Transcriptome comparison of the sex pheromone glands from two sibling *Helicoverpa* species with opposite sex pheromone components

Zhao-Qun Li1,2, Shuai Zhang1, Jun-Yu Luo1, Chun-Yi Wang1, Li-Min Lv1, Shuang-Lin Dong2 & Jin-Jie Cui1

1State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang 455000, China, 2College of Plant Protection, Nanjing Agricultural University/Key Laboratory of Integrated Management of Crop Diseases and Pests (Nanjing Agricultural University), Ministry of Education, Nanjing 210095, China.

Differences in sex pheromone component can lead to reproductive isolation. The sibling noctuid species, *Helicoverpa armigera* and *Helicoverpa assulta*, share the same two sex pheromone components, Z9-16:Ald and Z11-16:Ald, but in opposite ratios, providing an example of such reproductive isolation. To investigate how the ratios of the pheromone components are differently regulated in the two species, we sequenced cDNA libraries from the pheromone glands of *H. armigera* and *H. assulta*. After assembly and annotation, we identified 108 and 93 transcripts putatively involved in pheromone biosynthesis, transport, and degradation in *H. armigera* and *H. assulta*, respectively. Semi-quantitative RT-PCR, qRT-PCR, phylogenetic, and mRNA abundance analyses suggested that some of these transcripts involved in the sex pheromone biosynthesis pathways perform. Based on these results, we postulate that the regulation of desaturases, KPSE and LPAQ, might be a key factor regulating the opposite component ratios in the two sibling moths. In addition, our study has yielded large-scale sequence information for further studies and can be used to identify potential targets for the bio-control of these species by disrupting their sexual communication.

In insects, species-specific behaviors elicited by sex pheromones play a key role in reproduction and are associated with reproductive isolation1. The regulation of sex pheromone-related enzymes leads to speciation by changing mate recognition systems. In moths, most sex pheromone components are C10–C18 long-chain unsaturated alcohols, aldehydes or acetate esters that are produced de novo via a modified fatty-acid biosynthesis pathway in the sex pheromone glands (PGs) by acetylation, desaturation, chain shortening, reduction, and oxidation either separately or in combination2,3. Different combinations of these reactions produce unique species-specific pheromone blends in different species.

Sex pheromone biosynthesis in moths starts with the production of the saturated fatty-acid precursor, malonyl-CoA, from acetyl-CoA and is catalysed by acetyl-CoA carboxylase (ACC) and fatty acid synthase (FAS)4. Then, the fatty chain is modified to introduce a double bond by specific desaturases (DESs), and shortened by β-oxidation5. Thus far, six types of DES have been functionally characterized, including Δ5, Δ6, Δ9, Δ11, Δ10–12, and Δ14. After the production and release of the sex pheromone components by females, the pheromone molecules are captured by odorant binding proteins (OBPs)12–14 or chemosensory proteins (CSPs)15 and transported to membrane-bound olfactory receptors (ORs)16–18. After OR activation, the pheromone molecules are rapidly removed by odorant degrading enzymes (ODEs), such as carboxylesterase19 and aldehyde oxidases (AOXs)20 to restore the sensitivity of the sensory neuron. Analysing these genes involved in the production of specific pheromone components will provide insights into the regulation of the pheromone component and thereby the evolution of moth sexual communication.

The lepidopterans, *Helicoverpa armigera* and *Helicoverpa assulta* are two sympatric sibling species that are morphologically indistinguishable in the egg, larval, and pupal stages21. Furthermore, these two species share the common sex pheromone components, Z9-16:Ald and Z11-16:Ald, but the ratios between the two components is completely reversed22,23, 100:7 in *H. armigera* and 7:100 in *H. assulta*. It is plausible that this difference likely contributes to the reproductive isolation of the two species. Some studies have been carried out to explore the
regulatory mechanisms that determine these species-specific ratios\textsuperscript{22}, but the mechanisms remain not well known especially from the molecular perspective. Therefore, we constructed and sequenced cDNA libraries from the PGs isolated from \textit{H. armigera} and \textit{H. assulta} to investigate the genetic factors associated with sex pheromone biosynthesis in these two species.

After analysis, we identified 108 and 93 putative pheromone biosynthesis, transport, and degradation transcripts in the PGs of \textit{H. armigera} and \textit{H. assulta}, respectively. Our results together with previous studies\textsuperscript{22,24} support the conjecture that the regulation of DESs is likely to play an important role in determining the opposite sex pheromone components ratios in the two species. In addition, our results also provide large-scale sequence information for further studies and identification of potential targets to disrupt sexual communication in \textit{H. armigera} and \textit{H. assulta} for the control of these lepidopterans.

**Results**

**Overview of the PG transcriptomes.** PGs from \textit{H. armigera} and \textit{H. assulta} were collected as previously described for the \textit{Heliothis virescens} PG transcriptome\textsuperscript{25} (Fig. 1) followed by construction of the corresponding cDNA libraries. Large-scale transcripts were assembled and annotated in the PG transcriptomes from \textit{H. armigera} and \textit{H. assulta} (Supplementary Table S1 online).

