 (Illumina), mRNA was enriched using magnetic beads crosslinked to Oligo (dT), and fragmented into small pieces using the fragmentation buffer. First-strand cDNA was synthesized using small mRNA fragments with random primers and reverse transcriptase, and second-strand cDNA synthesis was conducted by adding dNTPs, DNA polymerase I and RNase H. Next, double stranded cDNA was purified with AMPure XP beads (Beckman Coulter, USA), and treated for end-repairing, Poly-A tailing and sequencing adapters linking. The size of the fragment was chosen using AMPure XP beads and the cDNA library was constructed by PCR amplification (Veriti<sup>®</sup> 96-Well Thermal Cycle, Applied Biosystems, USA). The concentration and insert size of the cDNA library were detected using Qubit 2.0 and Agilent 2100, and quantified with q-PCR (CFX-96, Bio-Rad, USA).

Sequencing was performed by Genomics Services Lab of the Beijing Novogene Technologies Co., Ltd. (Beijing, China) using the Illumina HiSeq<sup>™</sup> 4000 platform to generate 150 bp pair-end reads. The raw data processing and base calling were performed using the Illumina instrument software.

Homologous sequences were searched using BLASTx and BLASTn against the Nr (non-redundant protein database) and Nt (non-redundant nucleotide sequence database) in NCBI with an E-value cut-off of  $1.0 \times 10^{-5}$ .

**Identification OBP of and CSP genes.** Putative TeleOBP and TeleCSP genes were identified by searching odorant binding protein and chemosensory protein keywords in the annotated unigenes, and by using BLAST. Annotated OBP and CSP genes from other Hemipteran species, such as *A. suturalis* Jakovlev, *A. lineolatus* Goeze,

*A. lucorum* Meyer-Dür, *N. ericae* Schilling, *Halyomorpha halys* Stål, and *L. lineolaris* Palisot de Beauvois were used as references. The putative *T. elegans* OBP and CSP genes were confirmed by searching against the NCBI non-redundant (nr) protein database using BlastX with cut-off *E*-value of 10<sup>-5</sup>.

**Bioinformatic analysis.** Signal peptides were predicted using SignalP 4.1 server<sup>47</sup> (<http://www.cbs.dtu.dk/services/SignalP/>). The similarity searches were performed using NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/>). Multiple sequence alignment was conducted using DNAMAN 6.0. Amino acid sequence alignments of the matured OBPs and CSPs from *T. elegans* and other Hemipteran species were performed using MAFFT (<http://mafft.cbrc.jp/alignment/server/clustering.html>), and phylogenetic trees were constructed using PhyML<sup>48</sup> in Seaview v.4 using the Jones-Taylor-Thomton (JTT) model with 1000-fold bootstrap replication in neighbor-joining method (NJ).

**Tissue specific expression of OBPs and CSPs.** Antennae (300 pairs), heads (without antennae) (180), thoraxes (120), abdomens (50), legs (300 pairs) and wings (200 pairs) of male and female adults at 3-day after eclosion were excised and immediately frozen in liquid nitrogen. All total RNA samples were extracted using the RNAsiso Plus kit (TaKaRa, Dalian, China) and the isolated RNA was transcribed to first-strand cDNA by PrimeScript<sup>TM</sup> RT reagent Kit with gDNA Eraser (TaKaRa, Dalian, China) following the manufacturer's instructions. The nucleotide sequences of all 19 TeleOBPs and 7 TeleCSPs were confirmed by cloning and sequencing (Figure S1). Real-time quantitative PCRs (RT-qPCRs) were performed with SYBR<sup>®</sup> Premix Ex Taq<sup>TM</sup> II (TaKaRa, Dalian, China). The *T. elegans* actin gene (Genbank accession no. MG322127) was used as control. Each reaction was performed with 200 ng/μl cDNA sample. Primers for RT-qPCR were designed using Primer Premier 5.0 software and are listed in Table S2. The RT-qPCR reactions were conducted in 20 μL reaction mixtures containing 10 μL SYBR Premix Ex Taq II, 20 ng cDNA templates, 0.2 μM of each primer, and nuclease-free water. The cycling conditions were: one cycle of 95 °C for 5 min, followed by 40 cycles of 95 °C for 5 s and 55 °C for 30 s. Melt curve conditions were 95 °C for 10 s, 65 °C for 30 s. A no-template control (NTC) was also included to detect possible contamination. Three biological replicates were analysed and relative expression levels of OBP and CSP genes across the samples were measured using the 2<sup>-ΔΔCT</sup> method<sup>49</sup>. The differences in the expression of TeleOBP and TeleCSP genes between female and male tissues were compared by a one-way nested analysis of variance (ANOVA), followed by Tukey's honestly significance difference (HSD) test using SPSS (SPSS Institute 17.0, IBM, Chicago, IL, USA).

