

## Article

# Antennal Transcriptome of the Fruit-Sucking Moth *Eudocima materna*: Identification of Olfactory Genes and Preliminary Evidence for RNA-Editing Events in Odorant Receptors

Meenal Vyas , Kamala Jayanthi Pagadala Damodaram \*  and Gandham Krishnarao 

Division of Crop Protection, ICAR-Indian Institute of Horticultural Research (IIHR), Hessaraghatta Lake Post, Bengaluru 560089, Karnataka, India; meenalvyas77@gmail.com (M.V.); gandhamkrishna42@gmail.com (G.K.)

\* Correspondence: jaiinsect@gmail.com or kamalajayanthi.pd@icar.gov.in

**Abstract:** Unappealing shriveled fruits are a characteristic of one of the most elusive fruit pests. The perpetrator, *Eudocima materna*, attacks the fruit at a fully formed stage and, therefore, the antennal transcriptome for this insect was deduced to identify the molecular elicitors involved in the attraction to its host plants. A total of 260 olfactory genes, including 16 odorant-binding proteins (OBPs), four pheromone-binding proteins (PBPs), 40 antennal-binding proteins (ABPs), 178 odorant receptors (ORs), 17 chemosensory proteins (CSPs) and five sensory neuron membrane proteins (SNMPs) were identified. Phylogenetic analysis shows the divergence of *E. materna* proteins from closely related lepidopterans and provides insights on genes that have exclusively evolved in this insect. STRING network analysis revealed interactions of olfactory proteins among themselves and the proteins of other groups. Interestingly, online tools predicted RNA-editing events in the odorant receptor sequences, suggesting the possibility of multiple protein forms. Transcripts matching transposable element sequences were also detected in the dataset. Thus, the work reported here provides a valuable resource to design molecular methods for pest control.

**Keywords:** fruit-sucking moth; odorant-binding proteins; RNA editing; antennal transcriptomics; pheromone-binding proteins



**Citation:** Vyas, M.; Pagadala Damodaram, K.J.; Krishnarao, G. Antennal Transcriptome of the Fruit-Sucking Moth *Eudocima materna*: Identification of Olfactory Genes and Preliminary Evidence for RNA-Editing Events in Odorant Receptors. *Genes* **2022**, *13*, 1207. <https://doi.org/10.3390/genes13071207>

Academic Editor: Fabrizio Lombardo

Received: 28 April 2022

Accepted: 3 July 2022

Published: 6 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Most lepidopterans (moths and butterflies) are agricultural pests that only damage the crop in the larval stage, while the adults are usually short-lived and harmless, surviving on nectar/plant juices until they mate and lay eggs. Among the handful of moths that are actually damaging in the adult stage, the fruit-sucking moth (FSM), *Eudocima (Othreis) materna* L. stands out. Its ability to attack several commercial fruit crops in tropical and subtropical regions from Africa to the Pacific Islands results in huge economic losses across the globe [1–4]. Male and female adult moths of *E. materna* and related species such as *Eudocima phalonia*, *Eudocima salamina*, and *Eudocima homaena* are nocturnal feeders on fruits and cause extensive damage [1–3]. The moths attack fully developed fruits by piercing through the epicarp and sucking the juice, hence being commonly known as fruit-piercing or fruit-sucking moths [4]. The rotting of infested fruits is the collateral damage caused by secondary infestations by pathogens [5]. The female moths lay eggs on Menispermaceae plants such as *Tinospora cordifolia* that are abundant in forest and scrublands far away from the commercial fruit orchards [2], and the larvae feed on the foliage and develop into adults. Traditional management methods such as bagging, the hand-collection of moths, netting the orchards, night watching, insecticidal application, and bonfires to attract and kill the moths are of limited value in containing the fruit damage, as adult moths can fly long distances [6]. Therefore, the behavior-modifying strategies involving chemical cues such as

sex pheromones and host cues (kairomones) that target insect olfaction/chemoreception are of immense value, as these are eco-friendly and sustainable [6].

First described by Carl Linnaeus in 1767, the elusive nature of the FSM has kept its chemo-ecological behavior relatively unexplored. Specialized aspects of *E. materna*, such as the survival of adults exclusively on high-acid content, sugar-rich fruit juices and the ability of the larvae to feed on the Menispermaceae family of plants that have insecticidal properties [7] make them increasingly worthy of investigation. Therefore, exploring the specific antennal olfactory proteins that aid the moths while foraging for feeding and oviposition sites is imperative. Deciphering the molecular basis of olfaction and odor perception with regard to their host/mate location and oviposition site selection will help in formulating the most effective behavior-modifying management strategies.

Antennae are an important asset to an insect's olfactory system, facilitating chemoreception that detects cues for food location, mate identification, oviposition site selection, and threat detection [8,9]. The antennae bear sensilla [10] with a sensillum lymph consisting of olfactory proteins that can bind to the hydrophobic volatile cues in the environment [11]. Insects have evolved certain olfactory proteins such as chemosensory proteins (CSPs) and odorant-binding proteins (OBPs) that bind to the volatile cues from the environment and carry them further for recognition by the chemosensory neurons, which are detected by various olfactory receptors (ORs), gustatory receptors (GRs), or ionotropic receptors (IRs) [12–18]. Lepidopterans also have a few specialized OBPs, commonly referred to as PBP, that detect pheromones [19].

Recent transcriptome (whole-body and antennal) studies have led to the identification of several OBPs and ORs across the insect community. However, transcriptome data of lepidopteran pests such as *Chilo suppressalis* (Walker) [20], *Conogethes punctiferalis* (Guenée) [21], *Plodia interpunctella* (Hübner) [22], *Spodoptera exigua* (Hübner) [23], *Cnaphalocricis medinalis* (Guenée) [24–26], and *Plutella xylostella* (Linn.) [27] are insufficient to explore the olfactory genes of FSMs that are distinct from other lepidopterans in their behavior, as explained earlier. Significant details of a de novo antennal transcriptome were reported recently in another fruit-piercing noctuid moth, *Oraesia marginata* Fab. The larvae of this moth develop on the plants of Menispermaceae and the adults are pests of several commercial fruit crops [28], much like the *E. materna* larval stages. Despite the similarity, there could be specifics that are not common to the two FSMs mentioned and several others that have not been discussed. Interestingly, a study comparing the OBPs from 12 *Drosophila* spp. genomes has shown that species that are specialized with regard to the ecological system have rapid evolution in the OBPs [29].

The behavioral chemical ecology requires the involvement of several receptors that are processed in the brain [16]. Although RNA-editing events are a well-known phenomenon that brings about protein diversity from a single gene, its importance in insects has been discovered in the last decade. RNA editing has been identified as a mechanism that ensures proper functioning of the brain in some model systems [30–32]. Transcript-specific modifications are brought about by the RNA-editing molecular mechanism, resulting in protein diversity [30–32]. The adenosine-to-inosine RNA editing mediated by adenosine deaminases acting on RNA or ADARs [33,34] is one such mechanism reported across several eukaryotes. Studies in different systems have shown that significant RNA editing exists in the brain [31]. More recently, significant RNA-editing events have been reported in bumblebees that manifest into specialized tasks [35]. Consequently, it would be interesting to know whether such RNA-editing events are a part of the machinery in insect pests that helps them adapt to a larger host range. Significant findings such as the one stated above underscore the importance of elucidating the olfactory genes of *E. materna* through antennal transcriptomics, which will eventually help comprehend its olfactory mechanism in detail and provide details on its commonalities with other fruit-sucking/piercing moths.

In the present study, the antennal transcriptome of *E. materna* was obtained and assembled to identify the various genes involved in its olfaction/chemosensation. Additionally, the phylogenetic analysis of the identified gene classes highlights the relationship of FSM

with other closely related lepidopterans. STRING analysis using online tools helped to identify interactions between the different olfactory genes. Preliminary bio-informatic studies were also carried out to identify RNA-editing sites in some categories of the olfactory proteins that would help gain insights into how they govern insect behavior.

## 2. Materials and Methods

### 2.1. Insects

The insects were collected during larval development directly from the *Tinospora cardifolia* blocks of ICAR-IIHR, Bengaluru, Karnataka. The larvae were reared in the laboratory on natural *T. cardifolia* cuttings to complete development under controlled conditions ( $26 \pm 1$  °C, 75 to 80% RH). The emerged adults (one pair) were placed in oviposition cages (30 × 45 cm) lined with paper towels as oviposition substrate and covered with nylon mesh on the upper side. The moths were fed on ripe banana, guava, and pomegranate in small plastic caps inside the cages and the food was replaced daily. Eggs laid were collected and kept in a circular insect-breeding dish until hatching. The eggs were then examined at an interval of 12 h and the emerging larvae were reared individually for the collection of antennae. For antennal transcriptomics, 50 antennae each from unmated male and female moths were excised one day after eclosion using a pair of sterile sharp micro scissors and were immediately frozen in liquid nitrogen, and stored at  $-80$  °C until the RNA was extracted.

### 2.2. RNA Isolation

The RNA was extracted from samples stored using a Direct-zol™ RNA Mini Prep (Cat# R2050, Zymo research, Irvine, CA, USA) kit following the manufacturer's protocol. In brief, 1 mL of TRI Reagent (Sigma-Aldrich, St. Louis, MO, USA) was added to 100 mg of tissue ground to a fine powder in a DEPC-treated, sterilized mortar and pestle using liquid nitrogen and was vortexed for 30 s. 1 mL of 100% ethanol was then added to the homogenized sample. The sample was loaded into a Zymo-Spin™ IIC column and centrifuged at  $12,000 \times g$ . The initial flow-through was discarded and the column was washed with 500 µL of Direct-zol™ RNA pre-wash buffer. Further, 700 µL of RNA wash buffer was added to the spin tube and centrifuged to discard the flow-through. The RNA was finally eluted in 50 µL of nuclease-free water. The RNA quantity was estimated using Qubit RNA BR assay (Thermo Fisher Scientific, Waltham, MA, USA) and the quality was checked using an Agilent Bio analyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and RNA Nano kit (Agilent Technologies, Santa Clara, CA, USA).

### 2.3. cDNA Library Construction and RNA Sequencing

Sequencing libraries were prepared from 1 µg of RNA using NEB Next Ultra RNA Library Prep Kit for Illumina (NEB #E7770, New England Biolabs, Ipswich, MA, USA) after depleting the rRNA using an NEB Next rRNA Depletion Kit (Catalog #: E7400S, New England Biolabs, Ipswich, MA, USA). In brief, probes were hybridized to the RNA, treated with RNase H and DNase 1 to degrade the rRNA and DNA probes, and the RNA was purified using Agencourt RNA clean XP. The purified RNA was fragmented and used for cDNA synthesis, which was subsequently PCR enriched, adapter ligated, and purified using 1.8× Agencourt AMPure XP beads. The library quantity was checked using a Qubit dsDNA HS kit (Thermo Fisher Scientific, Waltham, MA, USA). The library quality was checked using an Agilent Bioanalyzer 2100 and Agilent DNA 7500 Kit (Agilent Technologies, Santa Clara, CA, USA).

### 2.4. Illumina Sequencing

After quality inspection, the cDNA libraries were denatured and diluted to Illumina-recommended 1.8 pM concentrations using 0.2 nM NaOH and HT1 reagent, respectively. The denatured and diluted libraries were loaded onto a NexSeq reagent cartridge. Cluster generation and sequencing were performed on an Illumina Next Seq

500 to generate  $2 \times 150$  paired-end reads. Soon after completion of the run, demultiplexed sequence data was collected in FASTQ format and the data quality was checked using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 14 June 2017) and MultiQC [36] for base quality distribution, % reads with average Q30 and Q20, and % GC and sequencing adapter contamination. Raw sequence reads were processed to remove adapter sequences and low-quality bases using the Trim-galore ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), accessed on 14 June 2017) and Trimmomatic [37] packages. The raw sequence reads are available under the SRA accession PRJNA658216. This Transcriptome Shotgun Assembly project was deposited at DDBJ/EMBL/GenBank under the accession GIUU00000000.