GO annotation was used to classify the PG transcripts into functional categories. GO terms were represented in all three major GO categories: biological process, cellular component, and molecular function. The most represented sub-category in the biological process category was cellular process, in the cellular component category it was cell and cell part, and in the molecular functions category, binding and catalytic activity were the most represented (Fig. 2).

**Identification of putative genes involved in pheromone biosynthesis, transport, and degradation in the two \textit{Helicoverpa} species.** After removal of repetitive sequences following blastX against the NCBI Nr database and alignment with ClustalX 2.0, we identified a total of 108 and 93 putative transcripts involved in the pheromone biosynthesis, transport, and degradation in \textit{H. armigera} and \textit{H. assulta} PGs, respectively (Tables 1 and 2). These transcripts belonged to gene families represented by multiple transcripts in these two moth species. For example, \textit{ACC} had 2 members in the 2 species each, \textit{alcohol dehydrogenase} (\textit{ALR}) was represented by 17 and 18 sequences in \textit{H. armigera} and \textit{H. assulta}, \textit{DES} with 7 and 8, \textit{FAS} with 3 and 3, \textit{FAR} with 18 and 13, \textit{CSP} with 19 and 16, \textit{OBP} with 26 and 23, \textit{aldehyde dehydrogenase} (\textit{AD}) with 9 and 6, and \textit{AOX} with 7 and 4 members respectively, in \textit{H. armigera} and \textit{H. assulta} (Tables 1 and 2, Supplementary Tables S2–S5 online).

**Tissue expression profile and mRNA abundance of the sex pheromone biosynthesis putative genes.** We further characterized the expression levels and tissue expression pattern of the transcripts putatively involved in pheromone biosynthesis by semi-quantitative RT-PCR and qRT-PCR. Transcript abundance in the PG was also calculated as RPKM (reads per kilobase per million mapped reads).

We identified two \textit{ACCs} from the PGs of both \textit{H. armigera} and \textit{H. assulta} (Tables 1 and 2). Semi-quantitative RT-PCR and qRT-PCR results revealed that \textit{HarmACC2} and \textit{HassACC2} were highly expressed in PGs compared to the female body without the PGs.

![Figure 1](https://example.com/image1.png) **Figure 1 | Dissection of \textit{Helicoverpa armigera} and \textit{Helicoverpa assulta} sex pheromone glands.** The pheromone glands in \textit{H. armigera} (a) and \textit{H. assulta} (b) were squeezed out from the abdomen using forceps (the gland is similarly inflated when the female calls). The abdomen of \textit{H. armigera} (c) and \textit{H. assulta} (e) were cut at the sclerotized cuticle from the 8th abdominal segment, and the sclerotized cuticle was removed (\textit{H. armigera} (d) and \textit{H. assulta} (f)) before immersing the glands in liquid nitrogen. 1: Sclerotized ovipositor valves; 2: Pheromone gland; 3: Sclerotized cuticle that was removed.
(Figs. 3 and 4). Their transcript abundance was also markedly higher (40.9 and 41.5 RPKM) than HarmACC1 and HassACC1 (0.4 and 7.7 RPKM) in the transcriptomes.

Three FASs were identified in the PGs from H. armigera and H. assulta (Tables 1 and 2). Semi-quantitative RT-PCR revealed that all three transcripts were expressed at higher levels in the female body when compared to the PGs (Figs. 3 and 4). However, the RPKM values indicated that both HarmFAS2 (237.5) and HassFAS2 (345.7) were abundant in the PG transcriptomes. The RPKM values of HarmFAS2 and HassFAS2 were 3- and 82-fold higher than the other transcripts in the PG transcriptomes.

Seven and eight DESs were identified in the PGs of H. armigera and H. assulta, respectively (Tables 1 and 2). Semi-quantitative RT-PCR and qRT-PCR results showed that HarmLPAQ, HarmGATD, HassLPAQ, HassGATD, and HassKPSE had robust expression in the PGs when compared to the female body (Figs. 3 and 4).

To evaluate transcript expression abundances, the RPKM values of DESs, HarmKPSE in the H. assulta was 39-fold higher than HarmKPSE in H. armigera, while HassLPAQ was 30-fold lower than HarmLPAQ. HarmGATD and HassGATD had lower abundances in PG transcriptomes compared to HarmLPAQ, HassKPSE and HassLPAQ.

There were 18 and 13 fatty acyl-CoA reductases (FAR) in the H. armigera and H. assulta PG transcriptomes, respectively (Tables 1 and 2). Among the 18 FARs in H. armigera, HarmFAR12 (FPKM = 414.1) was more abundant in the PG transcriptome than the other 11 PG-biased FARs (RPKM < 70) (Table 1, and Figs. 3 and 4). In H. assulta, HassFAR6 (RPKM 960.1) was more abundant in the PG transcriptome than the other PG-biased FARs (RPKM < 102) (Table 2, and Figs. 3 and 4).

ALR is involved in converting an alcohol to an aldehyde. We identified 17 and 18 ALRs in the H. armigera and H. assulta PG transcriptomes, respectively (Tables 1 and 2). Semi-quantitative RT-PCR and qRT-PCR results indicated that HarmALR15, HarmALR11, and HarmALR2 in H. armigera, and HassALR5 and HassALR15 in H. assulta had PG-biased expression (Figs. 3 and 4). In H. armigera, HarmALR2 was highly abundant (RPKM = 323.6) in the PG transcriptome than the other PG-biased ALRs (RPKM < 35) (Tables 1 and 2, and Fig. 3). In H. assulta, HassALR15 had a higher RPKM (66.1) value than HassALR5 (10.7).

