

# SCIENTIFIC REPORTS



OPEN

## Sex-specific spatial and temporal gene expressions of Pheromone biosynthesis activating neuropeptide (PBAN) and binding proteins (PBP/OBP) in *Spoladea recurvalis*

Rajendran Senthilkumar & Ramasamy Srinivasan 

*Spoladea recurvalis* is one of the most destructive insect pests of amaranth, a leafy vegetable in both Asia and Africa. The present study characterized the pheromone biosynthesis-activating neuropeptide (DH-PBAN) and pheromone/odorant binding proteins in *S. recurvalis*. The open reading frame of 600 base pairs encodes a 200-amino acid protein possessing five neuropeptide motifs (DH, PBAN,  $\alpha$ -,  $\beta$ -, and  $\gamma$ -subesophageal ganglion neuropeptides) and shares a characteristic conserved C-terminal pentapeptide fragment FXPRL. The full-length genome of Spre-DH-PBAN was 4,295 bp in length and comprised of six exons interspersed by five introns. Sequence homology and phylogenetic analysis of Spre-DH-PBAN have high similarity to its homologs in Crambidae of Lepidopteran order. We quantitatively measured the relative expression level (qRT-PCR) of Spre-DH-PBAN gene, the binding proteins such as odorant binding proteins (OBPs) and pheromone binding protein (PBPs) at different developmental stages. The results confirmed their role in recognition and chemoreception of sex pheromone components, and they were distinct, tissue- and sex-specific. This is the first report on the molecular analysis of PBAN gene and binding proteins, which can improve the understanding of molecular mechanisms of growth, development, and reproductive behavior of *S. recurvalis*, and may become effective targets for controlling this insect.

Physiological processes such as diapause, mating, eclosion, reproduction and metamorphosis are determined by neurohormones, known as insect neuropeptides<sup>1</sup>. Female insects produce pheromones to entice sexual communication between male and female moths over long distances, which are recognized by a distinct neural arrangement in male moths, and eventually processed for mating<sup>2</sup>. Mostly, species-specific sex pheromones are produced in female pheromone glands (PG) located on the terminal abdominal segments in most of the moth species, which are released to attract conspecific males for mating<sup>2,3</sup>. The Pyrokinin (PK)/pheromone biosynthesis-activating neuropeptide (PBAN) helps in regulating the sex pheromone production in insects, especially lepidopteran species in particular<sup>4,5</sup>. Based on the studies of molecular characterization, DH-PBAN is a multigene family, encoded by five putative neuropeptides – PBAN, DH,  $\alpha$ -,  $\beta$ - and  $\gamma$ -subesophageal ganglion neuropeptides (SGNP). These are all well conserved C-terminal FXPRL amide motif, and are released through post-translational processing<sup>5,6</sup>. Although physiological functions such as melanization in moth<sup>7</sup>, regulation of pheromone biosynthesis<sup>8</sup>, and stimulation of embryonic diapause in insects like silkworm moth<sup>9</sup> were well documented, their functions were not aptly understood. The differing physiological functions, reported previously, were controlled by the wide distribution and expression of FXPRL family peptide encoded by PBAN<sup>10–12</sup>. PBAN serves to regulate the production and release of sex pheromones in female moths<sup>13</sup>. In *Helicoverpa zea*, a 33 amino acid C-terminal amidated peptide of PBAN was first identified<sup>8</sup>, whereas the determination of its amino acid

World Vegetable Center, Shanhua, Tainan, 74151, Taiwan. Correspondence and requests for materials should be addressed to R.S. (email: [srini.ramasamy@worldveg.org](mailto:srini.ramasamy@worldveg.org))

sequence was done by direct isolation and purification from the *Bombyx mori*'s head<sup>14</sup>. Later, PBAN encoding genes were identified in *Aedes aegypti*<sup>15</sup>, *Ostrinia nubilalis*<sup>16</sup>, *Chlumetia transversa*<sup>17</sup>, *Maruca vitrata*<sup>18</sup>, *Manduca sexta*<sup>19</sup>, and *B. mori*<sup>20</sup>. As a result, DH-PBAN proteins have been reported for more than 30 insect species spanning four orders including Lepidoptera and Hymenoptera<sup>21</sup>.

The pheromone binding proteins (PBPs) penetrate into an aqueous layer of sensillum lymph, transported to olfactory receptor neurons, where signal transduction is initiated<sup>22–24</sup> and thus mediating perceptions of pheromones in the moths. In fact, PBPs are the sub-class of insect odorant binding proteins (OBPs)<sup>25</sup>. PBPs were initially discovered in Lepidoptera and later characterized in a series of insect orders including Diptera, Hymenoptera, Hemiptera and Coleoptera<sup>26,27</sup>. Studies on PBPs indicated that they concentrate odorants in the sensillum lymph, act as solubilizers and carriers of the hydrophobic pheromones, involves in odorant molecule deactivation, serve as co-factors in the activation of pheromone receptors, and protect the pheromone from enzymatic degradation<sup>28</sup>. In fact, absence of PBPs affects the sensitivity of the neuron to its odor ligand as reported previously<sup>29</sup>. OBP sequences possess conserved cysteine residues, and are classified based on their number and location such as “classic”, “plus-C”, “minus-C” and “duplex”<sup>30</sup>. However, most pattern of the cysteine residues in the amino acid sequences pose a signature for moth OBPs, and are differing<sup>31</sup>. Among the many OBP classes, the six-cysteine residues are the most pertinent in the process of PBPs folding<sup>26,32</sup>. Therefore, it is evident that the olfactory recognition system serves an important role in many processes like orientation, oviposition, mating, feeding, and in searching for hosts, rather than its primary role of ensuring reproduction and insect survival<sup>33</sup>.

The beet webworm, *Spoladea recurvalis* Fab. (Lepidoptera: Crambidae) mainly inhabits the tropical and sub-tropical areas of the world, where it causes severe damage to crops including bean, amaranth, cucurbits, beetroot, sweet potato, and eggplant<sup>34</sup>. It was shown that *S. recurvalis* can migrate long distances causing serious crop destruction in new locations<sup>35</sup>. Studies on host-plant interactions and bio-ecology of *S. recurvalis* showed serious consequences in amaranth<sup>36</sup>. Contrary to the use of chemical pesticides that pose serious risks to environment and human health, use of pheromones to interrupt the mating behavior is deemed to be the most appropriate component in integrated pest management (IPM) packages.

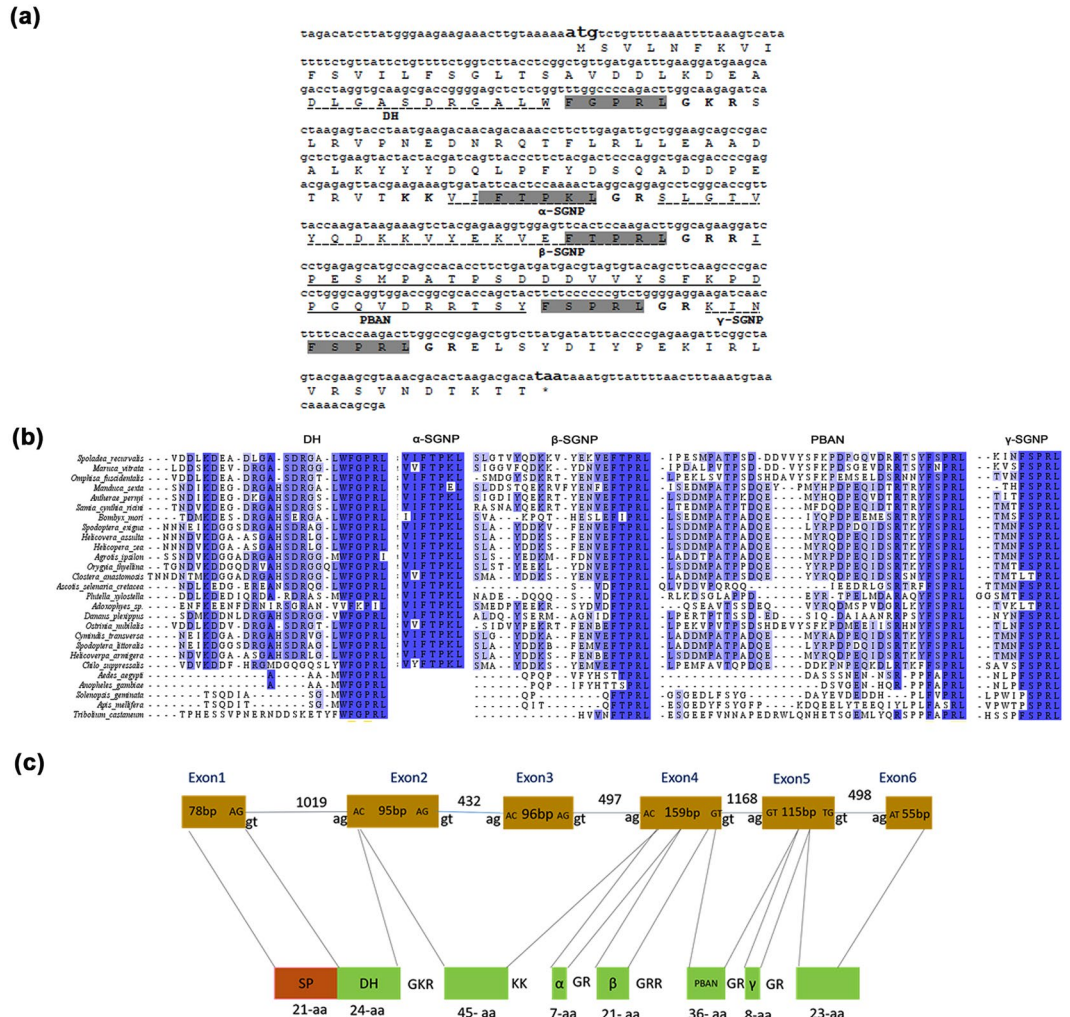
In light of this, we characterized the pheromone biosynthesis activating neuropeptide and binding proteins of *S. recurvalis* to perform the tissue-specific DNA expression analyzes in both male and female moths with the aim of using the proteins as primary molecular targets for designing and developing new pest management strategies in future studies.