### 2.5. De Novo Transcriptome Assembly and Functional Annotation

Quality-passed reads were assembled using Trinity v2.4.0 [38] with default parameters (Kmer = 25). The assembled contigs were clustered using CD-HIT-EST [39] with 90% sequence similarity. The longest sequence of the clustered contigs was considered as a transcript. The transcripts were annotated by mapping onto NCBI nr, Uniprot–Swissprot, and COG databases using Diamond BLASTp with an e-value cutoff of  $1 \times 10^{-5}$ . The transcripts that mapped onto the database with an e-value  $> 1 \times 10^{-5}$  and sequence identity  $< 70\%$  were rejected. Transcripts with confident annotations were considered as unigenes and were used for downstream analysis. Additionally, KEGG Orthology (KO) annotations were assigned using KAAS (KEGG Automatic Annotation Server) [40].

### 2.6. Gene Ontology and KEGG Pathway Enrichment Analysis for Antennal Transcriptomics

Gene ontology (GO) information for the unigenes was retrieved using the idmapping web application in Uniprot [41]. To check the Biological Process, Cellular Component, and Molecular Functions of the transcripts, we performed GO term enrichment analysis using BINGO [42]. A GO distribution graph was drawn using WEGO [43]. The expressed transcripts were submitted as a multiple FASTA file and the available insect gene sets in KEGG were used as trainer gene sets for annotation using the BBH (bi-directional best hit) method. COG terms annotation was performed by importing the NCBI-NR annotation of each gene into MEGAN.

### 2.7. Data Collection of Olfactory and Chemosensation-Related Sequences of Other Lepidopteran Species

The NCBI protein sequence tool was used to download the sequences of related species. Initially, “Lepidoptera” as a search term was used along with the specific protein type, “odorant-binding proteins”/“odorant receptor”/“pheromone-binding proteins”/“chemosensory proteins”/“antennal-binding proteins”/“sensory neuronal proteins”, at the NCBI database. The list of proteins obtained was further classified based on the other species of the family Noctuidae and closely related families. The entire protein information along with sequences in the FASTA format were downloaded and saved in an Excel sheet. The data were sorted based on the specific protein types, and partial sequences and unrelated sequences were removed. Sequences that had short length were also removed. The insect species were arranged, and the sequences obtained after cleaning were aligned with MUSCLE [44] using the MEGA software [45] and the aligned sequences were saved as MAS files.

### 2.8. Olfactory Gene Identification and Phylogenetic Analyses of Antennal Transcriptome

All candidate OBPs, ORs, CSPs, and SNMPs, and PBP were manually checked by the BLASTx program at the National Center for Biotechnology Information (NCBI). For contigs with hits against genes of interest, open reading frames (ORFs) were identified and the annotation verified OBP, OR, CSP, SNMP and PBP protein sequences. Orthologs in closely related species of *E. materna* were used to analyze the characteristics of olfactory genes in *E. materna*. Nucleotide sequences of all olfactory genes that were identified from

the *E. materna* antennal transcriptomes were named according to sequence homology analysis and numbered arbitrarily. The proteins were aligned using MUSCLE [44] and phylogenetic analysis was done with MEGA using the maximum likelihood method. The test of phylogeny was done using the bootstrap method with 500 replications. The method followed was the Jones–Taylor–Thornton (JTT) model. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches is the support value for the cluster. The initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with a superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 200 amino acid sequences and there was a total of 638 positions in the final dataset for the FSMCSP tree, 342 amino acid sequences with a total of 796 positions in the final dataset for the FSMOBP tree, 530 amino acid sequences and a total of 1043 positions in the final dataset for the FSMOR tree, and 27 amino acid sequences and a total of 544 positions in the final dataset for FSMSNMP. For the FSMPB tree, this analysis involved 94 amino acid sequences with a total of 178 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [45]. The Newick tree generated was then transferred to Fig Tree 4 (available from <http://tree.bio.ed.ac.uk/software/figtree/>, accessed on 29 June 2022) using the bootstrap values for further analysis and figure preparation. The branches were assigned a color based on the clade and the radial phylogeny was obtained. A bootstrap analysis of 500 iterations was performed to evaluate the branch strength of each tree. The support values represented on the branches for a cluster are in percentages based on the bootstrap values. They signify the number of times the cluster was generated in a specified number of iterations. In this case, it would be 500.

### 2.9. Identification of Transposable Element-Related Transcripts

Insects are known to harbor transposable elements, which could integrate and be a part of the genome. We used the transcriptome data generated from the antennal tissue to identify presence of transposable elements in *E. materna*. The annotated transcripts were filtered to identify such elements using text filters “transposon” and “transposable element”. The transcripts identified were sorted based on the read counts.

### 2.10. CDNA Synthesis for Sequence Validation

Antennae were collected from male/female FSM moths as explained earlier, and the RNA was extracted from them using a Qiagen Qiaquick RNA extraction kit. The RNA was quantified and about 1 µg of the total RNA was processed to synthesize cDNA (reverse transcriptase cDNA synthesis kit—Qiagen, Germantown, MD, USA) after DNaseI treatment. The cDNA was diluted 1:100 and used for PCR. Specific primers were used for each of the genes selected using transcriptome data. The PCR cycling parameters were set at 40 cycles of 95 °C for 1 min, 60 °C for 30 s, and 72 °C for 30 s. The products obtained were confirmed by visualization on 1% agarose gel and a column purified before sequencing.

Eleven genes were chosen from the transcriptome data annotations and these were used for sequence validation by RT-PCR. The transcript sequences were used to design primers, and then PCR was carried out with the cDNA synthesized above. The 11 products obtained were column-purified and Sanger-sequenced using specific primers. The sequences obtained were blasted and we could indeed identify the specific protein hits. The sequences were submitted to NCBI with the accession IDs MW186442–MW186452.

### 2.11. STRING Protein Interacting Network Analysis

To perform the STRING analysis of the FSM olfactory proteins, the translated sequences of the transcripts identified from the antennal transcriptome of FSMs were submitted online using the freely available tool at <https://string-db.org/>, accessed on 1 June 2021. The scoring parameters are described in [46]. The resulting network was analyzed using

the *Bombyx mori* L. genome as the reference. The interactions were obtained and interpreted based on the legends provided in the database. A medium confidence level of 0.4 was used for the settings. The non-interacting nodes were then removed to obtain a clear picture of the interactions.

### 2.12. RNA-Editing Site Prediction

cDNA/transcript sequences for the FSM odorant receptors were obtained from the TRINITY transcriptome sequencing files. They were converted to FASTA format and submitted to two different freely available online prediction tools, namely, AIRLINER [47] (<http://upalpha.dmi.unict.it/airliner/>, accessed on 1 June 2021) and InosinePredict [48] (<http://www.biochem.utah.edu/bass/inosinepredict>, accessed on 1 June 2021). AIRLINER is a mathematical modeling-based software program that predicts ADAR-based RNA-editing sites using information from several experimentally determined edited sites. The predictions are probability values assigned to the base with respect to the model. Any value > 0.5 is considered to have a high probability of being edited. The data were further sorted in Excel to obtain the number of sites that had a >0.5 probability. Simple functions such as COUNTIF and sort were used. Similarly, InosinePredict is also a freely available online tool that predicts RNA-editing sites based on the possible outcomes of the four human-based ADARs (hADAR1, hADAR2, hADARd1, and hADARd2) in vivo data. It provides a percent edited possibility—the higher the percentage, the higher the probability that the site will undergo RNA editing. The results obtained were analyzed, and many sites predicted by both of the tools overlapped.

## 3. Results

### 3.1. Sequencing and Unigene Assembly

A total of 32,660,232 raw reads were obtained from the antenna samples, and the removal of low-quality reads yielded 31,416,578 clean reads with lengths ranging from 100 to 136 bp (Q20 content at 99.31% and Q30 content at 93.48%). The clean reads were assembled into 57,600 transcripts. The minimum and maximum length of the assembled transcripts was 301 nt and 10,724 nt, respectively, with the average length of transcripts at 930.96 nt. The GC content was 44.50% and the number of N50 and N10 transcripts was 1324 and 3311, respectively (Table 1). Among the 68,986 assembled *E. materna* transcripts, a total of 16,476 unigenes could be generated and were within a length range of 200–500 bp (Supplementary Figure S1a). The greatest number of hits to closely related species was to *Amyelois transtiella* W., followed by *B. mori* and *Papilio xuthus* L. (Supplementary Figure S1b).

### 3.2. Functional Annotation

Among the 68,986 assembled *E. materna* transcripts, 19,445 unigenes could be generated and a total of 15,718 unigenes showed hits to unique NCBI non-redundant proteins. Finally, 16,476 unigenes were obtained for 57,600 transcripts from the antennae of *E. materna*. At the “Cellular Components” level, the Cell and Cell Part categories were the most abundant, whereas Symplast, Virion, Synapse, Virion Part, and Cell Junction were the least abundant. The Binding and Catalytic Activity categories accounted for majority of the unigenes at the “Molecular Function” level. The Biological Process category showed maximum assignments to the subcategories of Cellular Process and Metabolic Process (Supplementary Figure S2). Further, the annotation of unigenes using the COG database revealed a total of 8559 unigenes that could be annotated and categorized into 25 functional groups. The two largest groups were Post-Translational Modification, Protein Turnover, Chaperones and Intracellular Trafficking, Secretion, and Vesicular Transport, accounting for 2196 unigenes. The smallest groups were Cell Motility and Nuclear Structure with 10 unigenes each (Supplementary Figure S3).

**Table 1.** Validated unigene sequence details.

SNO.	Accession ID	Description	Primer Name	Primer Sequence
1	MW186442	F-BAR domain-only protein 2-like	FSM1F FSM1R	AACCTACGCATCCCCTACAC GCGCAGTACTTCAATGTGGA
2	MW186443	Predicted esterase B-1-like	FSM2F FSM2R	AAGCAACACCAAGGCCTCTA GATACCACTTTGCGCCATT
3	MW186444	Carboxylic ester hydrolase	FSM3F FSM3R	CACCTGTTGGTGGTCACAAG CGTGGCTCACGACTGACTTA
4	MW186445	Esterase B1	FSM4F FSM4R	TAGCGTGTGGTATGCCTCTG GAAGCAGTCCGTTTCTTGC
5	MW186446	Acetate esterase-6	FSM5F FSM5R	GCCCACTGAAGAGTCAGGAG CTTTAAAAGCGCTGGATTCCG
6	MW186447	Chemosensory ionotropic receptor75q2	FSM6F FSM6R	TGCGTGTCAACAAGGAAGTC CTCCGGTCTCCATGTGAAAT
7	MW186448	Odorant receptor 4-like	FSM8F FSM8R	CGACGCAGTTCAATTTGCTA AGCGTCACCTATTCCCACAC
8	MW186449	Odorant-binding protein 4	FSM11F FSM11R	ATTGGGACGATTTGAAGCAG GGCATTGATTTTCGTCCAGT
9	MW186450	Uncharacterized protein	FSM12F FSM12R	CTAGAGCTACGCCCGTGAAC TCCCTGCTGCGAAAGTTATT
10	MW186451	Odorant-binding protein	FSM13F FSM13R	TTTCTGGGAACAGGTTGTC TGTTTCGGACGTCGTTGTAA
11	MW186452	Odorant-binding protein 1	FSM14F FSM14R	GGCCGTGATCAAATACAGGT CTCAGCCTCGGATTGTAGC

KAAS server based analysis to explore the unigenes involved in different biological and biochemical pathways revealed 5423 unigenes that could be assigned to various KEGG pathways. Further, KEGG Orthology classifications of the unigenes mapped to five categories, i.e., Cellular Process, Environmental Information Processing, Genetic Information Processing, Human Diseases, and Metabolism and Organismal Systems. Among the KEGG Orthology classifications, Signal Transduction was the most abundant group, whereas Cellular Community-Prokaryotes was the least abundant group (Supplementary Figure S4).

