



OPEN

SUBJECT AREAS:
TRANSCRIPTOMICS
BIOPHYSICAL CHEMISTRYReceived
3 July 2014Accepted
4 December 2014Published
20 January 2015

Large number of putative chemoreception and pheromone biosynthesis genes revealed by analyzing transcriptome from ovipositor-pheromone glands of *Chilo suppressalis*

Yi-Han Xia^{1*}, Ya-Nan Zhang^{1,2*}, Xiao-Qing Hou¹, Fei Li¹ & Shuang-Lin Dong¹

Correspondence and requests for materials should be addressed to F.L. (lifei@njau.edu.cn) or S.-L.D. (sldong@njau.edu.cn)

* These authors contributed equally to this work.

¹Education Ministry Key Laboratory of Integrated Management of Crop Diseases and Pests, College of Plant Protection, Nanjing Agricultural University, Nanjing, China, ²College of Life Sciences, Huaibei Normal University, Huaibei, China.

The chemoreception role of moth ovipositor has long been suggested, but its molecular mechanism is mostly unknown. By transcriptomic analysis of the female ovipositor-pheromone glands (OV-PG) of *Chilo suppressalis*, we obtained 31 putative chemoreception genes (9 *OBPs*, 10 *CSPs*, 2 *ORs*, 1 *SNMP*, 8 *CXEs* and 1 *AOX*), in addition to 32 genes related to sex pheromone biosynthesis (1 *FAS*, 6 *Dess*, 10 *FARs*, 2 *ACOs*, 1 *ACC*, 4 *FATPs*, 3 *ACBPs* and 5 *ELOs*). Tissue expression profiles further revealed that *CsupCSP2* and *CsupCSP10* were OV-PG biased, while most chemoreception genes were highly and preferably expressed in antennae. This suggests that OV-PG employs mostly the same chemoreception proteins as in antennae, although the physiological roles of these proteins might be different in OV-PG. Of the 32 pheromone biosynthesis related genes, *CsupDes4*, *CsupDes5* and *CsupFAR2* are strongly OV-PG biased, and clustered with functionally validated genes from other moths, strongly indicating their involvement in specific step of the pheromone biosynthesis. Our study for the first time identified a large number of putative chemoreception genes, and provided an important basis for exploring the chemoreception mechanisms of OV-PG in *C. suppressalis*, as well as other moth species.

Olfaction plays a critical role in guiding insect behaviors, such as finding of mating partners, food plants and oviposition sites. The periphery process of insect olfaction is thought to involve several major steps. Firstly, external chemical volatiles enter into the chemosensilla and then are captured by odorant binding proteins (OBPs)^{1,2} or chemosensory proteins (CSPs)³⁻⁵. Secondly, the OBP or CSP bound chemical volatiles are transported to the olfactory receptors (ORs)⁶ located on dendrite membrane of sensory neurons, triggering the transduction of chemical signals to electric signals. In addition, sensory neuron membrane proteins (SNMPs) may also participate in the chemoreception. Two subtypes of SNMP proteins, SNMP1 and SNMP2, have been identified from different insects⁷. After the transduction, odorant molecules will be rapidly deactivated to resume the sensitivity of the sensory neurons, by odorant degrading enzymes (ODEs) such as carboxylesterases (CXEs)^{8,9} and aldehyde oxidases (AOXs)⁵. Antennae are the primary olfactory organs in insects, but other organs such as mouthpart appendages, legs, wings and female ovipositors also bear some olfactory sensilla, and thus play some distinct roles in insect behaviors. In particular, ovipositor in moths has long been proposed to play important chemoreception roles in oviposition site selection. In female *Monopis crocicapitella*, multipotous sensilla on the ovipositor were observed and supposed to play role in perception of volatiles for the general assessment of the oviposition site¹⁰. In Pyralidae, sensilla on ovipositor were supposed to be involved in the perception of the oviposition-detering pheromone secreted by the larvae and other volatiles in *Homoiosoma nebulella*¹¹ and *Ephestia kuehniella*²⁰. Furthermore, expression of pheromone binding protein 2 (PBP2) (a sex pheromone specific OBP) and the sex pheromone specific OR were detected in the female ovipositor of a noctuid



*Heliothis virescens*¹², suggesting a possible role of the ovipositor in feedback regulation of the female sex pheromone biosynthesis and emission from the sex pheromone gland. However, the chemoreception related proteins expressed in moth ovipositors and their functions are mostly unknown.

In most female moths, the ovipositor is anatomically in close connection onto the sex pheromone gland (PG), the site of sex pheromone biosynthesis and emission¹³. The ovipositor-sex pheromone glands (OV-PG) together provide an important role for the reproductive behavior of moths. Sex pheromone components in moth species are primarily C10–C18 long straight-chain, unsaturated derivatives of fatty acids, with the carbonyl carbon modified to form an oxygen-containing functional group such as alcohol, aldehyde, or acetate ester¹⁴. Their biosynthesis starts by acetyl-CoA carboxylases (ACCs) and fatty acid synthetases (FASs) catalyzing the saturated fatty acid precursor malonyl-CoA from acetyl-CoA in the first committed biosynthesis step^{15,16}. Similar fatty-acid metabolism enzymes function in desaturation, chain-shortening by β -oxidation, functional-group modifications by reduction and finally producing the pheromone components by acetylation or oxidation¹⁷. Different combinations of these enzymes can produce unique species-specific pheromone blends in different species. So far, two classes of essential enzymes involved in moth pheromone synthesis have been functionally identified. The first is the desaturases (Dess), functioning to introduce double bonds into pheromone precursors. Dess are the most intensively studied class of enzymes involved in sex pheromone biosynthesis¹⁸. Based on the position that double bond is introduced, 5 types of desaturases ($\Delta 9$, $\Delta 10$, $\Delta 11$, $\Delta 12$ and $\Delta 14$) have been identified^{19,20}. The second is the fatty acid reductases (FARs), responsible for reducing fatty acids to alcohols, and have been functionally identified in a couple of moth species, such as *pgFAR-Z/E* in *Ostrinia nubilalis*²¹ and *pgFAR* in *Bombyx mori*²². Besides, some other important genes are postulated to be involved in moth pheromone production and remain to be characterized. Acyl-CoA oxidases (ACOs), are responsible for lipid metabolism by catalyzing the conversion of acyl-CoA into trans-2-enoyl-CoA during fatty acid β -oxidation; fatty acid transport proteins (FATPs) are integral membrane-bound proteins found in both the plasma membrane and endoplasmic reticulum, several of which facilitate the uptake and activation of exogenous long chain fatty acids²³; acyl-CoA binding proteins (ACBPs) bind acyl-CoA esters with high specificity and affinity and thus are thought to act as intracellular transporters of acyl-CoA esters between different enzymatic systems²⁴.

The rice stem borer, *Chilo suppressalis* (Walker) (Lepidoptera: Pyralidae), is a notorious rice pest in East Asian countries, causing great economic losses to rice crops²⁵. It also oviposites on and damages variety of other crops, such as corn and cane²⁶, suggesting an important role of chemoreception in oviposition and food plant selection. The sex pheromones of female *C. suppressalis* were initially identified as a binary mixture of (Z)-11-hexadecenal (Z11-16:Ald) and (Z)-13-octadecenal (Z13-18:Ald)^{27,28}. Later on, the third component (Z)-9-hexadecenal (Z9-16:Ald) was discovered²⁹, and the tertiary blend of Z11-16:Ald, Z13-18:Ald and Z9-16:Ald at the ratio of 48:6:5 had the maximum attraction²⁹. However, genes involved in the pheromone biosynthesis of *C. suppressalis* have not been explored. Our present study reported a genetic database of the genes expressed in the female OV-PG of *C. suppressalis* by using the Illumina HiSeq^(TM) 2000 sequencing platform. Totally, 63 putative genes related with chemoreception and pheromone biosynthesis were identified. The tissue expression profile investigation showed that some of those genes might play important roles in oviposition behavior and sex pheromone biosynthesis of *C. suppressalis*.

Results

Transcriptome sequencing and sequence assembly. Sequencing of a cDNA library prepared from mRNA of the OV-PG of *C. suppressalis*

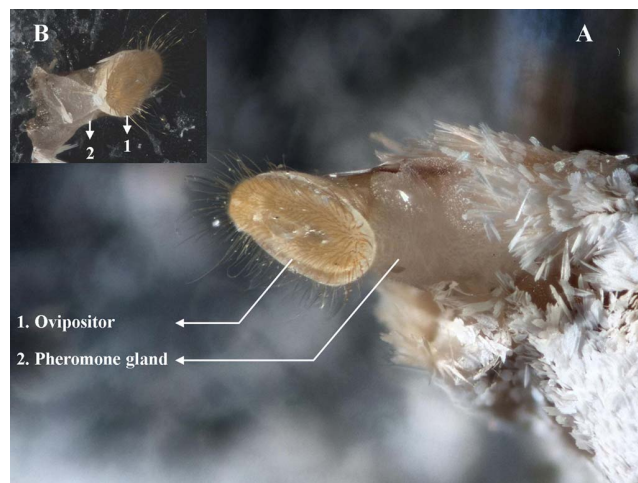


Figure 1 | Dissection of *C. suppressalis* ovipositor-pheromone gland for RNA extraction. (A). Ovipositor-pheromone gland was forced out by squeezing the abdomen. (B). Ovipositor-pheromone gland was cut from the 8th abdominal segment.

(Fig. 1) provided about 65 million raw reads ($\gg 4$ Gb). After trimming adaptor sequences and removing low quality sequences, about 54 million clean reads remained, and were assembled into 66,971 contigs (≥ 75 bp) with a mean length of 359 bp and the N50 length of 579 bp. After clustering and redundancy filtering, we finally acquired 37,619 longer sequences (≥ 150 bp), including 5,942 Clusters (15.8%) and 31,677 Singletons (84.2%) with a mean length of 628 bp and the N50 length of 884 bp. We defined these 37,619 sequences as unigenes according to some recently published papers³⁰, although each of them may not necessarily represents a unique gene. Of the 37,619 unigenes, those ≥ 1000 bp accounted for 16.26% of the transcriptome assembly (Fig. 2).

Homology analysis and Gene Ontology (GO) annotation. Of the 37,619 unigenes, 18,037 were matched by the BLASTX homology search to the entries in NCBI non-redundant (nr) protein database with a cut-off E-value of 10^{-5} . The highest percentage of matched sequences is to *Danaus plexippus* (57.97%), followed by *Bombyx mori* (6.49%), *Papilio xuthus* (4.09%), *Tribolium castaneum* (2.57%), *Papilio polytes* (1.29%), *Acyrtosiphon pisum* (0.91%), and *Manduca sexta* (0.83%). The remaining 25.8% sequences were matched to other insects (Fig. 3).

