A novel domain duplication SlitFAR3 involved in sex pheromone biosynthesis in Spodoptera litura

Biyun Zhang
Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Institute of Biotechnology, Shanxi University

Fengqi Li
Institute of Plant Protection, Beijing Academy of Agriculture and Forestry Sciences

Cheng Qu
Institute of Plant Protection, Beijing Academy of Agriculture and Forestry Sciences

Yuejun Fu
Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Institute of Biotechnology, Shanxi University

Chen Luo (luochen1010@126.com)
Institute of Plant Protection, Beijing Academy of Agriculture and Forestry Sciences

Research Article

Keywords: Spodoptera litura, pheromone gland, fatty acyl reductase, yeast expression, domain duplication

Posted Date: May 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1625040/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Fatty acyl reductases (FARs) are key enzymes that participate in sex pheromone biosynthesis by reducing fatty acids to fatty alcohols. Lepidoptera typically harbor numerous FAR gene family members. While the FAR gene of moth is involved in the biosynthesis of sex pheromone, the key FAR gene of Spodoptera litura is still unclear. In this work, we predicted 30 FAR genes from S. litura genome. A special domain duplication was found with gene SlitFAR3, which exhibited high and preferential expression in the sexually mature female pheromone glands, and its expression pattern was rhythmic during the scotophase of sex pheromone production. Functional expression in yeast cells combined with comprehensive gas chromatography (GC) indicates that SlitFAR3 gene was able to catalyze four methyl ester precursors into corresponding fatty alcohol products. The domain duplication FAR genes were further found to exist in 11 species of the other 19 Lepidoptera insect. The phylogenetic tree showed SlitFAR3 was grouped with the other six FAR with domain duplication under the pgFAR subfamily clade, which is important candidate FAR genes participating in sex pheromone biosynthesis, and the other eight FAR domain duplication genes of 7 species were scattered in different clades. Domain duplications may facilitate the evolutionary diversification of protein sequences, which played diverse roles. This study is the first to focus on the special phenomenon of FAR domain duplication, which will advance understanding the biosynthesis-related genes from the perspective of evolutionary biology.

1. Introduction

Fatty acyl reductases (FAR) are key enzymes involved in fatty alcohol synthesis in plants, animals, and microorganisms (Hofvander et al. 2011; Doan et al. 2016). Fatty alcohols are precursor of lipid production, which has a wide range of biological functions. Fatty alcohols are precursors involved in the synthesis of waterproof in plants, insects and other animals (Vioque and Kolattukudy 1997; Hellenbrand et al. 2011; Jaspers et al. 2014; Wang et al. 2017; Li et al. 2020); the precursors of the synthesis of energy storage in certain bacteria (Fixter et al. 1986; Ishige et al. 2002; Kalscheuer and Steinbuchel 2003) and the components of ether lipids abundant in cell membranes (Nagan and Zoeller 2001). In insects, fatty alcohols are vital precursors of wax esters, cuticular hydrocarbons and sex pheromones. FAR gene function has been extensively studied in the PGs of moths. The substrate specificity of Bombyx mori pgFAR determines the exact proportion of components in the mixture of sex pheromones (Moto et al. 2003). A single FAR gene was found in Yponomeuta spp., which played a pivotal role in the production of noctuid multi-component pheromones (Lienard et al. 2010). The step of fatty alcohol production catalyzed by FAR is considered to be a potential reason for the differences of pheromones in Ostrinia spp. (Lassance et al. 2013). Therefore, FAR plays an essential role in the biosynthesis of moth sex pheromone.

FAR genes belong to a multigene family, which has experienced the evolutionary model of gene birth-and-death in the process of evolution. The random occurrence of gene replication and loss will lead to the change of gene copy numbers or homologous family members in the genome (Eirín-López et al. 2012). FAR gene loss was found in the lineage of nematodes, which only involves the loss of a single gene. In contrast, most species had a dynamic gain pattern with FAR genes. The FAR gene was very few in fungi,
vertebrates and non-insect invertebrates, while plants and insects have a large number of FAR gene family members in the process of divergence (Eirín-López et al. 2012). FAR genes in insects have been widely screened. For example, there are 17 FAR genes in *Nilaparvata lugens* (Hemiptera) (Li et al. 2020), 12 FAR genes in *Bombus lapidarius*, 35 FAR genes in *B. terrestris* and 16 FAR genes in *B. lucorum* (Hymenoptera), respectively (Tupec et al. 2019). A total of 206 FAR genes have been identified in 12 *Drosophila* species (Diptera), ranging from 14 to 21 (Finet et al. 2019). However, direct experimental characterization for the function of FAR gene family is still limited.

