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

Yi-Han Xia¹*, Ya-Nan Zhang¹,²*, Xiao-Qing Hou¹, Fei Li¹ & Shuang-Lin Dong¹

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)¹² or chemosensory proteins (CSPs)³⁵. 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)⁸⁹ and aldehyde oxidases (AOXs)³. 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 croccapitella, 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-deterring pheromone secreted by the larvae and other volatiles in Homoeosoma nebulella¹¹ and Ephestia kuehniella¹². 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...
**Results**

**Transcriptome sequencing and sequence assembly.** Sequencing of a cDNA library prepared from mRNA of the OV-PG of *C. suppressalis* (Walker) (Lepidoptera: bombycidae) was carried out using the Illumina HiSeq(TM) 2000 sequencing platform. Totally, 63 putative genes 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*.

**Homology analysis and Gene Ontology (GO) annotation.** Of the 37,619 unigenes, those ≥1000 bp accounted for 16.26% of the transcriptome assembly (Fig. 2).

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 assigned to 208 GO terms, and 1,258 of these terms were located in the biological process category. Among these GO terms, metabolic process (1,154) and cellular process (1,041) were the most abundant. The most represented GO terms were those related to biological regulation (1,045), cellular component organization or biogenesis (1,024), and biological regulation (1,014). The biological process category was split into five subcategories: molecular function, cellular component, organismal process, physiological process, and regulation of biological processes. The molecular function category was the most abundant, with 1,045 GO terms, followed by the cellular component category, with 1,024 GO terms. The organismal process category had 141 GO terms, the physiological process category had 137 GO terms, and the regulation of biological processes category had 136 GO terms.

**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.

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

*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 emission15. 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 ester15. 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 step15,16. 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 sex pheromone biosynthesis18. Based on the position that double bond is introduced, most intensively studied class of enzymes involved in sex pheromone biosynthesis step15,16. 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 sex pheromone biosynthesis18.
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\textsuperscript{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.
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 (GOBP4: GOBP2; OBPG9: 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 CsupOBP2 and CsupOBP7 were not expressed in olfactory tissues but in OV-PG 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).