The Gene Ontology (GO) annotation was used to classify the 37,619 unigenes into different functional groups using BLAST2GO. Based on the sequence homology, 8,506 unigenes (22.61%) could be

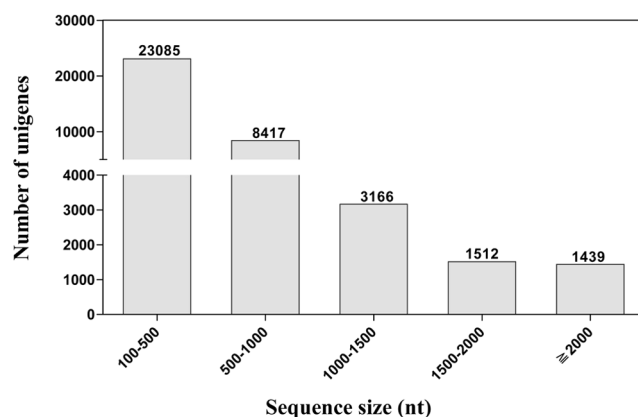


Figure 2 | Distribution of Unigenes size in the *C. suppressalis* transcriptome assembly.

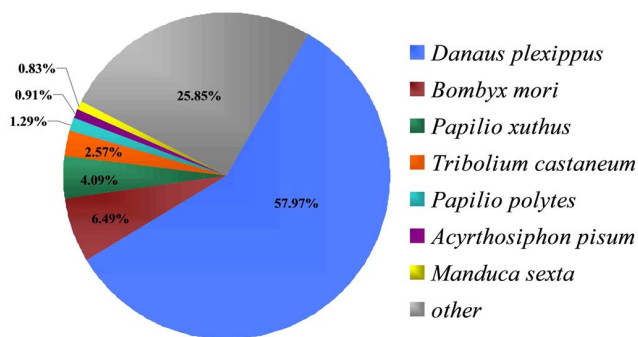


Figure 3 | Percentage of homologous hits of the *C. suppressalis* transcripts to other insect species. The *C. suppressalis* transcripts were searched by BLASTx against the non-redundancy protein database with a cutoff E-value 10^{-5} . Species which have more than 0.8% matching hits to the *C. suppressalis* transcripts are shown.

annotated, and each unigene was classified into one or more functional groups of the three biological processes (Fig. 4). Of the 8,506 annotated unigenes, more than half could align to “cellular process” (62.53%), “metabolic process” (52.92%), “binding” (51.98%), “cell” (50.48%), and “cell part” (50.48%). In total, 31,874 annotation hits (unigene-functional group) come to the biological process, 18,952 to the cellular component and 9,639 to the molecular function.

Identification of putative genes related to chemoreception and sex pheromone biosynthesis. Moth ovipositors bear some chemoreception sensilla^{10,31}, and the sex pheromone gland may also express some

chemoreception proteins such as OBPs and CSPs that are postulated to facilitate the transportation of the sex pheromones and their precursors. By homologous searches, a total of 31 putative chemoreception genes were identified, including 9 OBPs, 10 CSPs, 1 SNMP, 2 ORs and 9 ODEs (Tab. 1). In addition, 32 putative genes related to the sex pheromone biosynthesis were also obtained, which includes 1 FAS, 6 *Dess*, 10 *FARs*, 2 *ACOs*, 1 *ACC*, 4 *FATPs*, 3 *ACBPs* and 5 elongation of very long chain fatty acids (*ELOs*) (Tab. 2).

Among the 63 identified genes (Tab. 1 and 2), 6 genes were the same as sequences already deposited in the GenBank: 5 *CsupOBPs* (GenBank accession number: AGK24577.1, ACJ07120.1, AGM38609.1, AGK24580.1, ACJ07126.1) and 1 *CsupSNMP* (GenBank accession number: AFS50074.1), while the other 57 transcripts found in the current study were new in *C. suppressalis*. The abundances of the 63 genes in the transcriptome were shown in Fig. 5. Three classes of genes (*Dess*, *FARs* and *CSPs*) showed higher abundances than the others, with the *CSPs* being the highest. To validate these sequences, RT-PCR validation experiments were conducted, and the results showed that all these 63 sequences were consistent with that of the PCR products.

Expression profile of the putative genes in chemoreception and sex pheromone biosynthesis. To provide functional clues, expression profiles of the genes were investigated by RT-PCR for all 63 genes and by Quantitative Real Time RT-PCR (qPCR) measurements for selected 3 genes (Fig. 6 and 7). The results showed that only 2 chemoreception related genes (*CsupCSP2* and *CsupCSP10*) and 3 sex pheromone biosynthesis related genes (*CsupDes4*, *CsupDes5* and *CsupFAR2*) displayed higher expression in the OV-PG complex than in other tissues of *C. suppressalis*.

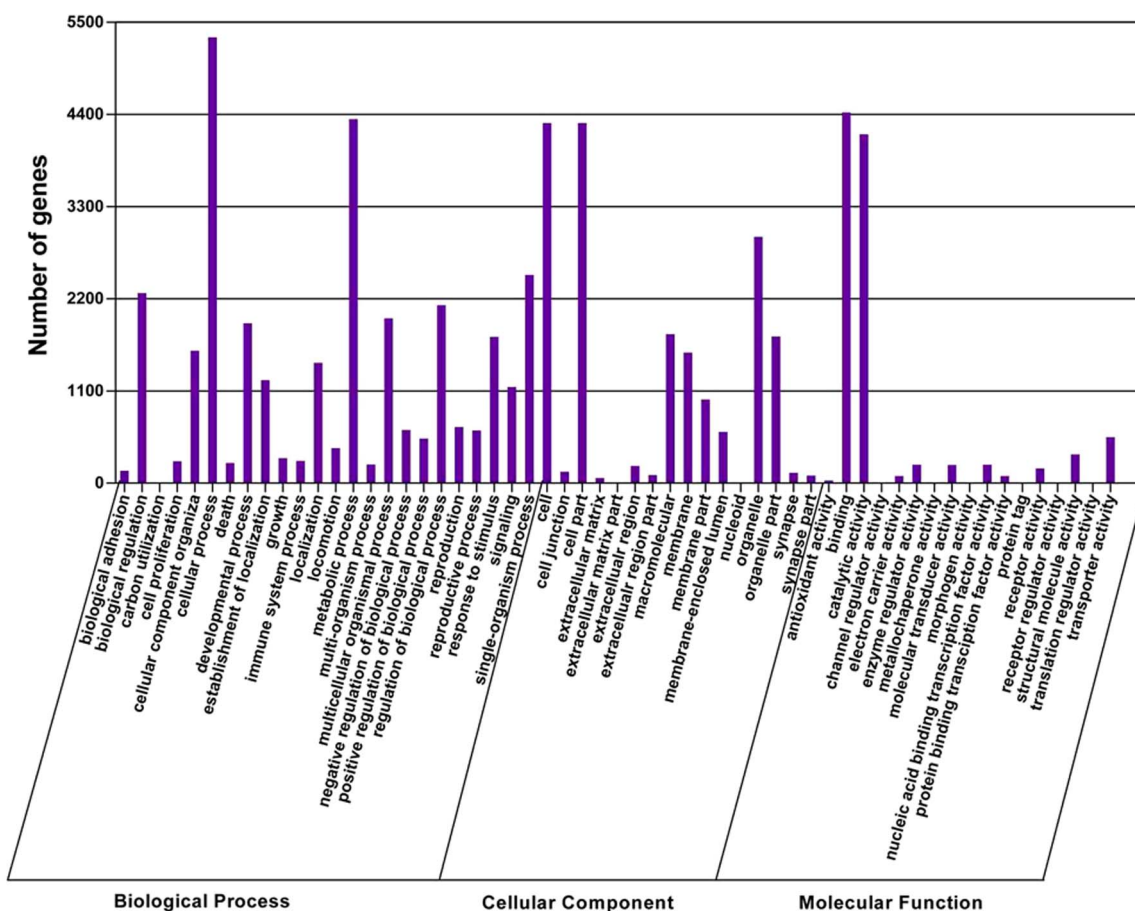


Figure 4 | Gene ontology (GO) classification of the *C. suppressalis* transcripts with Blast2GO program. One unigene could be annotated into more than one GO term.



Table 1 | Putative chemoreception transcripts in the female ovipositor-pheromone glands of *C. suppressalis*. Nucleotide sequences for the identified transcripts are given in Table S1