### 3.3. RT-PCR Validation of Transcripts

Sequencing results for products obtained from the cDNA using specific primers designed against transcript sequences generated from the data confirmed that the products were, indeed, transcript sequences obtained from FSM. The blast results were recorded and the identified sequences corroborated with the annotations. We successfully identified OBPs from FSM antennae and were able to successfully validate 11 unigenes by RT-PCR and sequencing (MW186442–MW186452) (Table 1).

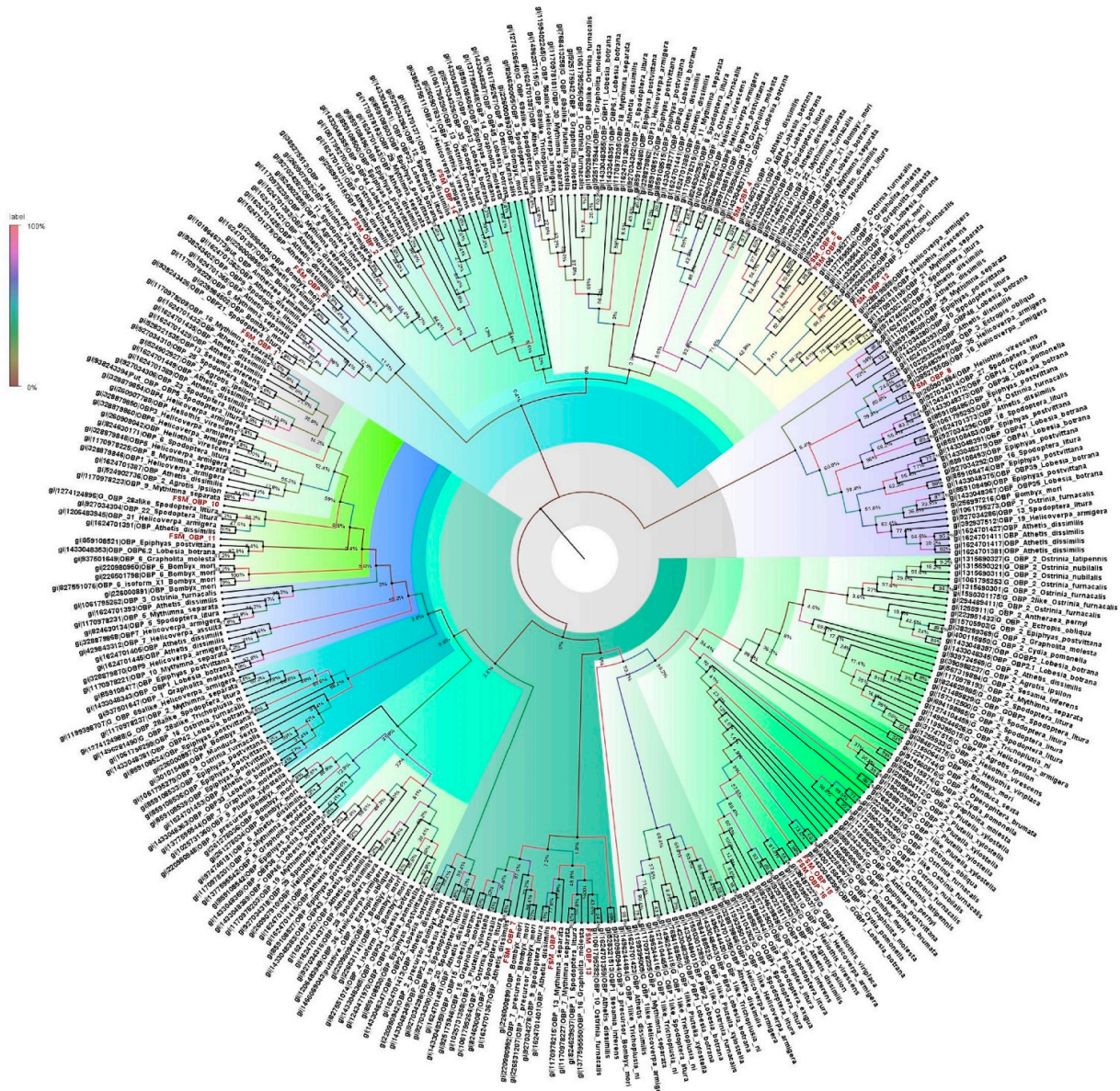
### 3.4. Identification of Olfactory Genes

A total of 260 olfactory genes, including 16 odorant-binding proteins (OBPs), some specialized OBPs such as four pheromone-binding proteins (PBPs), 40 antennal-binding proteins (ABPs), including two general odorant-binding proteins (GOBPs), 178 odorant receptors (ORs), 17 chemosensory proteins (CSPs), and five sensory neuron membrane proteins (SNMPs) were identified from the antennal transcriptome of *E. materna*.

### 3.5. Odorant-Binding Proteins

Unigenes corresponding to the annotated olfactory genes (16 OBPs, four PBPs, 40 ABPs) showed similarities to lepidopteran OBPs. All 16 OBPs were almost full length, ranging from 364 to 1540 bp. Besides the entrant OBP genes, some unigenes were identified as GOBP genes in *E. materna* (Supplementary Table S1). The phylogeny obtained from the OBP sequences of several different Lepidoptera species, including members of the Noctuidae family, namely, *E. materna*, *C. punctiferalis*, *Athetis dissimilis* (Hampson), *Helicoverpa armigera* (Hübner), *Heliothes virescens* (Hübner), *Spodoptera litura* (Fab.) *Mythimna separata* (Walker)

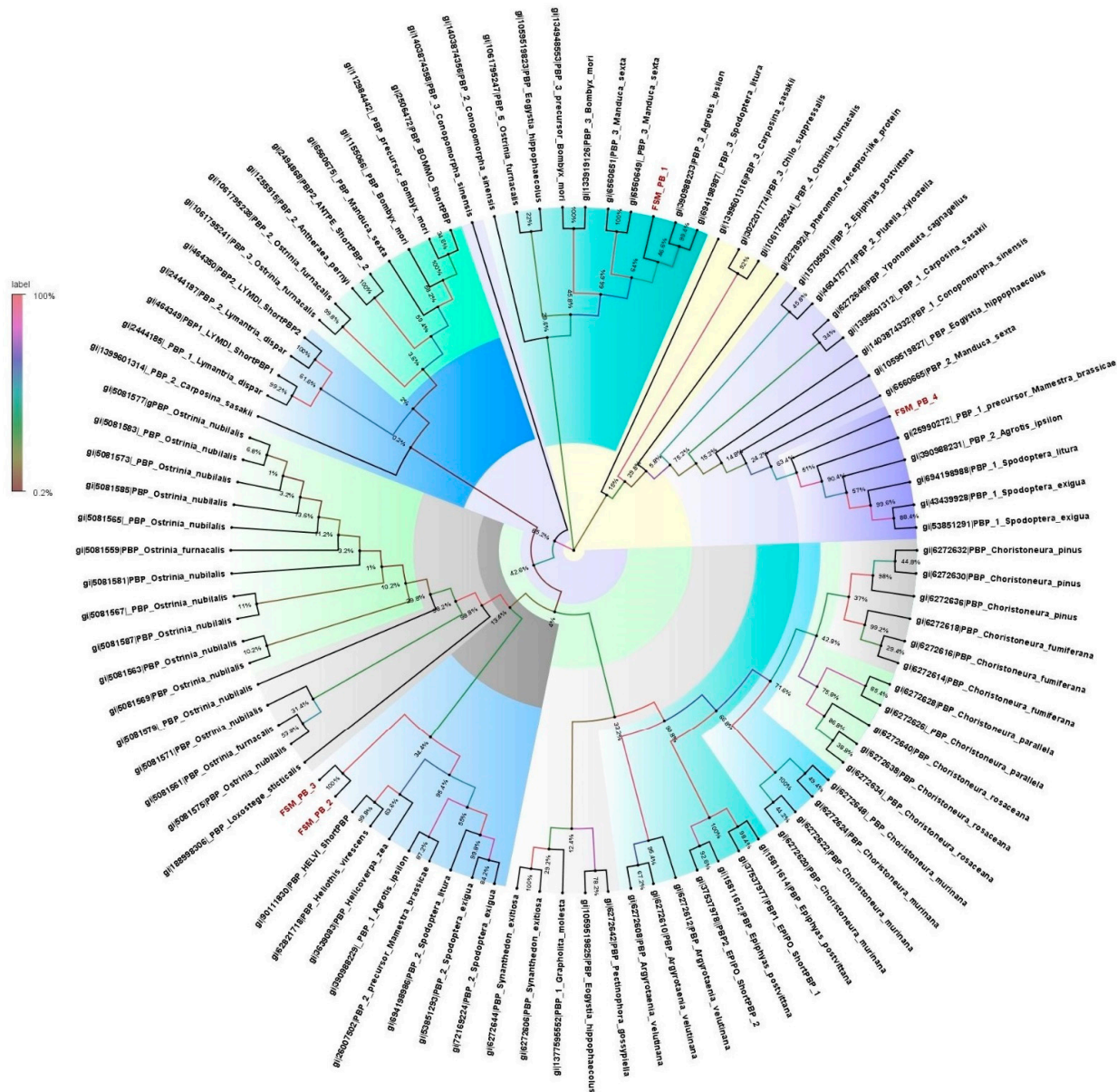
and others, as well as *Epiphyas postvittana* (Walker), *B. mori* and *Manduca sexta* (Linn.) of related families, indicated that most of the time, the identified OBPs of the *E. materna* transcriptome are usually associated with the noctuids. However, there are instances when some of the FSM OBPs (such as FSMOBP5, FSMOBP6, FSMOBP8, and FSMOBP12) were associated with related family moths like *Ostrinia furnacalis* (Guenée). Interestingly, OBPs such as FSMOBP15 and FSMOBP16 were closely associated. Overall, the FSMOBPs are quite diversified and may have developed into specialized proteins responsible for the unusual feeding habits of the moth (Figure 1).



**Figure 1.** Neighbor-joining dendrogram based on protein sequences of candidate odorant-binding proteins (OBPs) in fruit-sucking moth (FSM) *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related lepidopteran species OBPs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSM OBPs show diversification from each other and other lepidopterans. The colors used were to make the clades more prominent and do not bear any significance to clustering. Confidence values for each tree are represented as percentage values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.



