


# SCIENTIFIC REPORTS



OPEN

## Identification of Genes Involved in Chemoreception in *Plutella xylostella* by Antennal Transcriptome Analysis

Shiyong Yang<sup>1</sup>, Depan Cao<sup>2</sup>, Guirong Wang<sup>2</sup> & Yang Liu<sup>1</sup> 

Perception of environmental and habitat cues is of significance for insect survival and reproduction. Odor detection in insects is mediated by a number of proteins in antennae such as odorant receptors (ORs), ionotropic receptors (IRs), odorant binding proteins (OBPs), chemosensory proteins (CSPs), sensory neuron membrane proteins (SNMPs) and odorant degrading enzymes. In this study, we sequenced and assembled the adult male and female antennal transcriptomes of a destructive agricultural pest, the diamondback moth *Plutella xylostella*. In these transcriptomes, we identified transcripts belonging to 6 chemoreception gene families related to odor detection, including 54 ORs, 16 IRs, 7 gustatory receptors (GRs), 15 CSPs, 24 OBPs and 2 SNMPs. Semi-quantitative reverse transcription PCR analysis of expression patterns indicated that some of these ORs and IRs have clear sex-biased and tissue-specific expression patterns. Our results lay the foundation for future characterization of the functions of these *P. xylostella* chemosensory receptors at the molecular level and development of novel semiochemicals for integrated control of this agricultural pest.

Olfaction plays a pivotal role in intra- and inter-specific interactions by directing insects towards food or prey, mating partners, oviposition sites, and away predators as well as toxic compounds<sup>1</sup>. The specialized organ for olfaction in insects is the antenna, on which hair-like, multi-pore sensilla are situated and peripheral olfactory signaling events occur. Olfactory receptor neurons (ORNs) and their auxiliary structures are located at the roots of the antennae<sup>2</sup>, and the entire olfactory system is dependent to a great extent on receptors expressed at the peripheral ORNs. Starting with perception of semiochemicals and ultimately ending with the translation of olfactory signals into behavior, the entire process requires orchestration of the insect's sophisticated olfactory system at various levels. Several types of olfactory proteins are believed to participate in the selective detection and, once they have conveyed information, the rapid inactivation of trace amount of odorants, i.e. odorant receptors (ORs), ionotropic receptors (IRs), gustatory receptors (GRs), odorant binding proteins (OBPs), chemosensory proteins (CSPs) and sensory neuron membrane proteins (SNMPs)<sup>3</sup>.

Insect ORs are seven-transmembrane domain proteins with a reversed topology compared to the G-protein coupled ORs in vertebrates<sup>4,5</sup>. ORs play a central role in converting semiochemicals into electrical signal, functioning as a heterodimer with a divergent, conventional ORx and a highly conserved noncanonical OR co-receptor Orco in fruit fly, OR2 in moths and OR7 in mosquitoes<sup>3</sup>. The OR genes are expressed in the olfactory neurons housed within the olfactory sensilla (found mainly on the antenna)<sup>6</sup>.

GRs are also seven-transmembrane domain proteins, but they are more ancient than ORs. GR genes are expressed in the gustatory neurons housed within the gustatory sensilla (found on the labia, maxillary palps, antennae, legs and genitalia)<sup>7</sup>. GRs can respond to tastants such as sugars, bitter substances, CO<sub>2</sub> and some contact pheromones<sup>8–11</sup>.

IRs belong to the ionotropic glutamate receptor (iGluR)-like protein family and can be activated by small molecules like acetates and amine-like volatile compounds<sup>12,13</sup>. It has been proven that IRs are involved in chemosensation<sup>14,15</sup> and other functions, i.e. regulation of the circadian clock in *Drosophila melanogaster*<sup>16</sup> and induction

<sup>1</sup>Anhui Provincial Key Laboratory of the Conservation and Exploitation of Biological Resources, College of Life Sciences, Anhui Normal University, Wuhu, 241000, China. <sup>2</sup>State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, China. Correspondence and requests for materials should be addressed to Y.L. (email: [yangliu@ippcaas.cn](mailto:yangliu@ippcaas.cn))

	Sample	Total Number	Total Length (nt)	Mean Length (nt)	N50 (nt)	Consensus Sequences	Distinct Clusters	Distinct Singletons
Contig	Female	124,488	34,667,373	278	403	—	—	—
	Male	132,190	35,402,665	268	369	—	—	—
Unigene	Female	62,278	34,543,989	555	829	62,278	16,328	45,950
	Male	63,928	33,941,348	531	761	63,928	15,969	47,959
	All	59,844	39,492,885	660	979	59,844	18,570	41,274

**Table 1.** Summary of the *Plutella xylostella* transcriptome assembly.

of physical defense in *Daphnia pulex*<sup>17</sup>. IRs usually contain three transmembrane domains (TMDs), a bipartite ligand-binding domain with two lobes and one ion channel, and have been proposed to act as dimmers or trimers of subunits coexpressed in the same neuron<sup>12</sup>. However, they aren't expressed in chemosensory neurons that express ORs or Orco<sup>14</sup>.

OBPs are the liaisons between external cues and ORs<sup>18</sup>, and they selectively bind hydrophobic odorant chemicals and transport them to the surface of the dendrites of ORNs<sup>19–21</sup>. OBPs also function in the recognition of specific odors through activation of the ORx/Orco complex<sup>20</sup>. Another class of odorant binding proteins, CSPs, are small soluble proteins expressed predominantly in the sensillum lymph as well as in non-olfactory tissues. It is clear that CSPs bind odorant or pheromone compounds<sup>22–24</sup>, but their olfactory mechanisms areas yet poorly studied.

SNMPs are insect membrane proteins that are known to associate with pheromone sensitive ORNs in Lepidoptera and Diptera<sup>25</sup>. There are two types of SNMPs, SNMP1 and SNMP2<sup>25</sup>. In moth, the subtype SNMP1 is coexpressed with pheromone receptors (PRs) in pheromone-responsive neurons<sup>25</sup>, whereas the subtype SNMP2 is confined to sensilla support cells<sup>25–28</sup>.

The diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae), is a destructive insect pest distributed worldwide that can cause considerable damage in cruciferous crops. It is estimated that the total loss caused by *P. xylostella* is about US\$4–5 billion annually<sup>29</sup>. Although a bioinformatics analysis of the whole-genome sequence has explained the evolutionary success of *P. xylostella* with regard to its expansion in gene families associated with the perception and detoxification of plant defense compounds/insecticides at the genetic and molecular levels<sup>30</sup>, the peripheral olfactory mechanisms that contribute to the fitness of this insect pest remain poorly understood. Identification of genes expressed in the antennae will supply baseline information to understand their likely function in odorant perception in *P. xylostella* and insects adaptation to various host plants.

In the present study, we sequenced and analyzed the antennal transcriptome of *P. xylostella* adults using second-generation high-throughput Illumina RNA sequencing (RNA-seq). The purpose of our study was to identify olfaction-related genes which might be targets as a part of pest control strategies of this insect pest species that devastates cruciferous vegetables. We identified 118 candidate chemosensory genes encoding 54 ORs, 16 IRs, 7 GRs, 15 CSPs, 24 OBPs and 2 SNMPs. The sex-biased and tissue-specific expression patterns of 54 ORs and 16 IRs was also determined by semi-quantitative reverse transcription PCR. We reported the protein sequences of these chemosensory genes in Supplementary Dataset File.

## Results

**Sequencing and unigene assembly.** By using Hiseq. 2000 sequencing approach, a total of 60,041,232 and 59,753,272 raw reads were obtained from the *P. xylostella* female and male antennae samples, respectively. After removing low quality and adaptor reads, female and male antennae yielded 54,430,716 and 54,059,300 clean reads and 4,898,764,440 nt and 4,865,337,000 nt clean nucleotides, respectively. After initial assembly, 124,488 (mean length 278 nt) and 132,190 contigs (mean length 268 nt) were obtained from the female and male antennae libraries, respectively. Next, 62,278 female (mean length 555 nt) and 63,928 male unigenes (mean length 531 nt) were generated after contig connecting. These two unigene sets were then pooled together for further clustering, which yielded a final set of 59,844 unigenes consisting of 18,570 distinct clusters and 41,274 distinct singletons. The mean length of these unigenes was 660 nt, and N50 was 979 nt (Table 1).

**Identification of candidate chemosensory receptors: ORs and GRs.** All the unigenes were searched by blastx against nr database and further by tblastn using 63 ORs from *B. mori* as queries, 54 candidate OR genes were identified (Table 2). Of these, 23 were predicted to have full-length open reading frames (ORFs). The length of these 23 OR genes ranges from 376 to 473 amino acid residues, and the encoded proteins are estimated to have 5–7 TMDs, which is characteristic of typical insect ORs. The remaining 31 OR genes code for at least 163 amino acids and are predicted to have more than one TMD. A phylogenetic analysis was then performed using our candidate ORs and the ORs from other Lepidopteran insects including *H. armigera*, *H. virescens* and *B. mori* (Fig. 1).

The OR co-receptor gene was easily identified because of extremely high conservation among species compared to other chemosensory receptors. Similar to other insect ORs, most *P. xylostella* (Pxyl) ORs are highly divergent and share low similarity with other Lepidopteran insect ORs, including ORs from *H. armigera*, *H. virescens* and *B. mori*. However, nine PxylORs had 33%–100% identity to previously characterized PRs from *P. xylostella* and *B. mori*. They formed a single subgroup in a phylogenetic tree of Lepidopteran ORs (Fig. 1). Seven of these nine PxylORs (PxylOR1 and PxylOR3–8) were predicted to have full-length ORFs. Two short sequences (PxylOR41 and PxylOR45) were also clustered in the PR branch. PxylOR41 has high similarity to PxylOR4, and PxylOR45 has relatively high similarity to BmorOR6. 12 of the remaining PxylORs were clustered with their