| Name | Gene ID | ORF (bp) | Best Blastx Match | | | | |
|---|--------------|----------|---|----------------|----------------------------------|-----------|--------------|
| | | | Name | Acc. number | Species | E value | Identity (%) |
| Odorant binding protein (OBP) | | | | | | | |
| OBP1 | CL2370 | 585 | odorant binding protein | EJ77172.1 | [<i>Danaus plexippus</i>] | 8.00E-49 | 49 |
| OBP2 | Unigene12040 | 423 | odorant-binding protein 1 | AGK24577.1 | [<i>Chilo suppressalis</i>] | 4.00E-96 | 99 |
| OBP3 | Unigene13030 | 444 | odorant binding protein | BAI44701.1 | [<i>Bombyx mori</i>] | 7.00E-46 | 55 |
| OBP4 (GOBP2) | Unigene14429 | 489 | general odorant binding protein 2 | ACJ07120.1 | [<i>Chilo suppressalis</i>] | 1.00E-110 | 98 |
| OBP5 | Unigene16075 | 729 | odorant binding protein fmxg18C17 precursor | NP_001157372.1 | [<i>Bombyx mori</i>] | 3.00E-52 | 44 |
| OBP6 | Unigene31091 | 120 | odorant binding protein 7 | AGH70103.1 | [<i>Spodoptera exigua</i>] | 8.00E-18 | 100 |
| OBP7 | Unigene3584 | 417 | odorant-binding protein | AGM38609.1 | [<i>Chilo suppressalis</i>] | 1.00E-94 | 99 |
| OBP8 | Unigene7265 | 441 | odorant-binding protein 4 | AGK24580.1 | [<i>Chilo suppressalis</i>] | 3.00E-101 | 99 |
| OBP9 (GOBP1) | Unigene19049 | 522 | general odorant binding protein 1 | ACJ07126.1 | [<i>Chilo suppressalis</i>] | 7.00E-124 | 100 |
| Chemosensory protein (CSP) | | | | | | | |
| CSP1 | CL2012 | 390 | chemosensory protein-14 | AGR44885.1 | [<i>Bombyx mori</i>] | 2.00E-62 | 84 |
| CSP 2 | CL908 | 366 | chemosensory protein 10 | AFR92094.1 | [<i>Helicoverpa armigera</i>] | 7.00E-55 | 82 |
| CSP 3 | Unigene11917 | 570 | chemosensory protein | BAF91711.1 | [<i>Papilio xuthus</i>] | 4.00E-43 | 64 |
| CSP 4 | Unigene13679 | 381 | chemosensory protein | BAF91714.1 | [<i>Papilio xuthus</i>] | 6.00E-59 | 72 |
| CSP 5 | Unigene17503 | 396 | chemosensory protein | ABB91378.1 | [<i>Helicoverpa assulta</i>] | 4.00E-64 | 75 |
| CSP 6 | Unigene28065 | 333 | chemosensory protein CSP1 | EJ76401.1 | [<i>Danaus plexippus</i>] | 5.00E-40 | 71 |
| CSP 7 | Unigene3176 | 330 | chemosensory protein | EJ67380.1 | [<i>Danaus plexippus</i>] | 3.00E-51 | 97 |
| CSP 8 | Unigene5529 | 363 | chemosensory protein 7 precursor | NP_001037068.1 | [<i>Bombyx mori</i>] | 2.00E-27 | 53 |
| CSP 9 | Unigene7291 | 372 | chemosensory protein | BAF91716.1 | [<i>Papilio xuthus</i>] | 7.00E-50 | 70 |
| CSP 10 | Unigene7788 | 198 | putative chemosensory protein | AGY49255.1 | [<i>Sesamia inferens</i>] | 2.00E-15 | 48 |
| Odorant receptor (OR) | | | | | | | |
| OR1 | CL95 | 114 | putative odorant receptor OR43, partial | AFC91751.1 | [<i>Cydia pomonella</i>] | 9.00E-09 | 69 |
| OR2 | Unigene29549 | 267 | putative odorant receptor OR12 | AFC91721.1 | [<i>Cydia pomonella</i>] | 8.00E-28 | 53 |
| Sensory neuron membrane protein (SNMP) | | | | | | | |
| SNMP | Unigene13832 | 1549 | sensory neuron membrane protein 2 | AFS50074.1 | [<i>Chilo suppressalis</i>] | 0.0 | 99 |
| Carboxyl esterase (CXE) | | | | | | | |
| CXE5 | CL2230 | 1686 | antennal esterase CXE5 | ADR64702.1 | [<i>Spodoptera exigua</i>] | 0.0 | 62 |
| CXE9 | Unigene14020 | 1677 | antennal esterase CXE9 | ACV60236.1 | [<i>Spodoptera littoralis</i>] | 0.0 | 63 |
| CXE11 | Unigene13082 | 1209 | antennal esterase CXE11 | ACV60238.1 | [<i>Spodoptera littoralis</i>] | 1.00E-167 | 60 |
| CXE14 | Unigene15349 | 381 | antennal esterase CXE14 | AEJ38205.1 | [<i>Spodoptera exigua</i>] | 1.00E-34 | 53 |
| CXE16 | CL564 | 237 | antennal esterase CXE16 | ACV60243.1 | [<i>Spodoptera littoralis</i>] | 8.00E-39 | 86 |
| CXE17 | Unigene19468 | 582 | antennal esterase CXE17 | EJ64436.1 | [<i>Danaus plexippus</i>] | 2.00E-48 | 51 |
| CXE18 | Unigene13662 | 1014 | antennal esterase CXE18 | AEJ38204.1 | [<i>Spodoptera exigua</i>] | 3.00E-160 | 70 |
| CXE20 | Unigene9656 | 378 | antennal esterase CXE20 | ADR64703.1 | [<i>Spodoptera exigua</i>] | 4.00E-41 | 55 |
| Aldehyde oxidase (AOX) | | | | | | | |
| AOX1 | CL2736 | 1116 | antennae-specific aldehyde oxidase 2 | AGQ43599.1 | [<i>Amyelois transitella</i>] | 8.00E-178 | 66 |

Odorant binding protein (OBP) and chemosensory protein (CSP). The tissue expression profiles of *CsupOBPs* are shown in Fig. 6A. The 9 *OBPs* identified in this study included 2 *GOBPs* (*OBP4*: *GOBP2*; *OBP9*: *GOBP1*) and 7 other *OBPs*. Among the 9 *CsupOBPs*, 3 genes (*CsupOBP1*, *CsupOBP2* and *CsupOBP6*) displayed a very wide range of tissue distribution in all 6 tissues, while 4 genes (*CsupOBP3*, *CsupOBP4*, *CsupOBP7* and *CsupOBP9*) were expressed specifically or predominately in adult antennae and legs and 1 (*CsupOBP8*) was specifically expressed in abdomen. In addition, 4 genes (*CsupOBP1*, *CsupOBP2*, *CsupOBP3* and *CsupOBP6*) were detected in the OV-PG.

Compared to *CsupOBPs*, the abundances of *CsupCSPs* were much higher in OV-PG of *C. suppressalis*. Most *CsupCSPs* were expressed in similar levels between olfactory and non-olfactory tissues, while *CsupCSP2* and *CsupCSP10* were not expressed in olfactory tissues but in OV-PG and legs (Tab. 1 and Fig. 6A).

Odorant receptor (OR) and sensory neuron membrane protein (SNMP). 2 putative *CsupORs* were identified from the *C. suppressalis* OV-PG. *CsupOR2* encoding a protein with 53% identity to putative odorant receptor OR12 of *Cydia pomonella* (GenBank accession number: AFC91721.1) was detected in OV-PG, adult antennae and

abdomens. In contrast, *CsupOR1* encoding a protein with 69% identity to putative odorant receptor OR43 of *C. pomonella* (GenBank accession number: AFC91751.1) was expressed in all chemosensory tissues at low levels and not expressed in OV-PG and abdomens (Tab. 1 and Fig. 6A).

1 putative *CsupSNMP2* was obtained from *C. suppressalis*. However, it was strongly expressed in adult antennae, but not detected in OV-PG (Tab. 1 and Fig. 6A).

Aldehyde oxidase (AOX) and carboxyl esterase (CXE). AOXs catalyze the oxidation of aldehydes to carboxylic acids and they may also be involved in the degradation of sex pheromone compounds specifically in the conversion of aldehydes to carboxylic acids^{14,22}. Only 1 AOX homolog was obtained from *C. suppressalis* transcriptome, which shared 66% identity to the antennae-specific AOX2 of *Amyelois transitella* (Tab. 1). *CsupAOX* was antennae-predominantly expressed, with very weak expression in OV-PG (Fig. 6A).

Esterases are hydrolases, and hydrolysis of esters occurs during plant volatile and sex pheromone degradation⁹. We totally obtained 8 *CsupCXEs* from the *C. suppressalis* OV-PG transcriptome. Compared to the other putative chemoreception genes, all *CsupCXEs* were



Table 2 | Putative transcripts related to sex pheromone biosynthesis in the ovipositor-pheromone glands of *C. suppressalis*. Nucleotide sequences for the identified transcripts are given in Table S1

| Name | Gene ID | ORF (bp) | Best Blastx Match | | | | |
|---|--------------|----------|---|----------------|----------------------------------|-----------|--------------|
| | | | Name | Acc. number | Species | E value | Identity (%) |
| Fatty acid synthase(FAS) | | | | | | | |
| FAS1 | CL973 | 660 | fatty acid synthase | AGR49310.1 | [<i>Agrotis ipsilon</i>] | 0.0 | 74 |
| Desaturase(DES) | | | | | | | |
| Des1 | CL1833 | 1116 | acyl-CoA delta-9 desaturase | CAJ27975.1 | [<i>Manduca sexta</i>] | 0.0 | 91 |
| Des2 | Unigene14403 | 774 | desaturase HassGATD | EHJ71380.1 | [<i>Danaus plexippus</i>] | 3.00E-166 | 80 |
| Des3 | Unigene14404 | 471 | desaturase | AAQ74260.1 | [<i>Spodoptera littoralis</i>] | 3.00E-143 | 82 |
| Des4 | Unigene16774 | 356 | delta11 desaturase | AAF81787.1 | [<i>Helicoverpa zea</i>] | 1.00E-60 | 71 |
| Des5 | Unigene16772 | 270 | acyl-CoA delta-11 desaturase | AAK21863.1 | [<i>Trichoplusia ni</i>] | 1.00E-10 | 53 |
| Des6 | Unigene18619 | 885 | predicted delta11-like | XP_004925563.1 | [<i>Bombyx mori</i>] | 6.00E-139 | 68 |
| Fatty-acyl reductase(FAR) | | | | | | | |
| FAR1 | CL2119 | 1266 | fatty-acyl CoA reductase 5 | EHJ72233.1 | [<i>Danaus plexippus</i>] | 0.0 | 77 |
| FAR2 | Unigene15762 | 1407 | fatty-acyl reductase | AGG19592.1 | [<i>Ostrinia latipennis</i>] | 1.00E-122 | 47 |
| FAR3 | CL442 | 1374 | fatty-acyl CoA reductase 3 | ADI82776.1 | [<i>Ostrinia nubilalis</i>] | 0.0 | 83 |
| FAR4 | Unigene10614 | 1388 | fatty-acyl CoA reductase 3 | XP_004930522.1 | [<i>Bombyx mori</i>] | 0.0 | 82 |
| FAR5 | Unigene13646 | 1297 | fatty-acyl CoA reductase 4 | ADI82777.1 | [<i>Ostrinia nubilalis</i>] | 0.0 | 73 |
| FAR6 | Unigene14049 | 1875 | fatty-acyl CoA reductase 2 | ADI82775.1 | [<i>Ostrinia nubilalis</i>] | 0.0 | 82 |
| FAR7 | Unigene14605 | 1614 | putative fatty acyl-CoA reductase CG5065-like | XP_004926017.1 | [<i>Bombyx mori</i>] | 0.0 | 72 |
| FAR8 | Unigene16587 | 564 | putative fatty acyl-CoA reductase CG8306-like | AGR49318.1 | [<i>Agrotis ipsilon</i>] | 8.00E-98 | 81 |
| FAR9 | Unigene28511 | 330 | fatty-acyl CoA reductase 1 | ADI82774.1 | [<i>Ostrinia nubilalis</i>] | 9.00E-59 | 78 |
| FAR10 | Unigene7020 | 644 | fatty-acyl CoA reductase 6 | EHJ76493.1 | [<i>Danaus plexippus</i>] | 6.00E-69 | 52 |
| Acyl-CoA oxidase(ACO) | | | | | | | |
| ACO1 | Unigene10405 | 414 | PREDICTED: probable peroxisomal acyl-coa enzyme oxidase 1-like | XP_004932400.1 | [<i>Bombyx mori</i>] | 1.00E-79 | 77 |
| ACO2 | Unigene14208 | 1059 | PREDICTED: probable peroxisomal acyl-coa enzyme oxidase 1-like | XP_004932404.1 | [<i>Bombyx mori</i>] | 0.0 | 83 |
| Acetyl-CoA Carboxylase(ACC) | | | | | | | |
| ACC1 | Unigene26870 | 363 | putative acetyl-CoA carboxylase | EHJ72299.1 | [<i>Danaus plexippus</i>] | 4.00E-59 | 77 |
| Fatty acid transport protein(FATP) | | | | | | | |
| FATP1 | Unigene14963 | 2061 | Fatty acid transport protein | BAJ33524.1 | [<i>Ostrinia scapularis</i>] | 0.0 | 86 |
| FATP2 | Unigene15353 | 744 | Fatty acid transport protein | ACT22576.1 | [<i>Manduca sexta</i>] | 0.0 | 84 |
| FATP3 | Unigene16281 | 1932 | long-chain fatty acid transport protein 4-like | XP_004929241.1 | [<i>Bombyx mori</i>] | 0.0 | 81 |
| FATP4 | Unigene22497 | 452 | long-chain fatty acid transport protein 4-like | XP_004929240.1 | [<i>Bombyx mori</i>] | 7.00E-73 | 70 |
| Acyl-CoA binding protein(ACBP) | | | | | | | |
| ACBP1 | CL2090 | 759 | acyl-CoA binding domain-containing protein 6-like | NP_001040308.1 | [<i>Bombyx mori</i>] | 1.00E-126 | 71 |
| ACBP2 | CL331 | 798 | acyl-CoA binding domain-containing protein 5-like | XP_004933263.1 | [<i>Bombyx mori</i>] | 3.00E-124 | 77 |
| ACBP3 | Unigene28350 | 182 | acyl-CoA binding protein-like protein | EHJ64012.1 | [<i>Danaus plexippus</i>] | 4.00E-23 | 85 |
| Elongation of very long chain fatty acids(ELO) | | | | | | | |
| ELO1 | CL1 | 933 | elongation of very long chain fatty acids protein AAEL008004-like | XP_004931946.1 | [<i>Bombyx mori</i>] | 3.00E-178 | 90 |
| ELO2 | CL2453 | 450 | elongation of very long chain fatty acids protein AAEL008004-like | XP_004931947.1 | [<i>Bombyx mori</i>] | 4.00E-92 | 86 |
| ELO3 | Unigene119 | 513 | elongation of very long chain fatty acids protein AAEL008004-like | XP_004924776.1 | [<i>Bombyx mori</i>] | 3.00E-94 | 81 |
| ELO4 | Unigene16941 | 939 | elongation of very long chain fatty acids protein AAEL008004-like | XP_004931951.1 | [<i>Bombyx mori</i>] | 0.0 | 87 |
| ELO5 | Unigene7895 | 708 | elongation of very long chain fatty acids protein 6-like | XP_004924792.1 | [<i>Bombyx mori</i>] | 1.00E-142 | 81 |