## Results

**Structure of *Spoladea recurvalis* PBAN (Spre-DH-PBAN) Gene.** The full-length transcript of 672 bp, which comprised of the 34 bp 5' Untranslated region (5'UTR), 600 bp Open reading frame (ORF), and 38 bp 3' Untranslated region (3'UTR) was identified from the *S. recurvalis* transcriptome library. The 600 bp nucleotides in the ORF encode 200 amino acids, possessing five putative neuropeptides namely,  $\beta$ -SGNP, PBAN, DH,  $\alpha$ -SGNP, and  $\gamma$ -SGNP from 5' end to 3' end (Fig. 1a). The polypeptide that comprises 21 amino acids was predicted to be a putative signal peptide with cleavage site between first M1 and V22 amino acids. Figure 1b indicated that the Spre-DH-PBAN showed high similarities of 80% to its orthologs of Crambidae from Lepidoptera, and has the lowest similarities of 32% to the DH-PBAN proteins of the family Culicidae from Diptera. The amino acid sequences of individual neuropeptides such as  $\gamma$ -SGNP,  $\alpha$ -SGNP, DH,  $\beta$ -SGNP, PBAN and the full length DH-PBAN showed 71–100%, 85–100%, 52–100%, 56–100%, 40–78%, and 30–80% identities, respectively, with all the compared insect species (Table 1). The highest percentage identity was shown for  $\alpha$ -SGNP and  $\gamma$ -SGNP, followed by DH and  $\beta$ -SGNP. Schematic comparison of the genomic DNA and the translated protein structure of *S. recurvalis* DH-PBAN are shown in Fig. 1c. The putative signal peptide and the N-terminal portion DH peptide were translated from exon 1 and 2 (95 bp) and encoded 45 amino acids. Markedly, neuropeptides  $\beta$ -SGNP, partial PBAN, and  $\alpha$ -SGNP, which were translated from the exon 4 (153 bp) encoded 51 amino acids. The  $\gamma$ -SGNP and PBAN were translated from exon 5, encoding 37 amino acids that cover the C-terminal ends (Fig. 1c). The intron-exon of Spre-DH-PBAN gene was approximately 4,295 bp in length containing 6 exons interspersed by 5 introns. The grouping of the corresponding genomic architecture and cDNA were summarized in Fig. 2. Untranslated regions (UTR) of 33 nucleotides start from upstream of 5', the start site ATG of exon 1 (81 bp) (Fig. 2). All progressions of the exon/intron borders were in agreement with the consent GT/AG series to both the acceptor and donor sites of the gene splicing. The intron sizes range from 432–1,168 nucleotide structural components, except for the intron between exon 4 and exon 5 with the largest proportion of 1,168 nucleotides in length, spanning PBAN neuropeptide (Fig. 2).

**Phylogenetic Relationship of Spre-DH-PBAN.** The interspecific phylogenetic relationships based on the DH-PBAN ORFs of *S. recurvalis* and the 25 other available insect DH-PBAN peptide sequences derived from NCBI GenBank were shown in Fig. 3a. The phylogram based on JTT matrix model was consistent with the taxonomic classification of insects. *Spoladea recurvalis* clustered together with *M. vitrata*, *C. suppressalis* and *O. fuscidentalis* (Lepidoptera: Crambidae). The members of the order Lepidoptera form a separate clade from the Diptera (*A. aegypti*), Coleoptera (*Tribolium castaneum*) and Hymenoptera (*Apis mellifera*).

**Gene Expression of DH-PBAN during Developmental Stages.** Tissue-specific gene-expression profile of DH-PBAN was done using samples obtained from various life-stages of *S. recurvalis* using qRT-PCR. The abdomen of the 5<sup>th</sup> instar larva showed higher expression of PBAN compared to the other tissues, for instance head and thorax (Fig. 3b). Then, PBAN expression was found to be high in the head and thorax tissues of male and female pupae. However, only a basal level of expression was observed in the thorax tissues of male and female pupae irrespective of their ages. Furthermore, PBAN gene expression was noticeably higher in the female and male head of unmated adults, from the day 1 to 5, respectively. The PBAN gene expression was also found to



**Figure 1.** Molecular characterization of gene encoding pheromone biosynthesis activating neuropeptide (DH-PBAN) gene in *Spoladea recurvalis*. **(a)** The putative start codon (ATG) and stop codon (TAA) are indicated in bold font. Amino acids of five putative FXPRL-NH<sub>2</sub> peptides are underlined: diapause hormone (DH), α,β,γ-suboesophageal ganglion neuropeptide (SGNP) and PBAN conserved C-terminal pentapeptide sequences are highlighted in bold letters with grey background. **(b)** Multiple sequence alignment of consensus sequences of diapause hormone-pheromone biosynthesis activating neuropeptide (DH-PBAN) precursor from *Spoladea recurvalis* and other lepidopteran species. Percentage of conserved amino acids are highlighted in blue. **(c)** DH-PBAN Exons are represented at the top solid boxes with boundary sequences. Horizontal lines between exons are represents intronic regions with boundary sequences. FXPRL-NH<sub>2</sub> neuropeptides along with signal peptide are represented at the bottom solid box. GKR;KK;GR;GRR denotes endoproteolytic cleavage site; SP-signal peptide; DH-diapause hormone; SGNP neuropeptides.

be abundant during day 1, and the level of expression was consistently higher from day 3 to 5 of the male and female head tissues in mated adults (Fig. 3b). On contrary, mated adult thorax tissues showed moderate PBAN expression in both male and female of the same age. There was no significant gene expression noticed in abdomen tissues virgin and mated adults of all the ages (Fig. 3b).

**Pheromone/Odorant Binding Proteins.** Based on the number and conserved cysteine location, *S. recurvalis* OBP11 and GOB1 belong to sub family “plus-C” class, whereas partial sequence of PBP is a typical odorant binding protein and OB4 (JHP) belongs to Takeout protein (Table 2). The open reading frame (ORF) of OBP11 was 471 bp, which encoded 157 amino acid residues, with predicted signal peptide cleavage site between 19 and residue 20 (Fig. 4a). Sequence alignment of Spre-OBP11 with corresponding proteins from other insects was listed in Supplementary Table S1. Among them, Spre-OBP11 showed highest expression primarily in thorax and abdomen of 4<sup>th</sup> instar larvae and higher in head of 5<sup>th</sup> instar larva, whereas moderately lower in thorax and abdomen of 5<sup>th</sup> instar larvae (Fig. 4b). Abundant gene expression was also found in thorax of male and female pupae of all ages, whereas its expression was moderate in male and female pupae head tissues. Furthermore, similar expression patterns were noticeably observed in unmated adult thorax tissue of all ages and moderately lower expression of Spre-OBP11 was observed in unmated male and female adult of antennal tissue. Spre-OBP11

Species	Neuropeptide (percentage % homology)						Accession Number
	DH	$\alpha$ -SGNP	$\beta$ -SGNP	PBAN	$\gamma$ -SGNP	FLDH-PBAN	
<i>Spoladea recurvalis</i>	—	—	—	—	—	—	
<i>Maruca vitrata</i>	82.3	85	71	77.7	75	80.2	AFX71575
<i>Omphisa fuscidentalis</i>	91.6	100	75	58	85	74.8	AFP87384
<i>Manduca sexta</i>	79.1	85	68	62	100	63.4	AAO18192
<i>Antheraea pernyi</i>	75	100	66	59	85	63.64	AAR17699
<i>Samia cynthia ricini</i>	75	100	75	55.5	—	62.7	AAP41132
<i>Bombyx mori</i>	78.2	85	66	50	100	62.6	AAB24327
<i>Spodoptera exigua</i>	51.7	100	75	52.7	100	64.3	AAT64424
<i>Helicoverpa assulta</i>	66.6	100	68	50	100	65.7	AAC64293
<i>Helicoverpa zea</i>	66.6	100	68	50	—	65.7	AAA20661
<i>Agrotis ipsilon</i>	66.6	100	62	50	100	62.5	CAA08774
<i>Orgyia thyellina</i>	61.5	100	62	52.7	100	63.6	BAE94185
<i>Clostera anastomosis</i>	73.9	85	75	50	—	60.2	ABR04093
<i>Ascotis selenaria cretacea</i>	55.5	100	85	40	100	53.8	BAF64458
<i>Plutella xylostella</i>	64	100	85	60	100	53.2	AAX99220
<i>Adoxophyes sp</i>	—	—	56	44.1	—	53.8	AAK72980
<i>Danaus plexippus</i>	64	100	71	47.2	100	52	OWR44299
<i>Ostrinia nubilalis</i>	83.3	85	58	59.4	100	70	AOY34029
<i>Chlumetia transversa</i>	66.6	100	66	50	100	63	AIY72749
<i>Spodoptera littoralis</i>	51.7	100	68	52.7	100	63	AKT95050
<i>Helicoverpa armigera</i>	66.6	100	62	50	100	65	AF492474
<i>Chilo suppressalis</i>	100	85	75	66.6	71	53	ALM30314
<i>Aedes aegypti</i>	87.5	—	100	—	100	32	XP001662212
<i>Anopheles gambiae</i>	87.5	—	—	—	100	30	XP307885
<i>Solenopsis geminata</i>	100	—	75	—	—	33	ADI88479
<i>Apis mellifera</i>	85.7	—	83	—	100	—	A8CL69
<i>Tribolium castaneum</i>	100	—	85	—	100	30	EFA11568

**Table 1.** Percentage of amino acid sequences homology of five neuropeptide and full-length DH-PBAN homolog to *Spoladea recurvalis* inferred from species of four orders (Lepidoptera, Coleoptera, Diptera and Hymenoptera).

was expressed higher in head and thorax of mated male on day 3 and 5, in the female abdomen on day 1 and 3, and female thorax on day 5 (Fig. 4b). However, there was no obvious Spre-OBP11 gene expression noticeable in antenna of mated male and female of all ages. Spre-OBP11 showed high amino acid identity with two Crambidae moths, *Cnaphalocrocis medinalis* (85%) and *Chilo suppressalis* (76%), followed by the Nymphalidae butterfly, *Danaus plexippus* (76%). Phylogenetic analysis showed that *S. recurvalis* had a high homology to *C. medinalis* belonging to the Plus-C OBP sub-family (Fig. 4c).

The full-length sequence of GOBP1 was 498 bp, encoded by 166 amino acids. Predicted signal peptide was found at the N-terminal position between 19 and 20 (Fig. 5a). The deduced amino acids were aligned, revealing conserved cysteine residues, and it belonged to class “Plus-C” sub family (Fig. 5a). We observed that Spre-GOBP1 showed significantly higher expression particularly in the head of 5<sup>th</sup> instar larvae and pupae of all ages (Fig. 5b). We observed significant Spre-GOBP1 gene expression in unmated and mated adult antennal and head tissues of male and female of all ages (Fig. 5b). Spre-GOBP1 has high amino acid sequence identity with three other Crambidae moths, *C. medinalis* (82%), *Conogethes punctiferalis* (77%) and *M. vitrata* (70%) (Supplementary Table S2). As indicated by the phylogenetic tree based on amino acid identity percentage, *C. medinalis*, *C. punctiferalis* and *M. vitrata* clustered together forming a single clade (Fig. 5c). The partial sequence (177 bp) of Spre-PBP was aligned and the sequence identity typically characterizes an odorant binding protein (Fig. 6a). Amino acid sequence alignment revealed high sequence similarity conserved among three other Crambidae moths, *C. medinalis* (70%), *Diaphantia indica* (66%), and *M. vitrata* (64%) (Supplementary Table S3). Spre-PBP gene expression showed that the expression was found to be higher in head of 4<sup>th</sup> instar larvae, and the level of expression was further observed to be higher in thoracic and abdomen region of 5<sup>th</sup> instar larvae (Fig. 6b). Higher expression of Spre-PBP gene was significantly found in the head of male pupae on day 5 (Fig. 6b). Furthermore, Spre-PBP gene expression was significantly observed in antennal and head tissues of unmated and mated adults of all ages (Fig. 6b). Phylogenetic analysis was shown in Fig. 6c. Spre-PBP was compared with PBPs of other insect species and the results indicated that significantly high degree of conservation among lepidopterans evolved possibly from a common ancestor.