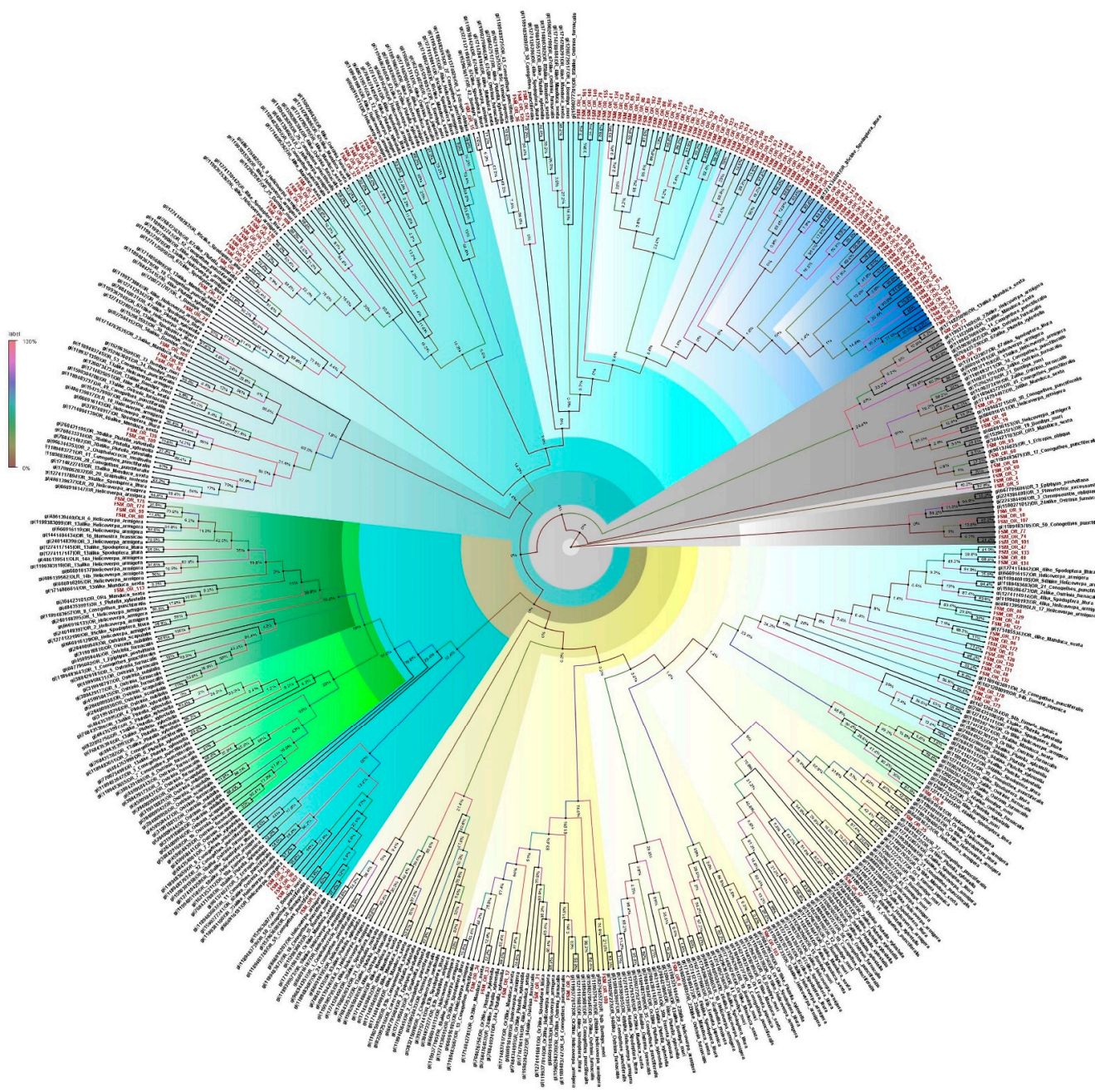
The four PBP identified in the antennal transcriptome of *E. materna* were in the range of 668 to 1517 bp, and a neighbor-joining tree of these sequences with related Lepidoptera species revealed that FSMPB1 and FSMPB4 are associated with PBP from *M. sexta*, since they occur in the same clade branching off from a distant common ancestor in *C. onopomorpha sinensis* B. The other PBPs, namely, FSMPB2 and FSMPB3, were closely associated to each other while occurring in the same clade as *Heliothis virescens* F. and *Loxostege sticticalis* L. (a Crambidae member), respectively (Supplementary Table S2 and Figure 2).



**Figure 2.** Neighbor-joining dendrogram based on protein sequences of candidate pheromone-binding proteins (PBPs) in fruit-sucking moth (FSM) *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related lepidopteran species PBPs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSM PBPs show diversification from each other and other lepidopterans. The colors used were to make the clades more prominent and do not bear any significance to the clustering. Confidence values for each tree are represented as percentage values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.

### 3.6. Odorant Receptors

Insect ORs are the most important receptors in sex pheromone and general odorant detection processes. A total of 178 ORs were identified from the antennal transcriptome of *E. materna*. The ORs identified are almost complete sequences in the range of 322 bp to 3251 bp. A phylogeny built for the OR sequences revealed that several ORs are closely associated with each other in the sub tree, but only show a relationship with *S. litura*. The tree also provides a display of multiple diversified FSMORs. The closest lepidopteran moth associated with FSM's ORs is *C. punctiferalis*. With respect to ORs, *E. materna* was found to be closely associated with other moth species such as *C. punctiferalis*, *M. sexta*, *S. litura*, and *O. furnacalis*. The multitude of ORs in FSM might provide the moth with an edge in detecting specific odors, especially at night, and may also help in mate selection (Supplementary Table S3 and Figure 3).



**Figure 3.** Neighbor-joining dendrogram based on protein sequences of candidate odorant receptors (ORs) in fruit-sucking moth *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related

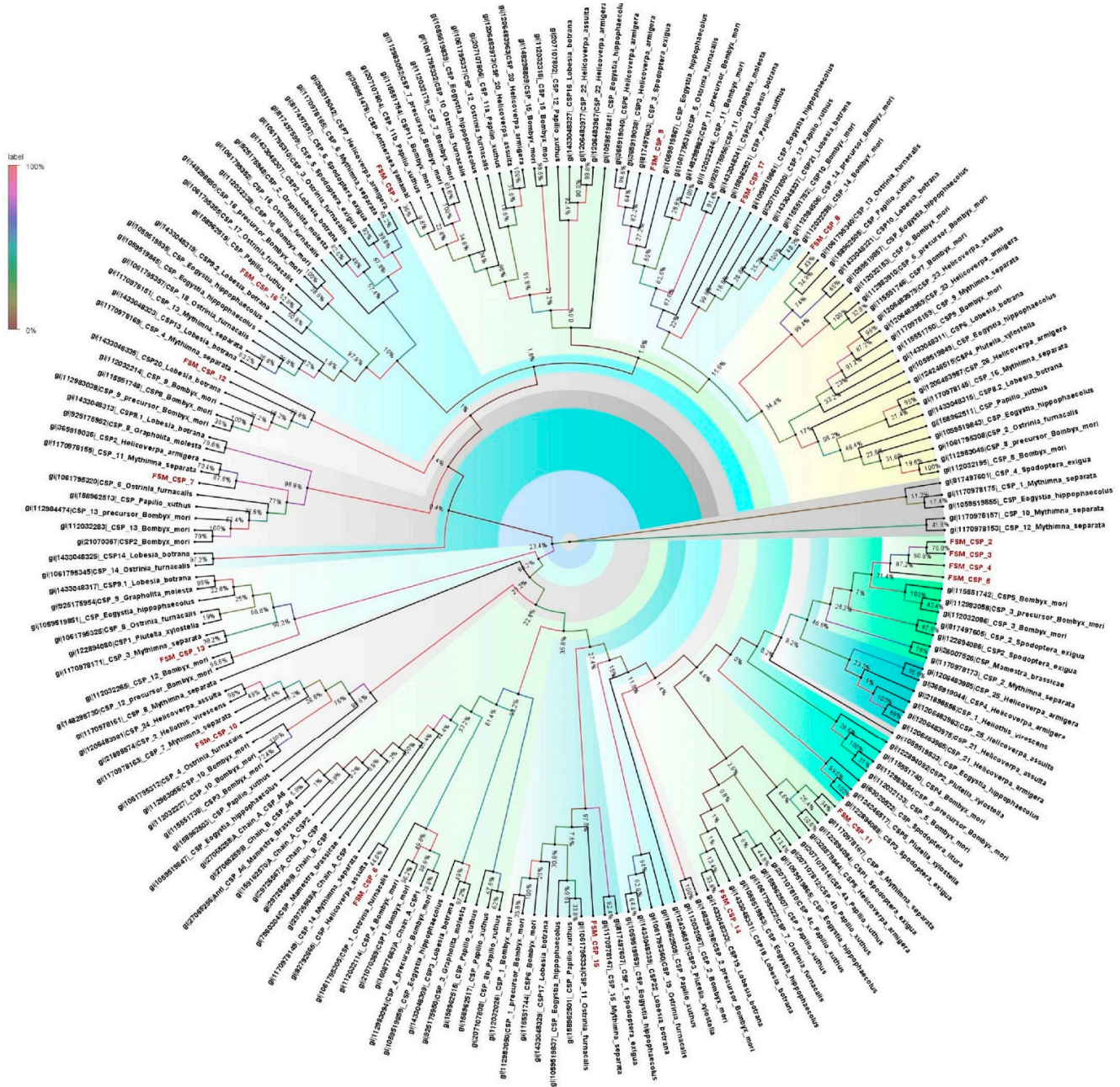
lepidopteran species ORs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSM ORs show diversification from each other and other lepidopterans. There are several ORs that cluster together and are specific to FSM. These could be the ones defining the unusual feeding behavior of FSM. The colors used were to make the clades more prominent and do not bear any significance to the clustering. Confidence values for each tree are represented as percent values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.