expressed at higher levels in male antennae than in female antennae, except for *CXE11* that displayed the contrary. 5 *CsupCXEs* (*CsupCXE11*, 14, 17, 18 and 20) were expressed weakly in OV-PG (Fig. 6A).

Fatty acid synthase (FAS), *Acetyl-CoA carboxylase (ACC)* and *Acyl-CoA oxidase (ACO)*. FAS has been reported to catalyze the conversion of malonyl-CoA and NADPH to saturated fatty acids¹⁵. 1 putative *CsupFAS* cDNA in full length was identified from the OV-PG complex of *C. suppressalis* (Tab. 2 and Fig. 6B). It contained an

ORF of 660 bp, encoding a protein with 74% amino acid identity to the FAS of *Agrotis ipsilon* (Gene Bank accession number: AGR49310.1). The RT-PCR revealed that *CsupFAS* (CL973) was expressed in all tested tissues of *C. suppressalis*. Similarly, the *CsupACC* gene was also expressed in all tissues. The gene encoded an ACC with 77% identity to ACC1 of *Danaus plexippus* (GenBank accession number: EHJ72299.1). As for the 2 *CsupACOs*, *CsupACO1* and *CsupACO2* shared 77% and 83% identity to ACO1 of *B. mori* (GenBank accession number: XP_004932400.1), respectively. Both *CsupACOs* were expressed in all tested tissues except for the fat body.

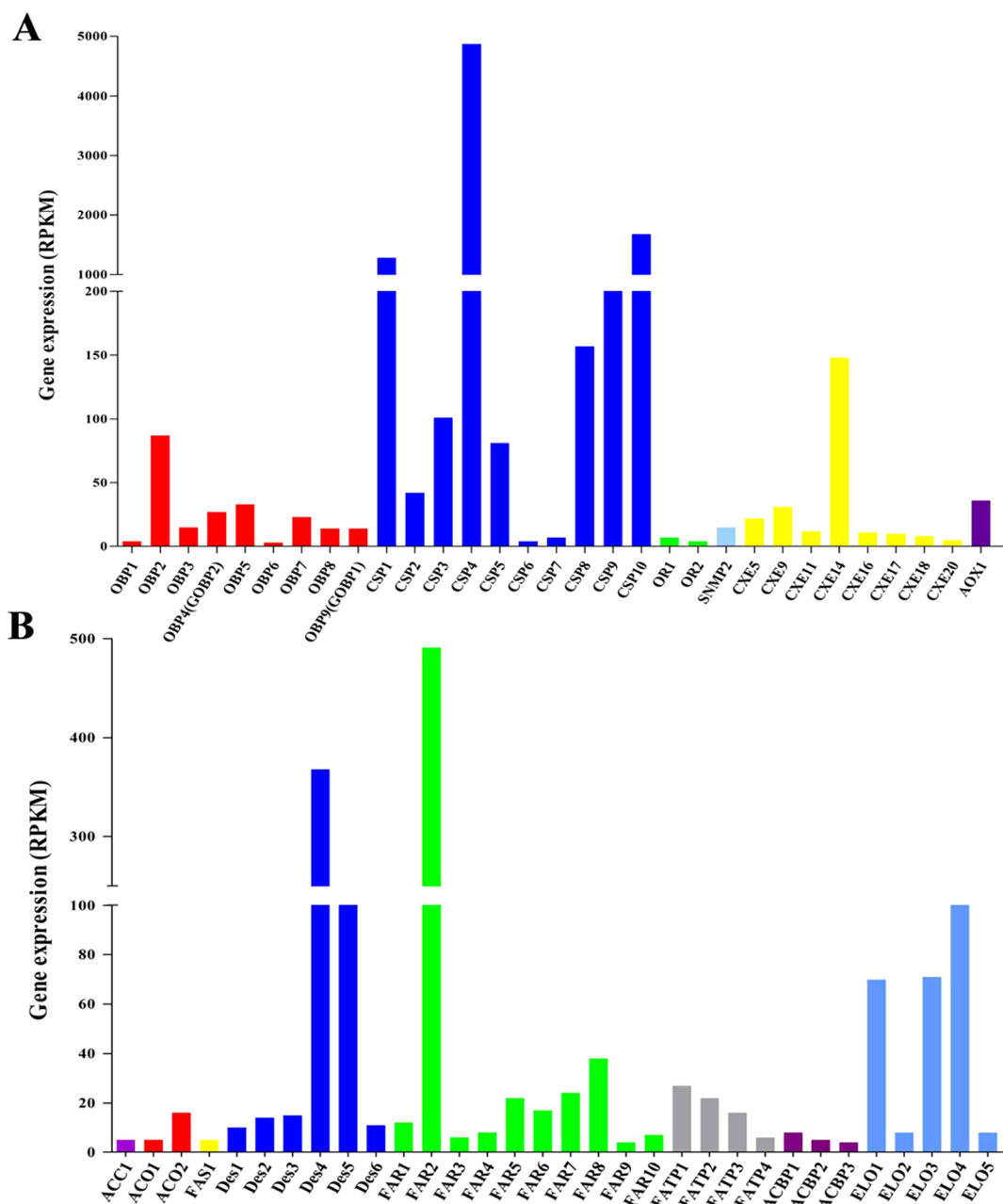


Figure 5 | Abundances of 63 transcripts which are putatively involved in chemoreception (A) and sex pheromone biosynthesis (B) in the transcriptome dataset of *C. suppressalis* ovipositor-pheromone gland. The genes expression abundance is indicated as the Reads Per Kilobase per Million mapped reads (RPKM) values.

Desaturase (Des). Among 6 *desaturases*, *CsupDes4* and *CsupDes5* displayed highly biased expression in OV-PG, while *CsupDes2* highly fat body biased in expression, *CsupDes3* and *CsupDes6* were detected weakly in epidermis, thoraxes, legs and wings. All the later 3 genes were not detected to be expressed in OV-PG by the RT-PCR.

Fatty acyl-CoA reductase (FAR). Of the 10 newly identified *CsupFARs*, only *CsupFAR2* was predominantly expressed in OV-PG. *CsupFAR10* was only detected in the thoraxes of female *C. suppressalis*, while 8 genes (*CsupFAR1*, 3, 4, 5, 6, 7, 8 and 9) displayed a wide range of tissue distribution.

Fatty acid transport protein (FATP), Acyl-CoA binding protein (ACBP) and Elongation of very long chain fatty acids (ELO). All 4 *CsupFATPs* (except for *CsupFATP3*) and 3 *CsupACBPs* were expressed at very low levels in the OV-PG, and were not OV-PG

biased (Fig. 6B). *CsupELO1-5* encoded proteins with high identity (90%, 86%, 81%, 87% and 81%) to *ELOs* of *B. mori* (GenBank accession number: XP_004931946.1, XP_004931947.1, XP_004924776.1, XP_004931951.1 and XP_004924792.1), respectively. *CsupELO1*, 3, and 5 were expressed both in high levels in OV-PG as well as several other tissues, the other 2 genes were very weakly expressed in the OV-PG complex (Fig. 6B).