Similarly, the full-length open reading frame of Spre-OBP4 was 762 bp, encoding 253 amino acids. Predicted signal peptide and cleavage site position lies on 20 and 21 (Fig. 7a). Multiple sequence alignment and phylogenetic analysis showed that Spre-OBP4 has close homologues with two Crambidae moths, *C. medinalis* (86%), and

>*Spoladea recurvalis*\_PBAN Sequence 4295 BP; 1389 A; 826 C; 835 G; 1245 T; 0 other;

```

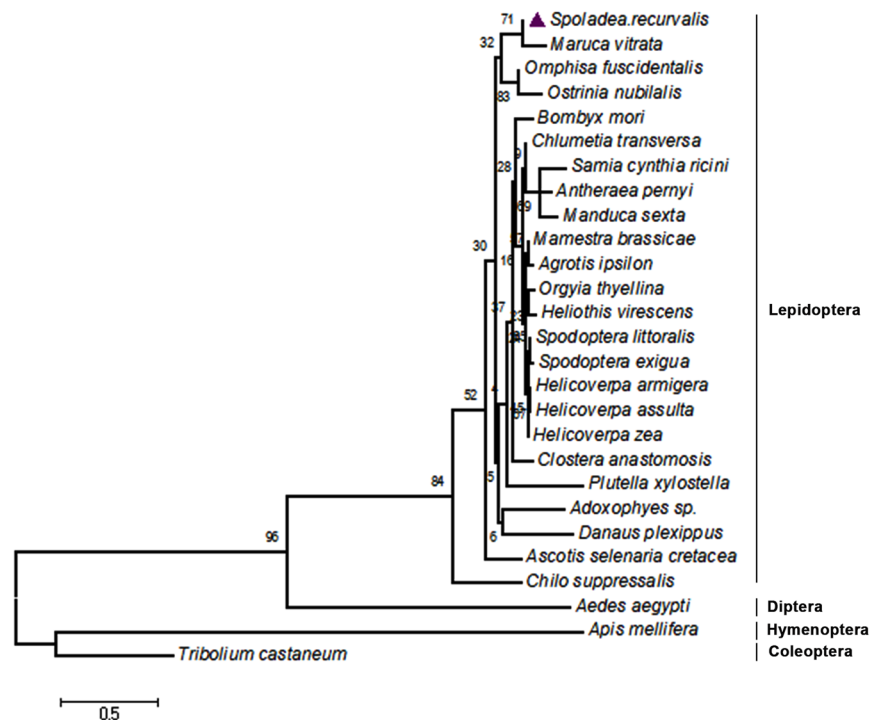
TAGACATCTT ATGGGAAGAA GAAACTTGTA AAAATGAAA GGTCTGTTTT ATTTTAAAGT 60
CATATTTTCT GTTATTTCTGT TTTCTGGTCT TACCTCGGCT GTTGATGATT TGAAGGTAAT 120
TGCATATTTT ATTTATTTAT TATGATGATG ATGATGATTG TATGATGTA TGTTCAGTTA 180
CTCCAATAAT ATCATCAGCA CGATTCAACG GAGATGGGCT CACCCCTGT CACGTGTGCA 240
AATTTCTGTAT TAAAAATGGT AGAATGGGGA TGGGAAGAAT ACGATATTTT TTGTCTCTCG 300
CTACCCCTCA GAGAAGAGAG GCTTGAATTT ATGTATGTAT TTTTATATGT TAACCATCCC 360
TTCATAGGAG TCCTATTTGA CATAAGCGTT CAGAACTTTA CATAAGTACG AGTTATACTC 420
GTATATTGAA GACACTAACA AACAAATGTC GATGATTTCT CGTGATCTCC CTAGCGTTCA 480
TCCCGGTCTA CGCCAGGTTT CGCTTTGACC ATGATCCTGG ATTCAGCGAG TCTGGCTATT 540
ACATTTAAAG GACTACCTAT CTGACCTTCG TCACCCATAG AAAGGAGAAA AAGACAAATT 600
GGATCATGTT TACACATCCA GAATAAAGCT AATAACTGAA TGTGCAGGTT GCCCAACGAT 660
GTTTAAACCAC ACCGAAAAGC CAGTTATGTT TTTACAAATC GTACCTATAT GTATAAGTTG 720
AAAAACAACCT CATTGATACG ATTTCGGTCCA TCTGGATTGG AACCCACGCC ACTCGTGAAA 780
GAGGCGACAT TCTTAGCCAC ACAACCCACA CGACTTAGTT TTTTCATGAT TCTCGTGAAT 840
TAAAAATAAAA ATGAATTCGT GTAAAATCTC ATACTAAGTT GCTAAAGCAC TTGTATTACA 900
GCTAATAGAC AGCAGAGGAA CAGAAATATA ATACTTGAAG ACATAGGTAT ATTACACCCA 960
GACATAGAAC AAACACAAAT GTCCATCACA TGATTTTTTT TCCCTGGCCG AAAAAAGCAA 1020
CGACGAAAAA AATTTGAACC CACGACCGGT AGCTTGCAAA GGCAAGGTCA CTACCTATTG 1080
CGCAACGAG CCCGTC AATT TGTATAATG TACAGTAAAT TATATATTTC TCAGGATGAA 1140
GCAGACCTAG GTGCAAGCGA CCGGGGAGCA CTCTGGTTCG GTCCAAAGCT TGGCAAGAGA 1200
TCTCTAAGAG TACCTAATGA AGACAACAGG TTAGTAAAAG AAAGAAATAA TTTTATTGTA 1260
AATAAATACGT ACTTTTGTCA TAACCATAAG ACAGATTTTC GAATTTATAC ATATACCTAC 1320
AATACAATTA AATATTTACC GATGAGATCA TAATCTCATC ACAAATGAA TGCTCTTATA 1380
CAAGTGTATT AATATATCTG GACAATGAAG TATCATGTAT AAATTTGTAC AATAGACATA 1440
CTATTCCTTC AATCACCTCG AATTACACAT ACACATCACG CACTTAAAC AAAAAATAAA 1500
AGAACTCTGT CATACCTAAA CTTATACACA AACACAAAT AAAGAAATAA AATAAATTAT 1560
ATAATTTTAC AGGGTGAACC GTATAAATAA ATATAAAAAA ATACCTAAGA AAATTAGCTT 1620
ATAAAGGCTT TCTACCAATT CATCTAAGTC AATTTGTTTCA CAGACAAACC TTCTTGAGAT 1680
TACTGGGAGC AGCCGACGCT CTGAAGTACT ACTACGATCA GCTACCCCTC TACGACTCCC 1740
AGGCTGACGA CCCAGAGACG AGAGGTAATA TTAATTACTA AAAGTAAAGT ATTACCTAAG 1800
TATCAGCCTA TACAGGGTGT ATCGTGAAC AGTTACATCC CTTCAGAAC GTATTGTACT 1860
CATCAGTATT AACAACTTTT GTTCTACAAC TATTTGAAAT TCGTGAAAAA AAATCCGACT 1920
TCCCATACAA AAGTGCTTAA GCGACTGATT CAGTTTCTAA GCATGAAAAG TCTAAAAAAA 1980
AGTTGTTTTG ATTGACTAAG TCGTAGAACA AAAGTTGTTT GTAGTGACAA GTACAAATACG 2040
AATAATATAA CGGGGATTA CGCGACGGTT AAATTTAAG TACATGATTA TGCATAATCA 2100
AGCAGTAAAG GCATAAATTT AACCGTCGCG TTAATCCCGG TTATATTTAT TGTAGTTATA 2160
ATCAACCGCG ATAATCTAAA GAACTATTTA AGTACAATAC GGTCCTGAAG GGATGCAGCT 2220
GACCAAAAAG GGTGGTCGTT ATGACTGTTT ATTTTCTACC CACAGTTACG AAGAAAGTGA 2280
TATTCCTACC AAAACTAGGC AGGAGCCTCG GCACCGTTTA CCAAGATAAG AAAGTCTACG 2340
AGAATGTGGA GTTTACTCCA AGACTTTGGCA GAAGAAATCCC TGAGAGCATG CCAGCCACAC 2400
CATCAGATGA TGACGTAGTG TAAGTACAGA GCAATCGCGC TAAAAATAAC GTGTCAACCA 2460
TAACTGAACC TTAGACGTTT CACATCTTTT TTGACATTTA GTATCGAGGT AGTGGTTTGA 2520
CAGCTAGCTG AAACGTAAC AAACCTGAAC TAGAACGTTT CTATAGCATT ACTGAATGGT 2580
TTAGTACCAC TAACGTGATA AAAAAAATAT GTCAAAAAAA TGTGAAACAT CTGAGGTTCA 2640
GTTACGGGGT ACACGTTATT TTTAGCGGGA TTTGCCCTCA GGCTCCTATT ATTTACCTTT 2700
TACGGTACCT ATGTTTGAAT ACCTTCAGGG GTATTTAAGT ATTAATTTTG GGTTCCTCCG 2760
TTGTCGGAAA TCCGGGACAC AATTGCTTAA GCTACTTAGT ACGGAGTCAG GTCAATGTAT 2820
GAAATGTTGG GGAATTTTTT TTTTCTAAG GGCCGGGGGG TTTCCACTAA CCCATCTCAC 2880
CCCGGGGGTT TTGTAAAACC CAACAAGGGA ATTTAACCTC TACAATTTT CACGTTTCAA 2940
CGGCAATAGG TTTACCACCG GTCTTTGGGG CAATTTCTGC ATTTGAAGGG GATCTGGGAA 3000
TGGAAAAACC ACAATTTTTT CTGTCTCTGT CTTCAAAGGA AAAAGGGGAG ATTTTATGTA 3060
TGACGACCTT ATAATAGCTT TTAATGGGAA TCCATATTTG CATAGGGTAC CAAAATTTCA 3120
TTTTGCACTA CTAATACTCG ACATTCAGGG GGATTACAAT TCAAAAGTCA AATTCCTCCC 3180
GTCTAGACCT CAAACTAGAT CCGCTCCGTG TATTTTCGGC ATAAAGACGG TCTGATCCAG 3240
TACATCTTAA TACAACCTCG GGTGAGTGAT CCGGTACATC TAATTGGAGG GTTGTAAATA 3300
CCAGTGGGCT GGACACTTCC CATTTCCTGA AATAGAGAGA GGAGTGATTT AGTATCTCAT 3360
TTGGGACGGG CACCAATFAA CATTCTGTAT AGAGAAGAGC AAGGTCCTGC GTTGGCGGATA 3420
AAATTATGAA AAAATGTTAT TTTTATACCT CCCGCTTAAG AGAAAAAGGG CCCAATCTTT 3480
TAATAATGAT GGTTCCTGTT TTTTGTATTG ACTACCCCTC AGGCACCGTC AGTGGGGTGT 3540
AATAATGATA AGTAGTATTA CTGTTGTTAA ACTCCAGGCT GTTCCAGGTA CAGCTTCAAG 3600
CCCGACCTTG GGCAGTTGGT CCGGCGCTCC AGCTACTTCT CCCCCGCTC TGGGGAGGAG 3660
ATCAACTTTT CACCGAGACT TGGCCGCGAG CTGTCTTATG GTGAGCACAA GTTTCTATAT 3720
AAAGAGGTTA GGGGACGTTT TACACCTTTT GCTGTCACTG TCACTGAAAA CTGTGCCGTG 3780
TGGTTTCCGG CACTTAAAAG AATTGAACCA ACCCATCTCT TCCCGTGGGT GTTGTAAAAG 3840
GCCACTAAGG GAAATCAACT ACAATAACAT CTACGGATTT GCACGTTTCA ACGGCGATGG 3900
GCTCACCACC TGTCACGTGT GCAATTTCTG TATTCAGGAA TGGTGAAATG TGGATGGAGG 3960
GATCGCAGTA TTCTTCTGTC TCTGCCACCC TCTTAGGGGA GAAAGCCGTG ATTTTATGTA 4020
TGTATGTGT TATGTATGTC ACTGAAGAAAT GTGAAGAATG TATGTAGCGT GGCAGACATT 4080
TTCAGGTTGA GCACAGTGA TGCACAAGCG GCACGGCAAC TGAATTTCTA TATATTCTTT 4140
CTACGGAAAC ATTAACACTG TAAAAGTAGG GAACTAAAAA AGTTTTTATT TTTTCAAGAT 4200
ATTTACCCCG AGAAGATTCG GCTAGTACGA AGCGTAAACG ACACTAAGAC GACATAATAA 4260
ATGTTATTTT AACTTTAAAT GTAACAAAAC AGCGA ★ 4295