## References

- Field, L. M., Pickett, J. A. & Wadhams, L. J. Molecular studies in insect olfaction. *Insect Mol. Biol.* **9**, 545–551 (2000).
- Zhang, J., Walker, W. B. & Wang, G. Pheromone reception in moths: from molecules to behaviors. *Prog. Mol. Biol. Transl. Sci.* **130**, 109–128 (2015).
- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391 (2013).
- Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell Mol. Life Sci.* **63**, 1658–1676 (2006).
- Zhou, J. J. Odorant-binding proteins in insects. *Vitam. Horm.* **83**, 241–272 (2010).
- Zhou, W. *et al.* Identification and expression profiling of putative chemosensory protein genes in two rice planthoppers, *Laodelphax striatellus* (Fallén) and *Sogatella furcifera* (Horváth). *J. Asia Pac. Entomol.* **18**, 771–778 (2015).
- Tegoni, M., Campanacci, V. & Cambillau, C. Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem. Sci.* **29**, 257–264 (2004).
- Hull, J. J., Perera, O. P. & Snodgrass, G. L. Cloning and expression profiling of odorant-binding proteins in the tarnished plant bug, *Lygus lineolaris*. *Insect Mol. Biol.* **23**, 78–97 (2014).
- Yuan, H. B. *et al.* Molecular characterization and expression profiling of odorant-binding proteins in *Apolygus lucorum*. *PLoS ONE* **10**, e0140562 (2015).
- Li, S. *et al.* Multiple functions of an odorant-binding protein in the mosquito *Aedes aegypti*. *Biochem. Biophys. Res. Commun.* **372**, 464–468 (2008).
- Sun, Y. L., Huang, L. Q., Pelosi, P. & Wang, C. Z. Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling Helicoverpa species. *PLoS ONE* **7**, e30040 (2012).
- Pelosi, P., Iovinella, I. & Felicioli, A. Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* **5**, 320 (2014).
- McKenna, M. P., Hekmat-Scafe, D. S., Gaines, P. & Carlson, J. R. Putative Drosophila pheromone-binding proteins expressed in a subregion of the olfactory system. *J. Biol. Chem.* **269**, 16340–16347 (1994).
- Pikielny, C., Hasan, G., Rouyer, F. & Rosbash, M. Members of a family of Drosophila putative odorant-binding proteins are expressed in different subsets of olfactory hairs. *Neuron*. **12**, 35–49 (1994).
- Sun, H. Z., Song, Y. Q., Du, J., Wang, X. D. & Cheng, Z. J. Identification and tissue distribution of chemosensory protein and odorant binding protein genes in *Athetis dissimilis* (Lepidoptera: Noctuidae). *Appl. Entomol. Zool.* **51**, 409–420 (2016).
- Maleszka, R. & Stange, G. Molecular cloning, by a novel approach, of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Cactoblastis cactorum*. *Gene* **202**, 39–43 (1997).
- Nagnan-Le Meillour, P. *et al.* Chemosensory proteins from the proboscis of *Mamestra brassicae*. *Chem. Senses*. **25**, 541–553 (2000).
- Jin, X. *et al.* Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell Mol. Life Sci.* **62**, 1156–1166 (2005).
- Zhou, S. H., Zhang, J., Zhang, S. G. & Zhang, L. Expression of chemosensory proteins in hairs on wings of *Locusta migratoria* (Orthoptera: Acrididae). *J. Appl. Entomol.* **132**, 439–450 (2008).
- Ban, L., Zhang, L., Yan, Y. & Pelosi, P. Binding properties of a locust's chemosensory protein. *Biochem. Biophys. Res. Commun.* **293**, 50–54 (2002).
- Jacquin-Joly, E., Vogt, R. G., François, M. C. & Nagnan-Le Meillour, P. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. *Chem. Senses* **26**, 833–844 (2001).
- Dani, F. R. *et al.* Odorant-binding proteins and chemosensory proteins in pheromone detection and release in the silkworm *Bombyx mori*. *Chem. Senses* **36**, 335–344 (2011).
- Dyanov, H. & Dzitoeva, S. Method for attachment of microscopic preparations on glass for *in situ* hybridization, PRINS and *in situ* PCR studies. *Biotechniques* **18**, 822–826 (1995).
- Zhang, Y. N. *et al.* De novo assembly and characterization of antennal transcriptome reveal chemosensory system in *Nysius ericae*. *J. Asia Pac. Entomol.* **19**, 1077–1087 (2016).