Phylogenetic analyses of the DESs. To further investigate the function of the DESs from H. armigera and H. assulta, 15 candidate DESSs from these two species were phylogenetically analysed with other lepidopteran DESSs (Fig. 5). In the resulting
phylogenetic tree, we observed three well-supported clades including Δ9-desaturases (16C > 18C), Δ9-desaturases (16C < 18C), and Δ11-desaturases. The five PG-biased transcripts from *H. armigera* and *H. assulta* were well separated from each other, with many other lepidopteran DEs interspersed among them. HarmLPAQ was very close to HassLPAQ in the Δ11-desaturases clade, and HassKPE was a member of the Δ9-desaturases (16C > 18C) group. Interestingly, HarmKPE, did not show PG-biased expression (Figs. 3 and 4) although it was present in the same clade as HassKPE and the two proteins shared high amino acid identity (99.72%). Similarly, HarmLPAQ and HassLPAQ also shared high amino acid identity (99.70%). It is notable that two transcripts with PG-biased expression, HarmGATD and HassGATD, did not belong to any of the three main clades.

**Tissue expression profiles of the sex pheromone transport putative genes.** We identified 19 and 16 CSPs, and 26 and 23 OBPs in *H. armigera* and *H. assulta*, respectively (Supplementary Tables S2 and S3 online). Semi-quantitative RT-PCR results indicated that the orthologous transcripts had similar expression profiles (Fig. 6). Most of the OBPs were highly expressed in antennae and/or PGs, indicating their function in the detection and protection of plant volatiles, oviposition-deterring pheromones, and sex pheromones. Most CSPs were expressed in a range of tissues,
HarmAOX7 online). Semi-quantitative RT-PCR revealed that HarmAD4 in Helicoverpa assulta and H. armigera were highly expressed in antennae and/or PGs (Fig. 6).

Table 2 | BLASTX results for putative sex pheromone biosynthesis transcripts in Helicoverpa assulta pheromone glands

| Transcript | Name | ID | ORF | RPKM |
|------------|------|----|-----|------|
| Acetyl-CoA carboxylase (ACC) | ACC1 | U22914 | 510 | 7.7 |
| | ACC2 | CL1044-1 | 4983 | 41.5 |
| Aldo-Ketone Reductase (ALR) | ALR1 | U4829 | 320 | 4.0 |
| | ALR2 | CL3549-1 | 483 | 8.8 |
| | ALR3 | CL3700-1 | 813 | 7.2 |
| | ALR4 | U1366 | 198 | 5.7 |
| | ALR5 | CL2452-2 | 1002 | 10.7 |
| | ALR6 | U18627 | 753 | 521.2 |
| | ALR7 | U6810 | 975 | 14.9 |
| | ALR8 | U4937 | 342 | 5.5 |
| | ALR9 | U12805 | 1017 | 9.6 |
| | ALR10 | U7365 | 305 | 2.9 |
| | ALR11 | U8744 | 813 | 3.3 |
| | ALR12 | U23789 | 483 | 3.8 |
| | ALR13 | CL115-1 | 759 | 0.0 |
| | ALR14 | U9712 | 1071 | 9.0 |
| | ALR15 | U19322 | 975 | 66.1 |
| | ALR16 | U4138 | 694 | 4.2 |
| | ALR17 | U1545 | 1131 | 106.3 |
| | ALR18 | U24329 | 366 | 3.2 |
| Desaturase (DES) | KPSE | U18841 | 1062 | 659.9 |
| | NPVE | U21938 | 1062 | 949.2 |
| | MPVE | U1020 | 1104 | 16.8 |
| | GATD | U18038 | 1119 | 40.0 |
| | LPAQ | U21077 | 1017 | 132.9 |
| | KSVE | U21918 | 1119 | 26.9 |
| | KSPP | U12152 | 892 | 6.5 |
| | TYSY | CL2025-1 | 966 | 20.3 |
| Fatty acid synthase (FAS) | FAS1 | U13060 | 910 | 4.2 |
| | FAS2 | U22164 | 7170 | 345.7 |
| | FAS3 | U2985 | 271 | 3.7 |
| Fatty acyl-CoA reductase (FAR) | FAR1 | CL3772-1 | 1614 | 7.8 |
| | FAR2 | U795 | 1497 | 30.4 |
| | FAR3 | U1030 | 1488 | 101.1 |
| | FAR4 | U1584 | 1575 | 55.6 |
| | FAR5 | U18296 | 1557 | 26.0 |
| | FAR6 | U20971 | 1371 | 960.1 |
| | FAR7 | U22269 | 1569 | 94.4 |
| | FAR8 | CL283-1 | 1533 | 14.7 |
| | FAR9 | U25153 | 244 | 2.7 |
| | FAR10 | U25265 | 305 | 2.0 |
| | FAR11 | CL598-1 | 1572 | 0.3 |
| | FAR12 | CL1250-1 | 1617 | 0.1 |
| | FAR13 | CL1309-1 | 488 | 16.7 |