Unigene reference	Name	Length(bp)	ORF(aa)	Blastx best hit (Reference/ Name/Species)	E value	Identity	TMD (No)	Status
<b>Co-receptor</b>								
Unigene25399	PxylOR2	2187	473	dbj BAG71421.2  olfactory receptor-2 [ <i>Plutella xylostella</i> ]	0	1	7	Complete
<b>Pheromone receptors</b>								
CL4851.Contig2	PxylOR1	1800	422	dbj BAG71420.1  olfactory receptor-1 [ <i>P. xylostella</i> ]	0	1	6	Complete
CL902.Contig17	PxylOR3	1650	402	dbj BAG71425.2  olfactory receptor [ <i>P. xylostella</i> ]	0	0.99	5	Complete
CL902.Contig2	PxylOR4	1595	402	dbj BAG71426.1  olfactory receptor [ <i>P. xylostella</i> ]	0	0.95	7	Complete
CL902.Contig3	PxylOR5	1630	404	dbj BAG71426.1  olfactory receptor [ <i>P. xylostella</i> ]	0	0.82	6	Complete
Unigene18038	PxylOR6	1584	424	dbj BAG71426.1  olfactory receptor [ <i>P. xylostella</i> ]	3.00E-129	0.48	7	Complete
CL3732.Contig1	PxylOR7	1415	424	dbj BAG71425.2  olfactory receptor [ <i>P. xylostella</i> ]	5.00E-107	0.42	7	Complete
CL3275.Contig3	PxylOR8	1717	427	dbj BAG71425.2  olfactory receptor [ <i>P. xylostella</i> ]	3.00E-129	0.63	6	Complete
CL902.Contig18	PxylOR41	580	193	dbj BAG71426.1  olfactory receptor [ <i>P. xylostella</i> ]	1.00E-83	0.77	1	5', 3' lost
Unigene8020	PxylOR45	568	189	ref NP_001036928.1  olfactory receptor 6 [ <i>Bombyx mori</i> ]	3.00E-27	0.33	3	5', 3' lost
<b>Olfactory receptors</b>								
CL1915.Contig1	PxylOR9	1466	449	ref NP_001116817.1  olfactory receptor-like [ <i>B. mori</i> ]	5.00E-145	0.59	6	Complete
CL1947.Contig5	PxylOR10	1602	428	gb AFC91732.1  putative odorant receptor OR24 [ <i>Cydia pomonella</i> ]	4.00E-127	0.45	7	Complete
Unigene8291	PxylOR11	1369	421	ref NP_001166621.1  olfactory receptor 64 [ <i>B. mori</i> ]	2.00E-73	0.5	6	Complete
Unigene25275	PxylOR12	1340	420	gb AFC91725.1  putative odorant receptor OR17 [ <i>C. pomonella</i> ]	1.00E-97	0.51	6	5' lost
CL6791.Contig2	PxylOR13	1396	415	emb CAD31949.1  putative chemosensory receptor 8 [ <i>Heliothis virescens</i> ]	1.00E-124	0.49	7	5' lost
CL6176.Contig1	PxylOR14	1451	412	emb CAG38121.2  putative chemosensory receptor 20 [ <i>H. virescens</i> ]	1.00E-137	0.53	7	Complete
CL3142.Contig2	PxylOR15	1579	409	ref NP_001091789.1  olfactory receptor 15 [ <i>B. mori</i> ]	4.00E-76	0.39	7	5' lost
CL2401.Contig2	PxylOR16	1257	405	gb AFC91721.1  putative odorant receptor OR12 [ <i>C. pomonella</i> ]	2.00E-166	0.58	6	Complete
Unigene19920	PxylOR17	1722	399	gb AFC91726.1  putative odorant receptor OR18 [ <i>C. pomonella</i> ]	1.00E-120	0.45	7	Complete
Unigene3520	PxylOR18	1367	396	tpg DAA05974.1  TPA_exp: odorant receptor 15 [ <i>B. mori</i> ]	3.00E-94	0.4	7	Complete
Unigene5731	PxylOR19	1294	395	ref NP_001166617.1  olfactory receptor 56 [ <i>B. mori</i> ]	8.00E-145	0.53	7	Complete
CL6714.Contig1	PxylOR20	1362	393	ref NP_001091789.1  olfactory receptor 15 [ <i>B. mori</i> ]	1.00E-80	0.37	6	Complete
CL2099.Contig4	PxylOR21	1751	393	ref NP_001166892.1  olfactory receptor 36 [ <i>B. mori</i> ]	4.00E-34	0.24	7	Complete
CL2099.Contig5	PxylOR22	1606	393	ref NP_001166892.1  olfactory receptor 36 [ <i>B. mori</i> ]	9.00E-39	0.26	7	Complete
CL2363.Contig1	PxylOR23	1265	392	tpg DAA05974.1  TPA_exp: odorant receptor 15 [ <i>B. mori</i> ]	5.00E-90	0.4	7	Complete
CL918.Contig2	PxylOR24	1222	391	ref NP_001166892.1  olfactory receptor 36 [ <i>B. mori</i> ]	5.00E-35	0.27	7	Complete
Unigene25128	PxylOR25	1219	389	ref NP_001166892.1  olfactory receptor 36 [ <i>B. mori</i> ]	5.00E-47	0.3	6	Complete
Unigene5953	PxylOR26	1156	385	gb EHJ78030.1  olfactory receptor 29 [ <i>Danaus plexippus</i> ]	6.00E-141	0.63	6	3' lost
Unigene5680	PxylOR27	1314	376	gb EHJ64733.1  olfactory receptor 18 [ <i>D. plexippus</i> ]	2.00E-136	0.55	7	Complete
Continued								

Unigene reference	Name	Length(bp)	ORF(aa)	Blastx best hit (Reference/ Name/Species)	E value	Identity	TMD (No)	Status
CL1359.Contig2	PxylOR28	1737	359	ref NP_001091790.1  candidate olfactory receptor [B. mori]	1.00E-71	0.33	6	5' lost
CL6074.Contig2	PxylOR29	1214	356	emb CAG38113.1  putative chemosensory receptor 12 [H. virescens]	9.00E-65	0.38	6	5', 3'lost
CL2099.Contig6	PxylOR30	1140	301	ref NP_001166892.1  olfactory receptor 36 [B. mori]	2.00E-34	0.28	5	5' lost
Unigene14039	PxylOR31	949	279	ref NP_001166611.1  olfactory receptor 59 [B. mori]	3.00E-56	0.38	2	5' lost
Unigene11354	PxylOR32	835	277	gb EHJ65925.1  olfactory receptor 12 [D. plexippus]	8.00E-62	0.45	4	5', 3'lost
CL741.Contig1	PxylOR33	927	272	gb AFC91717.1  putative odorant receptor OR7, partial [C. pomonella]	2.00E-41	0.4	4	5' lost
Unigene600	PxylOR34	862	270	tpg DAA05988.1  TPA_exp: odorant receptor 32 [B. mori]	2.00E-30	0.33	4	3' lost
CL4545.Contig1	PxylOR35	824	269	tpg DAA05974.1  TPA_exp: odorant receptor 15 [B. mori]	3.00E-56	0.39	5	5' lost
Unigene17021	PxylOR36	768	252	gb ACH69152.1  olfactory receptor 49 [B. mori]	8.00E-120	0.68	5	5' lost
Unigene21064	PxylOR37	706	235	gb AFC91721.1  putative odorant receptor OR12 [C. pomonella]	4.00E-32	0.39	4	5', 3' lost
CL7033.Contig1	PxylOR38	646	215	ref NP_001166892.1  olfactory receptor 36 [B. mori]	1.00E-28	0.37	3	5', 3' lost
Unigene25541	PxylOR39	613	204	gb AFC91719.1  putative odorant receptor OR10 [C. pomonella]	3.00E-69	0.55	3	5', 3' lost
Unigene3305	PxylOR40	601	200	gb AFC91724.1  putative odorant receptor OR16 [C. pomonella]	6.00E-70	0.66	4	5', 3' lost
Unigene21899	PxylOR42	581	193	ref NP_001104832.2  olfactory receptor 16 [B. mori]	5.00E-70	0.66	3	5', 3' lost
CL4065.Contig1	PxylOR43	578	192	tpg DAA05974.1  TPA_exp: odorant receptor 15 [B. mori]	6.00E-24	0.36	2	5', 3' lost
Unigene7439	PxylOR44	570	190	gb ACC63240.1  olfactory receptor 20, partial [Helicoverpa armigera]	8.00E-32	0.37	4	5', 3' lost
Unigene21835	PxylOR46	654	187	gb EFA09245.1  odorant receptor 14 [Tribolium castaneum]	1.00E-08	0.23	2	5' lost
Unigene9201	PxylOR47	545	181	gb ACM18061.1  putative odorant receptor OR3 [Manduca sexta]	8.00E-21	0.36	3	5', 3' lost
CL764.Contig1	PxylOR48	544	180	ref NP_001091791.1  candidate olfactory receptor [B. mori]	2.00E-12	0.27	3	5', 3' lost
CL3314.Contig3	PxylOR49	797	177	ref NP_001166611.1  olfactory receptor 59 [B. mori]	1.00E-17	0.31	3	5', 3' lost
Unigene27391	PxylOR50	531	177	gb EHJ78030.1  olfactory receptor 29 [Danaus plexippus]	2.00E-38	0.49	3	5', 3' lost
Unigene23191	PxylOR51	522	174	ref NP_001166893.1  olfactory receptor 27 [B. mori]	4.00E-65	0.55	4	5', 3' lost
Unigene5685	PxylOR52	809	170	dbj BAH66323.1  olfactory receptor [B. mori]	3.00E-34	0.55	2	5' lost
Unigene28136	PxylOR53	491	164	gb AEF32141.1  odorant receptor [S. exigua]	5.00E-26	0.51	3	5', 3' lost
Unigene11787	PxylOR54	490	163	ref NP_001166616.1  olfactory receptor 54 [B. mori]	1.00E-30	0.47	1	5', 3' lost
<b>Gustatory receptors</b>								
Unigene22668	PxylGR1	1588	392	ref XP_001848097.1  gustatory receptor 22 [Culex quinquefasciatus]	0	0.71	7	Complete
Unigene15579	PxylGR2	958	227	dbj BAK52798.1  gustatory receptor 66 [B. mori]	9.00E-32	0.35	4	5' lost
CL3914.Contig2	PxylGR3	507	168	gb ABY40622.1  gustatory receptor [T. castaneum]	2.00E-50	0.62	2	5', 3' lost
Unigene32005	PxylGR4	343	114	ref NP_001233217.1  gustatory receptor 68 [B. mori]	3.00E-14	0.38	1	5', 3' lost
Continued								

Unigene reference	Name	Length(bp)	ORF(aa)	Blastx best hit (Reference/Name/Species)	E value	Identity	TMD (No)	Status
Unigene6419	PxylGR5	328	109	emb CAD31850.1  putative chemosensory receptor 1 [ <i>H. virescens</i> ]	2.00E-21	0.48	2	5', 3' lost
Unigene34245	PxylGR6	264	88	dbj BAK52798.1  gustatory receptor 66 [ <i>B. mori</i> ]	3.00E-10	0.49	0	5', 3' lost
Unigene19491	PxylGR7	723	240	emb CAD31850.1  putative chemosensory receptor 1 [ <i>H. virescens</i> ]	8.00E-31	0.35	3	5', 3' lost

**Table 2.** Candidate olfactory receptor and gustatory receptor unigenes.

Lepidopteran orthologous genes in the phylogenetic tree. But most PxylORs appeared to be distantly related to the known insect ORs (Fig. 1). We named the Orco unigene PxylOR2 and the 7full-length candidate PR unigenes PxylOR1 and PxylOR3-PxylOR8. The other 46 OR unigenes were ranked in order of decreasing ORF length and named PxylOR9-PxylOR54. We also identified 7 candidate GRs and named them as PxylGR1-PxylGR7.