Over a long period of evolution, each insect has formed its own specific chemical communication system (Pasteels and Daloze 2003; Smadja and Butlin 2009). In Lepidoptera, female moths release sex pheromones to attract males of the conspecific to mate in the scotophase. *Spodoptera litura* (Lepidoptera, Noctuidae) is a notorious crop pest, which is widely distributed throughout the world, including the South Asia, East Asia, Oceania, and Pacific Islands (Fu et al. 2015; Efsa et al. 2019). The main pheromone components of *S. litura* have been identified as (Z,E)-9,11-tetradecadienyl acetate (Z9E11-14:OAc), (Z,E)-9,12-tetradecadienyl acetate (Z9E12-14:OAc), cis-9-tetradecenyl acetate (Z9-14:OAc) and trans-11-tetradecenyl acetate (E11-14:OAc) (Tamaki et al. 1973; Tamaki et al. 1976). Sex pheromones of most moths are synthesized de novo in pheromone glands (PGs). The biosynthesis of this type of sex pheromone begins with palmitic acid and stearic acid, followed produced double bonds by desaturase (Des) at specific positions in the carbon chain, fatty acid β-oxidation shortens or extends the carbon chain to form the quired carbon chain skeleton. Finally, the functional groups are modified at the end of the carbon chain by fatty acid reductase (FAR) and acyltransferase (ACT) (Tillman et al. 1999; Matsumoto 2010).

The different positions of unsaturated double bonds and terminal functional groups lead to the differences of sex pheromones among different species. The enzymes Des and FAR affecting sex pheromone biosynthesis have been widely studied (Moto et al. 2003; Lienard et al. 2008; Albre et al. 2012; Lassance et al. 2013; Bucek et al. 2015). The ACT of the last step in sex pheromone biosynthesis is rarely identified. In *S. litura*, the multifunctional *SlitDes5* produces unsaturated double bonds in the carbon chain (Xia et al. 2019), while the key FAR gene for the carbon chain terminal modification of sex pheromone precursor compound is unknown. Thus far, only 13 FAR genes have been identified from the pheromone gland transcriptome of *S. litura*, and these FAR genes have not been further studied and functionally characterized (Zhang et al. 2015). Following the genome and transcriptomes was sequenced from organisms spanning the entire tree of life in *S. litura* (Cheng et al. 2017; Liu et al. 2019), tracing the evolution of gene families and correlating them with the evolution of phenotypic traits has been facilitated.

In this study, we investigated the FAR genes family from the genome of *S. litura* with analysis of transcriptome from tissues of antennae, legs, proboscis and genitalia. A special domain duplication was found with gene *SlitFAR3*, and expression profiles of *SlitFAR3* was clarified by real-time quantitative PCR on pheromone glands (PGs) as well. A heterologous gene expression study in yeast combined with comprehensive gas chromatography (GC) analysis demonstrated that *SlitFAR3* gene can transform
precursors to produce fatty alcohols and plays an important role in the process of sex pheromone biosynthesis. We further analyzed domain duplication FAR genes in Lepidoptera by a schematic phylogenetic tree. Our findings will contribute to elucidate the sex pheromone biosynthesis mechanism and understand the evolutionary diversification of protein sequences with domain duplication.

2. Materials And Methods

2.1 Insect rearing

*S. litura* pupae of both sexes were obtained from Henan Jiyuan Baiyun Industrial Co., Ltd, China. The pupae were sexed in cages (30 cm × 30 cm × 30 cm) under controlled conditions at 25°C with 70% RH and a photoperiod of 16h: 8h (L:D). Adults were provided with 10% honey solution that was renewed daily. Adults on the first day of eclosion were designated as 0-day-old individuals. Pheromone glands (PGs) of the 2-day-old female were removed and flash-frozen in liquid nitrogen and stored at -80°C until used to isolate RNA. Additionally, the abdomen without PG and PGs of female virgin *S. litura* were used as reference material.

2.2 Chemicals

Methyl-(Z,E)-9,11-tetradecadienoate (Z9E11-14:Me), Methyl-(Z,E)-9,12-tetradecadienoate (Z9E12-14:Me), Methyl-(E)-11-tetradecenoate (E11-14:Me) and Methyl-(Z)-9-tetradecenoate (Z9-14:Me) were chemically synthesized by Pherobank (Duurstede, Netherlands). The synthesis of methyl ester compound was performed according to the method previously reported (Jurenka et al. 1994). The synthesized compounds were confirmed by using various spectral techniques like mass spectrometry. These methyl ester compounds as precursor for functional exploitation which were dissolved in ethanol to form a stock concentration of 500 mM. (Z, E)-9,11-tetradecadienol (Z9E11-14:OH), (Z, E)-9,12-tetradecadienol (Z9E12-14:OH), (E)-11-tetradecenol (E11-14:OH) and (Z)-9-tetradecenol (Z9-14:OH) were purchased from Nimrod Inc. (Changzhou, China). (Z9, E12)-tetradecadien-1-yl acetate (Z9E12-14:OAc) was purchased from BIOBERRY (Delaware, United States). The standard compounds were diluted in hexane (1:100,000).