Phylogenetic analyses. In order to assign putative functions to *CsupOBPs*, *CsupCSPs*, *CsupCXEs*, *CsupDess* and *CsupFARs*, phylogenetic analyses were conducted for each group of the genes. The OBP phylogenetic tree revealed that *CsupOBP4* (*CsupGOBP2*) and *CsupOBP9* (*CsupGOBP1*) were clustered into PBP/GOBP clade, and other *CsupOBPs* were distributed in five different groups (Fig. 8). In addition, the 10 putative *CsupCSPs* were clustered with at least one lepidopteran orthologous gene (Fig. S1). The 8 *CsupCXEs* were

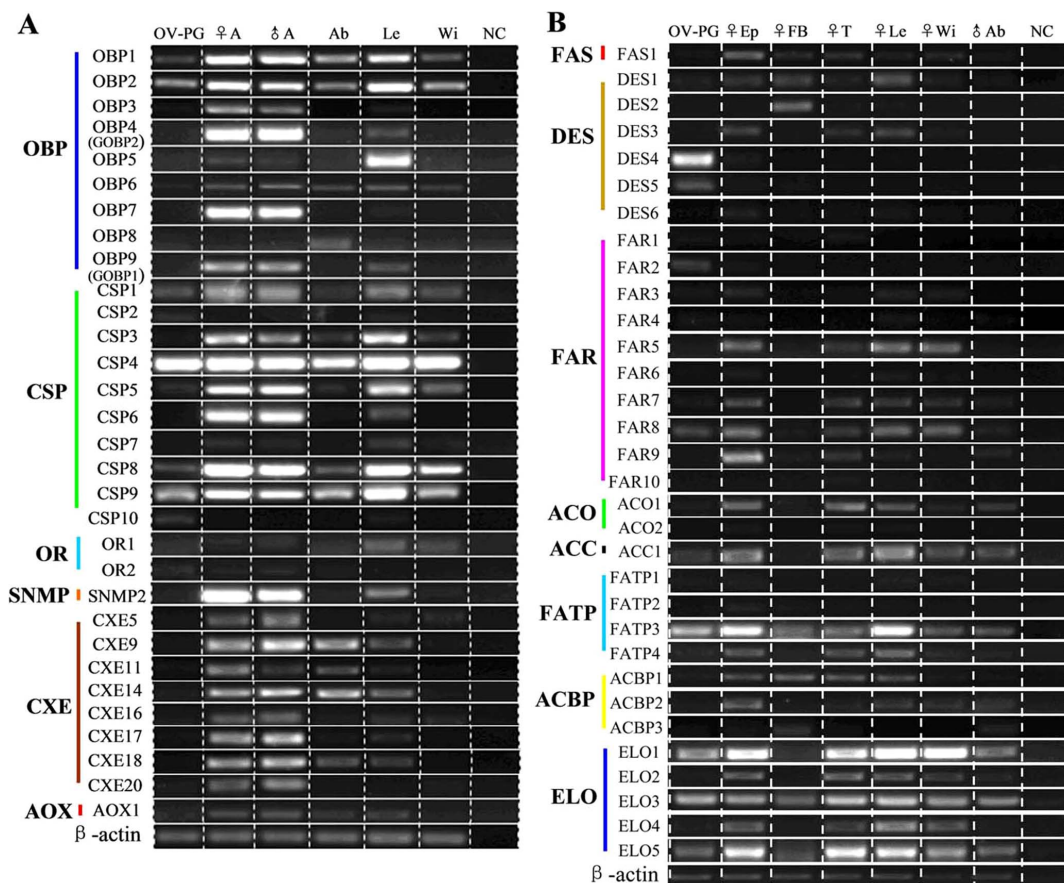


Figure 6 | Expression of *C. suppressalis* chemoreception and pheromone biosynthesis related transcripts in different adult tissues, determined by semi-quantitative RT-PCR. β -actin gene was used as the positive control and NC (no cDNA template) as the negative control. OV-PG, female ovipositor-pheromone glands; A, antennae; Ab, abdomens (mixture of equivalent male and female cDNA in panel A); Le, legs (mixture of equivalent male and female cDNA in panel A); Wi, wings (mixture of equivalent male and female cDNA in panel A); Ep, epidermis; Fb, fat body; T, thoraxes. ♀, female; ♂, male. Primers used for the RT-PCR were listed in Table S3.

distributed into three groups, with *CsupCXE14*, *CsupCXE17*, *CsupCXE18* and *CsupCXE20* in group (A) (mitochondrial, cytosolic, and secreted esterases), *CsupCXE9* and *CsupCXE11* in group (B&C) (higher and lower dipteran microsomal α -esterases) and *CsupCXE5* and *CsupCXE16* in group (D) (integumental/antennal esterases) (Fig. S2).

A phylogenetic tree using Des protein sequences from *Clubiona parallela*, *Lampronia capitella*, *O. nubilalis*, *Ostrinia scapularis*, *Spodoptera littoralis* and some other lepidopteran species (Fig. 9) showed that *CsupDes6* was clearly clustered into the clade of $\Delta 14$ desaturase, while *CsupDes1*, *CsupDes2* and *CsupDes4* were clustered in the clade of $\Delta 9$ ($16C > 18C$), $\Delta 9$ ($14C - 26C$) and $\Delta 11$ desaturase, respectively. Interestingly, the *CsupDes3* and *CsupDes5* were clustered in two dif-

ferent subclasses from other identified Dess (Fig. 9). In the FAR phylogenetic tree, *CsupFAR2* was clustered with other lepidopteran pgFAR sequences^{21,22,32-35}. Most *CsupFAR* sequences were clustered with at least one lepidopteran orthologous gene (Fig. S3), with only one gene (*CsupFAR9*) having no counterpart.

Discussion

It has long been reported that some olfactory sensilla are distributed on the ovipositor^{10,31}, and thus ovipositor may play olfactory roles in detection of plant odors, ovipositor-detering pheromones (ODEs) and sex pheromones¹⁰. To better understand the olfactory role of the moth ovipositor, we conducted investigations of chemoreception genes by analyzing the transcriptomic data of *C. suppressalis* OV-

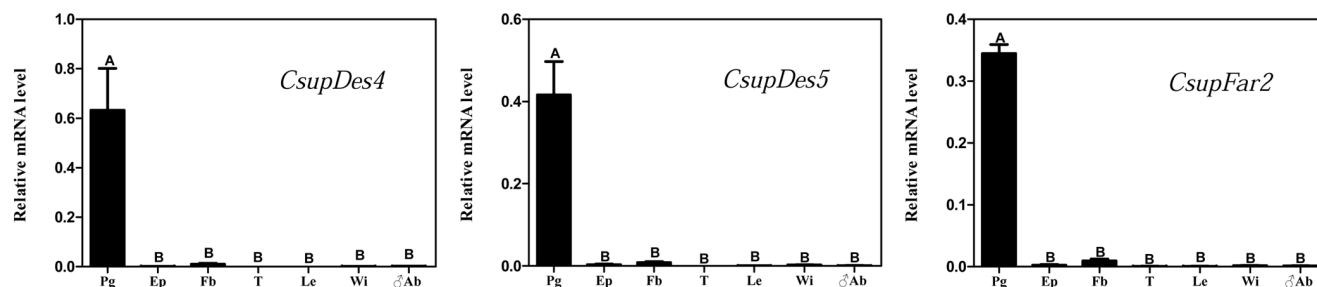


Figure 7 | Relative expression levels of 3 putative pheromone biosynthesis related transcripts in different female tissues, determined by qPCR. OV-PG, ovipositor-pheromone glands; Ep, epidermis; Fb, fat body; T, thoraxes; Le, legs; Wi, wings; ♂Ab, male abdomens. Primers used for qPCR were listed in Table S3.

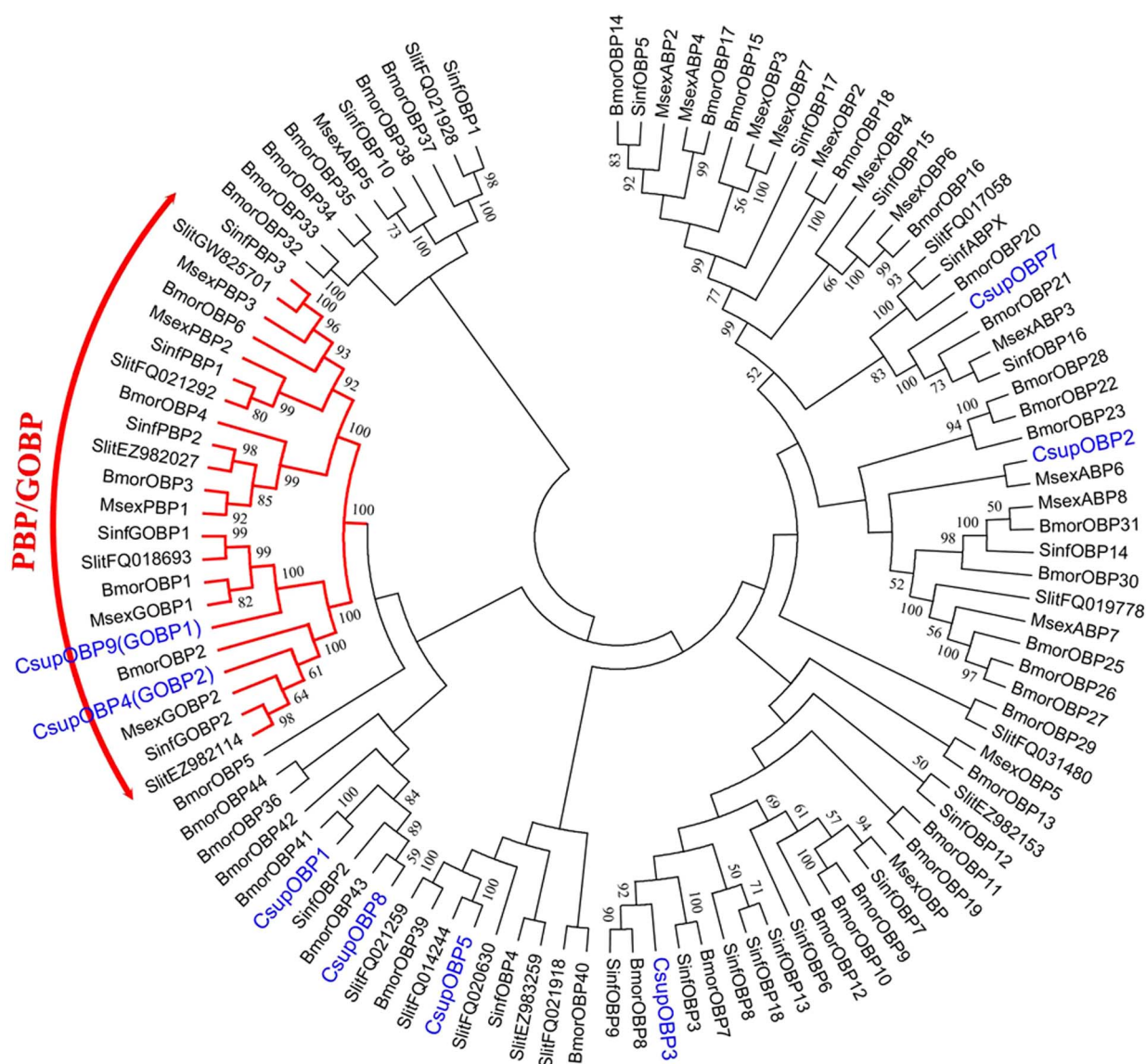


Figure 8 | Phylogenetic tree of putative OBPs from *C. suppressalis* and some other Lepidoptera species. The tree was constructed with MEGA5.0, using the Neighbour-joining method. Values indicated at the nodes are bootstrap values based on 1000 replicates, and the bootstrap values < 50% are not shown. Amino acid sequences of the species and their accession numbers are given in Table S2.