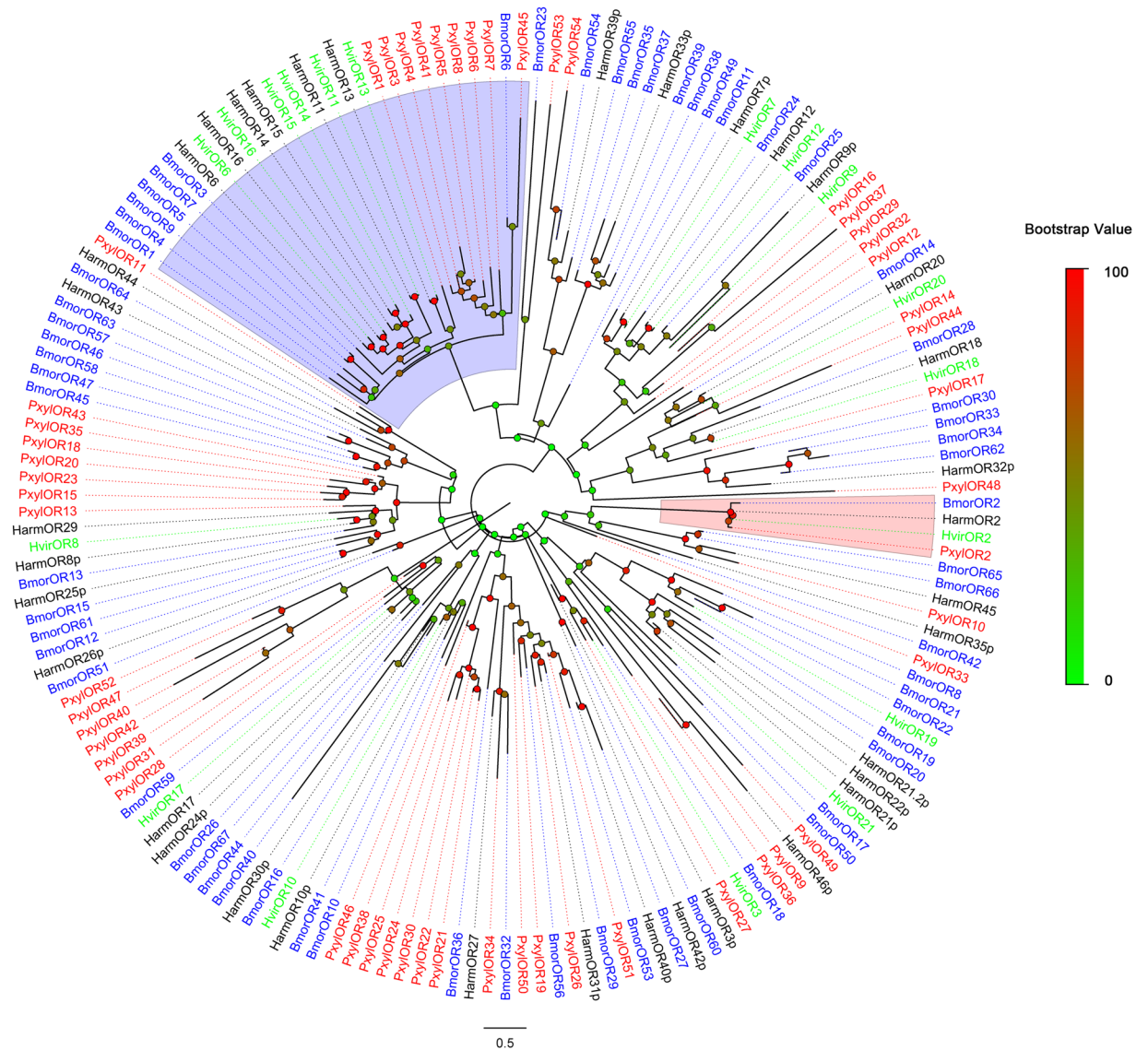
**Identification of candidate IRs.** IR sequences in the *P. xylostella* antennal transcriptome were identified based on similarity to known IRs of Lepidopteran insects, *B. mori*, *C. pomonella*, *H. armigera*, *H. virescens* and *S. littoralis*. Sixteen candidate IRs were identified by bioinformatic analysis, and five unigenes were predicted to have a full-length ORFs. The insect IRs typically have three TMDs. Of the 16 candidates IRs, 15 are predicted to have at least one TMD (Table 3). Twelve of the 16 putative IRs are at least 48% identical to the corresponding Lepidopteran orthologous IRs in *S. littoralis* and *C. pomonella*. The remaining four unigenes have relatively low similarity to other insect IRs: CL2177.Contig2 has 35% identity with IR1 of *S. littoralis*, unigene13888 has 31% identity with IR75 of *C. pomonella*, CL4692.Contig1 has only 25% identity with IR60a of *D. melanogaster*, and CL5979.Contig2 has only 24% identity with IR7c of *D. melanogaster* (Table 3). Phylogenetic analyses suggested that the prediction of IRs was credible. In a neighbor-joining tree of insect IRs, all candidate PxylIRs were clustered in a separate clade with their Lepidopteran orthologs (Fig. 2). All of these 16 candidate IR unigenes were named based on their homology to known IRs. For example, the IR Unigene 19385 has 55% similarity with IR75q2 and CL1791. Contig1 had 64% similarity to IR75q2. So, we named Unigene 19835 PxylIR75q2.2.

**Identification of putative OBPs.** We identified 24 unigenes encoding OBPs from the antennal transcriptome of *P. xylostella*, including 3pheromone binding proteins (PBPs) and 3 general odorant binding proteins (GOBPs) (Table 4). Twenty-two of these 24 unigenes were predicted to have signal peptides, and 19 have full length ORFs. Signal peptide sequences were not detected in the remaining two putative OBPs due to incomplete N-terminal sequences. All 24 putative OBPs had high similarity to known Lepidopteran OBPs. The PBP and GOBP sequences were clustered in a separate clade in the OBP neighbor-joining tree (Fig. 3). Three candidate OBPs were classified into a PBP subgroup in the phylogenetic tree. They share 66%~100% similarity with previously characterized Lepidopteran PBPs and thus were named PBPs. We also found two GOBPs in the antennal transcriptome of *P. xylostella* and named them PxylGOBP1 and GOBP2. A new GOBP (PxylGOBP1.2) was identified that has 77% identity with PxylGOBP1. It was clustered in the GOBP clade and distinguished from other OBPs in the phylogenetic tree. The other 18 candidate OBPs are obviously distinct from the PBP and GOBP clades and have relatively lower similarity to OBPs from other Lepidopteran insects. Most candidate OBP sequences, such as PxylOBP2, PxylOBP3, and PxylOBP7, are closely clustered with at least one Lepidopteran ortholog, in congruence with the blastx results. Some candidate OBP sequences such as PxylOBP6, PxylOBP9, PxylOBP11 and PxylOBP17 are not clustered with OBPs from other Lepidopteran insects (Fig. 3). A possible reason may be that the orthologs of these PxylOBPs have not been identified in other Lepidopteran insects.

**Identification of candidate CSPs.** Bioinformatic analysis led to the identification of 15 different sequences encoding candidate CSPs (Table 5). All 15 unigenes were predicted to have signal peptides and 14 have a full length ORFs. Four candidate PxylCSPs (PxylCSP1-4) match the previously identified *P. xylostella* CSP sequences<sup>31</sup>. The other 11 candidate CSP sequences have at least 35% identity with known CSPs from other insects, and we named them according to the length of the coding region in descending order. In a neighbor-joining tree, all 15 sequences form a cluster with Lepidopteran orthologous genes (Fig. 4).

**Identification of candidate SNMPs.** SNMPs were first identified in pheromone-sensitive neurons of Lepidoptera<sup>31</sup> and are thought to function in pheromone detection<sup>32</sup>. Two kinds of SNMPs (SNMP1 and SNMP2) have been identified in insects and transcripts corresponding to both were found in the *P. xylostella* transcriptome. The sequence of CL2414Contig2 is identical to the PxylSNMP1sequence published in Genbank. CL242. Contig4 has 70% identity with SNMP2 of *O. furnacalis*, and we annotated this sequence as *P. xylostella* SNMP2 (Table 6).

**Tissue- and sex-specific expression of candidate OR and IR genes in *P. xylostella*.** To investigate the expression profile of PxylORs and PxylIRs between sexes and tissues, we determined the expression patterns of the 54 candidate ORs and 16 candidate IRs genes in the antennae and legs of male and female adult *P. xylostella* by semi-quantitative reverse transcription PCR (RT-PCR) (Fig. 5). As shown in Fig. 5, all of these 54 ORs were expressed in the antennae. *PxylOR2* was expressed in male and female antennae and legs. The expression of the nine candidate PRs was observed only in antennae but not in legs. And of which, seven candidate PRs (PxylOR1, 3, 4, 5, 6, 7 and 41) had male-biased or male-specific expression patterns. Different from the other lepidopteran



**Figure 1.** Phylogenetic tree of candidate Lepidopteran ORs, including the PR and Orco clades. PxyI: *P. xylostella* (red), Harm: *Helicoverpa armigera* (black), Hvir: *Heliothis virescens* (green), Bmor: *Bombyx mori* (blue). The clade shaded in blue indicates the Orco clade. The clade shaded in red indicates the PBP clade. The bootstrap value for phylogenetic tree construction is 1000.

ORs, *PxyIOR8*, was only expressed in female antennae. *PxyIOR45* was expressed in both male and female at a similar level. In other 44 general ORs *PxyIOR54* expression was much higher in female than in male antenna and the remaining 43 ORs were expressed in both male and female antennae at a similar level. In contrast to ORs, the expression of all IRs did not differ significantly between males and females. All of these 16 *PxyIIRs* were expressed in the male and female antennae, but *PxyIIR7d.3* and *PxyIIR25a* were also expressed in legs.

## Discussion

In the present study, we profiled the antennal transcriptome of *P. xylostella* adults by RNA-seq technology and annotated 118 putative olfactory genes, including 54 putative ORs, 24 OBPs, 16 IRs, 15 CSPs, 7 GRs, and 2 SNMPs. Chemosensory genes have been identified in other Lepidopteran insects; 134 putative chemosensory unigenes were identified in the antennae of *H. armigera*, including 60 ORs, 34 OBPs, 19 IRs, 18 CSPs, 1 GR and 2 SNMPs, and 131 putative chemosensory unigenes were identified in *H. assulta* antennae, including 64 ORs, 19 IRs, 29 OBPs, 17 CSPs, and 2 SNMPs<sup>33</sup>. Our results are comparable with those from *H. armigera* and *H. assulta* in the number of genes identified. The identification of chemosensory genes from antennal transcriptomes was also reported for the moth *M. sexta* (91 genes, including 48 ORs, 18 OBPs, 21 CSPs and 4 IRs)<sup>34</sup> and *B. mori* (138 genes, including 71 ORs, 20 OBPs, 16 CSPs and 31 IRs)<sup>35–38</sup> and many other insect pests.