2.3 Prediction and Expression analysis of FAR genes

15 FAR genes from 10 Lepidoptera species (Supplementary Table S1) from the National Center for Biotechnology Information (NCBI) were used as a ‘query’ to identify candidate genes that encoded putative FARs in *S. litura* genome respectively. All candidate FARs were manually checked by using the local BLASTP program with an E-value threshold of $10^{-5}$. Then, the potential FAR genes further analyzed by the NCBI conserved domain Batch CD Search Tool (https://www.ncbi.nlm.nih.gov/Structure/bwrcsb/bwrcsb.cgi) to predict conserved NADB_Rossmann domain and a putative substrate binding site FAR-C domain. ExPASy Proteomics Server (http://cn.expasy.org/tools/pi_tool.html) was used to calculate the molecular weights and isoelectric points of the FAR protein sequences.
Transcriptome data of antennae, legs, proboscis and genitalia of female and male *S. litura* were downloaded from NCBI (accession numbers were enumerated in Supplementary Table S2) (Cheng et al. 2017). HISAT2 was used to compare the transcriptome data to the genome of *S. litura*. The reads count of each gene was calculated by featurecount, and the transcripts per kilobase per million mapped reads (TPM) value of each gene was calculated based on the length of each gene by TBtools (Chen et al. 2020). The gene expression levels were estimated using TPM. Heatmap analysis is based on TPM values. The TPM values and relative expression values were represented as means ± standard error of the mean (SEM) based on three biological replicates.

### 2.4 RNA isolation and cDNA synthesis

The frozen tissues of *S. litura* were ground to a fine powder using a pestle and mortar chilled with liquid nitrogen before use. RNA was separately extracted from different tissues (8 PGs, 1 abdomen without PG). The total RNA was extracted using TRIzol Reagent (Invitrogen, California, United States) according to the manufacturer's protocol. The quality and concentration of RNA samples were checked with a NanoDrop spectrophotometer (ThermoFisher Scientific, Wilmington, DE, USA). All samples were tested in three biological replicates. The first-strand cDNA was synthesized using the PrimeScript™ RT reagent Kit (TAKARA, Japan) with gDNA Eraser (perfect Real Time) according to provided protocol. The synthesized cDNA was stored at -20°C until use.

### 2.5 Specific Expression of FAR Genes of *S. litura*

Relative expression levels of the FAR genes in female PGs of 30 FARs was checked by qRT-PCR. And specific expression pattern of *SlitFAR3* in 2-day-old PGs during different times (L, light; D, dark) and different day ages were assessed by qRT-PCR as well. qRT-PCR was performed on ABI PRISM 7500 (Applied Biosystems, United States). The reaction consisted of 10 µL of GoTaq® qPCR Master Mix (Promega), 0.8 µL of primer (10 mM), 2 µL of sample cDNA, and 7.2 µL nuclease-free water. The cycling conditions were 95 ºC for 30 s, followed by 40 cycles of 95 ºC for 5 s, and 60 ºC for 34 s. Then a melting curve was conducted starting at 95 ºC for 15 s, 60 ºC for 60 s, 95 ºC for 15 s to determine the specificity of PCR products. Primers were listed in Supplementary Table S3. The housekeeping genes, *actin* and *gapdh*, were used as reference genes for quantifying the transcription level of the FAR genes in different tissues of *S. litura*.

Each sample was subjected to three biological replicates and three technical replicates, and qPCR data were analyzed by the $2^-\Delta\Delta Ct$ method (Schmittgen and Livak 2008). The differences in the transcript levels of FARs were compared by One-way ANOVA (SPSS 19.0, Chicago, IL, United States), followed by Tukey’s test. The relative gene expression values were visualized using GraphPad Prism 8 (GraphPad Sofware Inc., San Diego, CA; https://www.graphpad.com)

### 2.6 Gene cloning

*SlitFAR3* gene specific primers were designed as forward primer (5’-ATGGTTGTGTTTACCTTCAAAG-3’) and reverse primer (5’-TTAATACCATTTTCTAGGGTTTCAAATATT-3) based on the *SlitFAR3* gene sequence
(No. LOC111348489). The PCR reaction with a total volume of 50 µL containing: 0.5 µL of TaKaRa LA Taq (China, Beijing), 5 µL of 10 × LA Taq buffer, 8 µL of dNTP mixture, 2 µL of cDNA, 4 µL of primer, 30.5 µL of nuclease-free water. PCR thermal cycling conditions consisted of 94°C for 5 min, 35 cycles at 94°C for 30 s, 48°C for 30 s, 72°C for 2 min 40 s, and 72°C for 10 min. PCR products were separated on a 1% agarose gel.

### 2.7 Plasmid construction and yeast transformation

The open reading frame of *SlitFAR3* was ligated into the GAL1 promoter of pESC-URA plasmid through *Bam* and *Xho*. The recombinant plasmid was named pESC-URA-*SlitFAR3*. The pESC-URA-*SlitFAR3* recombinant plasmid was transformed into *Saccharomyces cerevisiae* WAT11 strain with a yeast plasmid transformation kit (Yeastmaker™ Yeast Transformation System 2, Takara) and cultured in SD-URA solid medium at 30 °C for 4 days. The monoclonal colonies were inoculated in SD-URA liquid medium at 30 °C, 200 rpm overnight. The plasmid was extracted with a plasmid extraction kit (Tiangen, Beijing, China) and verified by PCR. PCR was performed by using Premix Taq™ enzyme (TaKaRa Taq™ Version 2.0 plus dye, China, Beijing) with a total volume of 50 µL containing: 25 µL of Premix Taq™ enzyme, 5 µL of template (monoclonal colonies were lysed in a metal bath at 95 °C for 10 min as a template), 4 µL of primer, 16 µL of nuclease-free water, the conditions of which included 94°C for 5 min, 34 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 2 min 40 s, and 72°C for 10 min. 1 mL of verified correct bacterial solution was inoculated in 10 mL SC-URA medium at 30 °C, 200 rpm for 24 h, then diluted to 100 mL SC-URA medium (2% raffinose) in the ratio of 1:20, and continue grew to OD600 = 1.0.