### 3.7. Chemosensory Proteins

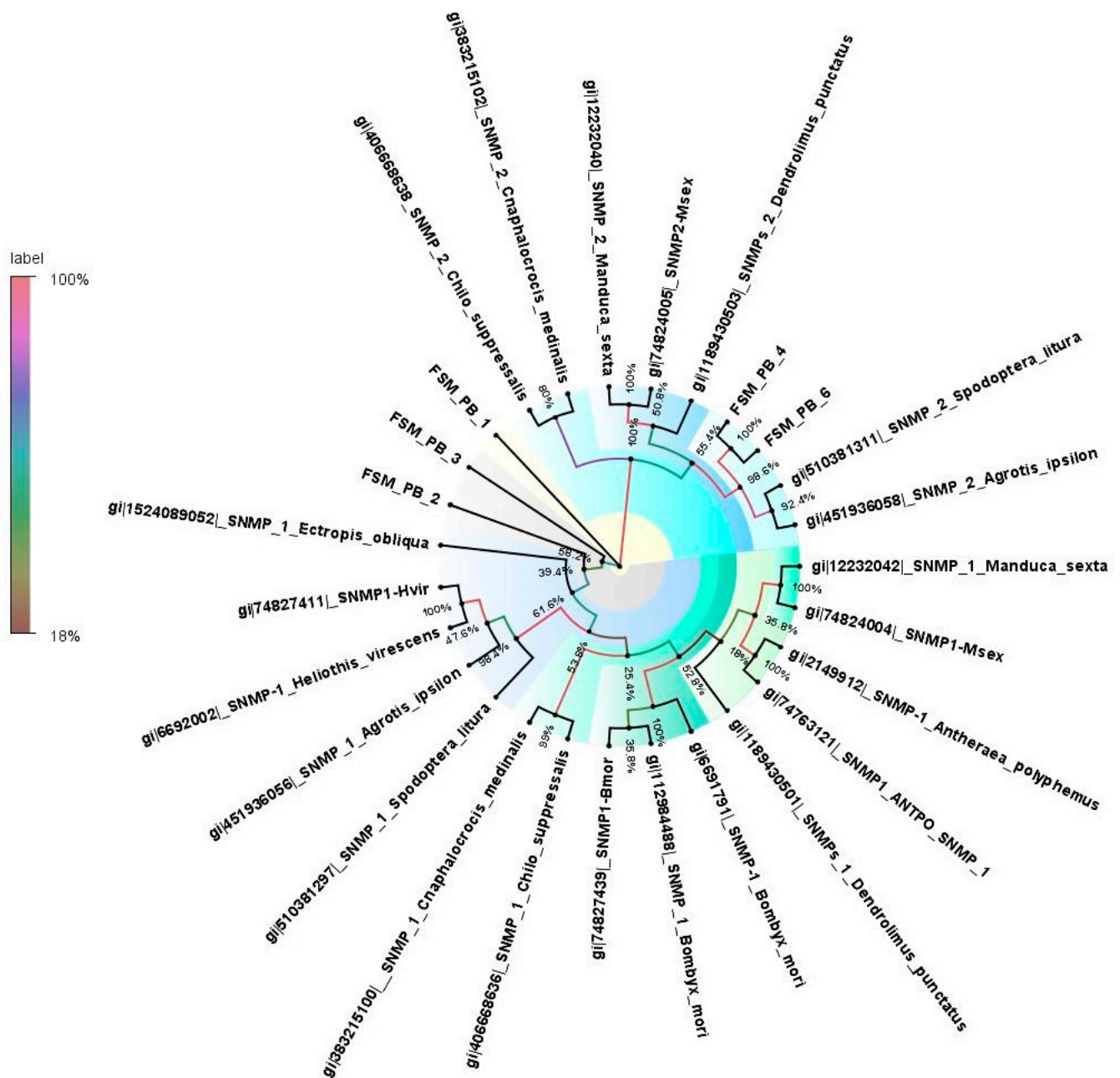
Insect chemosensory proteins (CSPs) are soluble proteins and widely distributed. CSPs are functionally important for sexual behavior and also the detection of volatiles. In the present study, a total of 17 CSPs were documented from the antennal transcriptome of *E. materna*. The length of the CSPs ranged from 310 bp to 2764 bp (Supplementary Table S4). A phylogeny of the CSP sequences of FSM along with several other Lepidopteran species, such as *B. mori*, *H. virescens*, *H. armigera*, *Helicoverpa assulta* (Guenée), *P. xylostella*, *O. furnacalis*, *Grapholita molesta* (Busck), *Papilio Xuthus* Linn., *S. exigua* (Hübner), *Lobesia botrana* (Denis and Schiffermüller), *Antheraea yamamai* (Guérin-Méneville), *Mamestra brassicae* (Linn.), *S. litura*, *Dendrolimus punctatus* Walk., *M. separata* Walk., *Eogystia hippophaecolus* (Hua, Chou, Fang and Chen), showed divergent relationships with each other. Some specific CSPs such as FSMCSP8 and FSMCSP17 are associated with *O. furnacalis*, *E. hippophaecolus*, *L. botrana* and *B. mori*, but in slightly divergent clades. The other CSPs, such as FSMCSP2, FSMCSP3, FSMCSP 4, and FSMCSP5, clustered together with some divergence (Figure 4). The support values for the tree vary from 100% to very low (4%) in some cases. This is mainly because there are several datapoints and the possibility of some clusters falling together is variable. The confidence would be considered for the clusters that have a support value greater than 50%.

### 3.8. Sensory Neuron Membrane Proteins

A total of five SNMPs were identified from the antennal transcriptome of *E. materna* with the length in the range of 332 bp to 3217 bp. A neighbor-joining tree of SNMP sequences showed divergent relationships with each other. Of all, FSMSNMP1 is the most likely divergent protein. The SNMPs, namely, FSMSNMP2 and FSMSNMP3, clustered together; however, these SNMPs look significantly divergent based on the branch length. The other SNMPs, namely, FSMSNMP4 and FSMSNMP6, showed associations with *S. litura* and *Agrotis ipsilon* (Hüfnagel) (Supplementary Table S5 and Figure 5). The support values for the tree vary from 100% to very low (4%) in some cases. This is mainly because there are several datapoints and the possibility of some clusters falling together is variable. Moreover, 40 ABPs were identified in the antennal transcriptome of *E. materna*, where ABP27 had the longest assembled sequence (4573 bp) and ABP 31 had the shortest (321 bp) (Supplementary Table S6).



**Figure 4.** Neighbor-joining dendrogram based on protein sequences of candidate chemosensory proteins (CSPs) in fruit-sucking moth (FSM) *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related lepidopteran species CSPs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSM CSPs show diversification from each other and other lepidopterans. The colors used were to make the clades more prominent and do not bear any significance to clustering. Confidence values for each tree are represented as percent values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.

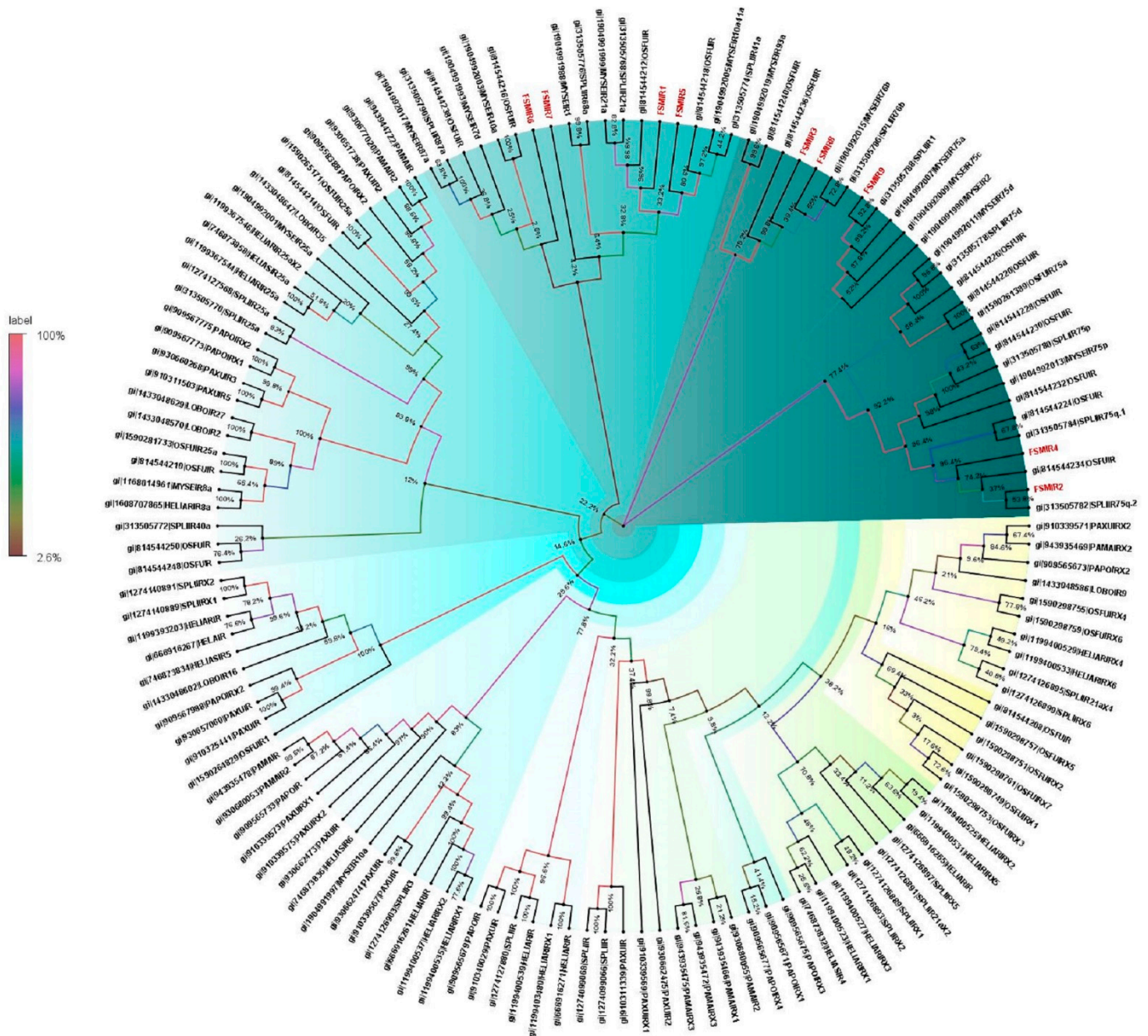


**Figure 5.** Neighbor-joining dendrogram based on protein sequences of candidate sensory neuron membrane proteins (SNMPs) in fruit-sucking moth (FSM) *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related lepidopteran species SNMPs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSM SNMPs show diversification from each other and other lepidopterans. The colors used were to make the clades more prominent and do not bear any significance to the clustering. Confidence values for each tree are represented as percent values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.

### 3.9. Ionotropic Receptors

Ionotropic receptors (IRs) have evolved from ionotropic glutamate receptors, which are a conserved group of synaptic ion ligand-gated ion channels. The antennal transcriptome of *E. materna* helped identify nine IRs. These, when phylogenetically analyzed with IRs from other related insect species, suggested that the FSM IRs clustered with those of *S. litura*

(Figure 6) and *O. furnacalis*. However, the individual IRs maintained their divergence from each other.



**Figure 6.** Neighbor-joining dendrogram based on protein sequences of candidate ionotropic receptors (IRs) in fruit-sucking moth (FSM) *E. materna*. Phylogenetic analysis of *E. materna* (red) and several related lepidopteran species IRs was done using the NJ method. The distances generated are based on the bootstrap values. The tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. FSMIRs show diversification from each other and other lepidopterans. The colors used were to make the clades more prominent and do not bear any significance to clustering. Confidence values for each tree are represented as percent values next to the cluster. Values greater than 50% would be considered to have higher confidence levels.

**3.10. Most-Represented Transcripts**

When the identified genes were sorted based on the raw counts, the most-represented members included the ribosomal proteins, followed by cytochrome oxidase subunits; chemosensory protein 3 was among the first 150 genes with high raw counts. Other

olfactory related proteins such as CSP5, sensory neuron membrane protein 1, sensory neuron membrane protein 2, take-out protein, and odorant-binding protein 7 were also found among the top 500 genes with high raw counts. The other related genes were general odorant-binding protein 72-like, odorant-binding protein 5, synaptic vesicle glycoprotein 2C-like, sensory neuron membrane protein 1, and synaptic vesicle glycoprotein 2C-like (Supplementary Table S7).

### 3.11. Transposable Elements

Most insects are known to have mobile genetic elements in their genome. Considering the possibility that the antennal tissue might show some interesting features with respect to the occurrence of such elements and their activity, we explored the presence of such elements in the antennal transcriptome of *E. materna*. Interestingly, we found hits against several transposable elements reported and identified in lepidopterans. The results are provided in the Supplementary Table S7.

### 3.12. General Transcripts (Non-Olfactory)

Some interesting non-olfactory-related proteins in the first 500 were circadian clock-controlled protein-like, PREDICTED: ecdysteroid-regulated 16 kDa protein-like, no child left behind, bax inhibitor 1 isoform X1, moesin/ezrin/radixin homolog 1 isoform X1, calmodulin, basigin, aquaporin AQPAe.a-like, titin, octopamine receptor in mushroom bodies, hypothetical protein OXYTRI\_14248, ribosome-associated membrane protein 4, stomatin-2, adiponectin receptor protein, ankyrin-3-like, glucose dehydrogenase, synaptobrevin-1, facilitated trehalose transported Tret1, and Apolipo proteins (Supplementary Table S7).