PG, and of expression profiles of these genes by semi-quantitative RT-PCR. Our results provide direct molecular evidences for the olfactory role of moth ovipositors, and also the important basis for further elucidation of the molecular mechanisms of the olfaction, as well as the sex pheromone biosynthesis in the OV-PG of *C. suppressalis*.

OBPs and CSPs are thought to be responsible for the binding and transport of the hydrophobic molecules including pheromones and plant volatiles^{1,3,4}. Among the 9 *CsupOBPs* identified in the present study, none was OV-PG specific or biased. In contrast, *CsupOBP3*, 4 (*GOBP1*), 7 and 9 (*GOBP2*) were strongly antennae-biased, indicating more important olfactory roles they play in antennae. Similar as OBPs, most *CsupCSPs* were expressed in multiple tissues, but *CsupCSP10* was weakly but more specifically detected in OV-PG (Fig. 6A), and deserves further studies with its function. Compared with 9 OBPs and 10 CSPs, only 2 *CsupORs* were identified in the OV-PG. It is noted that the Orco, acting as a chaperone and dimerization partner for other ORs^{5,36}, was not found in the OV-PG, possibly due to the lower expression levels.

PBP is a subclass of OBP, and is commonly thought to be involved in the reception of the sex pheromones³⁷. Very interestingly, a study

with *H. virescens*¹² reported that the PBP2 and pheromone receptor (PR) specific to Z11-16:Ald (the major sex pheromone component) were expressed in the sensilla on the ovipositors. Considering the close vicinity between the ovipositor and the sex pheromone gland, the authors proposed a negative feedback regulation of pheromone biosynthesis, in which sensilla on the ovipositor mediate the biosynthesis and emission of the sex pheromones. However, we did not detect any *CsupPBP* expression in OV-PG by transcriptomic analysis. Therefore, in *C. suppressalis*, antennae and other tissues that express PBPs³⁸ may play the role in the feedback regulation. In addition, the role of ovipositor in the regulation could not be excluded. The PBP expression in the ovipositor might be too low to be detected by the approach used in the present study in *C. suppressalis*. Alternatively, other OBPs especially the *GOBP2*, instead of PBPs, may function to transport the sex pheromone components in the ovipositor. *GOBP2* has been suggested playing roles in sex pheromone detection in some moth species, by ligand binding assays and in site hybridization experiments^{39,40}.

In our current study, 9 ODE candidate genes (8 *CXEs* and 1 *AOX*) were found to be expressed in the OV-PG, although in very lower

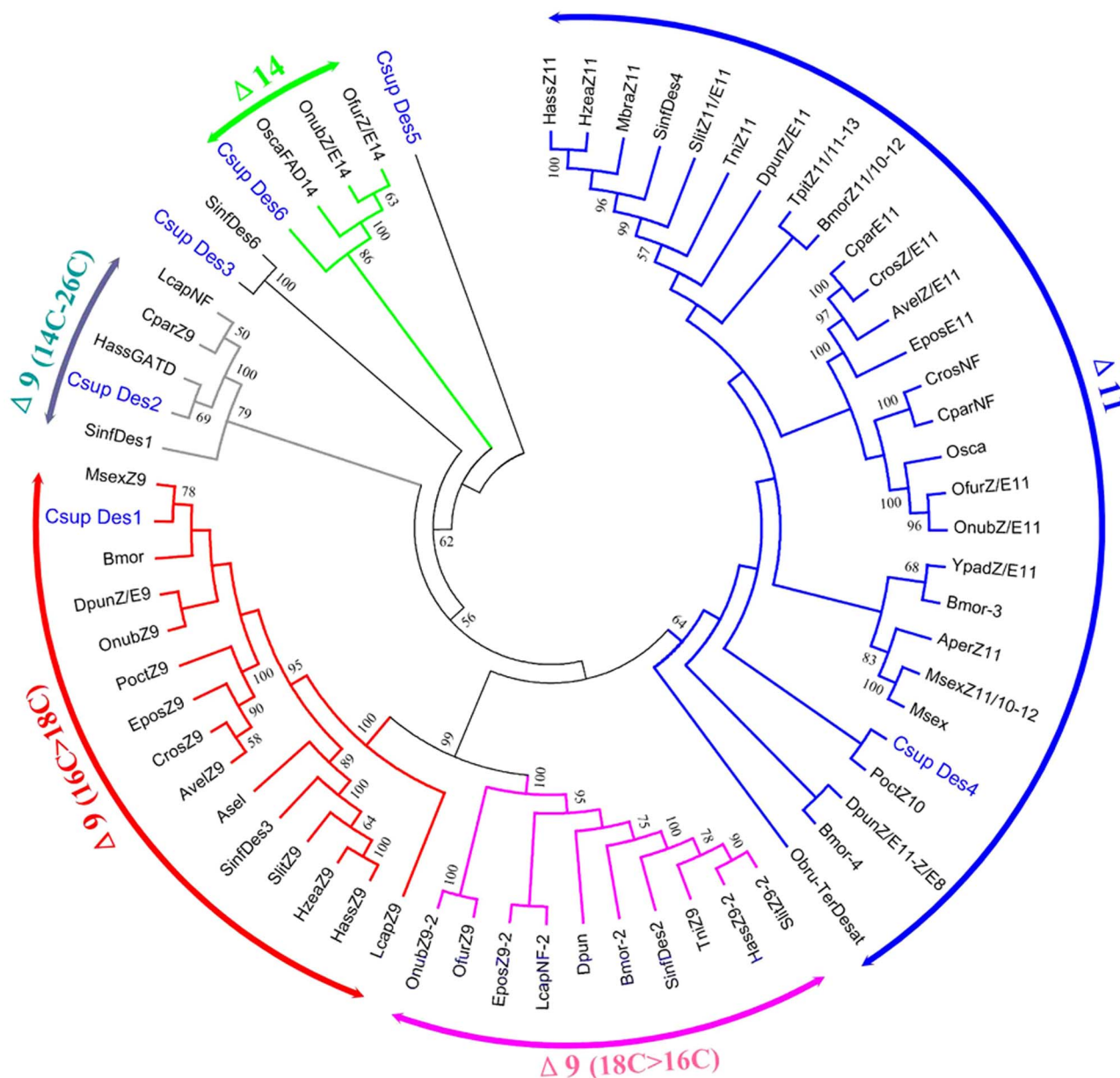


Figure 9 | Phylogenetic tree of putative Dess from *C. suppressalis* and some other Lepidoptera species. The tree was constructed with MEGA5.0, using the Neighbour-joining method. Values indicated at the nodes are bootstrap values based on 1000 replicates, and the bootstrap values < 50% are not shown. Amino acid sequences of the species and their accession numbers are given in Table S2.

level in relation to other tissues, especially the antennae. This may reflect that ovipositor is a less important olfactory organ than the antennae. In the CXEs phylogenetic tree, *CsupCXEs* were clustered into three insect CXE clades with other lepidopteran CXEs (Fig. S2), and at least one CXE in each of the three clades had been confirmed to act in the deactivation of ester plant volatiles and/or pheromone components^{8,41,42}, suggesting their importance in the degradation of ester odorants.

As for the 32 pheromone biosynthesis related genes expressed in the OV-PG, 2 *CsupDess* (*CsupDes4*, *CsupDes5*) and 1 *CsupFAR* (*CsupFAR2*) were highly expressed in OV-PG by transcriptomic analysis (Fig. 5B) and tissue expression investigation (Fig. 6B), suggesting that these 3 genes play important roles in *C. suppressalis* sex pheromone biosynthesis. In particular, *CsupDes4* shared an amino acid identity of 71% to a gene encoding a $\Delta 11$ desaturase in *Helicoverpa zea* (Genebank: AAF81787.1) (Tab. 2), and was clearly assigned to the $\Delta 11$ desaturase group in the phylogenetic tree (Fig. 9). It has been reported

that the synthesis of $\Delta 11$ -containing sex pheromones involves a step of $\Delta 11$ -desaturation catalyzed by a $\Delta 11$ desaturase^{43,44}. Considering that the $\Delta 11$ -containing Z11-16:Ald is the major sex pheromone component in *C. suppressalis*, *CsupDes4* is very likely involved in the desaturation step from saturated acids (16C) to unsaturated acids, with a double bond at 11th position of the carbon chain. In addition, *CsupDes4* may also play a role in the biosynthesis of *C. suppressalis* minor pheromone component Z13-18:Ald, if this component is formed from Z11-16:Ald by carbon chain elongation. The third pheromone component in *C. suppressalis* is Z9-16:Ald, suggesting a existence of a $\Delta 9$ desaturase in the sex pheromone glands. However, *CsupDes5*, with high and biased expression in the OV-PG, was not assigned to the $\Delta 9$ desaturase group, but a group without other members. On the other hand, *CsupDes1* and 2 were clustered into $\Delta 9$ desaturase group, but their expressions were not pheromone gland biased and very low in the OV-PG. Whether these 3 genes function to introduce the $\Delta 9$ double bond in Z9-16:Ald needs further investigation.



In the biosynthesis process of moth sex pheromones, once the specific unsaturated fatty acid precursors are produced, they will be converted into corresponding alcohols by FAR^{21,22,45}. In the current study, *CsupFAR2* was highly and more specifically expressed in OV-PG than other *CsupFARs* (Fig. 5B, 6B and 7), and the phylogenetic analysis indeed classified *CsupFAR2* into the pheromone gland FAR (pgFAR) group (Fig. S3), suggesting that *CsupFAR2* is responsible for the conversion of the unsaturated fatty acid precursors to corresponding alcohols in *C. suppressalis*.

In sex pheromone glands, aldehyde components are confirmed to be produced from alcohol oxidations by alcohol oxidase (AO)⁴⁶. Unfortunately, we did not find an AO from the transcriptomic database, similar as studies with sex pheromone gland transcriptome analysis in other species using aldehyde(s) as sex pheromone component(s), such as *H. virescens*⁴⁷, *Agrotis segetum*¹⁷ and *A. ipsilon*⁴⁸. In *Sesamia inferens*, 1 AO was found with very low expression level in the sex pheromone gland transcriptome⁴⁹. Therefore, more sensitive approaches are needed to identify the possible AO genes in moths. In addition to Des and FAR genes, we also found some genes of other classes that are thought to participate in the sex pheromone biosynthesis, such as FATP, ELO and ACBP. However, none of them were OV-PG-predominant in expression, suggesting their multiple functions in the insect physiology.