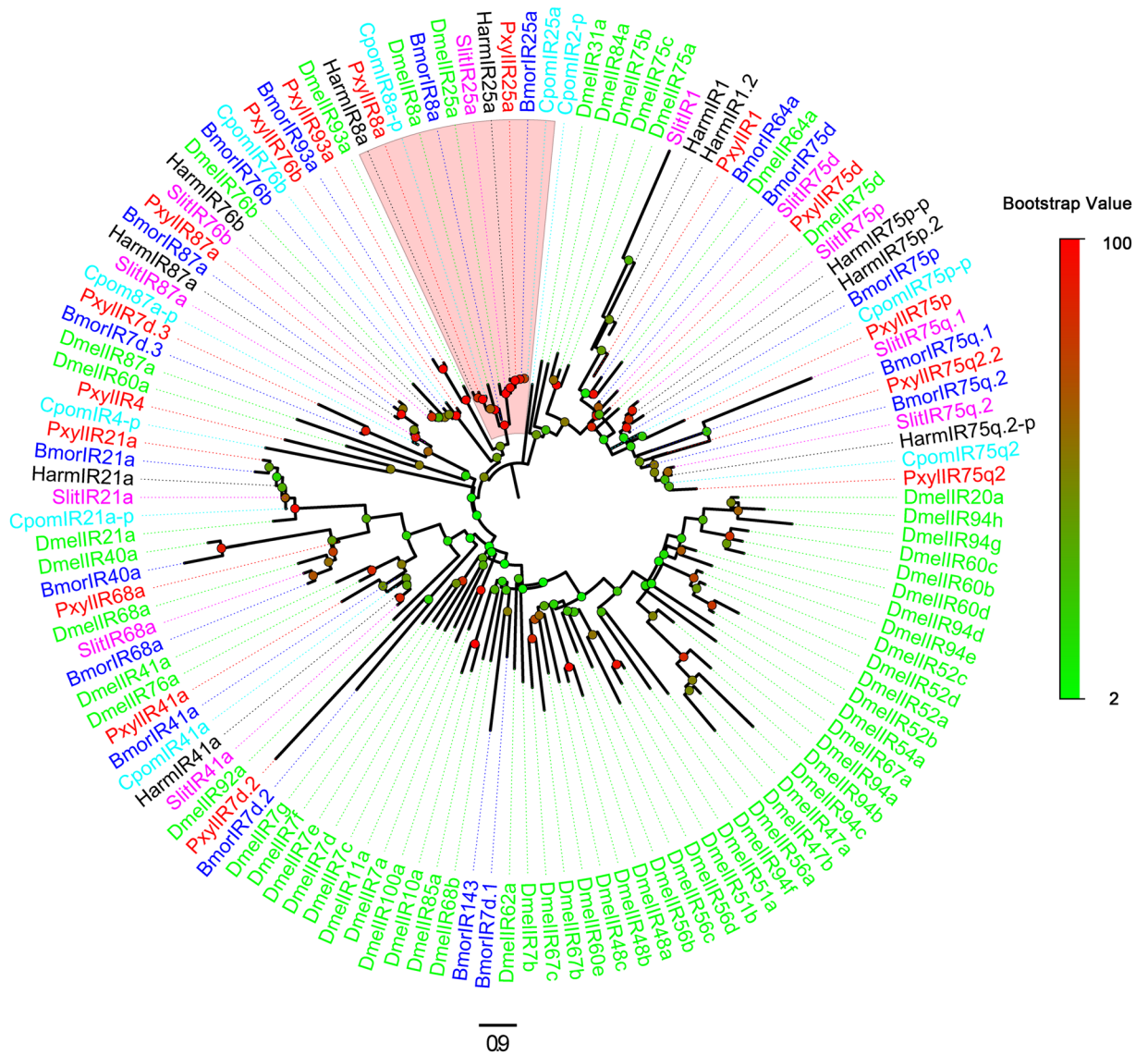
Insects utilize three groups of chemosensory receptors, ORs, IRs and GRs, to perform a variety of essential behaviors such as foraging, mating and oviposition. ORs are the centerpiece of peripheral olfactory reception

Unigene reference	Name	Length (bp)	ORF(aa)	Blastx best hit (Reference/Name/Species)	E value	Identity	TMD (No)	Status
CL2177.Contig2	PxylIR1	1559	483	gb ADR64688.1  putative chemosensory ionotropic receptor IR1 [ <i>Spodoptera littoralis</i> ]	5.00E-70	0.35	3	5' lost
Unigene13888	PxylIR4	1133	345	gb AFC91756.1  putative ionotropic receptor IR75, partial [ <i>Cydia pomonella</i> ]	6.00E-17	0.31	0	3' lost
CL4692.Contig1	PxylIR7d.2	1717	504	ref NP_611901.1  ionotropic receptor 60a [ <i>Drosophila melanogaster</i> ]	4.00E-31	0.25	3	3' lost
CL5979.Contig2	PxylIR7d.3	1624	330	gb AFC91764.1  ionotropic receptor 7c, isoform A [ <i>D. melanogaster</i> ]	1.00E-11	0.24	2	3' lost
Unigene18533	PxylIR8a	3047	907	gb AFC91764.1  putative ionotropic receptor IR8a, partial [ <i>C. pomonella</i> ]	0	0.79	4	Complete
CL721.Contig4	PxylIR21a	2576	858	gb ADR64678.1  putative chemosensory ionotropic receptor IR21a [ <i>S. littoralis</i> ]	0	0.65	4	5', 3' lost
Unigene25424	PxylIR25a	3139	932	gb AFC91757.1  putative ionotropic receptor IR25a [ <i>C. pomonella</i> ]	0	0.89	3	Complete
Unigene25124	PxylIR41a	994	330	gb AFC91758.1  putative ionotropic receptor IR41a [ <i>C. pomonella</i> ]	3.00E-102	0.53	1	5', 3' lost
Unigene255	PxylIR68a	869	289	gb ADR64682.1  putative chemosensory ionotropic receptor IR68a [ <i>S. littoralis</i> ]	4.00E-103	0.67	3	5', 3' lost
CL6386.Contig3	PxylIR75d	1884	593	gb ADR64683.1  putative chemosensory ionotropic receptor IR75d [ <i>S. littoralis</i> ]	4.00E-138	0.48	3	Complete
Unigene8511	PxylIR75p	1356	287	gb AFC91755.1  putative ionotropic receptor IR75p, partial [ <i>C. pomonella</i> ]	3.00E-127	0.79	3	5' lost
CL1791.Contig1	PxylIR75q2	1441	410	gb AFC91752.1  putative ionotropic receptor IR75q2 [ <i>C. pomonella</i> ]	1.00E-163	0.64	1	3' lost
Unigene19385	PxylIR75q2.2	1806	591	gb AFC91752.1  putative ionotropic receptor IR75q2 [ <i>C. pomonella</i> ]	0	0.55	3	5' lost
CL3281.Contig2	PxylIR76b	1790	551	gb AFC91765.1  putative ionotropic receptor IR76b [ <i>C. pomonella</i> ]	0	0.64	3	Complete
Unigene2044	PxylIR87a	1901	633	gb AFC91760.1  putative ionotropic glutamate receptor 87a, partial [ <i>C. pomonella</i> ]	5.00E-167	0.73	4	5', 3' lost
Unigene5567	PxylIR93a	2763	878	gb AFC91753.1  putative ionotropic receptor IR93a, partial [ <i>C. pomonella</i> ]	2.00E-174	0.74	3	Complete

**Table 3.** Candidate ionotropic receptor unigenes.

and determine the sensitivity and specificity of odorant reception<sup>3</sup>. Due to the availability of insect genome databases and progress in sequencing technology, increasing numbers of *OR* genes have been identified from many Lepidopteran species. To date, 68, 64, 70 *OR*s have been identified in the genome databases of *B. mori*<sup>38</sup>, *Danaus plexippus*<sup>39</sup> and *Heliocoonius Melpomene*<sup>40</sup>, respectively. Recently, by using next-generation sequencing technology the antennal transcriptome of *M. sexta* was profiled, and 48 *OR* genes were identified<sup>34,41</sup>. In this study, we identified 54 *OR*s in the antennal transcriptome of adult *P. xylostella*. The number of *OR*s identified in this paper is less than that identified by You *et al.*<sup>30</sup> in the genome database of *P. xylostella*. We might have missed some development-related *OR* genes because we only identified chemosensory genes in the adult antennae. Typical insect *OR*s are characterized by seven TMDs. We found less than seven TMDs in PxylORs, which is also observed in other Lepidopteran insects<sup>33,42,43</sup>. This is probably caused by the limited power of the software used for TMDs finding.

All of the PxylORs identified in the antennal transcriptome are highly divergent and share low similarity with other Lepidopteran insect *OR*s. A study showed that the common ancestor of Lepidopterans had fewer *OR* genes but that there were multiple gene gains and few gene losses during the evolution of Lepidoptera. This phenomenon of gene family expansion is suggested to be associated with the adaption of Lepidopteran species to host plants<sup>44</sup>. We also identified 9 (PxylOR1, PxylOR3-8, PxylOR41 and PxylOR45) candidate PRs based on their similarity to previously characterized PRs. The antennal expression pattern of PxylPRs is consistent with that of PRs in *H. armigera*<sup>42</sup> and *S. littoralis*<sup>45</sup>. Among these 9 candidate PRs, 7 showed male-biased expression, and PxylOR5 was only expressed in male antennae. In contrast, PxylOR8 was only expressed in female antennae. Sex



**Figure 2.** Phylogenetic tree of candidate IRs from *Plutella xylostella* and other insects. Pxyl: *P. xylostella* (red), Harm: *Helicoverpa armigera* (black), Hvir: *Heliothis virescens* (green), Bmor: *Bombyx mori* (blue), Slit: *Spodoptera littoralis* (purple), Cpom: *Cydia pomonella* (cyan). The clade shaded in red indicates the IR8a/IR25a clade. The bootstrap value for phylogenetic tree construction is 1000.

and tissue-specific expression of chemosensory genes is very common among Lepidopteran pests. It was found in *H. assulta*<sup>33</sup> and *H. armigera*<sup>42</sup> that some of their antennal OR genes showed sex-biased expression pattern. The male-specific expression of PxylOR5 probably plays a role in locating females, while the female-specific expression PxylOR8 likely also has ecological significance, i.e. optimization of pheromone production and spatial dispersion of females among host plants<sup>46,47</sup> and selection of oviposition sites.

We identified one Orco unigene, named PxylOR2, which has high similarity to HarmOR2, BmorOR2 and HvirOR2. Orco is highly conserved among all insect species<sup>3</sup> and carries out similar functions in different insects<sup>48</sup> by forming a ligand-gated ion channel<sup>49</sup>. Orco probably functions as a chaperone and forms a dimer with the other ORs in *P. xylostella*.

GRs can respond to tastants such as sugars, bitter substances, CO<sub>2</sub> and some contact pheromones<sup>50</sup>. Thus, GRs play very important roles in food selection and feeding behaviors in insects. The first insect GRs were identified in the fruit fly, *D. melanogaster*<sup>51</sup>. The number of Lepidopteran GRs varies greatly; there is one GR in *Cydia pomonella*<sup>52</sup> and *H. armigera*<sup>42</sup>, 2 in *M. sexta*<sup>34</sup>, 3 in *Heliothis virescens*<sup>53</sup> and 5 in *Spodoptera littoralis*<sup>45,54</sup>. In the antennal transcriptome of adult *P. xylostella* we identified 7 GRs, which is more than those in the Lepidopteran insects mentioned above, but far less than the number found in the silkworm *B. mori* (65 GRs)<sup>55</sup> and the oriental tobacco budworm *H. assulta* (18 GRs)<sup>56</sup>. GRs are mainly expressed in gustatory organs such as the proboscis and maxillary palps, rather than in antennae<sup>8</sup>. This is a possible reason why we identified only 7 GRs in *P. xylostella*. Two GR genes, GR21a and GR63a have been proved to be putative CO<sub>2</sub> receptors in the antennae of the fruit