Best Blastp Match

| Species | Identity | Acc. number |
|---------|----------|-------------|
| Culex quinquefasciatus | 2E-17 | XP_001848848 |
| Agrotis ipsilon | 3E-64 | AGR45608.1 |
| Danaus plexippus | 3E-62 | EHJ68420.1 |
| Agrotis ipsilon | 6E-18 | AGR45607.1 |
| Agrotis ipsilon | 7E-14 | AGR45608.1 |
| Danaus plexippus | 5E-133 | EHJ3729.1 |
| Danaus plexippus | 2E-21 | EHJ68420.1 |
| Bombbyx mori | 4E-176 | XP_004921845 |
| Danaus plexippus | 2E-25 | EHJ71310.1 |
| Danaus plexippus | 1E-127 | EHJ70606.1 |
| Danaus plexippus | 2E-118 | AGR45606.1 |
| Danaus plexippus | 0E-00 | AGR45608.1 |
| Bombbyx mori | 0E-00 | AGR45608.1 |
| Bombbyx mori | 2E-21 | 59% EHJ68420.1 |
| Bombbyx mori | 0E-00 | 74% EHJ3729.1 |
| Bombbyx mori | 0E-00 | 77% XP_004921850 |
| Danaus plexippus | 1E-49 | 82% EHJ71310.1 |
| Bombbyx mori | 0E-00 | 95% NP_001040507 |
| Bombbyx mori | 7E-36 | 78% NP_001037610 |
| Helicoverpa zeae | 0E-00 | 100% AAF81788.1 |
| Helicoverpa assulta | 0E-00 | 99% AAM28484.2 |
| Bombbyx mori | 1E-173 | 65% XP_004925564 |
| Helicoverpa assulta | 0E-00 | 99% AAM28480.2 |
| Helicoverpa assulta | 0E-00 | 99% AAM28483.2 |
| Heliothis virescens | 0E-00 | 98% AGR045842.1 |
| Bombbyx mori | 1E-160 | 75% NP_001274329 |
| Agrotis segetum | 0E-00 | 94% AID66588.1 |
| Bombbyx mori | 4E-82 | 48% XP_004927661 |
| Agrotis ipsilon | 0E-00 | 92% AGR49310.1 |
| Bombbyx mori | 8E-24 | 73% XP_004922804 |
| Bombbyx mori | 0E-00 | 75% XP_004926017 |
| Ostrinia nubilalis | 0E-00 | 69% ADI82777.1 |
| Agrotis ipsilon | 0E-00 | 94% AGR49319.1 |
| Danaus plexippus | 0E-00 | 63% EHJ6493.1 |
| Bombbyx mori | 0E-00 | 80% XP_004929961 |
| Helicoverpa attulla | 6E-108 | 100% AFD04727.1 |
| Ostrinia nubilalis | 0E-00 | 80% ADI82776.1 |
| Agrotis ipsilon | 0E-00 | 87% AGR49318.1 |
| Bombbyx mori | 1E-25 | 64% XP_004929587 |
| Agrotis ipsilon | 0E-00 | 71% AGR49316.1 |
| Ostrinia nubilalis | 0E-00 | 81% ADI82775.1 |

Suggesting common functions. Similar to OBPs, several CSPs in the Helicoverpa species were highly expressed in antennae and/or PGs (Fig. 6).

Tissue expression profiles of sex pheromone degradation putative genes. We identified nine and six ADs, and seven and four AOXs in H. armigera and H. assulta, respectively (Supplementary Tables S4 and S5 online). Semi-quantitative RT-PCR revealed that HarmAD4, HarmAOX7, HarmAOX2, HarmAOX3, HarmAOX5, and HarmAOX6 in H. armigera and HassAD9, HassAOX3, and HassAOX5 in H. assulta were mainly expressed in antennae and PGs (Fig. 6).

Discussion

Speciation in insects is often associated with changes in mate recognition systems. Particularly, sex pheromone-induced behaviours play crucial roles in insect reproduction and contribute significantly to reproductive isolation. In moths, sex pheromones are synthesized in the PGs. Both H. armigera and H. assulta, which are sibling noctuid species, use the sex-pheromone components, Z9-16:Ald and Z11-16:Ald. However, the components are present in opposite ratios in the two species. Intrigued by this, we investigated differences in the transcripts related to sex pheromone biosynthesis, transport and degradation in the two sibling species by sequencing the transcripts from the PGs of the two Helicoverpa species. A total of 108 and 93 putative pheromone biosynthesis, transport, and degradation transcripts were respectively identified in H. armigera and H. assulta PGs. Further characterization of these transcripts by semi-quantitative RT-PCR, qRT-PCR, phylogenetic, and mRNA abundance analyses revealed that some of the transcripts had three characteristics: 1) transcripts that are distinctly or highly expressed in
PGs than female body (without PG), 2) transcripts that are more abundant than the other transcripts in the PGs, and 3) transcripts that are homologous to other insect genes with demonstrated function in sex pheromone biosynthesis.