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 acids13,22. 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 degradation4. 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) | Name                          | Acc. number        | Species                      | E value | Identity (%) |
|---------------------------|------------------|----------|-------------------------------|--------------------|-------------------------------|---------|--------------|
| **Fatty acid synthase (FAS)** |                  |          |                               |                    |                               |         |              |
| FAS1                      | CL973            | 660      | fatty acid synthase           | AGR49310.1         | [Agrotis ipsilon]             | 0.0     | 74           |
| **Desaturase (DES)**      |                  |          |                               |                    |                               |         |              |
| Des1                      | CL1833           | 1116     | acyl-CoA delta-9 desaturase    | CAJ27975.1         | [Manduca sexta]               | 0.0     | 91           |
| Des2                      | Unigene14403     | 774      | desaturase HassGATD           | EHIJ1380.1         | [Danaus plexippus]            | 3.00E-166 | 80         |
| Des3                      | Unigene14404     | 471      | desaturase                    | AAQ74260.1         | [Spodoptera litoralis]        | 3.00E-143 | 82         |
| Des4                      | Unigene14677     | 356      | delta11 desaturase            | AAFB1787.1         | [Helicoverpa zea]             | 1.00E-60 | 70           |
| Des5                      | Unigene16772     | 270      | acyl-CoA delta-11 desaturase  | AAK21863.1         | [Trichoplusia ni]             | 1.00E-10 | 53           |
| Des6                      | Unigene18619     | 885      | predicted delta11-like       | XP_004925663.1     | [Bombus mori]                | 6.00E-139 | 68           |
| **Fatty-acyl reductase (FAR)** |              |          |                               |                    |                               |         |              |
| FAR1                      | CL2119           | 1266     | fatty-acyl CoA reductase 5    | EHIJ2233.1         | [Danaus plexippus]            | 0.0     | 77           |
| FAR2                      | Unigene15762     | 1407     | fatty-acyl reductase          | AGG19592.1         | [Ostrinia latipennis]         | 1.00E-122 | 47         |
| FAR3                      | CL442            | 1374     | fatty-acyl CoA reductase 3    | ADI82776.1         | [Ostrinia nubilalis]          | 0.0     | 83           |
| FAR4                      | Unigene10614     | 1388     | fatty-acyl CoA reductase 2    | XP_004930522.1     | [Bombus mori]                | 0.0     | 82           |
| FAR5                      | Unigene13646     | 1297     | fatty-acyl CoA reductase 4    | ADI82777.1         | [Ostrinia nubilalis]          | 0.0     | 73           |
| FAR6                      | Unigene14049     | 1875     | fatty-acyl reductase 2        | ADI82775.1         | [Ostrinia nubilalis]          | 0.0     | 82           |
| FAR7                      | Unigene14605     | 1614     | putative fatty acyl-CoA       | XP_004926017.1     | [Bombus mori]                | 0.0     | 72           |
| FAR8                      | Unigene16587     | 564      | putative fatty acyl-CoA       | AGR49318.1         | [Agrotis ipsilon]            | 8.00E-98 | 81           |
| FAR9                      | Unigene28511     | 330      | fatty-acyl CoA reductase 1    | ADI82774.1         | [Ostrinia nubilalis]          | 9.00E-59 | 78           |
| FAR10                     | Unigene7020      | 644      | fatty-acyl CoA reductase 6    | EHJ6493.1          | [Danaus plexippus]            | 6.00E-69 | 52           |
| **Acyl-CoA oxidase (ACO)** |                  |          |                               |                    |                               |         |              |
| ACO1                      | Unigene10405     | 414      | PREDICTED: probable peroxisomal acyl-CoA enzyme oxidase 1-like | XP_004932400.1 | [Bombus mori]                | 1.00E-79 | 77           |
| ACO2                      | Unigene14208     | 1059     | PREDICTED: probable peroxisomal acyl-CoA enzyme oxidase 1-like | XP_004932404.1 | [Bombus mori]                | 0.0     | 83           |
| **Acetyl-CoA Carboxylase (ACC)** |              |          |                               |                    |                               |         |              |
| ACC1                      | Unigene26870     | 363      | putative acetyl-CoA carboxylase | EHIJ2299.1 | [Danaus plexippus]            | 4.00E-59 | 77           |
| **Fatty acid transport protein (FATP)** |              |          |                               |                    |                               |         |              |
| FATP1                     | Unigene14963     | 2061     | Fatty acid transport protein  | BAJ33524.1         | [Ostrinia scapulalis]         | 0.0     | 86           |
| FATP2                     | Unigene15353     | 744      | Fatty acid transport protein  | ACT22576.1         | [Manduca sexta]              | 0.0     | 84           |
| FATP3                     | Unigene16281     | 1932     | long-chain fatty acid         | NP_004929241.1     | [Bombus mori]                | 0.0     | 81           |
| FATP4                     | Unigene22497     | 452      | long-chain fatty acid         | NP_004929240.1     | [Bombus mori]                | 7.00E-73 | 70           |
| **Acyl-CoA binding protein (ACBP)** |              |          |                               |                    |                               |         |              |
| ACBP1                     | CL2090           | 759      | acyl-CoA binding domain-containing protein 6-like | NP_001040308.1 | [Bombus mori]                | 1.00E-126 | 71         |
| ACBP2                     | CL331            | 798      | acyl-CoA binding domain-containing protein 5-like | NP_004933263.1     | [Bombus mori]                | 3.00E-124 | 77         |
| ACBP3                     | Unigene28350     | 182      | acyl-CoA binding protein-like  | EHIJ64012.1        | [Danaus plexippus]            | 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 | [Bombus mori]                | 3.00E-178 | 90         |
| ELO2                      | CL2453           | 450      | elongation of very long chain fatty acids protein AAEL008004-like | XP_004931947.1 | [Bombus mori]                | 4.00E-92  | 86         |
| ELO3                      | Unigene1119      | 513      | elongation of very long chain fatty acids protein AAEL008004-like | XP_004924776.1 | [Bombus mori]                | 3.00E-94  | 81         |
| ELO4                      | Unigene16941     | 939      | elongation of very long chain fatty acids protein AAEL008004-like | XP_004931951.1 | [Bombus mori]                | 0.0      | 87           |
| ELO5                      | Unigene7895      | 708      | elongation of very long chain fatty acids protein AAEL008004-like | XP_004924792.1 | [Bombus mori]                | 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 theOV-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: EHIJ2299.1). 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.
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, CsupDes 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 PBGOBP 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 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.
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 scapulalis, Spodoptera littoralis and some other lepidopteran species (Fig. 9) showed that CsupDes6 was clearly clustered into the clade of D14 desaturase, while CsupDes1, CsupDes2, and CsupDes4 were clustered in the clad of D9 (16C-18C), D9 (14C-26C) and D11 desaturase, respectively. Interestingly, the CsupDes3 and CsupDes5 were clustered in two different subclasses from other identified Dess (Fig. 9). In the FAR phylogenetic tree, CsupFAR2 was clustered with other lepidopteran pgFAR sequences21,22,32–35. 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 ovipositor10,31, and thus ovipositor may play olfactory roles in detection of plant odors, ovipositor-deterring 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-
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 \textit{C. suppressalis}.