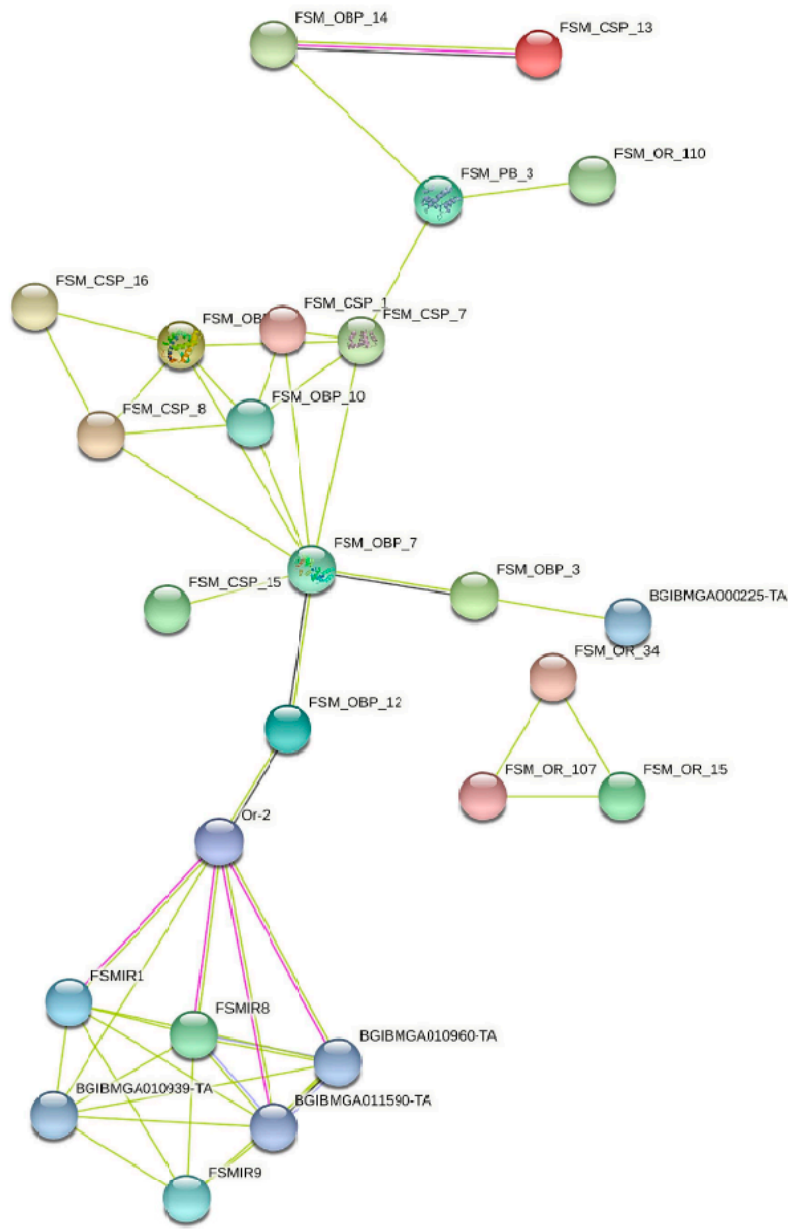
### 3.13. RNA-Editing Events in the FSM Odorant Receptors

Compared to the other related species, several odorant receptors were identified in *E. materna*, which is quite interesting. There is evidence in other protein families among insects of RNA-editing events that serves to give rise to protein diversity from a group of genes that have parallel functions, such as those in *B. mori*, mosquitoes, bumblebees, and *Drosophila*. Evidence of any such RNA-editing sites in the multiple ORs identified in the *E. materna* antennal transcriptome would suggest evolutionary significance of multiple ORs identified here. Therefore, freely available online tools such as AIRLINER and InosinePredict were used to predict such sites.

FASTA sequences of 124 FSMORs were submitted to AIRLINER and InosinePredict for the mathematical model-based prediction of RNA-editing sites in each sequence. Interestingly, in the 124 sequences used, the results from AIRLINER provided a total of 8895 sites that had an editing probability of >0.5 (Supplementary Table S8). There were 25 sequences with more than 100 edited sites with >0.5 editing probability. The hADAR1-based prediction in InosinePredict showed 5718 sites that had a prediction of >50% (Supplementary Table S9). Though preliminary, these results suggest that *E. materna* could possibly use the ADAR-dependent RNA-editing process to use ORs for different neurobehavioral, ecological, and chemo-ecological functions. Interestingly, four transcripts identified in the *E. materna* transcriptome had hits to adenosine deaminases that have a known function in RNA-editing events. One of the transcripts, when translated into a protein sequence using the ExPASy Translate tool, showed homology to adenosine deaminase protein of several moth species such as *M. sexta*.

STRING protein network analysis of some of the groups of sequences and also among themselves was carried out to explore the possible interactions. CSP proteins from FSM along with OBPs showed very few interactions. The most interesting interaction was between FSMOBP14 (annotated as *B. mori* sercotropin) and FSMCSP13. These proteins show evidence of co-expression and the interaction is experimentally determined based on the edges. The network also shows evidence of co-occurrence for FSM-OBP7 with FSMOBP3 and FSMOBP12. Allowing interacting partners with the *B. mori* proteins at the

next level revealed the interaction of FSMIR 8 and FSMIR1 with the *B. mori* Or2, which showed evidence for co-expression with FSMOBP12 (Figure 7).



**Nodes:**

Network nodes represent proteins  
*splice isoforms or post-translational modifications are collapsed, i.e. each node represents all the proteins produced by a single, protein-coding gene locus.*

**Node Color**

- colored nodes: query proteins and first shell of interactors
- white nodes: second shell of interactors

**Node Content**

- empty nodes: proteins of unknown 3D structure
- filled nodes: some 3D structure is known or predicted

**Edges:**

Edges represent protein-protein associations  
*associations are meant to be specific and meaningful, i.e. proteins jointly contribute to a shared function; this does not necessarily mean they are physically binding each other.*

**Known Interactions**

- from curated databases
- experimentally determined

**Predicted Interactions**

- gene neighborhood
- gene fusions
- gene co-occurrence

**Others**

- textmining
- co-expression
- protein homology

**Figure 7.** STRING protein interaction network analysis of fruit-sucking moth (FSM) *E. materna* transcripts translated into protein sequences.



#### 4. Discussion

Insect olfaction is the key area of focus for many agricultural researchers, since insects use olfactory senses to locate host plants, select mates, and detect threats. A molecular understanding of olfaction can help design customized solutions for pest control, particularly in devastating pests such as the fruit-sucking moth *E. materna*. The identification of 19,455 unigenes from 68,986 assembled transcripts and RT-PCR based sequence validation of the samples suggested the corroboration of the good quality of the data generated. The occurrence of a good number of transcripts from the antennal transcriptome suggests a high representation of actively transcribed genes reiterating the active role of the antenna. Antennal transcriptomes of other lepidopterans such as *C. suppressalis* [20] and *C. punctiferalis* [21] yielded a total of 66,560 and 47,109 unigenes, respectively.

The antennal transcriptome data for FSMs were generated with an aim to identify and explore the specific transcripts associated with the antennal tissue that aids the insects' response to chemo-ecological stimuli. In concurrence with the expectation, the COG categories that were most represented in the transcriptome were Post Translational Modification, Protein Turnover, Chaperones and Intracellular Trafficking, Secretion, and Vesicular Transport. The most-represented KEGG pathways were Environmental Information processing and, specifically, Signal Transduction, suggesting the antennal role in responding to environmental signals. Among the interesting genes identified, the most-represented transcript, Basigin, is known to be important for sensitivity to odors in *Drosophila* [49]. The interaction of the protein encoded by this gene with Apolipo protein D, a carrier for lipophilic molecules, has been documented [50]. Moreover, aquaporins, which have a suggested role in olfaction [51], were identified in the *E. materna* antennal transcriptome.

In addition to the olfactory-related transcripts and pathways, the identification of other interesting features specific to the FSM would help understand why some of the olfactory genes were found in unusual numbers. Therefore, the presence of transposon-related transcripts was explored in the dataset generated. The occurrence of transposable elements has been reported in several insects [52] and, not surprisingly, transcripts homologous to such elements were also identified in *E. materna*. PiggyBac and pogo-like elements were identified in the *E. materna* antennal transcriptome, which have also been reported in moths such as *P. xylostella* [53] and also some butterflies, including *Papilio* [54]. Surprisingly, transposable element-related results have not been reported in the transcriptomes published earlier for other fruit-sucking moths. The noteworthy aspect of this finding is that they are expressed in the antennal tissue. The occurrence of these transposable elements could reflect the possibility of the horizontal transmission of genetic material occurring in these moths [55]. The horizontal transmission of genetic material by transposable elements has been linked to facilitate genome evolution in insects [56]. Whether the occurrence of these transposable elements in the genome could have facilitated the evolution of olfactory genes is a question that requires insightful studies and deeper understanding.

Data generated in this study were specifically categorized to explore the olfactory-related genes, and different classes were generated including ORs, OBPs, CSPs, SNPs, and IRs. Transcriptome profiles generated for other lepidopteran insects seem to have a lot in common with *E. materna*. The antennal transcriptome of another fruit-sucking moth, *O. emarginata*, was released recently and identified 35 ORs, 41 OBPs, and 20 CSPs [28], against 178 ORs, 16 OBPs, and 17 CSPs in *E. materna*. Surprisingly, fewer ORs and a higher number of OBPs were identified in *O. emarginata*, another Noctuidae fruit-sucking moth, compared to *E. materna*. The number of *E. materna* ORs recognized in this study (178) was significantly more than those identified in other lepidopterans, viz., 46 and 59 in *C. punctiferalis* [21,57], 33 in *Dendrolimus kikuchii* G. [58], 47 in *H. armigera* [59], 39 in *Sesamia inferens* W. [60], and 72 in *B. mori* [61]. The OBPs, SNPs, and CSPs identified here were comparable to those identified in closely related moths [20,21,24,57–59,62–66]. To compare how the olfactory-related genes of *E. materna* fared with respect to other closely related insects/moths, phylogenetic analysis was carried out for every class using data available from public resources. The phylogenetic trees generated suggested the association of the

*E. materna* protein sequences with closely related moth species of the orders Noctuids and Cambridae. The ORs showed a stark difference in numbers, and those in *E. materna* were found to be associated with *Spodoptera litturis*' OR sequences. The phylogenetic analysis does provide some information about the closeness of *E. materna* to certain insect species, although nothing conclusive can be established in this regard since more detailed sequence information will be needed for the closely associated datasets. The trees generated for each of the olfactory genes show some unique clusters that have good support values and provide hints on the relatedness of the FSM to other moth species, and also provide insights on how different the moth is compared to other closely related moths.

However, the identification of several specialized ORs in the *E. materna* transcriptome could have a bearing on the behavioral patterns such as its unique feeding and egg laying behaviors. To assess this further, an explorative analysis of the OR sequences was carried out to look for any signatures of RNA-editing events that have been reported in the chemosensory proteins of *B. mori* [67]. Studies of *Drosophila* and other insects involving nAChRs suggest the involvement of RNA editing and highlight its evolutionary significance [68,69]. An RNA-editing mechanism involving the A-I modification has been shown to be important for the proper functioning of the nervous system and its integrity in *Drosophila* [32]. More recent studies with honey bees suggests the role of A-I RNA editing in adaptation and convergent evolution [70]. The results presented here for the A-I editing sites in FSMORs suggest a similar possibility with regard to odorant reception in *E. materna*. The RNA-editing predictions obtained for FSMORs using the AIRLINER and InosinePredict tools suggest a lot of modifications that occur in these sequences that could possibly result in functional gains. The confidence levels for about 600 sites among the 125 ORs tested was above 75%, and many of the sites overlapped in the results from both prediction tools. As a consequence, the resultant translated proteins could, therefore, have functional diversity and may help the insect specialize. Therefore, the occurrence of several ORs in *E. materna* could possibly be a result of RNA editing.