Generally, the pheromone biosynthesis pathway in moths begins with the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA catalysed by ACC 27. Compared to other ACCs, HarmACC2 and HassACC2 were highly expressed in PGs and were highly abundant than the other ACCs. In pheromone biosynthesis, FAS has been shown to use malonyl-CoA and NADPH to produce fatty acids 4. In this study, none of the FASs displayed PG-biased expression, although HarmFAS2 and HassFAS2 were highly abundant in the PG transcripts with high RPKM values compared to other FASs. Future studies on the functional characterization of the Helicoverpa ACCs and FASs may reveal their specific roles in pheromone biosynthesis.

During sex pheromone biosynthesis, DESs introduces double bonds at specific positions in fatty acid chains 28,29. Previous studies with labelled fatty acids demonstrated that different pathways are used in the pheromone biosynthesis in H. armigera and H. assulta 22 to achieve the markedly different ratios in the sex pheromone components, Z9-16:Ald and Z11-16:Ald. Among the DESs identified in our study, HarmGATD, HarmLPAQ, HassKPSE, HassGATD, and HassLPAQ displayed PG-biased expression compared with the adult female body. However, phylogenetic analyses showed that HarmGATD and HassGATD were not clustered into groups that were previously demonstrated to function in sex pheromone biosynthesis. On the other hand, HarmLPAQ and HassLPAQ were the members of Δ11-desaturases group, and HarmKPSE was closely related to HassKPSE in the Δ9-desaturases (16C.18C) group. These two groups of desaturases share a conserved biological function in sex pheromone biosynthesis 30. Previous studies on desaturases from H. assulta 24,31, Helicoverpa zea 32 and Trichoplusia ni 9,33 showed that HassKPSE encodes a Δ9-desaturase. The action of this Δ9-desaturase results in the production of higher amounts of Z9-16:Acid than Z9-18:Acid 24. HassLPAQ shown to encode a Δ11-desaturase that specifically produced Z11-16:Acid. HzPGDs 32, TniKPSE 33 and HassKPSE 24 have high amino acid identity, sharing the similar function. In addition, the function of HassLPAQ 24, HzPGDs1 12 and TniLPAQ 9 were similar to each other. Considering the amino acid high identity (about 99.7% with only one amino acid difference) between HarmKPSE and HassKPSE, and HarmLPAQ and HassLPAQ, it is likely that HarmKPSE and HassKPSE encode Δ9-desaturases, HarmLPAQ and HassLPAQ encode Δ11-desaturases in H. armigera and H. assulta.
Interestingly, HarmKPSE did not show PG-biased expression, suggesting that this gene is not involved in the sex pheromone biosynthesis. This results is well consistent with the previous labelling study\(^2\) with D\(^3\)-16:Acid and D\(^3\)-18:Acid showed that Z\(_{11}\)-16:Ald is produced by D\(_{11}\) desaturation of 16:Acid in both \(H.\) armigera and \(H.\) assulta. However, Z\(_{9}\)-16:Ald is produced by D\(_{11}\) desaturation of 18:Acid and one cycle of two-carbon chain shortening in \(H.\) armigera, while Z\(_{9}\)-16:Ald is mainly produced by D\(_{9}\) desaturation of

![Diagram showing relative expression levels of Helicoverpa armigera and Helicoverpa assulta transcripts with PG-biased expression in different female tissues. Expression levels of acetyl-CoA carboxylase, aldo-Ketose Reductase, desaturase and fatty acyl-CoA related transcripts were determined in female bodies without pheromone glands (B1) and PGs by qRT-PCR. Transcripts from \(H.\) armigera are labelled in red and \(H.\) assulta in blue. An asterisk indicates a significant difference between the expression levels in female body and PGs (\(P < 0.05\), Student’s t-test). “NS” indicates no significant difference (\(P > 0.05\)).](image-url)
16:Acid and by Δ11 desaturation of 18:Acid and one cycle of two-carbon chain shortening in *H. assulta*. Therefore, unlike the *HassKPSE*, *HarmKPSE* that encodes a Δ9-desaturase is not likely to be involved in sex pheromone biosynthesis.

On the other hand, PG abundance is another characteristic feature of the genes involved in sex pheromone biosynthesis. The high abundance of *HarmLPAQ*, *HassKPSE* and *HassLPAQ* in the PG transcriptomes suggest that these genes and PG biased transcripts may have a role in sex pheromone biosynthesis in the two *Helicoverpa* species. Furthermore, the abundance of *HassKPSE* (Δ9) was 7-fold higher than *HassLPAQ* (Δ11) in the *H. assulta* PG transcriptome consistent with the major pheromone component being Z9-16:Ald in *H. assulta*. As compared with *HarmLPAQ*, the lower abundance of *HarmKPSE* (about 239-fold) is consistent with that *HarmKPSE* is not likely to be involved in sex pheromone biosynthesis. Together our data along with others reported previously suggest that among the DESs identified in our study, only *HarmLPAQ* (Δ11) is likely involved in sex pheromone biosynthesis in *H. armigera*, while both *HassLPAQ* (Δ11) and *HassKPSE* (Δ9) may be involved in this process in *H. assulta*.