In order to study the function of *SlitFAR3 in vitro*, the 100 mL SC-URA medium of recombinant protein pESC-URA-*SlitFAR3* was centrifuged at 1100 g and resuspended with 100 mL fresh SC-URA medium (2% galactose) at 30 °C, 200 rpm for 24 h. The culture medium was diluted in a ratio of 1:5 to 5 ml of SC-URA medium containing 2% galactose, 1% tergitol (Nonidet P-40, sigma) and precursor compounds. Take 5 µL precursor compound into 5 ml (final concentration: 0.5 mM) culture medium, 24 h, 30 °C, 200 rpm. The cells were collected by centrifugation at 1100g, washed with sterile water for three times, and 200 µL hexane was extracted at 21 °C, 200 rpm for 1 h, and stored at -20 °C was subjected to GC-MS. Z9E12-14: OAc as internal standard.

### 2.8 Functional assay

Gas Chromatography/Mass Spectrometry (GC/MS) was used for determine the products. One microliter yeast extract was analyzed on an Agilent 7890A series gas chromatograph coupled to a mass-selective detector Agilent 5975C (Agilent Technologies, California, United States). The GC was equipped with either a HP-5MS capillary column (30 m×250 µm × 0.25 µm film thickness). Helium was the carrier gas (velocity 20 mL/min) and the injector was configured in splitless mode and maintained at 250°C injector temperature, 1 mL/min column flow. In order to obtain good separation of the different isomers, the oven temperature was held at 50°C for 2 min and rose at a rate of 5°C/min up to 150°C, 15°C/min up to 250°C with a final held for 3 min.
2.9 Phylogenetic analysis of FAR domain duplication genes in Lepidoptera

The genome and annotation file of 19 Lepidoptera species (Amyelois transitella, Bicyclus anynana, Bombyx mori, Calycopis cecrops, Danaus plexippus, Heliconius erato, Helicoverpa armigera, Heliothis virescens, Junonia coenia, Lerema accius, Manduca sexta, Papilio glaucus, P. machaon, P. polytes, P. xuthus, Phoebis sennae, Pieris napi, Plutella xylostella, S. frugiperda) were downloaded from Ensemble (http://ensembl.lepbase.org/index.html). Based on the annotation information and the genome to predict the FAR genes. Species evolution tree was generated using the TimeTree database (timetree.org). MEGA 7.0 (https://www.megasoftware.net/download_form) was used to align the protein sequences of FAR family identified in all species. IQ-TREE (http://iqtree.cibiv.univie.ac.at/) was used to estimate the best fit models (best-fit model: LG + I + G4 chosen according to BIC) and construct the gene family tree (Kalyaanamoorthy et al. 2017). The bootstrap parameter was set to 1000 replicates (Nguyen et al. 2015). FigTree (v1.4.4) was used to create the tree and jojoba FAR (AAD38039.1) was taken as the outgroup of rooted tree.

3. Results

3.1 The FAR gene family in S. litura

From a bioinformatic screening, 30 candidate genes were identified as putative FAR genes in S. litura, which is similar to that reported in the related insect species. The “30 new genes” were named after a four-letter code (first letter of the genus followed by the first three letters of the species name) + FAR + number according to their homology with the related species. 29 sequences were full length FAR genes that were characteristic of typical insect FARs because of their homologic sequences with other known FARs, while the N-terminal sequence of SlitFAR26 gene was incomplete. The sequence characteristics of all FAR genes indicated that the numbers of amino acid residues ranged from 428 to 868, with molecular weights ranging from 49.43 kDa to 99.13 kDa and isoelectric points of 6.21–9.39 (Supplementary Table S4). Protein family domain architecture (Pfam) revealed that FAR had a typical conserved domain NADB_Rossmann site at the N-terminal and a FAR-C domain at the C-terminal. Among these FAR genes, the SlitFAR3 gene had special domain duplication, containing two NADB_Rossmann domains and two FAR-C domains (Fig. 1).

3.2 Tissue expression specificity of FAR gene family in S. litura

To determine the tissue expression level of FAR genes, an intuitive heatmap was illustrated for the transcriptome data of antennae, legs, proboscis and genitalia of female and male S. litura. Subsequently, the FAR genes with similar expression patterns were clustered together. The heatmap results presented the SlitFAR3 was highly expressed in female genitalia, while SlitFAR21 was highly expressed in female and male antennae, with almost no expression in other tissues. The expression levels of the SlitFAR3,
SlitFAR6, SlitFAR11, SlitFAR16, SlitFAR23 and SlitFAR30 genes in female genitalia were higher than those in other tissues, and the other 10 FAR genes were hardly expressed in these tissues (Fig. 2).