```

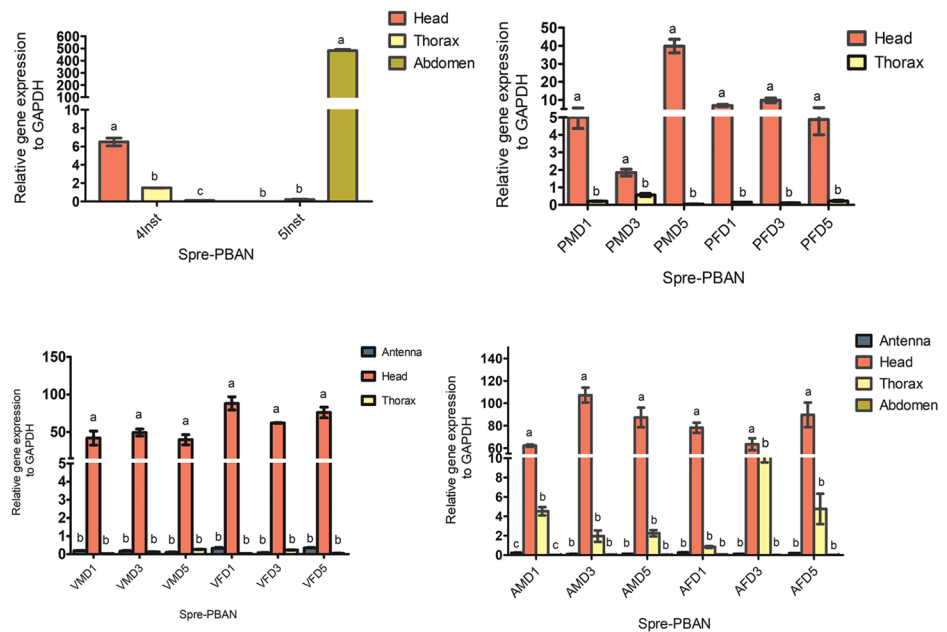
**Figure 2.** Exon-intron organization and complete nucleotide sequence of *Spoladea recurvalis* DH-PBAN gene. DNA sequences of 4.2 kb segment of DH-PBAN. Nucleotide sequences encoded by the exon are highlighted in blue highlights. Start codon indicates in box and stop codon indicated in ★.

*C. suppressalis* (79%), and a Pyralidae moth, *Amyelois transitella* (69%) (Supplementary Table S4). Spre-OBP4 sequence was similar to a juvenile hormone protein, which is a member of the odorant binding protein. Spre-OBP4 was primarily expressed in abdomen (Fig. 7b). The transcript levels were expressed in higher amounts in the head of pupae. More specifically, Spre-OBP4 was expressed in higher amounts in antenna of both unmated and mated adults of all ages (Fig. 7b). The resulting phylogenetic tree showed that Spre-OBP4 sequence clustered closely with the sequence of *C. medinalis* and *C. suppressalis* (Fig. 7c).

(a)



(b)



**Figure 3.** Phylogenetic analysis and Gene expression of PBAN during different developmental stages. (a) Maximum likelihood phylogeny tree was constructed using known and putative PBAN amino acid sequences inferred using four insect orders determined by using MEGA 6.0 with default settings model JTT + T inferred from 1000 replicates. (b) PBAN gene expression in the development of *S. recurvalis*. Total RNA was prepared at day 1, day 3 and day 5 of each stages Larvae, Pupae, Virgin (male and female), Adult (mated male and female) subjected to qRT-PCR. The results are shown as the Spre-PBAN gene expression level was normalized related to each development stages. GAPDH was used as internal control. Bar represents the  $\pm$  standard error. Different letters (a,b,c) shows the significant differences ( $P < 0.05$ ). (Inst-Instar; PMD-Pupae male day; PFD-Pupae female day; VMD-virgin male day; VFD-virgin female day; AMD-Mated adult male day; AFD-Mated adult female day).

<i>Spoladea recurvalis</i>	Subfamily	Signal Peptide	Sequence length	Protein core region
OBP11	Plus-C	Yes	471	C1-X <sub>8</sub> -C2-X <sub>4</sub> -C3-X <sub>43</sub> -C <sub>4</sub> -X <sub>12</sub> -C5-X <sub>7</sub> -C6-X <sub>8</sub> -C7-X <sub>6</sub> -C8
GOBP1	Plus-C	Yes	498	C1-X <sub>30</sub> -C2-X <sub>3</sub> -C3-X <sub>42</sub> -C <sub>4</sub> -X <sub>10</sub> -C5-X <sub>8</sub> -C6-X <sub>4</sub> -C7
PBP <sup>p</sup>	Classic OBP	—	177	C1-X <sub>30</sub> -C2-X <sub>3</sub> -C3-x <sub>7</sub>
OBP4 (JHP)	TOP	Yes	762	C1-X-C2

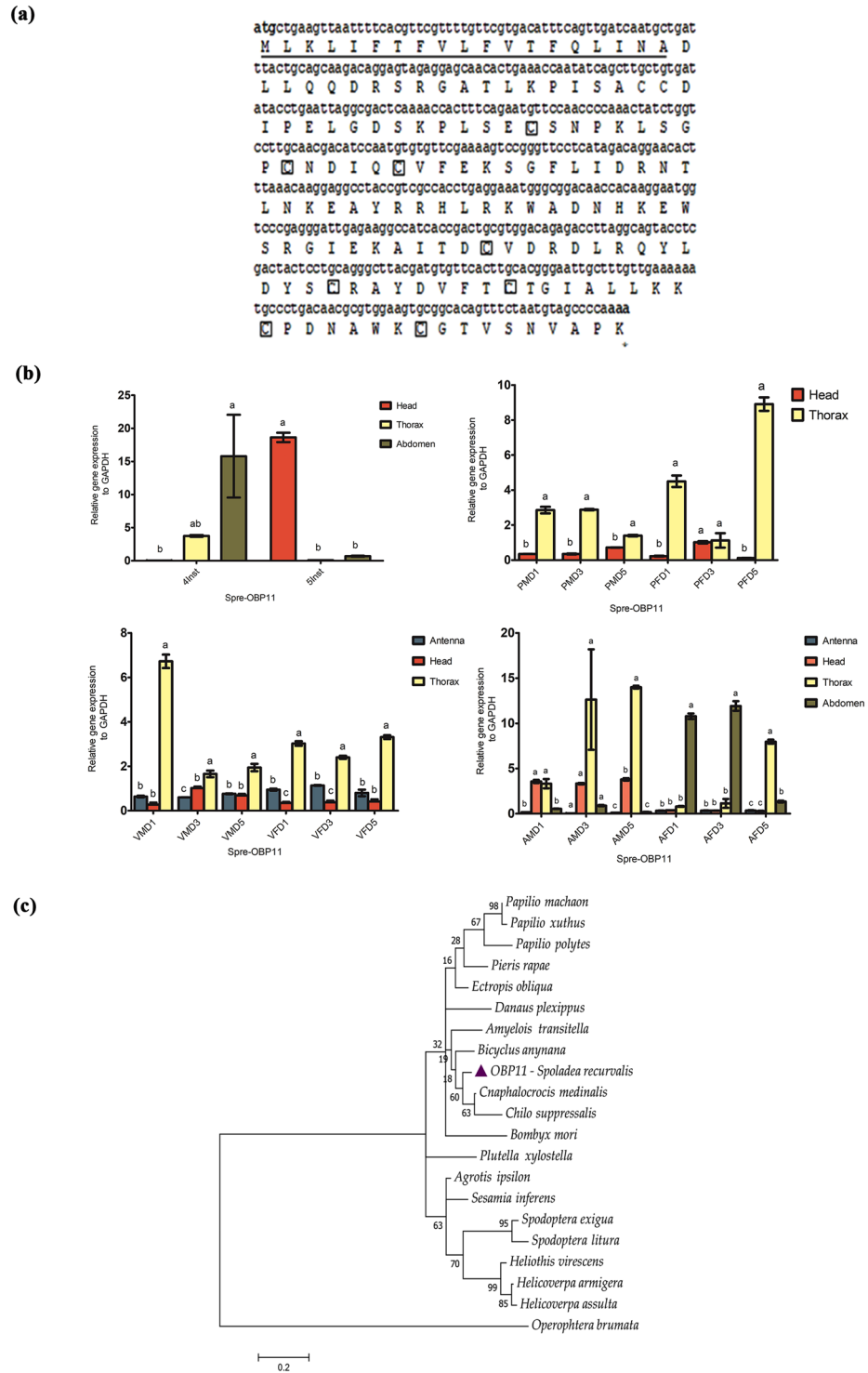
**Table 2.** Amino acid sequence characterization of putative PBP/OBPs identified from *S. recurvalis*. 'X' denotes any amino acid; superscript *p* refers partial sequence.

## Discussion

The production and release of sex pheromone in butterflies and moth species were mostly regulated by the neuropeptide, pheromone biosynthesis activating neuropeptide (PBAN) that plays various biological processes, potentially adapted to control insect pests<sup>4,37</sup>. For the first time, we are reporting here the characterization of Spre-DH-PBAN gene, binding proteins such as PBP and OBPs, and their differential and tissue-specific expression patterns. The Spre-PBAN flanked by three essential amino acids at its N terminus (G-K, G-R, G-R-R) and three other basic amino acids at its C terminus (G-R, G-R-K or G-R), which were believed to be the site for pro-hormone processing. The endoproteolytic cleavage, and glycine usually provide the amino group for amidation (Fig. 1)<sup>38–40</sup>, as reported for other DH-PBAN or PBAN that shared conserved pentapeptide sequence FXPR/KL at the C-termini<sup>16–18</sup>. Thus, amidation and physiological activity of DH and PBAN were essential, mainly for the members of Lepidoptera, Hymenoptera, Diptera, and Coleoptera<sup>41</sup>. Notably, the cDNA encoding all five neuropeptides (Spre-DH homologue, Spre- $\alpha$ -NP, Spre- $\beta$ -NP, Spre-PBAN, and Spre- $\gamma$ -NP) in this study were identical to the members of Lepidoptera<sup>17,18,27</sup>. The presence of 6 exons interrupted by five introns of DH-PBAN has been reported for other genera, including Marvi-DH-PBAN, since these gene products were the precursor of sex pheromone biosynthesis activating neuropeptide<sup>16,18</sup> (Fig. 1).