Methods

Insects rearing and collection. The rice stem borer *C. suppressalis* was originally collected from a rice field in the Jiangsu Provincial Academy of Agricultural Sciences, Nanjing, China (118.9°E, 32.0°N). The collected insects were reared in laboratory for several generations with rice seedlings according to the method reported by Shang et al.⁵⁰ until pupation and sexed as pupae. The rearing conditions were 28°C ± 1°C, 60%–70% relative humidity and a 16 h : 8 h light dark photoperiod. Adults were provided with a cotton swab dipped in 10% honey solution and renewed daily. For transcriptomic sequencing, OV-PG of 2-day-old female moths were collected, as 2- and 3-day-old moths showed highest mating activity⁵¹. For tissue expression study, tissues including male and female abdomens (for female abdomen, OV-PG was removed), legs and wings, and female OV-PG, epidermis, fat body and thoraxes were collected from 2-day-old virginal moths; while male and female antennae were collected from 2- and 3-day-old virginal moths, as more moths were needed for antennae collection. Two and three replications of tissue samples were collected for RT-PCR and qPCR measurements, respectively. All samples were collected during the 4–5 hour of the dark period and stored at –70°C until use.

cDNA library construction and Illumina sequencing. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The quantity of RNA was determined using a Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and 1.1% agarose gel electrophoresis. cDNA library construction and Illumina sequencing of the sample were performed at Beijing Genomics Institute (BGI)-Shenzhen, Shenzhen, China⁵². After the total RNA extraction and DNase I treatment, magnetic beads with Oligo (dT) are used to isolate mRNA (for eukaryotes) or by removing rRNAs from the total RNA (for prokaryotes). Mixed with the fragmentation buffer, the mRNA is fragmented into short fragments. Then cDNA is synthesized using the mRNA fragments as templates. Short fragments are purified and resolved with EB buffer for end reparation and single nucleotide A (adenine) addition. After that, the short fragments are connected with adapters. The suitable fragments are selected for the PCR amplification as templates. During the QC steps, Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System are used in quantification and qualification of the sample library. At last, the library could be sequenced using Illumina HiSeq™ 2000 platform.

De novo Assembly of Short Reads and Gene Annotation. Transcriptome de novo assembly was carried out with short reads assembling program Trinity⁵³. Trinity partitioned the sequence data into many individual de Bruijn graphs, each representing the transcriptional complexity at a given gene or locus, and then processed each graph independently to extract full-length splicing isoforms and to tease apart transcripts derived from paralogous genes. The result sequences of trinity were called Unigenes. The Unigenes larger than 150 bp were first aligned by BLASTX to protein databases, including Nr, Swiss-Prot, KEGG and COG (e-value < 10^{–5}), retrieving proteins with the highest sequence similarity with the given unigenes along with their protein functional annotations. Then, we used Blast2GO program⁵⁴ to get GO annotation of the unigenes, and got GO functional classification by using WEGO software⁵⁵.

Expression Abundance Analysis of the Unigenes. The expression abundance of these unigenes were calculated by the RPKM (Reads Per Kilobase per Million mapped reads) method⁵⁶, using the formula: $RPKM(A) = (1,000,000 \times C \times 1,000) / (N \times L)$. In the formula, RPKM (A) is the expression abundance of gene A; C is the number of

reads that uniquely aligned to gene A; N is total number of reads that uniquely aligned to all genes; and L is the number of bases on gene A. The RPKM method is able to eliminate the influence of different gene lengths and sequencing discrepancy on the calculation of expression abundance.

RNA isolation and cDNA synthesis for Reverse Transcription-PCR. Total RNA was extracted by SV 96 Total RNA Isolation System (Promega, Madison, WI, USA) following the manufacturer's instructions. RNA quality was checked with a spectrophotometer (NanoDrop™ 1000, Thermo Fisher Scientific, USA). The cDNAs from female OV-PG and other body parts (fat body, epidermis, thoraxes, antennae, legs, wings and abdomens (for female without the OV-PG)) were synthesized using the PrimeScript™ RT Master Mix (TaKaRa, Dalian, China).

Reverse Transcription-PCR Analysis. Gene specific primers across ORF of predicted chemosensory genes were designed using Beacon Designer 7.6 and Primer Premier 5.0 (PREMIER Biosoft International, CA, USA). The sequences of these primers were listed in Tab. S1. PCR experiments including negative controls (no cDNA template) were carried out in a MyCycler™ (Bio-Rad, USA) under the following conditions: 94°C for 3 min; 32 (35 for *CsupACO2*, *CsupACBP3*, *CsupFAR7*, *CsupELO2* and *CsupFATP4*) cycles at 94°C for 30 sec, 60°C for 40 sec, and 72°C for 50 sec, and final incubation for 10 min at 72°C. The reactions were performed in 25 µl with 1 µl of single-stranded cDNA, 4.0 mM MgCl₂, 0.4 mM dNTP, 0.8 mM for each primer and 2.5 U rTaq DNA polymerase (TaKaRa, Dalian, Liaoning, China). PCR products were analyzed by electrophoresis on 2.0% w/v agarose gel in TAE buffer (40 mmol/L Tris-acetate, 2 mmol/L Na₂EDTA·H₂O) and the resulting bands were visualized with ethidium bromide and digitized using Gel Capture (China).

In addition, all transcripts were chosen to perform a second biological replication in order to check the repeatability of the tissue expression. To validate the predicted sequences of chemoreception genes, the PCR products were purified by using the AxyPrep™ PCR Cleanup Kit (Axygen), and then were sub-cloned into a T/A plasmid using the pEASY-T3 cloning vector system (TransGene, China) following manufacturer's instructions. The plasmid DNA was used to transform into Trans1-T1 competent cells. Positive clones were checked by PCR and were sequenced by GenScript (Nanjing, China).

Quantitative Real Time-PCR Validation. The expression profiling of a total of 3 putative sex pheromone synthesis genes was carried out to validate the accuracy of the RT-PCR results using quantitative real-time-PCR (qPCR) experiments. The qPCR was performed on an ABI 7500 (Applied Biosystems, Foster City, CA, USA) using a mixture of 10 µl 2 × SYBR Green PCR Master Mix, 0.4 µl each primer (10 µM), 2.5 ng of sample cDNA, and 6.8 µl sterilized ultrapure H₂O. The reaction programs were 30s at 95°C, 40 cycles of 95°C for 5s and 60°C for 34s. The results were analyzed using the ABI 7500 analysis software SDS 1.4. The qPCR primers (Tab. S3) were designed using Beacon Designer 7.7 (PREMIER Biosoft International, CA, USA). The mRNA levels were measured by qPCR using the SYBR Premix ExTaq™ (TaKaRa, Dalian, Liaoning, China). This was followed by the measurement of fluorescence during a 55 to 95°C melting curve in order to detect a single gene-specific peak and to check the absence of primer dimmer peaks, and a single and discrete peak was detected for all primers tested. Negative controls were non-template reactions (replacing cDNA with H₂O).

Expression levels of 3 genes were calculated relative to the reference gene *Csupβ-action* and *CsupGAPDH* using the Q-Gen method in Microsoft Excel-based software of Visual Basic^{57,58}. For each sample, three biological replications were performed with each biological replication measured in three technique replications.

Phylogenetic Analyses. The phylogenetic trees were reconstructed for phylogenetic analyses of *CsupOBPs*, *CsupCSPs*, *CsupCXEs*, *CsupDESSs* and *CsupFARs* were based on the amino sequences (the signal peptides of sequences had been removed) of the putative genes and the sequences of other Lepidoptera insects. 8 OBP sequences from *C. suppressalis* (amino acids > 138 aa), 10 CSP sequences from *C. suppressalis* (amino acids > 65 aa), 8 CXE sequences from *C. suppressalis* (amino acids > 79 aa), 6 Des sequences from *C. suppressalis* (amino acids > 90 aa) and 10 FAR sequences from *C. suppressalis* (amino acids > 79 aa) were included in the phylogenetic tree data. The protein name and accession number of the genes used for phylogenetic tree building are listed in Tab. S2. Amino acid sequences were aligned with ClustalX 2.0⁵⁹ and unrooted trees were constructed by MEGA5.0⁶⁰ using the Neighbor-joining method, with Poisson correction of distances. Node support was assessed using a bootstrap procedure base on 1000 replicates.

- Vogt, R. G. [Biochemical diversity of odor detection:OBPs, ODEs and SNMPs.] *Insect pheromone biochemistry and molecular biology* [Blomquist G. J. (ed.)] [397–451] (Elsevier, London, 2003).
- Xu, Y. L. et al. *Large-scale identification of odorant-binding proteins and chemosensory proteins from expressed sequence tags in insects MC Genomics*. **10**, 632; DOI: 10.1186/1471-2164-10-632 (2009).
- Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cell Mol Life S-CMLS* **63**, 1658–1676 (2006).
- Pelosi, P., Calvello, M. & Ban, L. Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chem Senses* **30**, 291–292 (2005).
- Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* **58**, 373–391 (2013).