25. Cui, H. H. *et al.* Odorant-binding and chemosensory proteins identified in the antennal transcriptome of *Adelphocoris suturalis* Jakovlev. *Comp. Biochem. Phys. D* **24**, 139–145 (2017).
26. Gu, S. H. *et al.* Identification and tissue distribution of odorant binding protein genes in the lucerne plant bug *Adelphocoris lineolatus* (Goeze). *Insect Biochem. Molec.* **41**, 254–263 (2011).
27. Gu, S. H. *et al.* Functional characterizations of chemosensory proteins of the alfalfa plant bug *Adelphocoris lineolatus* indicate their involvement in host recognition. *PLoS ONE* **7**, e42871 (2012).
28. Wang, G. Y. *et al.* Identification of candidate odorant-binding protein and chemosensory protein genes in *Cyrtorhinus lividipennis* (Hemiptera: Miridae), a key predator of the rice planthoppers in Asia. *Environ. Entomol.* **46**, 654–662 (2017).
29. Bian, L. *et al.* Identification of the genes in tea leafhopper, *Empoasca onukii* (Hemiptera: Cicadellidae), that encode odorant-binding proteins and chemosensory proteins using transcriptome analyses of insect heads. *Appl. Entomol. Zool.* **53**, 93–105 (2018).
30. Wang, R. *et al.* Identification and expression profile analysis of odorant binding protein and chemosensory protein genes in *Bemisia tabaci* MED by head transcriptome. *PLoS ONE* **12**, e0171739 (2017).
31. Paula, D. P. *et al.* Identification and expression profile of odorant-binding proteins in *Halyomorpha halys* (Hemiptera: Pentatomidae). *Insect Mol. Biol.* **25**, 580–594 (2016).
32. Liu, N. Y., Zhu, J. Y., Ji, M., Yang, B. & Ze, S. Z. Chemosensory genes from *Pachypeltis micranthus*, a natural enemy of the climbing hemp vine. *J. Asia-Pac. Entomol.* **20**, 655–664 (2017).
33. Li, X. Y. *et al.* Recognition of the nymph instars of eight common harmful stinkbug species in the fields. *Plant Protection in China* **41**, 100–105 (2015).
34. Zhou, S. S., Sun, Z., Ma, W., Chen, W. & Wang, M. Q. De novo analysis of the *Nilaparvata lugens* (Stål) antenna transcriptome and expression patterns of olfactory genes. *Comp. Biochem. Phys. D* **9**, 31–39 (2014).
35. Shiojiri, K., Takabayashi, J., Yano, S. & Takafuji, A. Flight response of parasitoids toward plant-herbivore complexes: a comparative study of two parasitoid herbivore systems on cabbage plants. *Appl. Entomol. Zool.* **35**, 87–92 (2000).
36. Poivet, E. *et al.* A comparison of the olfactory gene repertoires of adults and larvae in the noctuid moth *Spodoptera littoralis*. *PLoS One* **8**, e60263 (2013).
37. Li, L., Liu, Z. & Sun, J. Olfactory cues in host and host-plant recognition of a polyphagous ectoparasitoid *Scleroderma guani*. *Bio. Control* **60**, 307–316 (2015).
38. Yasukawa, J., Tomioka, S., Aigaki, T. & Matsuo, T. Evolution of expression patterns of two odorant binding protein genes, Obp57d and Obp57e, in *Drosophila*. *Gene* **467**, 25–34 (2010).
39. Harada, E. *et al.* Functional evolution of duplicated odorant-binding protein genes, Obp57d and Obp57e, in *Drosophila*. *PLoS ONE* **7**, e29710 (2012).
40. He, M. & He, P. Molecular characterization, expression profiling, and binding properties of odorant binding protein genes in the white backed planthopper, *Sogatella furcifera*. *Comp. Biochem. Phys. B* **174**, 1–8 (2014).
41. Sun, Y. L., Huang, L. Q., Pelosi, P. & Wang, C. Z. Expression in antennae and reproductive organs suggests dual role of an odorant-binding protein in two sibling *Helicoverpa* species. *PLoS ONE* **7**, e30040 (2012).
42. Nomura, A., Kawasaki, K., Kubo, T. & Natori, S. Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Int. J. Dev. Biol.* **36**, 391–398 (1992).
43. Kitabayashi, A. N., Arai, T., Kubo, T. & Natori, S. Molecular cloning of cDNA for p10, a novel protein that increases in the regenerating legs of *Periplaneta americana* (American cockroach). *Insect Biochem. Mol. Biol.* **28**, 785–790 (1998).
44. Maleszka, J., Foret, S., Saint, R. & Maleszka, R. RNAi-induced phenotypes suggest a novel role for a chemosensory protein CSP5 in the development of embryonic integument in the honeybee (*Apis mellifera*). *Dev. Genes Evol.* **217**, 189–196 (2007).
45. Iovinella, I., Bozza, F., Caputo, B., della Torre, A. & Pelosi, P. Ligand-binding study of *Anopheles gambiae* chemosensory proteins. *Chem. Senses* **38**, 409–419 (2013).
46. Zhang, Y. N., Ye, Z. F., Yang, K. & Dong, S. L. Antenna-predominant and male-biased CSP19 of *Sesamia inferens* is able to bind the female sex pheromones and host plant volatiles. *Gene* **536**, 279–286 (2014).
47. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
48. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
49. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-ΔΔCt)</sup> method. *Methods* **25**, 402–408 (2001).

## Acknowledgements

This work was supported by Natural Science Foundation of China (31701788) and the Startup Project of Doctoral scientific research (4026-13480047) of Henan University of Science and Technology.

## Author Contributions

Conceived and designed the experiments: H.Z.S. Sample collection: J.D. Performed the experiments, data analysis and wrote the paper: Y.Q.S. and H.Z.S.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-26137-6>.

**Competing Interests:** The authors declare no competing interests.

**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) 2018