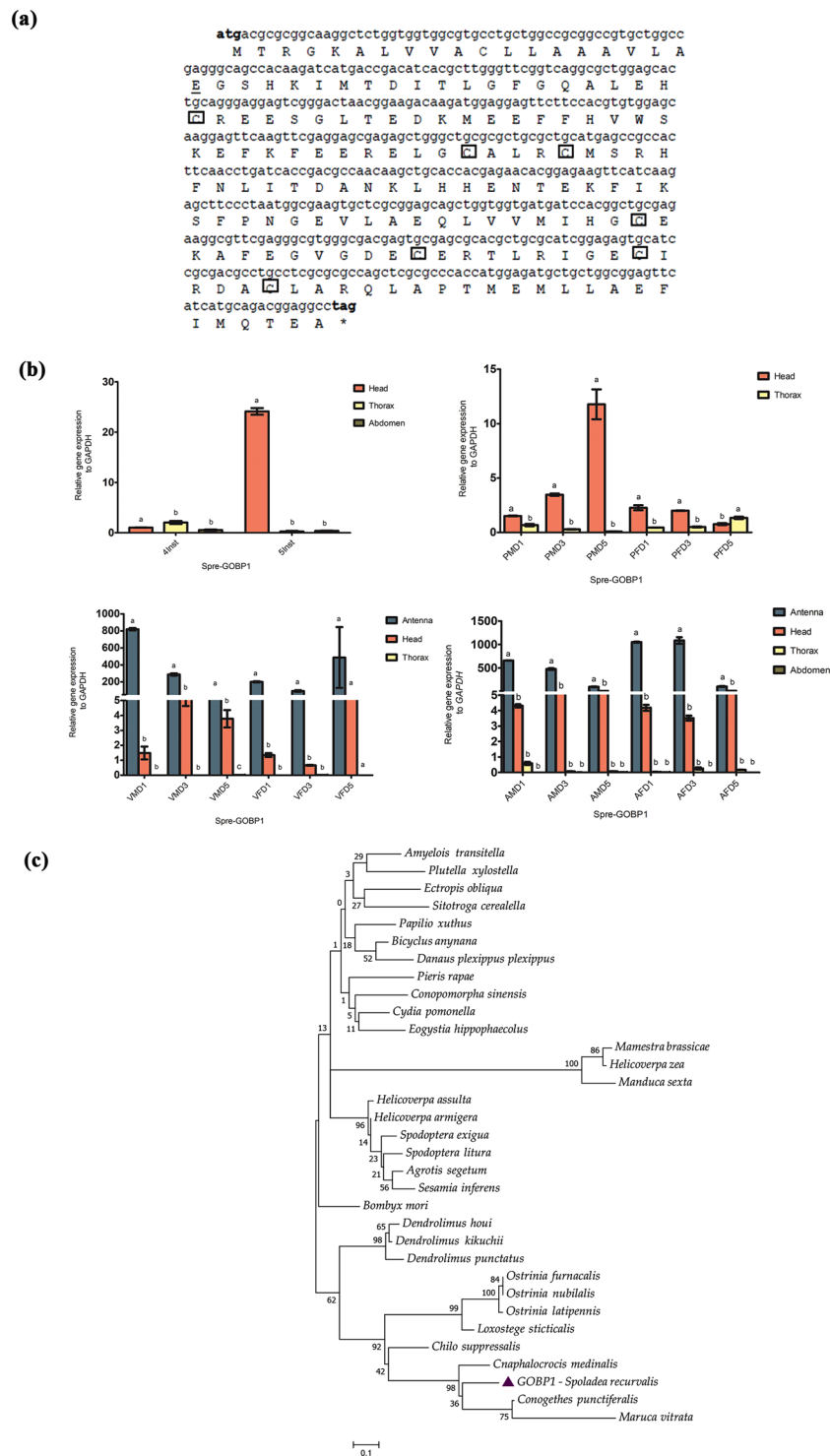
The DH neuropeptide that we observed contains a conserved C-terminal motif, WFGPRL, which has been reported to play the role of inducing embryonic diapause in *B. mori*, and impeding pupal diapause in moths of the Noctuidae<sup>42</sup>. It is more obvious that Crambidae including *S. recurvalis* and *M. vitrata* lacks histidine residues in the DH neuropeptide, which were common in Bombycoidea and Noctuoidea<sup>18</sup> (Fig. 1). As in other lepidopterans, *S. recurvalis* was also found to have  $\alpha$ -SGNP, which presents pheromonotropic activity, and can have interaction with the PBAN receptor, giving it similar functions as PBAN including diapause termination<sup>10</sup>. In addition,  $\alpha$ -SGNP has the function for hindgut muscle stimulation and contraction in Madeira cockroach, *Leucophaea maderae*. Thus,  $\alpha$ -SGNP was found to be present in insect orders besides Lepidoptera, contributing for other physiological functions, as well. Although  $\alpha$ -SGNP was missing in insect orders Diptera, Hymenoptera, and Coleoptera, it was purported that  $\alpha$ -SGNP domain can physically interact with  $\beta$ -SGNP to form one neuropeptide in fire ants, *Solenopsis invicta*<sup>10,43</sup>, but more evidences were needed to support this claim. Studies have shown that *M. vitrata*'s  $\beta$ -SGNP was different among the five peptides of the PBAN gene<sup>18</sup>, but not the most divergent as reported in *M. sexta* and *P. xylostella*<sup>19,44</sup>. Conversely in our study, *S. recurvalis* showed that  $\beta$ -SGNP has high similarity with *M. sexta*. Previous studies reported that the sequence of  $\gamma$ -SGNP has the higher resemblance with the likes of Noctuoidea, Pyraloidea, and Bombycoidea<sup>17,18</sup>. However, a variant C-terminal pentapeptide sequence LTPRLamide was identified in the study, and it was found to exist in *Clostera anastomosis* and *P. xylostella*. It has been shown that PBAN protein contains a highly conserved YFSPRL motif among the lepidopteran species constituting signature sequence for PK/PBAN family<sup>6,7</sup>. However, a variant C-terminal PBAN sequence showed asparagine substituted with serine in *M. vitrata*, revealing that this substitution was not commonly found in Pyraloidea super family (including this current study) (Fig. 1), which was associated with the pheromonotropic activity<sup>16,18</sup> (Fig. 1). The complete coding and non-coding sequences of Spre-DH-PBAN gene from this study revealed that approximately 4.2 kb contains six exons and five introns with conserved intron position within the codon regions (Fig. 2). Similar reports of DH-PBAN genomes from *B. mori*<sup>45</sup>, *M. vitrata*<sup>18</sup>, and *O. nubilalis*<sup>16</sup> had also been published, revealing that different pattern of exon and intron length and GC content could determine the evolutionary forces influenced by natural selection<sup>46</sup>.

The biosynthesis and release of DH-PBAN neuropeptides mostly arise from the sub-esophageal ganglion (SG) of moth species<sup>4</sup>. In our research, DH-PBAN was notably expressed in higher amounts in the abdominal region of 5<sup>th</sup> instar (5INS) larvae suggesting its essential role in preparation for pupation<sup>20</sup> (Fig. 3b). Previously, PBAN was widely recognized to accelerate pupation and cuticular tanning for color polymorphism<sup>7,11</sup>. Furthermore, PBAN has been shown to be associated with migrating ganglion cells in *Lymantria dispar* during development of larvae in an anteroposterior direction towards the abdomen<sup>47</sup>. Therefore, we hypothesized that Spre-PBAN was likely to function similarly during its development from larva to pupa stage. The Spre-PBAN gene expression during mid- and late-pupal stages implied physiological function of their role in stimulating the preparation for adult development, which was congruent with previous findings<sup>18</sup>, but contrary to *M. sexta* and *A. pernyi*<sup>19,48</sup>. In both virgin male and female moths, PBAN gene expressions were higher in head indicating the commencement of PBAN synthesis in sub-esophageal ganglion, followed by sex pheromone synthesis that attracts males for mating<sup>5,18</sup> (Fig. 3b). After mating, females maintain PBAN transcripts in their head and thoracic region at a high-level due to the necessity of keeping the sex pheromone biosynthesis for multiple mating, circadian regulation as it is that FXPR/L peptides in conjunction with the stimulation of pheromone can engage in other bioactivities<sup>11,49</sup>. However, evidences also suggested that mating may also inhibit pheromone production in females, which may be associated with cessation of PBAN release by hormonal and neural mechanism<sup>49</sup>. In *M. vitrata* and *C. transversa*, it was previously found that expression levels of PBAN in female were notably lower than those in male adults<sup>17,18</sup>.