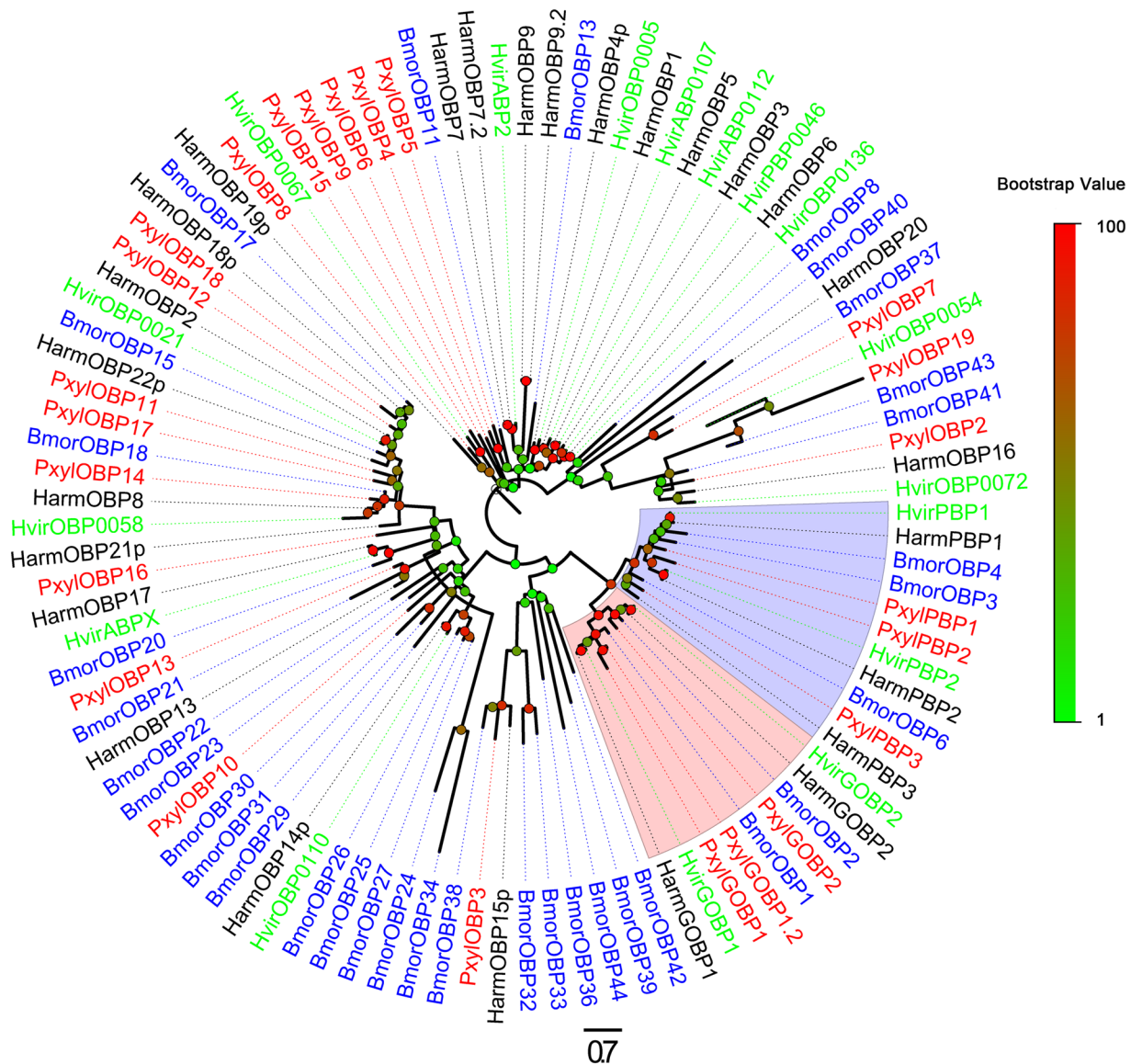


Unigene reference	Gene name	Length (bp)	ORF (aa)	Blastx best hit (Reference/Name/Species)	E value	Identity	Signal peptide	Status
<b>Pheromone binding protein</b>								
Unigene8499	PxylPBP1	761	164	dbj BAG71422.1  pheromone binding protein [ <i>Plutella xylostella</i> ]	5.00E-92	0.99	Yes	Complete
Unigene2096	PxylPBP2	845	172	gb AAF06143.1 AF177661_1 pheromone binding protein [ <i>Yponomeuta cagnagellus</i> ]	3.00E-63	0.66	Yes	Complete
CL3437.Contig1	PxylPBP3	1322	164	gb ACI28451.1  pheromone binding protein 1 [ <i>P. xylostella</i> ]	3.00E-88	0.95	Yes	Complete
<b>General odorant binding protein</b>								
CL5166.Contig1	PxylGOBP1	862	168	gb ABW05104.1  general odorant-binding protein 1 [ <i>P. xylostella</i> ]	4.00E-97	0.93	Yes	Complete
CL3061.Contig1	PxylGOBP1.2	1003	166	gb ABY71034.1  general odorant binding protein 1 [ <i>P. xylostella</i> ]	1.00E-70	0.77	Yes	Complete
CL3886.Contig3	PxylGOBP2	4230	163	gb ABY71035.2  general odorant binding protein 2 [ <i>P. xylostella</i> ]	1.00E-90	1.00	Yes	Complete
<b>Other odorant binding protein</b>								
CL6467.Contig2	PxylOBP2	811	190	gb EHJ77172.1  odorant binding protein [ <i>Danaus plexippus</i> ]	1.00E-40	0.41	Yes	Complete
Unigene10356	PxylOBP3	867	173	gb ACF48467.1  pheromone binding protein female 1 [ <i>Loxostege sticticalis</i> ]	2.00E-37	0.66	Yes	Complete
Unigene103	PxylOBP4	1894	161	gb AFD34177.1  odorant binding protein 1 [ <i>Argyresthia conjugella</i> ]	4.00E-30	0.48	Yes	Complete
Unigene6155	PxylOBP5	962	158	gb AFD34177.1  odorant binding protein 1 [ <i>A. conjugella</i> ]	1.00E-22	0.42	Yes	Complete
CL1521.Contig2	PxylOBP6	2242	153	gb ADK47525.1  odorant binding protein [ <i>Manduca sexta</i> ]	8.00E-23	0.40	Yes	Complete
Unigene25127	PxylOBP7	486	152	emb CAS90127.1  odorant binding protein 3 precursor [ <i>Bombyx mori</i> ]	5.00E-44	0.58	Yes	3' lost
CL5131.Contig2	PxylOBP8	531	149	gb AER27561.1  odorant binding protein [ <i>P. xylostella</i> ]	3.00E-38	0.99	Yes	Complete
CL4848.Contig1	PxylOBP9	570	148	gb EHJ67764.1  odorant-binding protein 5 [ <i>D. plexippus</i> ]	4.00E-15	0.37	Yes	Complete
CL2704.Contig3	PxylOBP10	736	143	gb ACX53795.1  odorant binding protein [ <i>Heliothis virescens</i> ]	1.00E-14	0.33	Yes	Complete
Unigene10167	PxylOBP11	582	143	gb AFD34180.1  odorant binding protein 3 [ <i>A. conjugella</i> ]	1.00E-42	0.60	Yes	Complete
CL4175.Contig1	PxylOBP12	1753	142	gb EHJ65653.1  odorant-binding protein 1 [ <i>D. plexippus</i> ]	6.00E-51	0.77	Yes	Complete
Unigene26843	PxylOBP13	1086	141	gb AFD34173.1  odorant binding protein 5 [ <i>A. conjugella</i> ]	6.00E-64	0.77	Yes	Complete
CL4228.Contig1	PxylOBP14	726	140	gb AFD34175.1  odorant binding protein 4 [ <i>A. conjugella</i> ]	3.00E-55	0.72	Yes	Complete
Unigene21533	PxylOBP15	422	140	gb ACX53756.1  odorant binding protein [ <i>H. virescens</i> ]	1.00E-37	0.52	Yes	5', 3' lost
Unigene15836	PxylOBP16	742	139	gb AFD34182.1  odorant binding protein 6 [ <i>A. conjugella</i> ]	2.00E-47	0.66	Yes	Complete
CL2382.Contig4	PxylOBP17	444	129	gb AFD34180.1  odorant binding protein 3 [ <i>A. conjugella</i> ]	9.00E-29	0.50	No	5' lost
CL4528.Contig1	PxylOBP18	502	97	gb AFG72998.1  odorant-binding protein 1 [ <i>Cnaphalocrocis medinalis</i> ]	2.00E-41	0.76	No	5' lost
Unigene37282	PxylOBP19	228	64	gb ACX53743.1  odorant binding protein [ <i>H. virescens</i> ]	2.00E-13	0.60	Yes	3' lost

**Table 4.** Candidate odorant binding protein unigenes.

fly<sup>57,58</sup>. And in mosquitos, 3 putative CO<sub>2</sub> receptor genes (GR22, 23 and 24) have been identified in the maxillary palps of different species<sup>59–61</sup>. The PxylGR1 was closely related to the GR22 in mosquito and GR21a in the fruit fly and predicted to be a candidate CO<sub>2</sub> receptor.

IRs belong to an ancient chemosensory receptor family, and two subfamilies of IRs have been identified recently, i.e. the conserved 'antennal IRs' and the species-specific 'divergent IRs'<sup>62</sup>. The first IR was identified in the coeloconic sensilla of *Drosophila*<sup>14</sup> and most *Drosophila* IRs have clear orthologs within the genus of Lepidoptera<sup>34,42,63</sup>. IRs are ligand-gated ion channels that mediate chemical communication between neurons<sup>14</sup>. In this study, we identified 16 IRs in the antennal transcriptome of *P. xylostella* and named them based on homologous sequences from other insects. Similar numbers of IRs have been identified from other Lepidopteran insects:



**Figure 3.** Phylogenetic tree of candidate Lepidopteran OBPs, including the GOBP and PBP clades. PxyI: *Plutella xylostella* (red), Harm: *Helicoverpa armigera* (black), Hvir: *Heliothis virescens* (green), Bmor: *Bombyx mori* (blue). The clade shaded in blue indicates the PBP clade. The clade shaded in red indicates the GOBP clade. The bootstrap value for phylogenetic tree construction is 1000.

19 IRs were identified in the antennal transcriptomes of *H. armigera* and *H. assulta*<sup>33</sup>, 15 IRs in *C. pomonella*<sup>52</sup>, 20 IRs in *Chio suppressalis*<sup>43</sup>, and 12 IRs in *S. littoralis*<sup>36</sup>. All of these IRs are expressed in antennae, but PxyIIR7d.3 and PxyIIR25a are also expressed in legs, which is different from the expression patterns of these genes in *H. assulta*<sup>33</sup>. Coincidentally, HarmIR25a, HarmIR75d, HarmIR75p and HarmIR76p are also expressed in the cotton bollworm legs<sup>42</sup>. The function of leg-expressed IRs remains unknown and deserves in-depth investigation.

OBPs are believed to be directly involved in the activation of the ORx/Orco complex in the recognition of specific odors<sup>20</sup>. A total of 24 OBPs were identified in the antennal transcriptome of *P. xylostella*, including three GOBPs and three PBPs. The number of OBPs identified in the present study was comparable to those identified in transcriptomic analyses of *H. armigera* (34) and *H. assulta* (29)<sup>33</sup>, *S. litura* (21)<sup>64</sup>, *S. littoralis* (26)<sup>54</sup>, but fewer than those identified in *B. mori* (44)<sup>37</sup>. OBPs showed lineage-specific expansion and diversification; therefore, it is not surprising that there are some differences, or even big differences, in the number of OBPs. Previous studies have also shown that some insect OBPs and CSPs are expressed exclusively in non-antennae tissues or in larvae<sup>65</sup>. Therefore, different sampling and sequencing strategies may lead to different results. In a previous study, two GOBPs, GOBP1 and 2, were identified in *P. xylostella* antennae<sup>66</sup>. GOBPs were also found in the antennae of *C. pomonella*<sup>67</sup> and *S. litura*<sup>68</sup>. The antennal *P. xylostella* GOBPs identified in this study have ecological significance, e.g. guiding *P. xylostella* to find better food<sup>69</sup>. The antennal *S. litura* GOBP1 can bind to plant odorants, while *S. litura* GOBP2 can bind to aldehyde-sex compounds and analogs<sup>68</sup>.