## 5. Conclusions

The current study is the first to report the antennal transcriptome profile of the fruit-sucking moth *E. materna* and we successfully identified 16 OBPs, four PBPs, 17 CSPs, 178 ORs, and other related transcripts including five SNMPs. Phylogenies constructed for the different categories of genes showed that *E. materna* has evolved olfactory-related genes in different clusters and showed divergence, even to closely related species, possibly reflecting its unique feeding and egg-laying habits. Despite the confidence/support values for some of the clusters being low, there were other prominent clusters that suggest that the moth has evolved to accommodate for its unique feeding habits and life cycle. The RNA-editing events identified also provide support in this regard. The transcriptome data provide information about the actively expressed genes in the antennae, which could be used for knockdown studies for further characterization and also to design customized control measures for other related fruit-sucking moths such as *E. phalonia*, *E. salamina*, and *E. homaena*. The data presented in this study show evidence, for the first time, for RNA editing in FSMORs. The RNA-editing site data highlight the fact that the occurrence of multiple ORs in FSM could be a result of post-transcriptional RNA-editing events and thus might lead to protein diversity. Such an occurrence could be significant in insect evolution and pest adaptations. The data also provide evidence for the expression of mobile element transcripts in the antennal tissue. A detailed analysis of FSM genes and other related moths will provide further insights into this aspect. The transcriptome of *E. materna* has helped establish a unique and useful resource of information that can support molecular efforts such as RNAi to design long-term solutions to control it. It would be interesting to look at the genes responsible for the oviposition and feeding behavior in female moths to understand the choice of host differentiation. Though the STRING-based network only provides a glimpse of the possibilities, the identification of these interactions, nevertheless, provides important information.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13071207/s1>, Figure S1a: The size distribution of the assembled unigenes from *Eudocima materna* L. male and female antennal transcriptome; Figure S1b: Top species distribution of unigenes in *E. materna* transcriptome; Figure S2: Gene ontology (GO) classification for the unigenes in male and female *E. materna* antennal transcriptome; Figure S3: Cluster of Orthologous groups (COG) classification. In total, 8559 unigenes were grouped into 25 COG classification Figure S4: KEGG orthology classification of unigenes, were mapped into five categories i.e., Cellular Process, Environmental Information Processing, Genetic Information Processing, Human diseases, Metabolism and Organismal systems; Table S1: Unigenes of candidate odorant binding proteins in *Eudocima materna* L. antennae transcriptome; Table S2: Unigenes of candidate pheromone binding proteins in *Eudocima materna* L. antennae; Table S3: Unigenes of candidate odorant receptors in *Eudocima materna* L. antennae transcriptome; Table S4: Unigenes of candidate chemo sensory proteins in *Eudocima materna* L. antennae transcriptome; Table S5: Unigenes of candidate sensory neuron membrane proteins in *Eudocima materna* L. antennae transcriptome; Table S6: Unigenes of candidate Antennal binding proteins in *Eudocima materna* L. antennae transcriptome; Table S7: Transcriptome annotations and raw counts; Table S8: Results of Airliner; Table S9: Results of Inosine Predict.

**Author Contributions:** The project was designed by K.J.P.D. and M.V. and the work was executed by M.V. and G.K. The analysis and interpretation of the data were carried out by M.V. and partly by G.K. The specific contributions are as follows: Conceptualization, K.J.P.D. and M.V.; methodology, M.V. and G.K.; software, M.V.; validation, M.V. and G.K. formal analysis, M.V. and G.K.; investigation, K.J.P.D. and M.V.; resources, K.J.P.D.; data curation, M.V. and G.K.; writing—original draft preparation, M.V. and G.K.; writing—review and editing, M.V., K.J.P.D. and G.K.; visualization, M.V.; supervision, K.J.P.D.; project administration, K.J.P.D.; funding acquisition, K.J.P.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** K.J.P.D. was supported by the Indian Council of Agricultural Research, New Delhi, India through National Fellow Project. The funders had no role in the study design, data collection, analysis, decision to publish, or the preparation of manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study has been deposited at the NCBI under the SRA accession PRJNA658216. This Transcriptome Shotgun Assembly project was deposited at DDBJ/EMBL/GenBank under the accession GIUU00000000.

**Acknowledgments:** The project was designed by K.J.P.D. and M.V. and the work was executed by M.V. and G.K. The analysis and interpretation of the data were carried out by M.V. and partly by G.K., M.V. and G.K. drafted the manuscript. K.J.P.D. and M.V. edited the manuscript. K.J.P.D. provided the funds, logistics, and lab support for this study. The authors would like to thank Rajshekhar Reddy for his input on the bioinformatic analysis of the data. The efforts of Sagar J in collecting the FSMs and maintaining the cultures are sincerely appreciated. The authors acknowledge financial support from the Indian Council of Agricultural Research through the ICAR National Fellow Project and infrastructural support by IIHR.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

1. Mote, U.; Tambe, A.; Patil, C. Observations on Incidence and Extent of Damage of Fruit Sucking Moths on Pomegranate Fruits. *J. Maharashtra Agric. Univ.* **1991**, *16*, 438–439.
2. Fay, H.C. Evolutionary and Taxonomic Relationships between Fruit-Piercing Moths and the Menispermaceae. *Aust. Syst. Bot.* **1996**, *9*, 227–233. [[CrossRef](#)]
3. Denton, G.; Muniappan, R.; Marutani, M.; McConnell, J.; Lali, T. *Biology and Natural Enemies of the Fruit-Piercing Moth Othreis Fullonia (Lepidoptera: Noctuidae) from Guam*; Research Extension Series-College of Tropical Agriculture and Human Resources; University of Hawaii, Cooperative Extension Service (USA): Honolulu, HI, USA, 1991.
4. Bhumannavar, B.S. Studies on Fruit Piercing Moths (Lepidoptera: Noctuidae)-Species Composition, Biology and Natural Enemies. Ph.D. Thesis, University of Agricultural Sciences, Bengaluru, India, 2000.
5. Sands, D.; Liebrechts, W.; Broe, R. Biological Control of the Fruit Piercing Moth, *Othreis Fullonia* (Clerck) (Lepidoptera: Noctuidae) in the Pacific. *Micronesica* **1993**, *4*, 25–31.

6. Bhumannavar, B.; Viraktamath, C. Biology, Ecology and Management of Fruit Piercing Moths (Lepidoptera: Noctuidae). *Pest Manag. Hortic. Ecosyst.* **2012**, *18*, 1–18.
7. Pandey, A.; Pandey, A.; Singh, M. Antifeeding, Repellent and Insecticidal Efficacy of Plant Products Against *Helicoverpa Armigera*. *Ann. Plant Prot. Sci.* **2010**, *18*, 304–306.
8. Field, L.; Pickett, J.; Wadhams, L. Molecular Studies in Insect Olfaction. *Insect Mol. Biol.* **2000**, *9*, 545–551. [[CrossRef](#)]
9. Pitts, R.J.; Rinker, D.C.; Jones, P.L.; Rokas, A.; Zwiebel, L.J. Transcriptome Profiling of Chemosensory Appendages in the Malaria Vector *Anopheles Gambiae* Reveals Tissue- and Sex-Specific Signatures of Odor Coding. *BMC Genom.* **2011**, *12*, 271. [[CrossRef](#)]
10. Schneider, D. Insect Antennae. *Annu. Rev. Entomol.* **1964**, *9*, 103–122. [[CrossRef](#)]
11. Steinbrecht, R.A. Pore Structures in Insect Olfactory Sensilla: A Review of Data and Concepts. *Int. J. Insect Morphol. Embryol.* **1997**, *26*, 229–245.
12. Pelosi, P.; Zhou, J.; Ban, L.; Calvello, M. Soluble Proteins in Insect Chemical Communication. *Cell. Mol. Life Sci. CMLS* **2006**, *63*, 1658–1676. [[CrossRef](#)]
13. Pelosi, P.; Iovinella, I.; Felicioli, A.; Dani, F.R. Soluble Proteins of Chemical Communication: An Overview Across Arthropods. *Front. Physiol.* **2014**, *5*, 320. [[CrossRef](#)] [[PubMed](#)]
14. Pelosi, P.; Iovinella, I.; Zhu, J.; Wang, G.; Dani, F.R. Beyond Chemoreception: Diverse Tasks of Soluble Olfactory Proteins in Insects. *Biol. Rev.* **2018**, *93*, 184–200. [[CrossRef](#)] [[PubMed](#)]
15. Clyne, P.J.; Warr, C.G.; Freeman, M.R.; Lessing, D.; Kim, J.; Carlson, J.R. A Novel Family of Divergent Seven-Transmembrane Proteins: Candidate Odorant Receptors in *Drosophila*. *Neuron* **1999**, *22*, 327–338. [[CrossRef](#)]
16. Leal, W.S. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annu. Rev. Entomol.* **2013**, *58*, 373–391. [[CrossRef](#)] [[PubMed](#)]
17. Wicher, D. Olfactory signaling in insects. In *Progress in Molecular Biology and Translational Science*; Anonymous; Elsevier: Amsterdam, The Netherlands, 2015; Volume 130, pp. 37–54.
18. Vogt, R.G.; Riddiford, L.M. Pheromone Binding and Inactivation by Moth Antennae. *Nature* **1981**, *293*, 161. [[CrossRef](#)] [[PubMed](#)]
19. Vogt, R.G.; Prestwich, G.D.; Lerner, M.R. Odorant-binding-protein Subfamilies Associate with Distinct Classes of Olfactory Receptor Neurons in Insects. *J. Neurobiol.* **1991**, *22*, 74–84. [[CrossRef](#)]
20. Depan Cao, Y.L.; Wei, J.; Liao, X.; Walker, W.B.; Li, J.; Wang, G. Identification of Candidate Olfactory Genes in *Chilo Suppressalis* by Antennal Transcriptome Analysis. *Int. J. Biol. Sci.* **2014**, *10*, 846.
21. Jia, X.; Wang, H.; Yan, Z.; Zhang, M.; Wei, C.; Qin, X.; Ji, W.; Falabella, P.; Du, Y. Antennal Transcriptome and Differential Expression of Olfactory Genes in the Yellow Peach Moth, *Conogethes Punctiferalis* (Lepidoptera: Crambidae). *Sci. Rep.* **2016**, *6*, 29067. [[CrossRef](#)]
22. Jia, X.; Zhang, X.; Liu, H.; Wang, R.; Zhang, T. Identification of Chemosensory Genes from the Antennal Transcriptome of Indian Meal Moth *Plodia interpunctella*. *PLoS ONE* **2018**, *13*, e0189889. [[CrossRef](#)]
23. Du, L.; Liu, Y.; Zhang, J.; Gao, X.; Wang, B.; Wang, G. Identification and Characterization of Chemosensory Genes in the Antennal Transcriptome of *Spodoptera Exigua*. *Comp. Biochem. Physiol. Part D Genom. Proteom.* **2018**, *27*, 54–65.
24. Liu, N.; Zhang, T.; Ye, Z.; Li, F.; Dong, S. Identification and Characterization of Candidate Chemosensory Gene Families from *Spodoptera Exigua* Developmental Transcriptomes. *Int. J. Biol. Sci.* **2015**, *11*, 1036. [[CrossRef](#)] [[PubMed](#)]
25. Zeng, F.; Zhao, Z.; Yan, M.; Zhou, W.; Zhang, Z.; Zhang, A.; Lu, Z.; Wang, M. Identification and Comparative Expression Profiles of Chemoreception Genes Revealed from Major Chemoreception Organs of the Rice Leaf Folder, *Cnaphalocrocis Medinalis* (Lepidoptera: Pyralidae). *PLoS ONE* **2015**, *10*, e0144267. [[CrossRef](#)] [[PubMed](#)]
26. Liu, S.; Wang, W.; Zhang, Y.; Zhang, B.; Rao, X.; Liu, X.; Wang, D.; Li, S. Transcriptome Sequencing Reveals Abundant Olfactory Genes in the Antennae of the Rice Leaf folder, *Cnaphalocrocis Medinalis* (Lepidoptera: Pyralidae). *Entomol. Sci.* **2017**, *20*, 177–188. [[CrossRef](#)]
27. Yang, S.; Cao, D.; Wang, G.; Liu, Y. Identification of Genes Involved in Chemoreception in *Plutella xylostella* by Antennal Transcriptome Analysis. *Sci. Rep.* **2017**, *7*, 11941. [[CrossRef](#)]
28. Feng, B.; Guo, Q.; Zheng, K.; Qin, Y.; Du, Y. Antennal Transcriptome Analysis of the Piercing Moth *Oraesia Emarginata* (Lepidoptera: Noctuidae). *PLoS ONE* **2017**, *12*, e0179433. [[CrossRef](#)]
29. Vieira, F.G.; Sánchez-Gracia, A.; Rozas, J. Comparative Genomic Analysis of the Odorant-Binding Protein Family in 12 *Drosophila* Genomes: Purifying Selection and Birth-and-Death Evolution. *Genome Biol.* **2007**, *8*, R235. [[CrossRef](#)]
30. Tonkin, L.A.; Saccomanno, L.; Morse, D.P.; Brodigan, T.; Krause, M.; Bass, B.L. RNA Editing by ADARs is Important for Normal Behavior in *Caenorhabditis Elegans*. *EMBO J.* **2002**, *21*, 6025–6035. [[CrossRef](#)]
31. Hoopengardner, B.; Bhalla, T.; Staber, C.; Reenan, R. Nervous System Targets of RNA Editing Identified by Comparative Genomics. *Science* **2003**, *301*, 832–836. [[CrossRef](#)]
32. Palladino, M.J.; Keegan, L.P.; O'Connell, M.A.; Reenan, R.A. A-to-I Pre-mRNA Editing in *Drosophila* is Primarily Involved in Adult Nervous System Function and Integrity. *Cell* **2000**, *102*, 437–449. [[CrossRef](#)]
33. Bass, B.L. RNA Editing by Adenosine Deaminases that Act on RNA. *Annu. Rev. Biochem.* **2002**, *71*, 817–846. [[CrossRef](#)]
34. Maas, S.; Rich, A.; Nishikura, K. A-to-I RNA Editing: Recent News and Residual Mysteries. *J. Biol. Chem.* **2003**, *278*, 1391–1394. [[CrossRef](#)] [[PubMed](#)]
35. Porath, H.T.; Hazan, E.; Shpigler, H.; Cohen, M.; Band, M.; Ben-Shahar, Y.; Levanon, E.Y.; Eisenberg, E.; Bloch, G. RNA Editing is Abundant and Correlates with Task Performance in a Social Bumblebee. *Nat. Commun.* **2019**, *10*, 1605. [[CrossRef](#)] [[PubMed](#)]