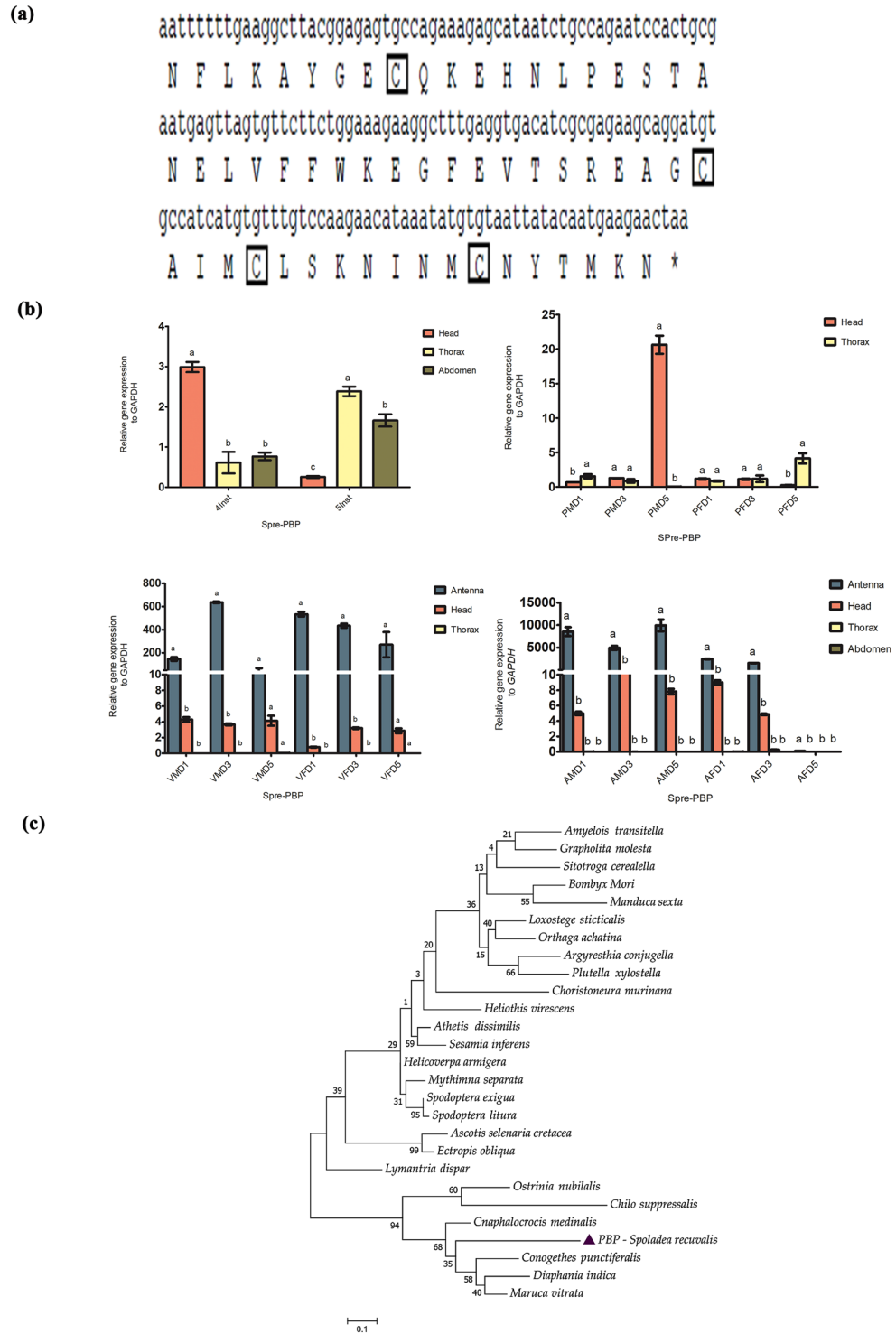


**Figure 4.** Schematic illustration of basic molecular characterization of odorant binding protein11 from *Spoladea recurvalis* (a) Nucleotide and deduced amino acid sequence of OBP11. The conserved cysteine residues are marked in square box. Predicted signal peptide was highlighted in underline. (b) qRT-PCR analysis shown the expression of Spre-OBP11 relative to each development stages (Inst-Instar; PMD-Pupae male day; PFD-Pupae female day; VMD- virgin male day; VFD- virgin female day; AMD-Mated adult male day; AFD-Mated adult female day). GAPDH was used as an internal control. Bar represents  $\pm$  standard error. Different letters (a,b,c) shows the significant differences ( $P < 0.05$ ). (c) Maximum Likelihood tree was constructed based on odorant-binding protein 11 the amino acid sequence using MEGA 6.0 with default settings model JTT + T inferred from 1000 replicates.

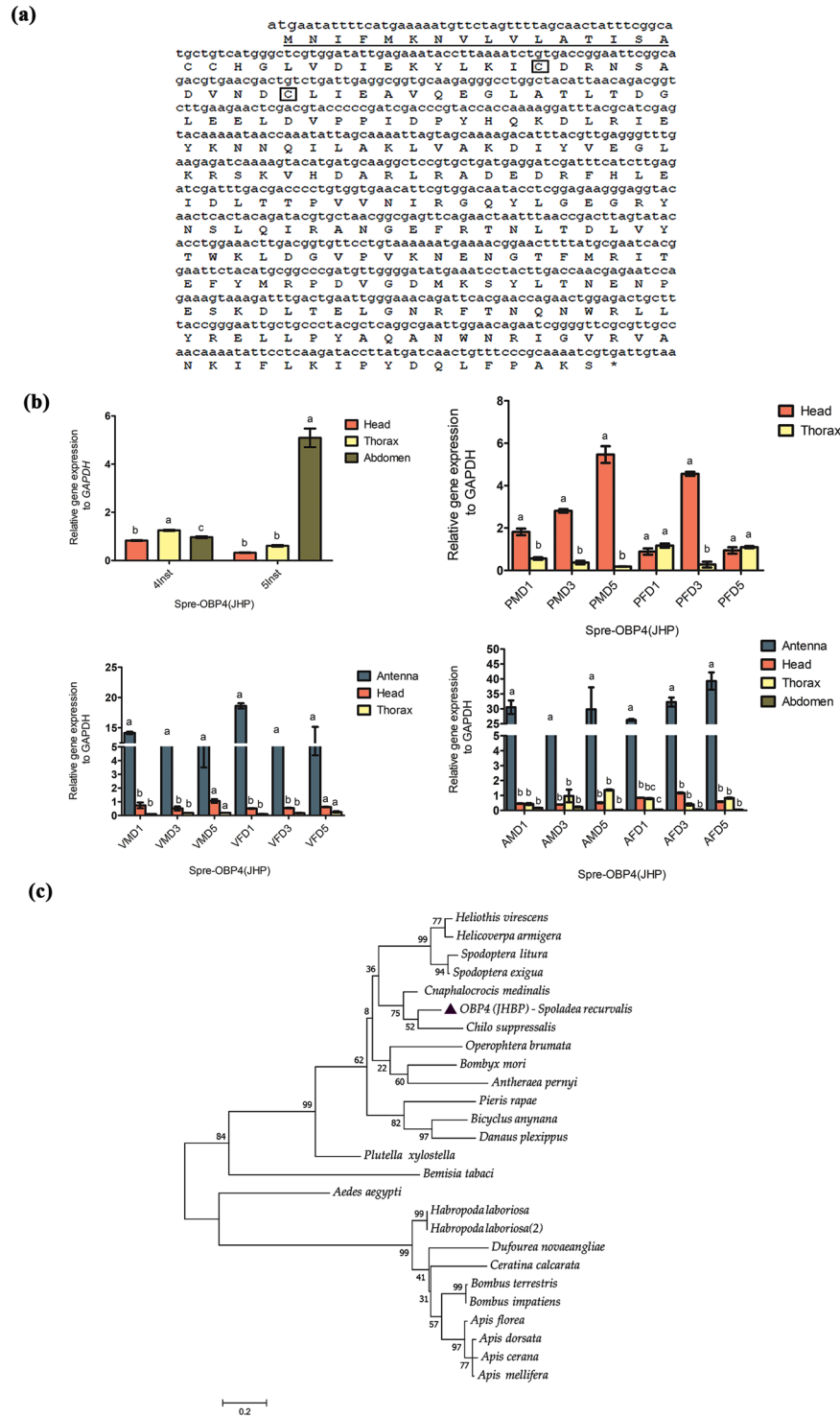




**Figure 5.** Schematic illustration of basic molecular characterization of general odorant binding protein1 (GOBP1) from *Spoladea recurvalis* (a) cDNA and derived amino acid sequence of Spre-GOBP1. The conserved cysteine residues are marked in square box. Predicted signal peptide was highlighted in underline. (b) qRT-PCR analysis shown the expression of Spre-GOBP1 relative to each development stages. (Inst-Instar; PMD-Pupae male day; PFD-Pupae female day; VMD-virgin male day; VFD-virgin female day; AMD-Mated adult male day; AFD-Mated adult female day). GAPDH was used as an internal control. Bar represents  $\pm$  standard error. Different letters (a,b,c) shows the significant differences ( $P < 0.05$ ). (c) Maximum Likelihood tree was constructed based on general odorant-binding protein 1 the amino acid sequence using MEGA 6.0 with default settings model LG + G inferred from 1000 replicates.



**Figure 6.** Schematic illustration of basic molecular characterization of pheromone binding protein (PBP) from *Spoladea recurvalis*. (a) 5' missing partial cDNA and derived amino acid sequence of Spre-PBP. The conserved cysteine residues are marked in square box. (b) qRT-PCR analysis shown the expression of Spre-PBP relative to each development stages. (Inst-Instar; PMD-Pupae male day; PFD-Pupae female day; VMD-virgin male day; VFD-virgin female day; AMD-Mated adult male day; AFD-Mated adult female day). GAPDH was used as an internal control. Bar represents  $\pm$  standard error. Different letters (a,b,c) shows the significant differences ( $P < 0.05$ ). (c) Maximum Likelihood tree was constructed based on Pheromone binding protein the amino acid sequence using MEGA 6.0 with default settings model LG + G + I inferred from 1000 replicates.



**Figure 7.** Schematic illustration of basic molecular characterization of odorant binding protein4 (OBP4) from *Spoladea recurvalis*. **(a)** cDNA and derived amino acid sequence of Spre-OBP4. The conserved cysteine residues are marked in square box. Predicted signal peptide was highlighted in underline. **(b)** qRT-PCR analysis shown the expression of Spre-GOBP1 relative to each development stages (Inst-Instar; PMD-Pupae male day; PFD-Pupae female day; VMD-virgin male day; VFD-virgin female day; AMD-Mated adult male day; AFD-Mated adult female day). GAPDH was used as an internal control. Bar represents  $\pm$  standard error. Different letters **(a,b,c)** shows the significant differences ( $P < 0.05$ ). **(c)** Maximum Likelihood tree was constructed based on odorant-binding protein 4 the amino acid sequence using MEGA 6.0 with default settings model LG + G inferred from 1000 replicates.

This low PBAN difference in females was thought to spring from the minimal need for the PBAN among the female pod borers. However, this is not observed in our current findings (Fig. 3b), and hence such variations could be species-specific.

In the present study, we have chosen four putative PBP/OBPs to understand the functional differentiation of PBPs and/or OBPs in *S. recurvalis*. Moreover, our study chose these PBP/OBPs given the fact that different studies had varying opinions on insect peripheral olfactory reception<sup>31</sup>, and we wanted to bridge the knowledge gap using molecular characterization. In our study, Spre-OBP11 showed differing expressions in the various developmental stages of the insect. In fact, these observations were not noted in the other studies (Fig. 4b). During larval stages, we presumed that Spre-OBP11 might assist in host plant discrimination and sense nutrient sources. The differences in Spre-OBP11 expression in antenna, head and thorax of virgin female and male tissues as well as post-mating head and thorax tissues may have plausibly been related to the perception of pheromone in addition to host plant volatile compounds. In post mating, female abdomen tissues were shown to abundantly express Spre-OBP11 compared to male tissues. Study from orthologue rice leaf-folder (*C. medinalis*) showed similar results that OBP was abundantly expressed in female tissues, in which CmOBP was used to find suitable host leaf surfaces prior to female oviposition<sup>50,51</sup>. It is conceivable that Spre-OBP11 may have a similar function in ovipositing *S. recurvalis* female moths. Hence, it is worth investigating such a behavior response in *S. recurvalis* to OBPs in future research.