Mutations that affect gene regulation could be more important in evolution than those changing the amino acid sequence of a protein. In our study, *HarmKPSE* and *HassKPSE* had high amino acid identity (99.72%) indicating similar function. But, their expression patterns were different, and the mRNA abundance of *HassKPSE* was 39-fold higher in the *H. assulta* PG than *HarmKPSE* in *H. armigera*.
PG, while HassLPAQ was 30-fold lower than HarmLPAQ. Therefore, we presume that the regulation of DESs in these two Helicoverpa species likely resulted in the evolution of different pathways in the sex pheromone biosynthesis resulting in the final opposite ratios between two sex pheromone components. Further studies on regulation of DESs and its function are needed to determine their specific roles in the biosynthesis pathways of these two Helicoverpa species.

After the introduction of a specific double bond in the sex pheromone biosynthesis pathway, the fatty acyl CoA pheromone precursors are further reduced to the corresponding alcohols by FAR and then catalysed by ALR. Among the FARs and ALRs identified in this study, HarmFAR12, HassFAR6, HarmALR2, and HassALR15 not only showed PG-biased expressions but also displayed a higher abundance than the others in the PGs suggesting their role in sex pheromone biosynthesis.

Some olfactory sensilla are distributed on the ovipositor, which may function in the olfactory detection of plant odours, ovipositor-deterrent pheromones, and sex pheromones. OBPs and CSPs are thought to be responsible for the binding and transport of hydrophobic molecules, including pheromones and plant volatiles. After sex pheromones have stimulated the olfactory receptor neurons, they must be rapidly removed by AD and/or AOX to restore the sensitivity of the sensory neuron. The OBPs and CSPs that are mainly expressed in antennae and PGs could play important roles in the binding and transport of plant volatiles, oviposition-deterrent pheromones, and sex pheromones. On the other hand, antennae and PGs highly expressed ADs and AOXs, which could be involved in degrading sex pheromone and aldehyde odorants.

In conclusion, we sequenced the PG transcriptomes in the two noctuid sibling species, H. armigera and H. assulta to identify the...
mechanisms regulating the opposite ratios of the sex pheromone components, Z9-16:Ald and Z11-16:Ald in the two species. Our analyses based on phylogeny, transcript expression, and transcript abundance indicates that a number of transcripts with PG-biased expression could be involved in the sex pheromone biosynthesis in the two species. Particularly, DESs seem to play a prominent role in the regulation of the component ratio in *H. armigera* and *H. assulta*. Additional functional analyses are needed to verify this conjecture in future.

### Methods

**Insect samples.** *Helicoverpa armigera* were collected from cotton fields in the Institute of Cotton Research at the Chinese Academy of Agricultural Sciences. *Helicoverpa assulta* were provided by the Henan University of Science and Technology in China. Both species were reared in the laboratory on an artificial diet at 25 ± 1 °C, 14:10 LD light cycle, and 65 ± 5% relative humidity. Pupae were sexed and kept separately in cages for eclosion. The pupae were checked daily for emergence and supplied with 10% honey solution as food for the emerging adults.

**Tissue collection.** To construct cDNA libraries, 15 Pgs from 3-day-old virgin females from each of the two species were collected at 5 h in scotophase (Fig. 1), immediately frozen in liquid nitrogen, and stored at −80 °C until further use. In addition, for semi-quantitative RT-PCR and qRT-PCR, female antennae (FA), male antennae (MA), pheromone glands (PGs), whole body without pheromone glands (B1), and whole body without pheromone glands and antennae (B2) were also collected from three-day-old virgin insects. These tissues were immediately frozen and stored at −80 °C until RNA isolation.

**cDNA library construction and Illumina sequencing.** Total RNA was extracted from the Pgs of *H. armigera* and *H. assulta* using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), and cDNA library construction and Illumina sequencing of the samples were performed at the Beijing Genomics Institute, Shenzhen, China. For each species, poly-adenylated RNA were isolated from 20 μg of pooled total RNA using eligo (dT) magnetic beads. Then, the mRNA from each species were fragmented into short pieces in the presence of divalent cations in fragmentation buffer at 94 °C for 5 min. Using the cleaved fragments as templates, random hexamer primers were used to synthesize first-strand cDNA using the. Second-strand cDNA was generated using the buffer, dNTPs, RnasaeH, and DNA polymerase 1. Following end repair and adaptor ligation, short sequences were amplified by PCR and purified with a QIAquick® PCR extraction kit (Qiagen, Venlo, The Netherlands), and sequenced on a HisSeq™ 2000 platform (Illumina, San Diego, CA, USA). The clean reads obtained in this study are available at the NCBI/SSA database under accession numbers SRX565435 and SRX1508989.

**Assembly and annotation.** The PG transcriptomes of *H. armigera* and *H. assulta* were assembled de novo using the short-read assembly program Trinity, which generated two classes of transcripts: clusters (prefix CL) and singletons (prefix U). Transcripts larger than 150 bp were compared to existing sequences in the protein databases, including NCBI NR, Swiss-Prot, KEGG®, and COG, using blastX. We then used the Blast2GO program for GO annotation of the transcripts and WEGO software to plot the GO annotation results.