OBPs and CSPs are thought to be responsible for the binding and transport of the hydrophobic molecules including pheromones and plant volatiles\textsuperscript{1,3,4}. Among the 9 \textit{CsupOBPs} identified in the present study, none was OV-PG specific or biased. In contrast, \textit{CsupOBP3}, 4 (GOBP1), 7 and 9 (GOBP2) were strongly antennae-biased, indicating more important olfactory roles they play in antennae. Similar as OBPs, most \textit{CsupCSPs} were expressed in multiple tissues, but \textit{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 \textit{CsupORs} were identified in the ovipositor. It is noted that the ORco, acting as a chaperone and dimerization partner for other ORs\textsuperscript{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 pheromones\textsuperscript{37}. Very interestingly, a study with \textit{H. virescens}\textsuperscript{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 \textit{CsupPBP} expression in OV-PG by transcriptomic analysis. Therefore, in \textit{C. suppressalis}, antennae and other tissues that express PBPs\textsuperscript{38} 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 \textit{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\textsuperscript{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 8** | Phylogenetic tree of putative OBPs from \textit{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 CXE 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, suggesting their importance in the degradation of ester odorants.

As for the 32 pheromone biosynthesis related genes expressed in the OV-PG, 2 CsupDes (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, CsupDes 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. 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 Z9-16: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 an 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.

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.
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 FAR1-3,11-14. 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)19. 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. virescens20, Agrotis segetum21 and A. ipsilon22. In Sesamia inferens, 1 AO was found with very low expression level in the sex pheromone gland transcriptome23. 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.24 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 activity25. 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 virgin moths; while male and female abdomens were collected from 2- and 3-day-old virgin moths, as more moths were needed for antennae collection. Two and three replicates 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, China26. 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 (adenine) (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 to quantitate, and quality control of the cDNA library. At last, the library could be sequenced using Illumina HisSeq™ 2000 platform.

De novo Assembly of Short Reads and Gene Annotation. Transcriptome de novo assembly was carried out with short reads assembling program Trinity27. 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 processed each graph independently to extract full-length splicing isoforms and to GO annotation of the unigenes, and got GO functional classification by using WEGO software28.