In this study, GOBP1 expressed abundantly in antenna and head of both virgin and mated female moths, which implies that Spre-GOBP1 likely involved in chemoreception to find suitable interspecific signals such as sex pheromone components and host plant volatiles (Fig. 5b). In *M. vitrata*, GOBP1 and GOBP2 were reported to be involved in chemoreception considering its broader role in detecting volatile odor molecules from host plants<sup>52</sup>. Another study revealed that two GOBPs of *Spodoptera exigua* were sex-equivalent supposing that Spre-GOBPs were likely to have similar role in host volatile recognition<sup>53</sup>. Our qRT-PCR results indicated that PBP was upregulated in head of 4<sup>th</sup> instar larvae and male pupae. Spre-PBP expressed in the antenna and heads of virgin moths of both sexes. However, Spre-PBP was highly expressed in mated males than in the females (Fig. 6b). We hypothesized that Spre-PBP involved in auto-detection of sex pheromone compounds by female moths in order to ensure multiple mating, which has been discussed in previous studies<sup>32</sup>. Some studies have suggested that dendrites that adjoining to the soluble PBPs/OBP in sensillum lymph were anticipated to relocate the dendrite membrane of the sensory neurons from the hydrophobic pheromone<sup>22</sup>. Other studies reported that PBPs attributed to certain receptions of sex-pheromone components, although PBPs were capable of binding to numerous pheromone components<sup>22,25,53–55</sup>. This was because of the fact that the complex peripheral mechanism in the olfactory system possesses the ability to purify the selection. Hence, the PBP/OBP family might have the additional function acting as transporters in the non-chemosensory tissues<sup>56</sup>. Thus, Spre-PBP may also have extensive functions besides sensing the pheromone compounds, which are yet to be studied. Similarly, results from Spre-OBP4 showed that the expression was abundant in tissues of 5<sup>th</sup> instar larval abdomen and pupal head stages of both sexes (Fig. 7b). We observed a higher expression pattern between pre-mated and post-mated adult tissues, especially mated male and female antennal tissues showing higher expression, whereas head and thorax tissue had moderately lower level of expression in all ages (Fig. 7b). It is interesting to note that Spre-OBP4 is a member of Takeout protein. It is unknown whether Spre-OBP4 has a binding property to small lipophilic juvenile hormone molecule, which might be regulated by starvation and timing or duration of host volatile odor signal<sup>57</sup>. However studies from *Phormia regina* and *B. mori* suggest that this odorant binding protein – juvenile hormone protein (OBP4-JHP) belongs to Take Out Protein subfamily particularly localized around axillary cell membranes and thought to be involved in early perception of the chemical signals in both gustatory and olfactory system<sup>58,59</sup>.

Overall, our results of selected PBP/GOBP/OBP prompted us to speculate that *S. recurvalis* exhibit some common biological functions when compared to other lepidopteran species to locate the host, recognition of odorous compounds including plant volatiles and pheromones, and circadian rhythm prior to oviposition. Considering that the selected *S. recurvalis* PBP and GOBP/OBP were close homologs to rice leaf-folder (*C. medinalis*), highly conserved among insect orders, presumably derived from single gene duplication event initiated from a common ancestor under environmental pressure across 250 million years<sup>60,61</sup>. Possibly, differential expression among the four putative binding protein genes can have developed some unique characteristics in chemosensory system of *S. recurvalis* to become adapted to different ecosystem in which host plants were present. Hence, characterization of these genes may be useful to investigate the population diversity in *S. recurvalis*, as well as to target the control of this pest based on genetic manipulation of olfactory system.

## Methods

**Insect Rearing and Tissue Collection.** *Spoladea recurvalis* moths were collected from amaranth crops in Shanhua, Tainan, Taiwan. The insects were mass-reared following the method described previously<sup>34</sup>, and tissue samples were collected as described in Chang *et al.*<sup>18</sup>. Whole insects in different life stages were collected as fourth-and-fifth instar larvae, pupae and virgin as well as mated male and female adults. Antenna, head without antenna, thorax, and abdomen from male and female moths were excised and stored in RNA later.

**RNA Extraction and 5' And 3' Rapid Amplification of cDNA Ends (5' and 3' RACE).** Owing to our great interest to understand the biological functions of PBAN and PBP/OBPs in *S. recurvalis* and molecular basis of pheromone olfactory system, all sequence information corresponding to PBAN and selected PBP/OBP transcripts were obtained from the transcriptome library<sup>62</sup> (Supplementary Table. S5). We included different life-stages (5<sup>th</sup> instar larvae, pupae, virgin and mated adults) for studying the gene expression of PBAN and PBP/OBPs. The sum of RNA was separated from the other tissues dissected from different life-stages described above using the “Total RNA mini kit” (Qiagen Ltd.) from Venlo-Netherlands according to the manufacturer’s protocol, with on-column DNase I treatment. A nano-drop spectrophotometer (Thermo Scientific) at 260 nm was

used to quantify RNA by absorbance measurement. First, sequences showing homology with OBP genes were used to design gene-specific and nested reverse and forward primers to obtain 5' and 3' ends of the full-length cDNAs by rapid amplification of cDNA ends (RACE method). The sequences and primers used for the 5' and 3' RACE are listed in Supplementary Table S6. All the steps of RACE were performed using 5'RACE RT-PCR system (Invitrogen) instructions. Total RNA from the fifth instar larval head was used as a template for generating 5' and 3' RACE ready cDNAs. Following ligation of the 5' RACE adapter oligomer, reverse transcription with Superscript III was performed using oligonucleotide as primer. Besides, a unique gene-specific DNA fragment was increased in each case. After amplification, the 5'RACE and 3'RACE PCR products were separated in a 2.5% agarose gel with ethidium bromide (EtBr) staining. The gel slices containing DNA fragments were purified using an agarose gel DNA purification kit (Geneaid, Taiwan), according to the manufacturer's instructions, and were subjected to sequencing.

**Quantitative real-time PCR (qRT-PCR) analysis.** Quantitative RT-PCR was carried out with a Bio-Rad CFX system in the final reaction volume of 10  $\mu$ L containing 2.5  $\mu$ L cDNA of nuclease-free water, 5x SYBER Green Real-time PCR Master Mix and 10  $\mu$ M each of forward and reverse primers (Supplementary Table. S6). The thermal cycling conditions were noted as follows: Initial denaturation 95 °C for 3 mins, 95 °C for 10 sec, Annealing 60 °C for 45 sec, following 40 cycles qRT-PCR was performed for three biological replicates. GAPDH was used to normalize the gene expression. The relative expression levels of the genes were calculated using the  $2^{-\Delta\Delta CT}$  method<sup>63</sup>, which determines the difference in the CT between the control GAPDH product and the target gene product. Relative fold change for each gene was compared to respective stage of tissue and day for normalization.

**Molecular and Phylogenetic Analyses of Sequence Data.** Nucleic acid and deduced amino acid sequences were determined by ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). To identify the putative orthologs of *S. recurvalis*, similarity sequence information was retrieved using BLASTn and BLASTp. The ClustalW<sup>64</sup> module in MEGA6<sup>65</sup> software was used to align multiple sequences. The cDNA encoding PBAN possess five neuropeptides: 24-amino acid residue of DH motif (V22 to L45), and the other four neuropeptides such as 36 aa PBAN (G175 to R166), three short SGNPs,  $\alpha$ -7-aa (K95 to G102),  $\beta$ -20 aa (R105 to G126), and  $\gamma$ -8 aa (R166 to G175), were individually aligned using Clustal W, and compared against NCBI database to identify homologs of other insects. The raw output of the multiple sequence alignment was refined to minimize insertion/deletion events. The evolutionary relationships were inferred using the Maximum-likelihood method for inter-specific and intraspecific species using MEGA6 with the best-fit phylogenetic model<sup>18</sup>. Branch supports were surveyed by bootstrapping 1000 times. Putative signal peptide was predicted using SIGNALP 4.1 program with default settings<sup>66</sup>.

**Statistical Analysis.** The statistical analysis was done using the Graph Pad Prism 5.0 (Graph Pad Software La Jolla, CA, USA). Results of qRT-PCR were analysed using  $2^{-\Delta\Delta Ct}$  method<sup>63</sup> to calculate the relative gene expression of three biological and three technical replicates ( $n = 6$ ) for each target gene. Values were given as mean  $\pm$  standard error (SE). Significant differences were indicated by different letters (a, b, c,  $P < 0.05$ ). Analysis of variance (ANOVA) was conducted to evaluate significant differences in expression level for each gene among different tissues and developmental stages of males and females, followed by Tukey's 'Honest Significance Difference (HSD) test with the critical level of  $\alpha = 0.05$ .

## Data Availability

All data generated or analysed during this study are included in the published article (and its Supplementary information files).

## References

- Nassel, D. R. Neuropeptides in the nervous system of *Drosophila* and other insects: multiple roles as neuromodulators and neurohormones. *Progress in neurobiology* **68**, 1–84 (2002).
- Sakurai, T., Namiki, S. & Kanzaki, R. Molecular and neural mechanisms of sex pheromone reception and processing in the silkworm *Bombyx mori*. *Frontiers in physiology* **5**, 125, <https://doi.org/10.3389/fphys.2014.00125> (2014).
- Matsumoto, S. Molecular Mechanisms Underlying Sex Pheromone Production in Moths. *Bioscience, Biotechnology, and Biochemistry adypub*, 0912241778–0912241778, <https://doi.org/10.1271/bbb.90756> (2010).
- Rafaeli, A. Pheromone biosynthesis activating neuropeptide (PBAN): regulatory role and mode of action. *General and comparative endocrinology* **162**, 69–78, <https://doi.org/10.1016/j.ygcen.2008.04.004> (2009).
- Choi, M. Y. & Vander Meer, R. K. Molecular Structure and Diversity of PBAN/pyrokinin Family Peptides in Ants. *Frontiers in endocrinology* **3**, 32, <https://doi.org/10.3389/fendo.2012.00032> (2012).
- Raina, A. K. & Kempe, G. T. Structure activity studies of PBAN of *Helicoverpa zea* (Lepidoptera: noctuidae). *Insect biochemistry and molecular biology* **22**, 221–225, [https://doi.org/10.1016/0965-1748\(92\)90058-M](https://doi.org/10.1016/0965-1748(92)90058-M) (1992).
- Raina, A. K., Kingan, T. G. & Kochansky, J. P. A pheromonotropic peptide of *Helicoverpa zea*, with melanizing activity, interaction with PBAN, and distribution of immunoreactivity. *Archives of insect biochemistry and physiology* **53**, 147–157, <https://doi.org/10.1002/arch.10094> (2003).
- Raina, A. K. *et al.* Identification of a neuropeptide hormone that regulates sex pheromone production in female moths. *Science* **244**, 796–798, <https://doi.org/10.1126/science.244.4906.796> (1989).
- Suwan, S., Isobe, M., Yamashita, O., Minakata, H. & Imai, K. Silkworm diapause hormone, structure-activity relationships indispensable role of C-terminal amide. *Insect biochemistry and molecular biology* **24**, 1001–1007 (1994).
- Choi, M. Y. & Vander Meer, R. K. Identification of a new member of the PBAN family of neuropeptides from the fire ant, *Solenopsis invicta*. *Insect molecular biology* **18**, 161–169, <https://doi.org/10.1111/j.1365-2583.2009.00867.x> (2009).
- Altstein, M., Hariton, A. & Nachman, R. J. In *Handbook of Biologically Active Peptides (Second Edition)* 255–266 (Academic Press, 2013).
- Jurenka, R. In *Advances in Insect Physiology* Vol. 49 (ed Russell Jurenka) 123–170 (Academic Press, 2015).
- Jurenka, R. A. In *Insect Pheromone Biochemistry and Molecular Biology* (ed Richard Vogt) 53–80 (Academic Press, 2003).