By applying qRT-PCR on Pheromone glands (PGs) of the 2-day-old female, we further explored the vital FAR genes involved in the biosynthesis of sex pheromone in S. litura. SlitFAR1,3–8,10–13,16,25–26,30 were expressed in female PGs, and the expression level of SlitFAR3 gene was the highest, which was 50 times higher than SlitFAR1 gene. Both SlitFAR8 and SlitFAR12 genes were expressed at more than 5 times higher than SlitFAR1. The expression levels of the other 15 FAR genes were nearly zero (Fig. 3). These results suggested that SlitFAR3 was one of the important genes related to sex pheromone biosynthesis.

### 3.3 Analysis of Expression characteristic of SlitFAR3 gene

We performed a detailed test for expression levels of SlitFAR3 gene in 2-day-old PGs during different times (L, light; D, dark) and different day ages. The results showed SlitFAR3 gene was significantly expressed in PGs during scotophase. And the expression level of SlitFAR3 gene peaked at 3 h and 6 h (Fig. 4A), which displayed a similar rhythm with the release of sex pheromone (Sun et al. 2002). In addition, the expression of SlitFAR3 gene expression was higher in 2-day-old female PGs than that of 0-day-old female PGs (Fig. 4B).

### 3.4 Cloning and Enzyme Function of SlitFAR3

The full length of SlitFAR3 gene was cloned by PCR from PG cDNA, and the length of the product was 2604 bp (Supplementary Figure S1). To test whether the PG-biosynthetic SlitFAR3 gene of S. litura has the fatty acyl reductase activity, we cloned the coding region of SlitFAR3 gene into Saccharomyces cerevisiae expression plasmid, heterologously expressed this SlitFAR3 in yeast, and analyzed fatty alcohols by GC. In the yeast cells transformed with recombinant plasmids, the functional research was carried out by adding each precursor compound (0.5 mM). It was apparent that SlitFAR3 can transform the pheromone precursors Z9E11-14: Me (Retention Times RT: 26.118), Z9E12-14: Me (RT: 25.771), E11-14: Me (RT: 25.724) and Z9-14:Me (RT: 25.649) into corresponding alcohol products Z9E11-14:OH (RT: 25.656), Z9E12-14:OH (RT: 25.265), E11-14:OH (RT: 25.206) and Z9-14:OH (RT: 25.130) (Fig. 5A), which are main precursor compounds for the formation of sex pheromones in S. litura.

To ensure that the produced fatty alcohol was caused by the expression of SlitFAR3 protein, the transformed empty vector pESC-URA was established as negative control yeast strains. And the function of SlittoFAR3 gene of S. littoralis was investigated as positive control group, which is reported previously in same method as in this study (Supplementary Figure S3). No alcohol products were obtained in the negative control group, which yeast transformed by empty plasmid, and only substrate can be detected (Fig. 5B). The corresponding fatty alcohol products could be detected in the positive control group after adding the substrate (Supplementary Figure S3). In general, SlitFAR3 determined in yeast has substrate reducibility.

### 3.5 Domain duplication FAR gene family in Lepidoptera
SlitFAR3 gene had domain duplication, which has never been reported before. To gain insight into this gene works in lepidopteran insects, we further predicted and analyzed 361 FAR genes from the genomes of 19 Lepidoptera insect species. The number of FAR genes in different species ranges from 11 to 35, and 11 insect species had FAR genes with duplication domains. One FAR gene with duplication domains was found with *H. erato*, *B. anynana*, *P. napi*, *L. accius*, *H. virescens*, *S. frugiperda*, *M. sexta*, *P. xylostella*, and two FAR genes with duplication domains was found with *P. machaon*, *P. Xuthus*, *P. polytes* (Fig. 6). We reconstructed a schematic phylogenetic tree of FAR genes from 20 Lepidoptera insect species. The genetic relationship between species in the species tree was consistent with previous report (Simon et al. 2021). These results suggested that the domain duplication of FAR gene was widespread in Lepidoptera species.

We also reconstructed a FAR gene tree using 391 FAR protein sequences of 20 Lepidoptera species including *S. litura*. As illustrated in Supplementary Figure S4, SlitFAR3 from *S. litura* under the Lepidopteran pheromone gland specific FAR (pgFAR) subfamily clade and was grouped with the other six FAR domain duplication from *H. erato*, *H. virescens*, *S. frugiperda*, *M. sexta*, *P. machaon*, *P. xuthus*, which is important candidate FAR genes participating in sex pheromone biosynthesis. The other eight FAR domain duplication genes of 7 species were scattered in different clades of the phylogenetic tree, including one FAR in *P. machaon*, one FAR in *P. xuthus* and two FAR genes in *P. polytes*. (Supplementary Figure S4).