Unigene reference	Gene name	Length (bp)	ORF (aa)	Blastx best hit (Reference/Name/Species)	E value	Identity	Signal peptide	Status
Unigene7305	PxylCSP1	732	152	gb ABM67686.1  chemosensory protein CSP1 [ <i>Plutella xylostella</i> ]	3.00E-84	0.99	Yes	Complete
Unigene12972	PxylCSP2	676	128	gb ABM67687.1  chemosensory protein CSP2 [ <i>P. xylostella</i> ]	2.00E-71	0.98	Yes	Complete
Unigene5262	PxylCSP3	425	122	gb ABM92663.1  chemosensory protein CSP3 [ <i>P. xylostella</i> ]	1.00E-65	0.99	Yes	3' lost
CL1074.Contig1	PxylCSP4	1010	126	gb ABM92664.1  chemosensory protein CSP4 [ <i>P. xylostella</i> ]	1.00E-66	1.00	Yes	Complete
Unigene1800	PxylCSP5	638	130	gb AAK53762.1 AF368375_1 chemosensory protein [ <i>Helicoverpa armigera</i> ]	4.00E-53	0.73	Yes	Complete
CL574.Contig2	PxylCSP6	1632	130	dbj BAF91712.1  chemosensory protein [ <i>Papilio xuthus</i> ]	8.00E-60	0.87	Yes	Complete
Unigene24730	PxylCSP7	552	127	dbj BAG71921.1  chemosensory protein 13 [ <i>P. xuthus</i> ]	2.00E-43	0.69	Yes	Complete
Unigene10872	PxylCSP8	520	127	gb ABM67689.1  chemosensory protein CSP2 [ <i>Spodoptera exigua</i> ]	9.00E-43	0.63	Yes	Complete
Unigene7440	PxylCSP9	657	123	gb ACX53825.1  chemosensory protein [ <i>Heliothis virescens</i> ]	5.00E-43	0.64	Yes	Complete
Unigene7557	PxylCSP10	1041	123	dbj BAF91711.1  chemosensory protein [ <i>P. xuthus</i> ]	4.00E-44	0.70	Yes	Complete
CL3090.Contig2	PxylCSP11	740	122	gb EHJ73330.1  chemosensory protein [ <i>Danaus plexippus</i> ]	2.00E-51	0.77	Yes	Complete
Unigene21123	PxylCSP12	667	122	gb AEX07265.1  CSP2 [ <i>H. armigera</i> ]	4.00E-42	0.56	Yes	Complete
CL1877.Contig3	PxylCSP13	886	120	gb EHJ73328.1  chemosensory protein 11b [ <i>D. plexippus</i> ]	3.00E-27	0.49	Yes	Complete
Unigene21118	PxylCSP14	548	111	dbj BAF91720.1  chemosensory protein [ <i>P. xuthus</i> ]	6.00E-45	0.81	Yes	Complete
CL2890.Contig2	PxylCSP15	672	110	ref XP_001844687.1  chemosensory protein 1 [ <i>Culex quinquefasciatus</i> ]	1.00E-09	0.35	Yes	Complete

**Table 5.** Candidate chemosensory protein unigenes.

CSPs are a class of small soluble proteins expressed highly in the chemosensilla lymph<sup>70</sup> and show high binding activity to odorants and pheromones<sup>71</sup>. We identified 15CSPs genes in the present study. The number of CSPs identified from *P. xylostella* was comparable to the number in *B. mori* (18)<sup>72</sup>, *H. armigera* (18) and *H. assulta* (17)<sup>33</sup> and *S. litura* (18)<sup>64</sup>, but fewer than the number in *M. sexta* (21)<sup>34</sup>, *Sesamia inferens* (24)<sup>63</sup> and *S. littoralis* (31)<sup>45</sup>. Because CSPs are also expressed in tissues other than antennae<sup>73,74</sup> and may participate in other physiological processes, it is possible that we have missed some CSPs in our antennal transcriptome analysis.

SNMPs are two-transmembrane domain proteins that share very high homology to members of the mammalian CD36 receptor family, which are thought to function in pheromone detection of Lepidopteran and Dipteran insects<sup>31</sup>. Two subtypes of SNMPs (SNMP1 and SNPM2) have been frequently identified in most insects, e.g. *Helicoverpa armigera*<sup>33,42</sup>, *Cnaphalocrocis medinalis*<sup>27</sup>, *S. exigua*<sup>75</sup>, *S. litura*<sup>28</sup>, *C. suppressalis*<sup>43</sup>, *H. assulta*<sup>33</sup>, and in this study, *P. xylostella*. The expression of antennal SNMPs in *P. xylostella* suggests their role in pheromone detection, similar to what has been reported in *D. melanogaster*<sup>32,76</sup>.

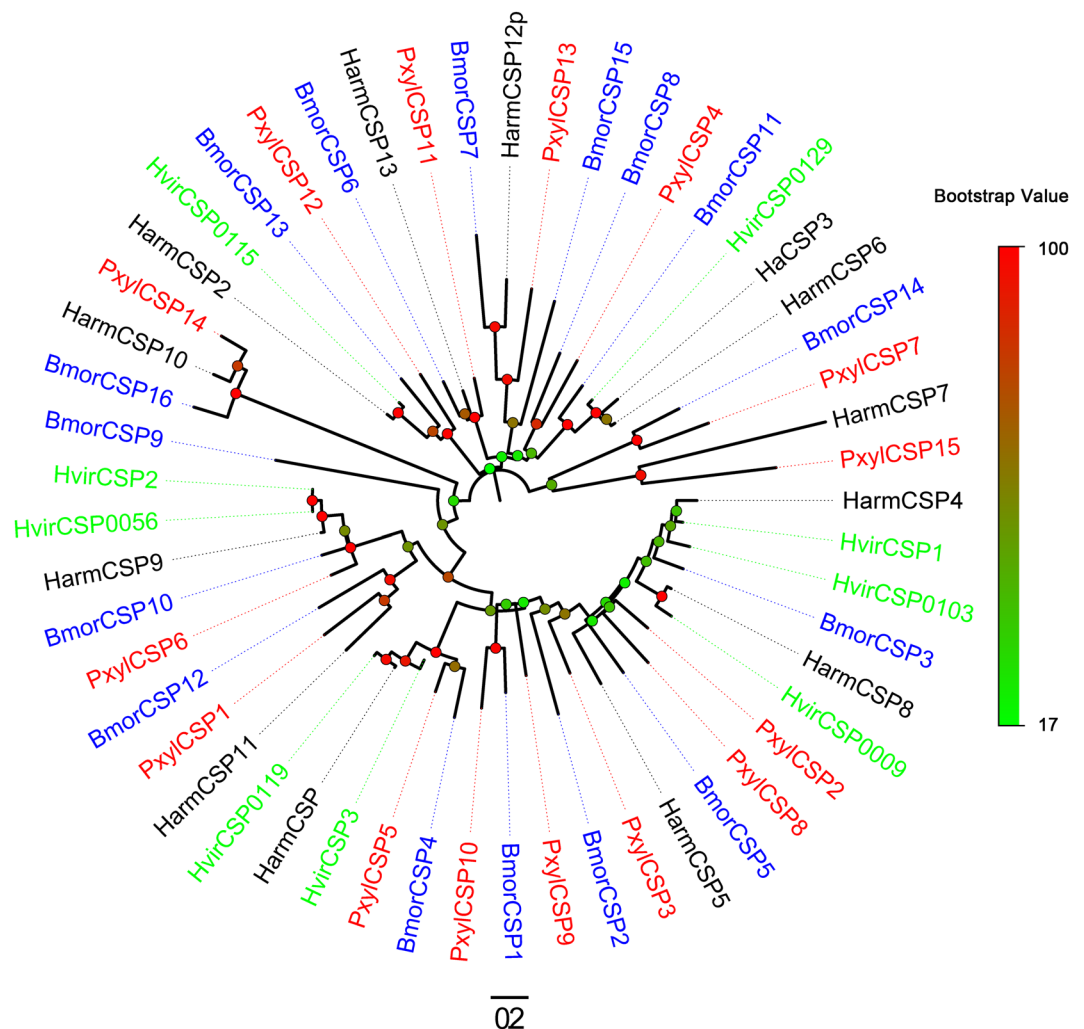
## Conclusions

In summary, we identified 118 candidate olfactory genes that may function in odorant perception in the diamond-back moth, *P. xylostella* by assembling and annotating transcriptomic sequence data. We carried out a comparative phylogenetic analysis to predict gene functions and examined the transcriptome patterns of the *P. xylostella* OR and IR genes. Genes with sex-biased and tissue-specific expression patterns, especially PxylOR5 and PxylOR8, are potential targets for environmentally-friendly management of this destructive insect pest. Our results lay the foundation for functional analysis of these receptors in both neurobiological and evolutionary studies.

## Materials and Methods

**Insect rearing.** The laboratory-maintained *P. xylostella* was reared in the Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China. The larvae and adults were fed on Chinese cabbage and kept in cages at 27 ± 1 °C under a 16: 8 (L: D) photoperiod and 65 ± 5% relative humidity. Male and female larvae were distinguished at the last instar and placed in separate cages. Antennae of female or male adults were dissected at 1–3 days after adult emergence, immediately frozen in liquid nitrogen, and then stored at –70 °C until use.

**Total RNA extraction.** The frozen antennae were transferred to a liquid nitrogen-cooled mortar and ground with a pestle. One mL of TRIzol reagent was pipetted to the homogenate (Invitrogen, Carlsbad, CA, USA) and total RNA was extracted following the manufacturer's instructions. Total RNA was resuspended in RNase-free