6. Nakagawa, T., Sakurai, T., Nishioka, T. & Touhara, K. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science* **307**, 1638–1642 (2005).
7. Vogt, R. G. *et al.* The insect SNMP gene family. *Insect Biochem Mol Biol* **39**, 448–456 (2009).
8. Merlin, C. *et al.* Antennal esterase cDNAs from two pest moths, *Spodoptera littoralis* and *Sesamia nonagrioides*, potentially involved in odourant degradation. *Insect Mol Biol* **16**, 73–81 (2007).
9. Prestwich, G. D., Vogt, R. G. & Riddiford, L. M. Binding and hydrolysis of radiolabeled pheromone and several analogs by male-specific antennal proteins of the moth *Antheraea polyphemus*. *J Biol Chem* **12**, 323–333 (1986).
10. Anderson, P. & Hallberg, E. Structure and Distribution of Tactile and Bimodal Taste Tactile Sensilla on the Ovipositor, Tarsi and Antennae of the Flour Moth, *Ephestia-Kuehniella* (Zeller) (Lepidoptera, Pyralidae). *Int J Insect Morphol* **19**, 13–23 (1990).
11. Faucheux, M. J. Morphology and Distribution of Sensilla on the *Cephalic Appendages*, Tarsi and Ovipositor of the European Sunflower Moth, *Homoeosoma-Nebulella* Den and Schiff (Lepidoptera, Pyralidae). *Int J Insect Morphol* **20**, 291–307 (1991).
12. Widmayer, P., Heifetz, Y. & Breer, H. Expression of a pheromone receptor in ovipositor sensilla of the female moth (*Heliothis virescens*). *Insect Mol Biol* **18**, 541–547 (2009).
13. Ando, T., Inomata, S. & Yamamoto, M. Lepidopteran sex pheromones. *Top Curr Chem* **239**, 51–96 (2004).
14. Jurenka, R. Insect pheromone biosynthesis. *Top Curr Chem* **239**, 97–132 (2004).
15. Volpe, J. J. & Vagelos, P. R. Saturated fatty acid biosynthesis and its regulation. *Annu Rev Biochem* **42**, 21–60 (1973).
16. Pape, M. E., Lopez-Casillas, F. & Kim, K. H. Physiological regulation of acetyl-CoA carboxylase gene expression: effects of diet, diabetes, and lactation on acetyl-CoA carboxylase mRNA. *Arch Biochem Biophys* **267**, 104–109 (1988).
17. Strandh, M., Johansson, T., Ahren, D. & Lofstedt, C. Transcriptional analysis of the pheromone gland of the turnip moth, *Agrotis segetum* (Noctuidae), reveals candidate genes involved in pheromone production. *Insect Mol Biol* **17**, 73–85 (2008).
18. Knipple, D. C., Rosenfield, C. L., Nielsen, R., You, K. M. & Jeong, S. E. Evolution of the integral membrane desaturase gene family in moths and flies. *Genetics* **162**, 1737–1752 (2002).
19. Matousova, P., Pichova, I. & Svatos, A. Functional characterization of a desaturase from the tobacco hornworm moth (*Manduca sexta*) with bifunctional Z11- and 10,12-desaturase activity. *Insect Biochem Mol Biol* **37**, 601–610 (2007).
20. Rosenfield, C. L., You, K. M., Marsella-Herrick, P., Roelofs, W. L. & Knipple, D. C. Structural and functional conservation and divergence among acyl-CoA desaturases of two noctuid species, the corn earworm, *Helicoverpa zea*, and the cabbage looper, *Trichoplusia ni*. *Insect Biochem Molecular Biol* **31**, 949–964 (2001).
21. Lassance, J. M. *et al.* Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* **466**, 486–489 (2010).
22. Moto, K. *et al.* Pheromone gland-specific fatty-acyl reductase of the silkworm, *Bombyx mori*. *P Natl Acad Sci USA* **100**, 9156–9161 (2003).
23. Anderson, C. M. & Stahl, A. SLC27 fatty acid transport proteins. *Mol Aspects Med* **34**, 516–528 (2013).
24. Burton, M., Rose, T. M., Faergeman, N. J. & Knudsen, J. Evolution of the acyl-CoA binding protein (ACBP). *Biochem J* **392**, 299–307 (2005).
25. Zhu, Z. R. *et al.* [Integrated management of rice stem borers in the Yangtze Delta, China.] In *Area-wide Control of Insect Pests* [Vreysen, M.J.B., Robinson, A.S., Hendrichs, J. (eds.)] [373–382] (Springer, Dordrecht, The Netherlands, 2007).
26. Li, A. X. & Li, C. H. [Host plants of *Chilo suppressalis*] In *Chilo suppressalis* (Walker) and its control [Li, A. X., Li, C. H. (eds.)] (Agricultural Science and Technology Press, China, 1996).
27. Brenda, F. *et al.* Identification of the female sex pheromones of the moth, *Chilo suppressalis*. *J Insect Phys* **21**, 1883–1886 (1975).
28. Ohta, K., Tztsuki, S., Uchiumi, K., Kurihara, M. & Fukami, J. I. Structures of sex pheromones of rice stem borer. *Agric. Biol. Chem.* **40**, 1897–1899 (1976).
29. Tatsuki, S. *et al.* Sex pheromone of the rice stem borer, *chilo suppressalis* (Walker) (Lepidoptera: Pyralidae): the third component, Z-9-hexadecenal. *Appl. Entomol. Zool.* **18**, 443–446 (1983).
30. 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**, e69715; DOI: 10.1371/journal.pone.0069715 (2013).
31. Faucheux, M. J. Multiporous Sensilla on the Ovipositor of Monopis-Crocicapitella Clem (Lepidoptera, Tineidae). *Int J Insect Morphol* **17**, 473–475 (1988).
32. Zhao, C. H., Lu, F., Bengtsson, M. & Lofstedt, C. Substrate specificity of acetyltransferase and reductase enzyme systems used in pheromone biosynthesis by Asian corn borer, *Ostrinia furnacalis*. *J Chem Ecol* **21**, 1495–1510 (1995).
33. Antony, B. *et al.* Pheromone-gland-specific fatty-acyl reductase in the adzuki bean borer, *Ostrinia scapularis* (Lepidoptera: Crambidae). *Insect Biochem Molecular Biol* **39**, 90–95 (2009).
34. Hagstrom, A. K., Walther, A., Wendland, J. & Lofstedt, C. Subcellular localization of the fatty acyl reductase involved in pheromone biosynthesis in the tobacco budworm, *Heliothis virescens* (Noctuidae: Lepidoptera). *Insect Biochem Molecular Biol* **43**, 510–521 (2013).
35. Lassance, J. M. *et al.* Functional consequences of sequence variation in the pheromone biosynthetic gene pgFAR for *Ostrinia* moths. *P Natl Acad Sci USA* **110**, 3967–3972 (2013).
36. Sato, K. *et al.* Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002–1006 (2008).
37. Xiu, W. M., Zhou, Y. Z. & Dong, S. L. Molecular characterization and expression pattern of two pheromone-binding proteins from *Spodoptera litura* (Fabricius). *J Chem Ecol* **34**, 487–498 (2008).
38. Cao, D. *et al.* Identification of Candidate Olfactory Genes in *Chilo suppressalis* by Antennal Transcriptome Analysis. *Int J Biol Sci* **10**, 846–860 (2014).
39. Gong, Z. J. *et al.* Cloning, expression and functional analysis of a general odorant-binding protein 2 gene of the rice striped stem borer, *Chilo suppressalis* (Walker) (Lepidoptera: Pyralidae). *Insect Mol Biol* **18**, 405–417 (2009).
40. He, X. *et al.* Binding of the general odorant binding protein of *Bombyx mori* BmorGOBP2 to the moth sex pheromone components. *J Chem Ecol* **36**, 1293–1305 (2010).
41. He, P. *et al.* Functional characterization of an antennal esterase from the noctuid moth, *Spodoptera exigua*. *Arch Insect Biochem Phys* **86**, 85–99 (2014).
42. He, P. *et al.* An antennae-enriched carboxylesterase from *Spodoptera exigua* displays degradation activity in both plant volatiles and female sex pheromones. *Insect Mol Biol* **23**, 475–86 (2014b).
43. Liu, W., Jiao, H., O'Connor, M. & Roelofs, W. L. Moth desaturase characterized that produces both Z and E isomers of delta 11-tetradecenoic acids. *Insect Biochem Mol Biol* **32**, 1489–1495 (2002).
44. Wang, H. L., Zhao, C. H. & Wang, C. Z. Comparative study of sex pheromone composition and biosynthesis in *Helicoverpa armigera*, *H. assulta* and their hybrid. *Insect Biochem Molecular Biol* **35**, 575–583 (2005).
45. Hagstrom, A. K., Lienard, M. A., Groot, A. T., Hedenstrom, E. & Lofstedt, C. Semi-selective fatty acyl reductases from four heliothine moths influence the specific pheromone composition *PLoS One* **7**, e37230; DOI:10.1371/journal.pone.0037230 (2012).
46. Wang, H. L., Lienard, M. A., Zhao, C. H., Wang, C. Z. & Lofstedt, C. Neofunctionalization in an ancestral insect desaturase lineage led to rare Delta6 pheromone signals in the Chinese tussah silkworm. *Insect Biochem Mol Biol* **40**, 742–751 (2010).
47. Vogel, H., Heidele, A. J., Heckel, D. J. & Groot, A. Transcriptome analysis of the sex pheromone gland of the noctuid moth *Heliothis virescens*. *BMC Genomics* **11**, 1471–2164 (2010).
48. Gu, S. H. *et al.* Identification of genes expressed in the sex pheromone gland of the black cutworm *Agrotis ipsilon* with putative roles in sex pheromone biosynthesis and transport *BMC Genomics* **14**, 636; DOI:10.1186/1471-2164-14-636 (2013).
49. Zhang, Y. N., Xia, Y. H., Zhu, J. Y., Li, S. Y. & Dong, S. L. Putative Pathway of Sex Pheromone Biosynthesis and Degradation by Expression Patterns of Genes Identified from Female Pheromone Gland and Adult Antenna of *Sesamia inferens* (Walker). *J Chem Ecol* **40**, 439–451 (2014).
50. Shang, Z. Z., Wang, Y. S. & Zou, Y. H. Study on rearing method of rice stem borer *Chilo suppressalis* Walker. *Acta Entomol Sin* **22**, 164–167 (1979).
51. Jiao, X. G., Xuan, W. J. & Sheng, C. F. Mating behaviour of the rice stem borer, *Chilo suppressalis* (Walker). *Acta Ecol Sin* **26**, 1110–1115 (2006).
52. Zhang, G. *et al.* Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome Res* **5**, 646–654 (2010).
53. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **10**, 1038–1083 (2011).
54. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
55. Ye, J. *et al.* WEGO: a web tool for plotting GO annotations. *Nucleic acids res* **34**, 293–297 (2006).
56. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621–628 (2008).
57. Simon, P. Q-Gene: processing quantitative real-time RT-PCR data. *Bioinformatics* **19**, 1439–1440 (2003).
58. Muller, P. Y., Janovjak, H., Miserez, A. R. & Dobbie, Z. Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* **32**, 1372–1374, 1376, 1378–1379 (2002).
59. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
60. 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).

Acknowledgments

We thank Master students He-Tan Chang, Ke Yang, Zhan-Feng Ye, Qiu-Pu Yang (Nanjing Agricultural University, China) for helping in collecting and feeding the rice stem borer. This work was supported by a grant from the National Natural Science Foundation (31372264) of China and a Special Fund for Agro-scientific Research in the Public Interest (201303017).



Author contributions

Y.-H.X., Y.-N.Z., F.L. and S.-L.D. conceived and designed the study. Y.-H. X. and X.-Q. H. performed the study. Y.-H.X., Y.-N.Z., F.L. and S.-L.D. analyzed and wrote the manuscript. All authors reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xia, Y.-H., Zhang, Y.-N., Hou, X.-Q., Li, F. & Dong, S.-L. Large number of putative chemoreception and pheromone biosynthesis genes revealed by analyzing transcriptome from ovipositor-pheromone glands of *Chilo suppressalis*. *Sci. Rep.* 5, 7888; DOI:10.1038/srep07888 (2015).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>