**Discussion**

Generally, there are a large number of FAR genes family members in insects, which have the function of reducing fatty acyl to produce fatty alcohol (Eirín-López et al. 2012; Tupec et al. 2019). As the main precursor of lipid, fatty alcohol is used in the biosynthesis of cuticular hydrocarbons, wax esters and sex pheromones in insects. FAR gene is involved in the biosynthesis of cuticular hydrocarbons in *N. lugens*, which is closely related to adult female fertility and cuticular development, and further affects survival (Li et al. 2020), The Chinese white wax scale insect, *Ericerus pela*, known for its wax production, relies on FAR genes biosynthesis wax (Yang et al. 2019); The first FAR gene involved in sex pheromone biosynthesis was identified in *B. mori* (Moto et al. 2003). Up to now, the FAR genes in Lepidoptera have been widely identified, such as *Yponomeuta* spp. (Lienard et al. 2010), *Ostrinia* spp. (Lassance et al. 2013), *S. exigua* and *S. littoralis* (Antony et al. 2016), *Helicoverpa zea* (Dou et al. 2020), which is mainly involved in the biosynthesis of sex pheromone. Previously, FAR genes of *S. litura* has been preliminarily screened from transcriptome, but further gene function has not been studied. Based on the genome data and annotation files of *S. litura*, we predicted 30 FAR genes comprehensively. Combining transcriptome and genomic data of *S. litura*, we obtained 17 new FAR genes. Normally, FAR gene is around 1500 bp, while the SlitFAR3 gene (NO. LOC111348489) has 2604 bp in this study, which is almost twice the 1365 bp of the previously reported SlitFAR3 gene (NO. KT261697) (Zhang et al. 2015). That means the result of transcriptome splicing could be limited without genome data and annotation files. We cloned its full-length sequence by PCR, and sequencing indicated it is a completely novel FAR gene with domain duplication (Supplementary Figure S1-S2).
To understand whether the FAR genes with special domain duplication also has the typical function of producing fatty alcohol, we analyzed the expression pattern of FAR genes in transcriptome of different tissues and the expression level in PGs. From the transcriptome of different tissues, SlitFAR3 was highly expressed in female genitalia. Furthermore, we revealed that 14 FAR genes were expressed in PGs by qPCR, and the expression level of SlitFAR3 gene was the highest. Therefore, we hypothesized that SlitFAR3 gene is related to the biosynthesis of sex pheromones. It has been measured that the sex pheromone content of female S. litura in the scotophase (Sun et al. 2002). The sex pheromone of S. litura was mainly released in the scotophase, and rested in the light phase (Li et al. 2012). There were two mating peaks, one at the beginning of the scotophase and one at 6–7 h after the scotophase (Wu et al. 2018). We discovered that the expression level of SlitFAR3 gene exhibited a fluctuated in the scotophase by qPCR, which was consistent with the release rhythm of sex pheromone. In Y. evonymellus, the pgFAR gene related to sex pheromone biosynthesis was highly expressed in sexually mature PGs and showed a 24 h cyclic fluctuation during the pheromone production period (Lienard et al. 2010). The key gene pgFAR of sex pheromone biosynthesis in B. mori is specifically expressed in PGs, and no detectable signal in other tissues (Moto et al. 2003). Hence, we supposed that SlitFAR3 was related to sex pheromone biosynthesis of S. litura and the gene function was performed through eukaryotic expression system in vitro.

Regarding expression and function of FAR genes involved in the biosynthetic pathway of sex pheromones has received a lot attention specially in Lepidoptera (Tillman et al. 1999; Matsumoto 2010; Antony et al. 2016). E10,Z12-16:acyl-CoA was reduced to E10,Z12-16:OH (bombykol) by FAR gene in B. mori (Moto et al. 2003). The orthologs pgFAR genes encoded by Y. evonymellus, Y. padellus and Y. rorellus had a wide range of substrates, which could effectively promote the conversion of C14- or C16-acyl precursors to corresponding fatty alcohols (Lienard et al. 2010). And the HzeaFAR1 could reduce precursors 16:Me, Z9-14:Me, Z9-16:Me, Z11-16:Me in H. zea (Dou et al. 2020). We speculated that the candidate SlitFAR3 gene may take the intermediate products of sex pheromones of S. litura Z9E11-14:Me, Z9E12-14:Me, E11-14:Me and Z9-14:Me as the precursor compounds of biosynthetic sex pheromones Z9E11-14:OH, Z9E12-14:OH, E11-14:OH and Z9-14:OH. The results displayed that SlitFAR3 has universal catalytic ability for these substrates. The enzymatic catalytic ability of SlitFAR3 in this work clearly illustrates that the sex pheromone component of S. litura is produced through the biosynthetic pathway involving SlitFAR3 (Fig. 7). Likewise, as a positive gene in this expression system, SlittoFAR gene also converts the substrate into the corresponding fatty alcohol product. Our work is the first research on FAR gene related to sex pheromone biosynthesis with domain duplication. The formation of FAR multiple family is driven by gene loss and gain model, whether the production of new types of domains duplication is the driving force to promote the generation of new genes remains to be explored.

Based on the large-scale study of genomic data, gene duplications and lineage-specific gene family expansion are considered to be an important mechanism for the evolution of new genes and new biochemical functions (Jordan et al. 2001; Lespinet et al. 2002). Protein domains represent a fundamental evolutionary unit forming proteins (Rossmann et al. 1974), and two proteins with common domains are likely to be evolutionarily related, originating from processes such as duplication and/or
shuffling of whole genes or exons encoding domains (Doolittle 1995). Domain duplication and/or shuffling are probably the most important forces driving protein evolution as well as proteome complexity. While duplication of entire genes as well as exons encoding domains increases the abundance of domains in the proteome, domain shuffling increases versatility, the number of different contexts in which domains can occur (Vogel et al. 2005). In the FAR multiple gene family of *S. litura*, different from the typical FAR gene domain, *SlitFAR3* has a special domain duplication, which has not been described before. Consequently, we predicted the FAR genes family of more species in Lepidoptera to determine whether this domain duplication phenomenon is accidental or common in evolution.