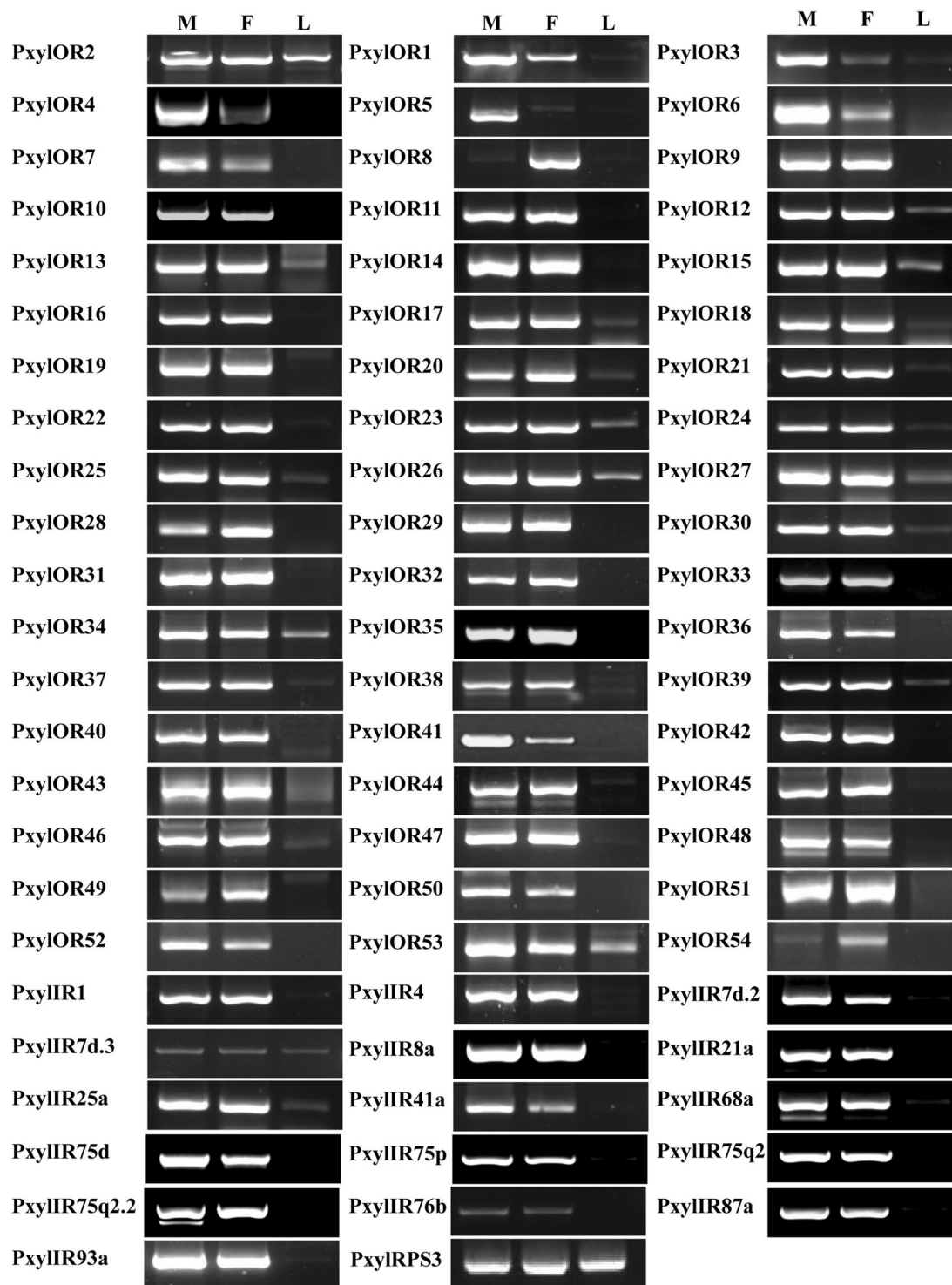
**Figure 4.** Phylogenetic tree of candidate Lepidopteran CSPs. PxyI: *Plutella xylostella* (red), Harm: *Heliocoverpa armigera* (black), Hvir: *Hethiothis virescens* (green), Bmor: *Bombyx mori* (blue). The bootstrap value for phylogenetic tree construction is 1000.

Unigene reference	Gene name	Length (bp)	ORF (aa)	BLASTx best hit (Reference/Name/Species)	E value	Identity	Status
CL2414.Contig2	PxyISNMP1	2408	522	gi 301153754 gb ADK66278.1  sensory neuron membrane protein 1 [ <i>Plutella xylostella</i> ]	0	1.00	Complete
CL242.Contig4	PxyISNMP2	2196	523	gi 312306074 gb ADQ73891.1  sensory neuron membrane protein 2 [ <i>Ostrinia furnacalis</i> ]	0	0.70	Complete

**Table 6.** Candidate sensory neuron membrane protein unigenes.

H<sub>2</sub>O, and RNA quantity was determined with a Nanodrop ND-2000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). RNA integrity was assessed using an Agilent 2100 BioAnalyzer (Agilent Technologies, Englewood, CO, USA).

**cDNA Library construction and Illumina sequencing.** Ten  $\mu$ g of total RNA, extracted from approximately 2000 antennae of 1–3 day old adult male or female moths. The cDNA library for each sample was prepared using the NEBNext<sup>®</sup> mRNA Library Prep Reagent Set for Illumina (NEB, Ipswich, MA, USA) following the manufacturer's instructions. Poly-A RNA for each sample was fragmented in fragmentation buffer to a length of 200 nt–700 nt. Random hexamers were used to generate first-strand cDNA, and second-strand cDNA was synthesized using RNaseH and DNA polymerase I. The double-strand cDNA (ds cDNA) samples were purified with the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and eluted with EB buffer. The short fragments were treated with T4 DNA Polymerase and T4 Polynucleotide Kinase for end-repair and dA-tailing, then sequencing



**Figure 5.** Tissue- and sex-specific expression patterns of candidate *PxyI ORs* and *PxyI IRs*. M: male antennae, F: female antennae, L: legs. *PxyI RPS3* is the reference.

adaptors with barcodes were ligated to the dA tail of ds cDNA using T4 DNA ligase. To select insert length, ds cDNA samples were separated by agarose gel electrophoresis and bands of approximately 200 bp were excised and purified with the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Paired-end sequencing of the library was performed on the Illumina HiSeq™ 2000 platform (Illumina, San Diego, CA, USA) at the Beijing Genome Institute (Shenzhen, China). The read length of each end was 90 bp. The male and female libraries were sequenced in one lane, and raw reads were then sorted by barcode sequence.

**Unigene generation.** Raw reads were pre-processed to remove low quality reads and reads containing adapter sequences and poly-A/T tails. The publicly available program Trinity was used to perform de novo assembly of clean reads to generate a set of transcripts<sup>77</sup>. The Trinity outputs were then clustered by TGICL (TGI Clustering tools)<sup>78</sup>. The final unigene dataset consists of uniformly clustered sequences and singletons.

**Gene identification and functional annotation.** Unigene sequences were first searched against protein databases like nr, Swiss-Prot, KEGG and COG, using blastx with an e-value cut-off of  $1e^{-5.79}$ . To identify more OR genes, 63ORs from *B. mori* were used as queries in tblastn searches of *P. xylostella* antennal unigenes. Unigene ESTs were predicted using ESTScan<sup>80</sup>. Signal peptides in the protein sequences were predicted using SignalP 4.0<sup>81</sup>. The TMDs of annotated genes were predicted using TMHMM Server Version2.0 (<http://www.cbs.dtu.dk/services/TMHMM>).

**Phylogenetic analyses.** Phylogenetic trees were constructed based on the amino sequences of the candidate olfaction genes and genes from the collected data sets. The OR datasets contained OR sequences identified from Lepidopteran insects (36 from *H. armigera*, 18 from *H. virescens* and 63 from *B. mori*)<sup>38,42,82,83</sup>. The IR datasets contained IR sequences from *H. armigera* (11), *S. littoralis* (11), *Cydia pomonella* (10), *B. mori* (18) and *D. melanogaster* (64)<sup>36,42,52,62</sup>. The OBP datasets contained sequences from *H. armigera* (26), *H. virescens* (17) and *B. mori* (34)<sup>37,42</sup>. The CSP data set contained sequences from *H. armigera* (13)<sup>42</sup>, *H. virescens* (9)<sup>84</sup> and *B. mori* (16)<sup>70</sup>. All amino acid sequences were aligned using ClustalW2<sup>85</sup>. The unrooted neighbor-joining trees were constructed by the Jones-Taylor-Thornton (JTT) method with 1,000 bootstrap replications as implemented in MEGA5 software<sup>86</sup>.

**Expression analysis of the candidate receptors by semi-quantitative reverse transcription PCR.** To illustrate and compare the expression patterns of candidate receptors in male and female antennae, semi-quantitative RT-PCR was performed using cDNA prepared from male antennae, female antennae and legs (male and female mixture). Legs were used as a control to confirm the antennae-enriched expression of candidate receptors. Total RNA was extracted as described above. Prior to cDNA synthesis, RNA was treated with DNase I (Fermentas, Vilnius, Lithuania) to remove trace amounts of genomic DNA. The cDNA was synthesized using the First Strand cDNA Synthesis Kit (Fermentas, Vilnius, Lithuania) and was used as a template in PCR reactions with gene-specific primers. The housekeeping gene RPS3 was used as a control<sup>87</sup>. Primers were designed using the Primer Premier 5 software (PREMIER Biosoft International), and the sequences are available in Supplementary Table S1. PCR was performed with the Veriti Thermal Cycler (Applied Biosystems, Carlsbad, CA, USA) under the following conditions: 94 °C for 2 min, 33 cycles of 94 °C for 30 s, 55–60 °C for 30 s, and 72 °C for 30 s, and 72 °C for 10 min. The cycle number was reduced to 27 and 30 for Actin and OR2 amplification because of their high expression level. The experiment was repeated three times using three independently isolated RNA samples. PCR amplification products were run on a 2% agarose gel and verified by DNA sequencing.

## References

- Sato, K. & Touhara, K. In *Chemosensory Systems in Mammals, Fishes, and Insects* 203–220 (Springer, 2009).
- Gao, Q. & Chess, A. Identification of candidate *Drosophila* olfactory receptors from genomic DNA sequence. *Genomics* **60**, 31–39 (1999).
- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391 (2013).
- Clyne, P. J. *et al.* A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**, 327–338 (1999).
- Benton, R., Sachse, S., Michnick, S. W. & Vosshall, L. B. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors *in vivo*. *PLoS Biol.* **4**, e20 (2006).
- Scott, K. *et al.* A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*. *Cell* **104**, 661–673 (2001).
- Chyb, S. *Drosophila* gustatory receptors: from gene identification to functional expression. *J. Insect Physiol.* **50**, 469–477 (2004).
- Sparks, J. T., Vinyard, B. T. & Dickens, J. C. Gustatory receptor expression in the labella and tarsi of *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **43**, 1161–1171 (2013).
- Jiang, X. J. *et al.* A gustatory receptor tuned to D-fructose in antennal sensilla chaetica of *Helicoverpa armigera*. *Insect Biochem. Mol. Biol.* **60**, 39–46 (2015).
- Mang, D. *et al.* Expression of a sugar clade gustatory receptor, BmGr6, in the oral sensory organs, midgut, and central nervous system of larvae of the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **70**, 85–98 (2016).
- Ning, C., Yang, K., Xu, M., Huang, L. Q. & Wang, C. Z. Functional validation of the carbon dioxide receptor in labial palps of *Helicoverpa armigera* moths. *Insect Biochem. Mol. Biol.* **73**, 12–19 (2016).
- Abuin, L. *et al.* Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* **69**, 44–60 (2011).
- Ai, M. R. *et al.* Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex *in vivo* in *Drosophila*. *J. Neuroscience* **33**, 10741–10749 (2013).
- Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136**, 149–162 (2009).
- Silbering, A. F. *et al.* Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neuroscience* **31**, 13357–13375 (2011).
- Chen, C. *et al.* *Drosophila* ionotropic receptor 25a mediates circadian clock resetting by temperature. *Nature* **27**, 516–U238 (2015).
- Miyakawa, H., Sato, M., Colbourne, J. K. & Iguchi, T. Ionotropic glutamate receptors mediate inducible defense in the water flea *Daphnia pulex*. *PLoS One* **10**, e0121324 (2015).
- Leal, W. S. Pheromone reception. In *The Chemistry of Pheromones and Other Semiochemicals II* 1–36 (Springer, 2005).
- Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* **63**, 1658–1676 (2006).
- Laughlin, J. D., Ha, T. S., Jones, D. N. M. & Smith, D. P. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell* **133**, 1255–1265 (2008).
- Zhou, J. J. Odorant-binding proteins in insects. *Vitam Horm* **83**, 241–272 (2010).