**Analysis of transcript expression in the pheromone glands.** Transcript expression abundances were calculated by the RPKM (reads per kilobase per million mapped reads) method, which can eliminate the influence of different transcript lengths and sequencing discrepancies in calculating expression abundance. RPKM was calculated from the equation (1):

\[
\text{RPKM}(A) = \frac{C \times 10^6}{N \times L}
\]

where RPKM (A) is the expression of transcript A; C is the number of reads uniquely aligned to transcript A; N is the total number of fragments uniquely aligned to all transcripts; and L is the number of bases in transcript A.

**Phylogenetic analysis.** To investigate the phylogenetic relationships between the two *Helicoverpa* species, we compared all putative transcripts involved in the pheromone biosynthesis, reception, and degradation in each of the two species using ClustalX2.0 with default settings. Since desaturases are the most extensively studied class of enzymes involved in sex pheromone biosynthesis, we identified 67 lepidopteran desaturases sequences from NCBI NR and those from *H. armigera* and *H. assulta*. All phylogenetic trees were constructed using the neighbour-joining method implemented in MEGA6 with default settings and 1000 bootstrap replicates.

**Semi-quantitative RT-PCR analysis.** Total RNA was isolated using the SV Total Isolation System (Promega, Madison, WI, USA) according to the manufacturer’s instructions. Single-stranded cDNA templates were synthesized using 1 μg of total RNA from various samples (FA, MA, PGs, B1 and B2) using the Reverse Transcription System (Promega) following the instructions in the manual.

Specific primers for the transcripts putatively involved in pheromone biosynthesis, reception, and degradation were designed using Beacon Designer 7.7 (Premier Biosoft, Palo Alto, CA, USA) (Supplementary Table S6 online). Semi-quantitative PCR experiments with negative controls (no cDNA template) were performed as follows: 94 °C for 2 min; followed by 28 cycles at 94 °C for 30 sec, 60 °C for 30 sec, and 72 °C for 30 sec. The reactions were performed in 20 μL PCR reactions containing 2 μL cDNA (10 μM), 0.8 μL of forward primer (10 μM), 0.8 μL of reverse primer (10 μM), 15 ng of single-stranded cDNA, and 0.2 μL Ex-Taq (5 U/μL). PCR products were analysed by electrophoresis on 2.0% w/v agarose gel in TAE buffer and the resulting bands were visualized with Glutaled according to the manufacturer’s instructions. The GAPDH-binding protein (GenBank No. AY757405) from *H. armigera* was used as an endogenous control. Each reaction had three independent biological replicates.

**Statistical Analysis of data.** Data (mean ± SE) from various samples were subjected to one-way nested analysis of variance (ANOVA) followed by a least significant difference test (LSD) for mean comparison. Two-sample analysis was performed by Student’s t-test using SPSS Statistics 17.0 (IBM, Chicago, IL, USA).

1. Wyatt, T. D. Fifty years of pheromones. Nature 457, 262–263 (2009).
2. Foster, S. P. Lipid analysis of the sex pheromone gland of the moth *Heliotropha* virescens. Archives of insect biochemistry and physiology 59, 80–90 (2005).
3. Jurenka, R. Insect pheromone biosynthesis. Topics in current chemistry 239, 97–142 (2004).
4. Volpe, J. J. & Vegeios, P. R. Saturated fatty acid biosynthesis and its regulation. Annu. Rev. Biochem. 42, 21–60 (1973).
5. Hoskovec, M., Luxova, A., Svatova, A. & Boland, W. Biosynthesis of sex pheromones in moths: stereochemistry of fatty alcohol oxidation in *Manduca sexta*. Tetraedron 58, 919–921 (2002).
6. Hagstrom, A. K. et al. A novel fatty acid desaturase from the pheromone glands of the cabbage looper *Ctenopseustis obliquana* and *C. herana* with specific Z5-desaturase activity on myristic acid. J. Chem. Ecol. 40, 63–70 (2014).
7. 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 tasar silkworm. Insect Biochem. Mol. Biol. 40, 742–751 (2010).
8. Park, H. Y., Kim, M. S., Paek, A., Jeong, S. E. & Knipple, D. C. An abundant acyl-CoA (Delta9) desaturase transcript in pheromone glands of the cabbage moth, *Mamestra brassicae*, encodes a catalytically inactive protein. Insect Biochem. Mol. Biol. 38, 581–595 (2008).
9. Knipple, D. C. et al. Cloning and functional expression of a cDNA encoding a pheromone gland-specific acyl-CoA Delta11-desaturase of the cabbage looper moth, *Trichoplusia ni*. Proc. Natl. Acad. Sci. U. S. A. 95, 15287–15292 (1998).
10. Moto, K. et al. Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkworm, *Bombyx mori*, sex pheromone. Proc. Natl. Acad. Sci. U. S. A. 101, 8631–8636 (2004).
11. Roelofs, W. L. & Rooney, A. P. Molecular genetics and evolution of pheromone biosynthesis in *Lepidoptera*. Proc. Natl. Acad. Sci. U. S. A. 100, 14599 (2003).
12. Zhou, J. J. Odorant-binding proteins in insects. Vitamins and hormones 83, 241–272 (2010).
13. Li, Z. Q. et al. Two Minus-C odorant binding proteins from *Helicoverpa armigera* display higher ligand binding affinity at acidic pH than neutral pH. J. Insect Physiol. 59, 263–272 (2013).
14. Rund, S. S. et al. Daily rhythms in antennal protein and olfactory sensitivity in the malaria mosquito *Anopheles gambiae*. Sci Rep 3, 2493 (2014).
15. Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. Cellular and molecular life sciences: CMLS63, 1658–1667 (2006).
16. Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu. Rev. Entomol. 57, 343–463 (2012).
17. Farhan, A. et al. The CCHamide 1 receptor modulates sensory perception and olfactory behavior in starved *Drosophila*. Sci Rep. 3, 2765 (2013).
18. Steck, K. et al. A high-throughput behavioral paradigm for *Drosophila* olfaction - The Flywalk. Sci Rep 2, 561 (2012).
19. Maren, C. et al. Antennal expression of CDSs from two pest moths, *Spodoptera littoralis* and *Sesiama nonagrioides*, potentially involved in odourant degradation. Insect Mol. Biol. 16, 73–81 (2007).
20. Leal, W. S. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annual review of entomology 58, 373–391 (2013).