From the genomic data of 20 Lepidoptera species, 391 FAR genes were predicted, of which 12 species had 15 domain duplication FAR genes. All predicted FAR genes were reconstructed and the phylogenetic tree exhibited that 47% domain duplication FAR encoded protein belonged to an identified Lepidoptera-specific pgFAR gene subfamily, pgFAR is a unique clade related to sex pheromone biosynthesis in Lepidoptera (Lassance et al. 2013). We speculated that the other FAR genes with domain duplication grouped in pgFAR subfamily clade may also participate in the biosynthesis of sex pheromones. FAR genes scattered in other clades may have diverse functions other than sex pheromone biosynthesis.

In summary, we predicted 30 fatty acyl reductase genes from the genome level of *S. litura*, and a novel domain duplication was found with *SlitFAR3* gene. *SlitFAR3* gene exhibited high and preferential expression in mature female PGs, its expression pattern in scotophase was consistent with the release rhythm of sex pheromone. The function of *SlitFAR3* reductase to produce fatty alcohol was further confirmed by *in vitro* yeast express system combined with GC-MS. The results highlight the importance of *SlitFAR3* with domain duplication in biosynthesis of sex pheromone, and provide a solid foundation for understand the evolutionary diversification of protein sequences with domain duplication.

Declarations

Author Contributions

CL and FQL conceived and designed research. BYZ and YJF conducted experiments. CQ contributed new reagents and/or analytical tools. BYZ and FQL analyzed data. BYZ wrote the initial draft. CL and YJF revised this manuscript. All authors read and approved the manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

All authors are in accord with the submission.

Availability of data and materials
The datasets generated during and/or analyzed during the current study are available on reasonable request.

**Competing interests**

The authors declare that no competing interests exist.

**Acknowledgements**

This work was supported by the National Key R & D Program of China (2019YFD1002100) and the China Agriculture Research System of MOF and MARA (CARS-24-C-03).

**References**


**Figures**
Figure 1

Phylogenetic relationships and conserved domains of the *S. litura* FAR genes family. The phylogenetic tree was constructed based on the sequences of *S. litura* FAR proteins using TBtools. The conserved domains are indicated by colored rectangles. The FAR protein structure includes an N-terminal NADB_Rossmann domain (green), and a C-terminal FAR-C domain (yellow). The scale line at the bottom shows the protein length of each FAR.
Gene expression level of FAR genes in different tissues of *S. litura*. The expression values represented by normalized log2(TPM values) mapping to the FAR coding regions. Gene names are listed on the right. Scale bar at the top right indicates the degree of expression, and the color range is from blue (low expression) to red (high expression). Tissue names are given at the bottom: F_An, female antennae; F_Le, female legs; F_Pr, female proboscis; F_Ge, female genitalia; M_An, male antennae; M_Le, male legs; M_Pr, male proboscis; M_Ge, male genitalia.