36. Ewels, P.; Magnusson, M.; Lundin, S.; Källér, M. MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report. *Bioinformatics* **2016**, *32*, 3047–3048. [[CrossRef](#)] [[PubMed](#)]
37. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
38. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. *Nat. Biotechnol.* **2011**, *29*, 644. [[CrossRef](#)]
39. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the Next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)]
40. Moriya, Y.; Itoh, M.; Okuda, S.; Yoshizawa, A.C.; Kanehisa, M. KAAS: An Automatic Genome Annotation and Pathway Reconstruction Server. *Nucleic Acids Res.* **2007**, *35*, W182–W185. [[CrossRef](#)]
41. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115–D119. [[CrossRef](#)]
42. Maere, S.; Heymans, K.; Kuiper, M. BiNGO: A Cytoscape Plugin to Assess Overrepresentation of Gene Ontology Categories in Biological Networks. *Bioinformatics* **2005**, *21*, 3448–3449. [[CrossRef](#)]
43. Ye, J.; Fang, L.; Zheng, H.; Zhang, Y.; Chen, J.; Zhang, Z.; Wang, J.; Li, S.; Li, R.; Bolund, L. WEGO: A Web Tool for Plotting GO Annotations. *Nucleic Acids Res.* **2006**, *34*, W293–W297. [[CrossRef](#)]
44. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
45. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis Across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
46. Von Mering, C.; Jensen, L.J.; Snel, B.; Hooper, S.D.; Krupp, M.; Foglierini, M.; Jouffre, N.; Huynen, M.A.; Bork, P. STRING: Known and Predicted Protein–protein Associations, Integrated and Transferred Across Organisms. *Nucleic Acids Res.* **2005**, *33*, D433–D437. [[CrossRef](#)] [[PubMed](#)]
47. Nigita, G.; Alaimo, S.; Ferro, A.; Giugno, R.; Pulvirenti, A. Knowledge in the Investigation of A-to-I RNA Editing Signals. *Front. Bioeng. Biotechnol.* **2015**, *3*, 18. [[CrossRef](#)]
48. Eggington, J.M.; Greene, T.; Bass, B.L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2*, 319. [[CrossRef](#)]
49. Curtin, K.D.; Meinertzhagen, I.A.; Wyman, R.J. Basigin (EMMPRIN/CD147) Interacts with Integrin to Affect Cellular Architecture. *J. Cell. Sci.* **2005**, *118*, 2649–2660. [[CrossRef](#)]
50. Najyb, O.; Brissette, L.; Rassart, E. Apolipoprotein D Internalization is a Basigin-Dependent Mechanism. *J. Biol. Chem.* **2015**, *290*, 16077–16087. [[CrossRef](#)]
51. Sørbø, J.G.; Moe, S.E.; Holen, T. Early Upregulation in Nasal Epithelium and Strong Expression in Olfactory Bulb Glomeruli Suggest a Role for Aquaporin-4 in Olfaction. *FEBS Lett.* **2007**, *581*, 4884–4890.
52. Gilbert, C.; Peccoud, J.; Cordaux, R. Transposable Elements and the Evolution of Insects. *Annu. Rev. Entomol.* **2021**, *66*, 355–372. [[CrossRef](#)]
53. Martins, S.; Naish, N.; Walker, A.; Morrison, N.; Scaife, S.; Fu, G.; Dafa’Alla, T.; Alphey, L. Germline Transformation of the Diamondback Moth, *Plutella xylostella* L., using the piggyBac Transposable Element. *Insect Mol. Biol.* **2012**, *21*, 414–421. [[CrossRef](#)]
54. Talla, V.; Suh, A.; Kalsoom, F.; Dincă, V.; Vila, R.; Friberg, M.; Wiklund, C.; Backström, N. Rapid Increase in Genome Size as a Consequence of Transposable Element Hyperactivity in Wood-White (Leptidea) Butterflies. *Genome Biol. Evol.* **2017**, *9*, 2491–2505. [[CrossRef](#)] [[PubMed](#)]
55. Peccoud, J.; Loiseaut, V.; Cordaux, R.; Gilbert, C. Massive Horizontal Transfer of Transposable Elements in Insects. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 4721–4726. [[CrossRef](#)] [[PubMed](#)]
56. Schrader, L.; Kim, J.W.; Ence, D.; Zimin, A.; Klein, A.; Wyschetzki, K.; Weichselgartner, T.; Kemena, C.; Stökl, J.; Schultner, E. Transposable Element Islands Facilitate Adaptation to Novel Environments in an Invasive Species. *Nat. Commun.* **2014**, *5*, 5495. [[CrossRef](#)] [[PubMed](#)]
57. Xiao, W.; Yang, L.; Xu, Z.; He, L. Transcriptomics and Identification of Candidate Chemosensory Genes in Antennae of *Conogethes punctiferalis* (Lepidoptera: Crambidae). *J. Asia-Pac. Entomol.* **2016**, *19*, 911–920. [[CrossRef](#)]
58. Zhang, S.; Zhang, Z.; Wang, H.; Kong, X. Antennal Transcriptome Analysis and Comparison of Olfactory Genes in Two Sympatric Defoliators, *Dendrolimus Houi* and *Dendrolimus Kikuchii* (Lepidoptera: Lasiocampidae). *Insect Biochem. Mol. Biol.* **2014**, *52*, 69–81. [[CrossRef](#)]
59. Liu, Y.; Gu, S.; Zhang, Y.; Guo, Y.; Wang, G. Candidate Olfaction Genes Identified within the *Helicoverpa armigera* Antennal Transcriptome. *PLoS ONE* **2012**, *7*, e48260. [[CrossRef](#)]
60. Zhang, Y.; Jin, J.; Jin, R.; Xia, Y.; Zhou, J.; Deng, J.; Dong, S. Differential Expression Patterns in Chemosensory and Non-Chemosensory Tissues of Putative Chemosensory Genes Identified by Transcriptome Analysis of Insect Pest the Purple Stem Borer *Sesamia inferens* (Walker). *PLoS ONE* **2013**, *8*, e69715.
61. Gong, D.; Zhang, H.; Zhao, P.; Xia, Q.; Xiang, Z. The Odorant Binding Protein Gene Family from the Genome of Silkworm, *Bombyx mori*. *BMC Genom.* **2009**, *10*, 332. [[CrossRef](#)]
62. Zhang, Y.; Qian, J.; Xu, J.; Zhu, X.; Li, M.; Xu, X.; Liu, C.; Xue, T.; Sun, L. Identification of Chemosensory Genes Based on the Transcriptomic Analysis of Six Different Chemosensory Organs in *Spodoptera exigua*. *Front. Physiol.* **2018**, *9*, 432. [[CrossRef](#)]

63. Zhang, T.; He, K.; Wang, Z. Transcriptome Comparison Analysis of *Ostrinia Furnacalis* in Four Developmental Stages. *Sci. Rep.* **2016**, *6*, 35008. [[CrossRef](#)]
64. Hu, P.; Tao, J.; Cui, M.; Gao, C.; Lu, P.; Luo, Y. Antennal Transcriptome Analysis and Expression Profiles of Odorant Binding Proteins in *Eogystia Hippophaecolus* (Lepidoptera: Cossidae). *BMC Genom.* **2016**, *17*, 651. [[CrossRef](#)] [[PubMed](#)]
65. Li, G.; Du, J.; Li, Y.; Wu, J. Identification of Putative Olfactory Genes from the Oriental Fruit Moth *Grapholita Molesta* Via an Antennal Transcriptome Analysis. *PLoS ONE* **2015**, *10*, e0142193. [[CrossRef](#)] [[PubMed](#)]
66. Zhu, J.; Ze, S.; Yang, B. Identification and Expression Profiling of Six Chemosensory Protein Genes in the Beet Armyworm, *Spodoptera Exigua*. *J. Asia-Pac. Entomol.* **2015**, *18*, 61–66. [[CrossRef](#)]
67. Xuan, N.; Bu, X.; Liu, Y.Y.; Yang, X.; Liu, G.X.; Fan, Z.X.; Bi, Y.P.; Yang, L.Q.; Lou, Q.N.; Rajashekar, B. Molecular Evidence of RNA Editing in *Bombyx* Chemosensory Protein Family. *PLoS ONE* **2014**, *9*, e86932. [[CrossRef](#)]
68. Grauso, M.; Reenan, R.; Culetto, E.; Sattelle, D. Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, D A 5, D A 6 and D A 7, in *Drosophila Melanogaster* Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing. *Genetics* **2002**, *160*, 1519–1533. [[CrossRef](#)]
69. Jin, Y.; Tian, N.; Cao, J.; Liang, J.; Yang, Z.; Lv, J. RNA Editing and Alternative Splicing of the Insect nAChR Subunit Alpha6 Transcript: Evolutionary Conservation, Divergence and Regulation. *BMC Evol. Biol.* **2007**, *7*, 98. [[CrossRef](#)]
70. Duan, Y.; Dou, S.; Porath, H.T.; Huang, J.; Eisenberg, E.; Lu, J. A-to-I RNA Editing in Honeybees Shows Signals of Adaptation and Convergent Evolution. *iScience* **2021**, *24*, 101983. [[CrossRef](#)]