Expression Abundance Analysis of the Unigenes. The expression abundance of these unigenes were calculated by the RPKM (Reads Per Kilobase per Million mapped reads) method29, using the formula: RPKM (A) = (1,000,000 × C × X/1,000) / (N × 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 the 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 (NanoDropTM 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, 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 TA plasmid using the pEASY-T3 cloning vector system (TransGen, 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 of 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: 30 s 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 (Table S3) were designed using Beacon Designer 7.7 (PREMIER Biosoft International, CA, USA). The mRNA levels were measured by qPCR using the SYBR Premix Ex Taq™ (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. All gene expression was normalized to 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 CsupGAPDH and CsupGAPDF using the Q-Gene method in Microsoft Excel-based software of Visual Basic30,31. For each sample, three biological replications were performed with each replication measured in three technical replicates.

Phylogenetic Analyses. The phylogenetic trees were reconstructed for phylogenetic analyses of CsupOBPs, CsupCSPs, CsupCXSs, CsupDEs 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. 250 sequences were obtained from C. suppressalis (amino acids:138 aa), 10 CSP sequences from C. sappanalis (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. sappanalis (amino acids:~90 aa) and sequences included from phylogeny analysis32. The protein name and accession number of the genes used for phylogenetic tree building are listed in Table S2. Amino acid sequences were aligned with Clustal X 2.0 and neighbor-joining method with Poisson correction of distances. Node support was assessed using a bootstrap procedure based on 1000 replicates.

1. Vogt, R. G. [Biochemical diversity of odor detection:OBPs, ODEs and SNMPs], Insect pheromone biochemistry and molecular biology [Blomquist G. J. (ed.)] 1, 3-49 (Elsevier, 2003).

2. 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).

3. Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. Cell Mol Life S-CMLS 63, 1658–1664 (2005).

4. Pelosi, P., Calvello, M. & Ban, L. Diversity of odorant-binding proteins and chemosensory proteins in insects. Chem Senses 30, 291–292 (2005).

5. Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu Rev Entomol 58, 373–391 (2013).
18. Knipple, D. C., Rosenfield, C. L., Nielsen, R., You, K. M. & Jeong, S. E. Evolution of
17. Strandh, M., Johansson, T., Ahren, D. & Lofstedt, C. Transcriptional analysis of
15. Volpe, J. J. & Vagelos, P. R. Saturated fatty acid biosynthesis and its regulation.
13. Ando, T., Inomata, S. & Yamamoto, M. Lepidopteran sex pheromones.
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).
11. Faucheux, M. J. Multiporous Sensilla on the Ovipositor of Monopis-Crocicapitella Clem (Lepidoptera, Tineidae).
10. Anderson, P. & Hallberg, E. Structure and Distribution of Tactile and Bimodal Taste Tactile Sensilla on the Ovispositor, Tarsi and Antennae of the Flour Moth, Ephesia kuehnellia (Zeller) (Lepidoptera, Pyralidae). Int J Insect Morphol 19, 13–23 (1990).
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 Antherea polyphemus. J Biol Chem 12, 323–333 (1986).
8. Merlin, C. et al. Antennal esterase cDNAs from two pest moths, Sesamia inferens (Fabricius) and the S. ou (Fabricius). Insect Mol Biol 34, 487–498 (2018).
7. Vogt, R. G. Arch Biochem Biophys 267, 104–109 (2014b).
6. Pak, 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).
5. 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 38, 484–494 (2008).
4. Xiu, W. M., Zhou, Y. Z. & Dong, S. L. Molecular characterization and expression pattern of two pheromone-binding proteins from Spodoptera litura (Fabrictius). Arch Biochem Biophys 34, 302–310 (1997).
3. 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 Bio 35, 573–583 (2005).
2. Hagstrom, A. K., Lienard, M. A., Groot, A. T., Hedenstrom, E. & Lofstedt, C. Selective fatty acyl reductases from four heliothine moths influence the specific pheromone composition. PLoS One 7, e37230; DOI:10.1371/journal.pone.0037230 (2012).
1. Shang, W. H., 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 tasar silkworm. Insect Biochem Mol Biol 40, 422–751 (2010).

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/