**Figure 2**

<table>
<thead>
<tr>
<th>Gene</th>
<th>F_An</th>
<th>M_An</th>
<th>F_Le</th>
<th>M_Le</th>
<th>F_Pr</th>
<th>M_Pr</th>
<th>F_Ge</th>
<th>M_Ge</th>
</tr>
</thead>
<tbody>
<tr>
<td>SlitFAR1</td>
<td>5.01</td>
<td>3.01</td>
<td>4.93</td>
<td>3.05</td>
<td>5.70</td>
<td>4.12</td>
<td>6.13</td>
<td>4.59</td>
</tr>
<tr>
<td>SlitFAR2</td>
<td>3.31</td>
<td>0.54</td>
<td>3.07</td>
<td>1.10</td>
<td>2.56</td>
<td>2.21</td>
<td>0.85</td>
<td>4.12</td>
</tr>
<tr>
<td>SlitFAR3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SlitFAR9</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SlitFAR6</td>
<td>3.90</td>
<td>3.90</td>
<td>4.26</td>
<td>3.78</td>
<td>3.75</td>
<td>3.78</td>
<td>0.47</td>
<td>3.89</td>
</tr>
<tr>
<td>SlitFAR23</td>
<td>1.64</td>
<td>1.64</td>
<td>1.93</td>
<td>1.90</td>
<td>0.52</td>
<td>0.52</td>
<td>1.29</td>
<td>1.29</td>
</tr>
<tr>
<td>SlitFAR11</td>
<td>1.14</td>
<td>1.14</td>
<td>1.23</td>
<td>1.20</td>
<td>1.00</td>
<td>1.00</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>SlitFAR16</td>
<td>3.70</td>
<td>3.70</td>
<td>4.09</td>
<td>3.97</td>
<td>3.17</td>
<td>3.17</td>
<td>3.52</td>
<td>3.52</td>
</tr>
<tr>
<td>SlitFAR21</td>
<td>7.03</td>
<td>7.03</td>
<td>7.07</td>
<td>7.07</td>
<td>1.25</td>
<td>1.25</td>
<td>1.16</td>
<td>1.16</td>
</tr>
<tr>
<td>SlitFAR4</td>
<td>4.85</td>
<td>4.85</td>
<td>5.27</td>
<td>5.12</td>
<td>3.91</td>
<td>3.91</td>
<td>4.75</td>
<td>4.63</td>
</tr>
<tr>
<td>SlitFAR13</td>
<td>1.85</td>
<td>1.85</td>
<td>2.29</td>
<td>2.20</td>
<td>2.16</td>
<td>2.16</td>
<td>2.32</td>
<td>2.32</td>
</tr>
<tr>
<td>SlitFAR7</td>
<td>6.85</td>
<td>6.85</td>
<td>7.04</td>
<td>7.00</td>
<td>3.57</td>
<td>3.54</td>
<td>4.25</td>
<td>4.25</td>
</tr>
<tr>
<td>SlitFAR25</td>
<td>3.11</td>
<td>3.11</td>
<td>3.64</td>
<td>3.60</td>
<td>1.66</td>
<td>1.66</td>
<td>1.91</td>
<td>1.91</td>
</tr>
<tr>
<td>SlitFAR26</td>
<td>0.19</td>
<td>0.19</td>
<td>0.05</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>SlitFAR8</td>
<td>5.60</td>
<td>5.60</td>
<td>5.27</td>
<td>5.23</td>
<td>2.27</td>
<td>2.27</td>
<td>4.13</td>
<td>4.13</td>
</tr>
<tr>
<td>SlitFAR15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SlitFAR19</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SlitFAR12</td>
<td>7.78</td>
<td>7.78</td>
<td>7.90</td>
<td>7.86</td>
<td>5.52</td>
<td>5.48</td>
<td>6.31</td>
<td>6.26</td>
</tr>
<tr>
<td>SlitFAR10</td>
<td>8.72</td>
<td>8.72</td>
<td>8.37</td>
<td>8.31</td>
<td>4.48</td>
<td>4.43</td>
<td>5.36</td>
<td>5.32</td>
</tr>
<tr>
<td>SlitFAR30</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.04</td>
<td>0.67</td>
<td>0.67</td>
<td>1.06</td>
<td>1.06</td>
</tr>
<tr>
<td>SlitFAR5</td>
<td>0.09</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SlitFAR20</td>
<td>1.09</td>
<td>1.09</td>
<td>1.02</td>
<td>1.01</td>
<td>0.42</td>
<td>0.42</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>SlitFAR24</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>SlitFAR22</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>SlitFAR27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>SlitFAR28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>SlitFAR29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>SlitFAR14</td>
<td>3.26</td>
<td>3.26</td>
<td>1.62</td>
<td>1.62</td>
<td>0.07</td>
<td>0.07</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>SlitFAR18</td>
<td>2.65</td>
<td>2.65</td>
<td>1.55</td>
<td>1.55</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Figure 3

Relative expression levels of 30 FAR genes in the PGs. Relative expression levels are indicated as mean±SEM of three biological replicates. The comparisons were analyzed using one-way ANOVA, and the different letters above each bar indicate significant differences (Tukey test, P<0.05).

Figure 4

Relative expression levels of \textit{SlitFAR3} genes in the PGs. (A) Relative expression levels of \textit{SlitFAR3} gene at different time. L, light; D, dark. (B) Relative expression levels of \textit{SlitFAR3} gene at different day age. Relative expression levels are indicated as mean±SEM of three biological replicates. The comparisons were
analyzed using one-way ANOVA, and the different letters above each bar indicate significant differences (Tukey test, P<0.05).

**Figure 5**

GC-MS analyses of the fatty alcohol products extract from yeast cells transformed with plasmid in presence of 0.5 mM precursor compounds. The chromatogram traces from top to bottom shows the fatty acyl standards, fatty alcohol standards and supplemented with substrates Z9E11-14:Me, Z9E12-14:Me, Z9-14:Me, E11-14:Me, respectively. The yeast cells transformed with (A) pESC-URA-SlitFAR3 and (B) pESC-URA. The x axes represent the retention time, the y axes represent the relative abundance. IS: internal standard (Z9E12-14:OAc).
Figure 6

Number of predicted FAR genes across 20 Lepidoptera species and domain duplication FAR genes. The schematic phylogenetic tree of Lepidoptera species was obtained from TimeTree (http://timetree.org/). The rightmost column indicates the number of FAR genes with domain duplication.
Sex pheromone biosynthesis pathway in *S. litura*. The biosynthetic pathway of sex pheromones involves a variety of enzymes, containing acetyl CoA carboxylase (ACC), fatty acid synthase (FAS), fatty acid β-oxidase, desaturase (Des), fatty acid reductase (FAR), acyltransferase (ACT). The de novo biosynthesis of all precursors starts from 16:CoA (palmitoyl-CoA). Fatty acyl-CoA (determined as methyl esters) was the substrate of SlitFAR3 enzyme. The schematic diagram of sex pheromone biosynthesis pathway was drawn with reference to the published article (Antony et al, 2016; Matsumoto 2010; Tillman JA et al. 1999).

### Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)