22. Briand, L. *et al.* Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera* L.) as a brood pheromone carrier. *Eur. J. Biochem.* **269**, 4586–4596 (2002).
23. Yi, X. *et al.* Involvement of a specific chemosensory protein from *Bactrocera dorsalis* in perceiving host plant volatiles. *J. Chem. Ecol.* **40**, 267–275 (2014).
24. Li, Z. Q. *et al.* Expression Analysis and Binding Assays in the Chemosensory Protein Gene Family Indicate Multiple Roles in *Helicoverpa armigera*. *J. Chem. Ecol.* **41**, 473–485 (2015).
25. Vogt, R. G. *et al.* The insect SNMP gene family. *Insect Biochem. Mol. Biol.* **39**, 448–456 (2009).
26. Gu, S. H. *et al.* Molecular identification and differential expression of sensory neuron membrane proteins in the antennae of the black cutworm moth *Agrotis ipsilon*. *J. Insect Physiol.* **59**, 430–443 (2013).
27. Liu, S. *et al.* Identification and characterization of two sensory neuron membrane proteins from *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae). *Arch. Insect Biochem.* **82**, 29–42 (2013).
28. Zhang, J., Liu, Y., Walker, W. B., Dong, S. L. & Wang, G. R. Identification and localization of two sensory neuron membrane proteins from *Spodoptera litura* (Lepidoptera: Noctuidae). *Insect Sci.* **22**, 399–408 (2015).
29. Furlong, M. J., Wright, D. J. & Dossdall, L. M. Diamondback moth ecology and management: problems, progress, and prospects. *Ann. Rev. Entomol.* **58**, 517–541 (2013).
30. You, M. *et al.* A heterozygous moth genome provides insights into herbivory and detoxification. *Nat. Genet.* **45**, 220–225 (2013).
31. Rogers, M. E., Krieger, J. & Vogt, R. G. Antennal SNMPs (sensory neuron membrane proteins) of *Lepidoptera* define a unique family of invertebrate CD36-like proteins. *J. Neurobiol.* **49**, 47–61 (2001).
32. 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).
33. Zhang, J. *et al.* Antennal transcriptome analysis and comparison of chemosensory gene families in two closely related noctuidae moths, *Helicoverpa armigera* and *H. assulta*. *PLoS one* **10**, e0117054 (2015).
34. Grosse-Wilde, E. *et al.* Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci. USA* **108**, 7449–7454 (2011).
35. Xia, Q. *et al.* A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* **306**, 1937–1940 (2004).
36. Olivier, V., Monsempes, C., Francois, M. C., Poivet, E. & Jacquin-Joly, E. Candidate chemosensory ionotropic receptors in a *Lepidoptera*. *Insect. Mol. Biol.* **20**, 189–199 (2011).
37. Gong, D. P., Zhang, H. J., Zhao, P., Xia, Q. Y. & Xiang, Z. H. The odorant binding protein gene family from the genome of silkworm *Bombyx mori*. *BMC Genomics* **10**, 332 (2009).
38. Tanaka, K. *et al.* Highly selective tuning of a silkworm olfactory receptor to a key mulberry leaf volatile. *Curr. Biol.* **19**, 881–890 (2009).
39. Zhan, S., Merlin, C., Boore, J. L. & Reppert, S. M. The monarch butterfly genome yields insights into long-distance migration. *Cell* **147**, 1171–1185 (2011).
40. Briscoe, A. D. *et al.* Female behaviour drives expression and evolution of gustatory receptors in butterflies. *Plos Genet.* **9**, e1003620 (2013).
41. Howlett, N. *et al.* Identification of chemosensory receptor genes in *Manduca sexta* and knockdown by RNA interference. *Bmc Genomics* **13**, 1 (2012).
42. Liu, Y., Gu, S., Zhang, Y., Guo, Y. & Wang, G. Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS One* **7**, e48260 (2012).
43. Cao, D. *et al.* Identification of candidate olfactory genes in *Chilo suppressalis* by antennal transcriptome analysis. *Int. J. Biol. Sci.* **10**, 846–860 (2014).
44. Engson, P., Sangket, U., Chotigeat, W. & Satasook, C. Molecular evolution of the odorant and gustatory receptor genes in *Lepidopteran* insects: implications for their adaptation and speciation. *J. Mol. Evol.* **79**, 21–39 (2014).
45. Legeai, F. *et al.* An Expressed Sequence Tag collection from the male antennae of the Noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research. *BMC Genomics* **12**, 86 (2011).
46. Palanaswamy, P. & Seabrook, W. Behavioral responses of the female eastern spruce budworm *Choristoneura fumiferana* (Lepidoptera, Tortricidae) to the sex pheromone of her own species. *J. Chem. Ecol.* **4**, 649–655 (1978).
47. Schneider, D., Schulz, S., Priesner, E., Ziesmann, J. & Francke, W. Autodetection and chemistry of female and male pheromone in both sexes of the tiger moth *Panaxia quadripunctaria*. *J. Comp. Physiol. a-Sensory Neural. Behav. Physiol.* **182**, 153–161 (1998).
48. Jones, W. D., Nguyen, T. A., Kloss, B., Lee, K. J. & Vosshall, L. B. Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Curr. Biol.* **15**, 119–121 (2005).
49. Nakagawa, T., Sakurai, T., Nishioka, T. & Touhara, K. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science's STKE* **307**, 1638 (2005).
50. Montell, C. A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol.* **19**, 345–353 (2009).
51. Clyne, P. J., Warr, C. G. & Carlson, J. R. Candidate taste receptors in *Drosophila*. *Science* **287**, 1830–1834 (2000).
52. Bengtsson, J. M. *et al.* Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis. *PLoS One* **7**, e31620 (2012).
53. Jorgensen, K., Kvellø, P., Almaas, T. J. & Mustaparta, H. Two closely located areas in the suboesophageal ganglion and the tritocerebrum receive projections of gustatory receptor neurons located on the antennae and the proboscis in the moth *Heliothis virescens*. *J. Comp. Neurol.* **496**, 121–134 (2006).
54. Jacquin-Joly, E. *et al.* Candidate chemosensory genes in female antennae of the noctuid moth *Spodoptera littoralis*. *Int. J. Biol. Sci.* **8**, 1036–1050 (2012).
55. Wanner, K. W. & Robertson, H. M. The gustatory receptor family in the silkworm moth *Bombyx mori* is characterized by a large expansion of a single lineage of putative bitter receptors. *Insect Mol. Biol.* **17**, 621–629 (2008).
56. Xu, W., Papanicolaou, A., Liu, N. Y., Dong, S. L. & Anderson, A. Chemosensory receptor genes in the Oriental tobacco budworm *Helicoverpa assulta*. *Insect Mol. Biol.* **24**, 253–263 (2015).
57. Jones, W. D., Cayirlioglu, P., Grunwald Kadow, I. & Vosshall, L. B. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature* **445**, 86–90 (2007).
58. Kwon, J. Y., Dahanukar, A., Weiss, L. A. & Carlson, J. R. The molecular basis of CO<sub>2</sub> reception in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **104**, 3574–3578 (2007).
59. Lu, T. *et al.* Odor coding in the maxillary palp of the malaria vector mosquito *Anopheles gambiae*. *Curr. Biol.* **17**, 1533–1544 (2007).
60. Robertson, H. M. & Kent, L. B. Evolution of the gene lineage encoding the carbon dioxide receptor in insects. *J. Insect Sci.* **9**, 14 (2009).
61. Kent, L. B., Walden, K. K. & Robertson, H. M. The Gr family of candidate gustatory and olfactory receptors in the yellow-fever mosquito *Aedes aegypti*. *Chem. Senses* **33**, 79–93 (2008).
62. Croset, V. *et al.* Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* **6**, e1001064 (2010).
63. Zhang, Y. N. *et al.* Differential expression patterns in chemosensory and non-chemosensory tissues of putative chemosensory genes identified by transcriptome analysis of insect pest the purple stem borer *Sesamia inferens* (Walker). *Plos One* **8**, e6971510 (2013).
64. Feng, B. *et al.* Transcriptome and expression profiling analysis link patterns of gene expression to antennal responses in *Spodoptera litura*. *BMC Genomics* **16**, 269 (2015).

65. 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).
66. Zhang, Z. C., Wang, M. Q., Lu, Y. B. & Zhang, G. Molecular characterization and expression pattern of two general odorant binding proteins from the diamondback moth, *Plutella xylostella*. *J. Chem. Ecol.* **35**, 1188–1196 (2009).
67. Garczynski, S. F. *et al.* Application of *Cydia pomonella* expressed sequence tags: Identification and expression of three general odorant binding proteins in codling moth. *Insect Sci.* **20**, 559–574 (2013).
68. Liu, N.-Y. *et al.* Two general-odorant binding proteins in *Spodoptera litura* are differentially tuned to sex pheromones and plant odorants. *Comp Biochem Physiol Part A: Moll Integr Physiol.* **180**, 23–31 (2015).
69. Zhu, J. *et al.* General odorant-binding proteins and sex pheromone guide larvae of *Plutella xylostella* to better food. *Insect Biochem. Mol. Biol.* **72**, 10–19 (2016).
70. 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).
71. Pelosi, P., Iovinella, I., Felicioli, A. & Dani, F. R. Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* **5**, 320 (2014).
72. Gong, D. P. *et al.* Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **37**, 266–277 (2007).
73. Jacquin-Joly, E., Vogt, R. G., Francois, 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).
74. Liu, Y. L., Guo, H., Huang, L. Q., Pelosi, P. & Wang, C. Z. Unique function of a chemosensory protein in the proboscis of two *Helicoverpa* species. *J. Exp. Biol.* **217**, 1821–1826 (2014).
75. Liu, C., Zhang, J., Liu, Y., Wang, G. & Dong, S. Expression of SNMP1 and SNMP2 genes in antennal sensilla of *Spodoptera exigua* (Hübner). *Arch. Insect Biochem. Physiol.* **85**, 114–126 (2014).
76. Jin, X., Ha, T. S. & Smith, D. P. SNMP is a signaling component required for pheromone sensitivity in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **105**, 10996–11001 (2008).
77. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnol.* **29**, 644–652 (2011).
78. Partea, G. *et al.* TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* **19**, 651–652 (2003).
79. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
80. Iseli, C., Jongeneel, C. V. & Bucher, P. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 138–148 (1999).
81. 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).
82. Krieger, J. Genes encoding candidate pheromone receptors in a moth (*Heliothis virescens*). *Proc. Natl. Acad. Sci. USA* **101**, 11845–11850 (2004).
83. Krieger, J. *et al.* A divergent gene family encoding candidate olfactory receptors of the moth *Heliothis virescens*. *Eur. J. Neurosci.* **16**, 619–628 (2002).
84. Picimbon, J. F., Dietrich, K., Krieger, J. & Breer, H. Identity and expression pattern of chemosensory proteins in *Heliothis virescens* (Lepidoptera, Noctuidae). *Insect Biochem. Mol. Biol.* **31**, 1173–1181 (2001).
85. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
86. 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).
87. Wanner, K. W. *et al.* Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol. Biol.* **16**, 107–119 (2007).

## Acknowledgements

This work was funded by Beijing Nova Program (Z161100004916119), the National Natural Science Foundation of China (31672095&31471833) and the Ministry of Education of Anhui Province (gxbjZD2016016). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

S.Y., G.W. and Y.L. designed the experiments. S.Y., D.C. and Y.L. performed the experiments. S.Y., D.C. and Y.L. contributed reagents/materials/gene identification. S.Y., D.C., G.W. and Y.L. analyzed the data. S.Y., G.W. and Y.L. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-11646-7

**Competing Interests:** The authors declare that they have no competing interests.

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



**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) 2017