www.nature.com/scientificreports
21. Li, H., Zhang, H., Guan, R. & Miao, X. Identification of differential expression genes associated with host selection and adaptation between two sibling insect species by transcriptional profile analysis. BMC Genomics 14, 582 (2013).
22. 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. Mol. Biol. 35, 575–583 (2005).
23. Wu, H., Hou, C., Huang, L. Q., Yan, F. S. & Wang, C. Z. Peripheral coding of sex pheromone blends with reverse ratios in two helicoverpa species. PLoS One 8, e70078 (2013).
24. Jeong, S. E., Rosenfield, C. L., Marsella-Herrick, P., Man You, K. & Knipple, D. C. Multiple acyl-CoA desaturase-encoding transcripts in spermatheca glands of Helicoverpa assulta, the oriental tobacco budworm. Insect Biochem. Mol. Biol. 33, 609–622 (2003).
25. Vogel, H., Heidel, A. J., Heckel, D. G. & Groot, A. T. Transcriptome analysis of the sex pheromone gland of the noctuid moth Heliothis virescens. BMC Genomics 11, 29 (2010).
26. Smadja, C. & Butlin, R. K. On the scent of speciation: the chemosensory system and its role in premating isolation. Heredity (Edinb) 102, 77–97 (2009).
27. 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).
28. Albre, J. et al. Sex Pheromone Evolution is Associated with Differential Regulation of the Same Desaturase Gene in Two Genera of Leafroller Moths. PLoS Genet 8 (2012).
29. Fuji, T. et al. Sex pheromone desaturase functioning in a primitive Ostrinia moth is cryptically conserved in congener’s genomes. Proc. Natl. Acad. Sci. U. S. A. 108, 7102–7106 (2011).
30. Hao, G., Liu, W., O’Connor, M. & Roolofs, W. Acyl-CoA Z9- and Z10-desaturases genes from a New Zealand leafloefr moth species, Planotortrix octo. Insect Biochem. Mol. Biol. 32, 961–966 (2002).
31. 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).
32. Rosenfield, C. L., You, K. M., Marsella-Herrick, P., Roolofs, 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. Mol. Biol. 31, 949–964 (2001).
33. Liu, W. et al. Cloning and functional expression of a cDNA encoding a metabolic acyl-CoA delta 9-desaturase of the cabbage looper moth, Trichoplusia ni. Insect Biochem. Mol. Biol. 29, 435–443 (1999).
34. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. Science 188, 107–116 (1975).
35. 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 (2012).
36. Lienard, M. A., Hagstrom, A. K., Lassance, J. M. & Lofstedt, C. Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. Proc. Natl. Acad. Sci. U. S. A. 107, 10955–10960 (2010).
37. Lienard, M. A. & Lofstedt, C. Functional flexibility as a prelude to signal diversity? Role of a fatty acyl reductase in moth pheromone evolution. Current Insect Biol 3, 586–588 (2010).
38. Fauchex, M. J. Multiporous sensilla on the ovipositor of Monopis crocicapitella Clem. (Lepidoptera: Tineidae) Int. J. Insect Morphol. Embryol. 17, 473–475 (1988).
39. Anderson, P. & Hallberg, E. Structure and distribution of tactile and bimodal taste/tactile sensilla on the ovipositor, tarsi and antennae of the flour moth, Ephesia kuehniella (Zeller) (Lepidoptera: Pyralidae). Int. J. Insect Morphol. Embryol. 19, 13–23 (1990).
40. Choo, Y. M., Pelletier, J., Atunugula, E. & Leal, W. S. Identification and characterization of an antennae-specific aldehyde oxidase from the navel orangeworm. PLoS One 8, e67794 (2013).
41. Wang, C. Z. & Dong, J. F. Interspecific hybridization of Helicoverpa armigera and H. assulta (Lepidoptera: Noctuidae). Chin. Sci. Bull. 46, 489–491 (2001).
42. Zhang, G. J. et al. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Res. 20, 646–654 (2010).
43. Grabher, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–U130 (2011).
44. Kanehisa, M. et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36, D480–484 (2008).
45. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676 (2005).
46. Ye, J. et al. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 34, W293–297 (2006).
47. 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).
48. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882 (1997).
49. Paffi, M. W. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e15 (2001).