14. Kitamura, A. *et al.* Amino acid sequence of pheromone-biosynthesis-activating neuropeptide (PBAN) of the silkworm, *Bombyx mori*. *Biochemical and Biophysical Research Communications* **163**, 520–526, [https://doi.org/10.1016/0006-291X\(89\)92168-2](https://doi.org/10.1016/0006-291X(89)92168-2) (1989).
15. Choi, M. Y., Estep, A., Sanscrainte, N., Becnel, J. & Vander Meer, R. K. Identification and expression of PBAN/diapause hormone and GPCRs from *Aedes aegypti*. *Molecular and cellular endocrinology* **375**, 113–120, <https://doi.org/10.1016/j.mce.2013.05.019> (2013).
16. Fodor, J. *et al.* Molecular cloning, mRNA expression and biological activity of the pheromone biosynthesis activating neuropeptide (PBAN) from the European corn borer, *Ostrinia nubilalis*. *Insect molecular biology* **26**, 616–632, <https://doi.org/10.1111/imb.12324> (2017).
17. Wei, H. *et al.* Identification and expression profiling of pheromone biosynthesis activating neuropeptide in *Chlumetia transversa* (Walker). *Pesticide biochemistry and physiology* **135**, 89–96, <https://doi.org/10.1016/j.pestbp.2016.05.005> (2017).
18. Chang, J. C. & Ramasamy, S. Identification and expression analysis of diapause hormone and pheromone biosynthesis activating neuropeptide (DH-PBAN) in the legume pod borer, *Maruca vitrata* Fabricius. *PLoS one* **9**, e84916, <https://doi.org/10.1371/journal.pone.0084916> (2014).
19. Xu, W. H. & Denlinger, D. L. Identification of a cDNA encoding DH, PBAN and other FXPRL neuropeptides from the tobacco hornworm, *Manduca sexta*, and expression associated with pupal diapause. *Peptides* **25**, 1099–1106, <https://doi.org/10.1016/j.peptides.2004.03.021> (2004).
20. Xu, W. H., Sato, Y., Ikeda, M. & Yamashita, O. Molecular characterization of the gene encoding the precursor protein of diapause hormone and pheromone biosynthesis activating neuropeptide (DH-PBAN) of the silkworm, *Bombyx mori* and its distribution in some insects. *Biochimica et biophysica acta* **1261**, 83–89 (1995).
21. Choi, M. Y., Meer, R. K., Shoemaker, D. & Valles, S. M. PBAN gene architecture and expression in the fire ant, *solenopsis invicta*. *Journal of insect physiology* **57**, 161–165, <https://doi.org/10.1016/j.jinsphys.2010.10.008> (2011).
22. Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nature* **293**, 161–163 (1981).
23. Vogt, R. G., Rogers, M. E., Franco, M. D. & Sun, M. A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). *The Journal of experimental biology* **205**, 719–744 (2002).
24. Gräter, F., Xu, W., Leal, W. & Grubmüller, H. Pheromone Discrimination by the Pheromone-Binding Protein of *Bombyx mori*. *Structure* **14**, 1577–1586, <https://doi.org/10.1016/j.str.2006.08.013> (2006).
25. Pelosi, P., Zhou, J. J., Ban, L. P. & Calvello, M. Soluble proteins in insect chemical communication. *Cellular and molecular life sciences: CMLS* **63**, 1658–1676, <https://doi.org/10.1007/s00018-005-5607-0> (2006).
26. Leal, W. S. In *The Chemistry of Pheromones and Other Semiochemicals II: —/—* (ed Stefan Schulz) 1–36 (Springer Berlin Heidelberg, 2005).
27. Liu, N. Y., Liu, C. C. & Dong, S. L. Functional differentiation of pheromone-binding proteins in the common cutworm *Spodoptera litura*. Comparative biochemistry and physiology. Part A, *Molecular & integrative physiology* **165**, 254–262, <https://doi.org/10.1016/j.cbpa.2013.03.016> (2013).
28. Zhang, J., Walker, W. B. & Wang, G. In *Progress in Molecular Biology and Translational Science* Vol. 130 (ed Richard Glatz) 109–128 (Academic Press, 2015).
29. Laughlin, J. D., Ha, T. S., Jones, D. N. M. & Smith, D. P. Activation of pheromone-sensitive neurons is mediated by conformational activation of pheromone-binding protein. *Cell* **133**, 1255–1265, <https://doi.org/10.1016/j.cell.2008.04.046> (2008).
30. Gong, D. P., Zhang, H. J., Zhao, P., Xia, Q. Y. & Xiang, Z. H. The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*. *BMC genomics* **10**, 332, <https://doi.org/10.1186/1471-2164-10-332> (2009).
31. Zhou, J.-J. In *Vitamins & Hormones* Vol. 83 (ed Gerald Litwack) 241–272 (Academic Press, 2010).
32. Mao, A. *et al.* Sex pheromone recognition and characterization of three pheromone-binding proteins in the legume pod borer, *Maruca vitrata* Fabricius (Lepidoptera: Crambidae). *Scientific Reports* **6**, 34484, <https://doi.org/10.1038/srep34484> <https://www.nature.com/articles/srep34484#supplementary-information> (2016).
33. Clark, J. T. & Ray, A. Olfactory Mechanisms for Discovery of Odorants to Reduce Insect-Host Contact. *Journal of chemical ecology* **42**, 919–930, <https://doi.org/10.1007/s10886-016-0770-3> (2016).
34. Hsu, Y.-c. & Ramasamy, S. Desert horse purslane weed as an alternative host for amaranth leaf webber, *Hymenia recurvalis* in Taiwan. Vol. 32 (2012).
35. Miyahata, Y. Immigration of the Hawaiian beet webworm moth, *Hymenia recurvalis*, during the Bai-u season. 2. *The arrival of the moths in early June 1992*. Vol. 39 (1993).
36. Halder, J., Kushwaha, D., Yadava, R. B., Rai, A. & Singh, D. B. *Incidence of Amaranthus foliage feeders in relation to different organic soil amendments*. Vol. 21 (2015).
37. Scherckenbeck, J. & Zdobinsky, T. Insect neuropeptides: structures, chemical modifications and potential for insect control. *Bioorganic & medicinal chemistry* **17**, 4071–4084, <https://doi.org/10.1016/j.bmc.2008.12.061> (2009).
38. Ma, P. W., Knipple, D. C. & Roelofs, W. L. Structural organization of the *Helicoverpa zea* gene encoding the precursor protein for pheromone biosynthesis-activating neuropeptide and other neuropeptides. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6506–6510 (1994).
39. Arabanian, A., Mohammadnejad, M. & Balalaie, S. A novel and efficient approach for the amidation of C-terminal peptides. *Journal of the Iranian Chemical Society* **7**, 840–845, <https://doi.org/10.1007/BF03246077> (2010).
40. Southey, B. R., Sweedler, J. V. & Rodriguez-Zas, S. L. Prediction of neuropeptide cleavage sites in insects. *Bioinformatics* **24**, 815–825, <https://doi.org/10.1093/bioinformatics/btn044> (2008).
41. Choi, M. Y., Vander Meer, R. K. & Valles, S. M. Molecular diversity of PBAN family peptides from fire ants. *Archives of insect biochemistry and physiology* **74**, 67–80, <https://doi.org/10.1002/arch.20356> (2010).
42. Yamashita, O. Diapause hormone of the silkworm, *Bombyx mori*: Structure, gene expression and function. *Journal of insect physiology* **42**, 669–679, [https://doi.org/10.1016/0022-1910\(96\)00003-0](https://doi.org/10.1016/0022-1910(96)00003-0) (1996).
43. Zhao, J. Y., Xu, W. H. & Kang, L. Functional analysis of the SGNP I in the pupal diapause of the oriental tobacco budworm, *Helicoverpa assulta* (Lepidoptera: Noctuidae). *Regulatory peptides* **118**, 25–31, <https://doi.org/10.1016/j.regpep.2003.10.024> (2004).
44. Lee, D. W. & Boo, K. S. Molecular characterization of pheromone biosynthesis activating neuropeptide from the diamondback moth, *Plutella xylostella* (L.). *Peptides* **26**, 2404–2411, <https://doi.org/10.1016/j.peptides.2005.04.016> (2005).
45. Xu, W. H., Sato, Y., Ikeda, M. & Yamashita, O. Stage-dependent and temperature-controlled expression of the gene encoding the precursor protein of diapause hormone and pheromone biosynthesis activating neuropeptide in the silkworm, *Bombyx mori*. *The Journal of biological chemistry* **270**, 3804–3808 (1995).
46. Zhu, L. *et al.* Patterns of exon-intron architecture variation of genes in eukaryotic genomes. *BMC genomics* **10**, 47, <https://doi.org/10.1186/1471-2164-10-47> (2009).
47. Giebultowicz Jadwiga, M., Raina Ashok, K., Uebel Edward, C. & Ridgway Richard, L. Two-step regulation of sex-pheromone decline in mated gypsy moth females. *Archives of insect biochemistry and physiology* **16**, 95–105, <https://doi.org/10.1002/arch.940160203> (1991).
48. Wei, Z. J., Hong, G. Y., Jiang, S. T., Tong, Z. X. & Lu, C. Characters and expression of the gene encoding DH, PBAN and other FXPRLamide family neuropeptides in *Antheraea pernyi*. *Journal of Applied Entomology* **132**, 59–67, <https://doi.org/10.1111/j.1439-0418.2007.01249.x> (2008).

49. Lu, Q. *et al.* Sex pheromone titre in the glands of *Spodoptera litura* females: circadian rhythm and the effects of age and mating. *Physiological Entomology* **42**, 156–162, <https://doi.org/10.1111/phen.12185> (2017).
50. Mack, P. D., Kapelnikov, A., Heifetz, Y. & Bender, M. Mating-responsive genes in reproductive tissues of female *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10358–10363, <https://doi.org/10.1073/pnas.0604046103> (2006).
51. Zeng, F. F. *et al.* Identification and Comparative Expression Profiles of Chemoreception Genes Revealed from Major Chemoreception Organs of the Rice Leaf Folder, *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae). *PLoS one* **10**, e0144267, <https://doi.org/10.1371/journal.pone.0144267> (2015).
52. Zhou, J. *et al.* Identification of Host-Plant Volatiles and Characterization of Two Novel General Odorant-Binding Proteins from the Legume Pod Borer, *Maruca vitrata* Fabricius (Lepidoptera: Crambidae). *PLoS one* **10**, e0141208, <https://doi.org/10.1371/journal.pone.0141208> (2015).
53. Liu, N. Y. *et al.* Two subclasses of odorant-binding proteins in *Spodoptera exigua* display structural conservation and functional divergence. *Insect molecular biology* **24**, 167–182, <https://doi.org/10.1111/imb.12143> (2015).
54. Zhang, T. T. *et al.* Characterization of three pheromone-binding proteins (PBPs) of *Helicoverpa armigera* (Hubner) and their binding properties. *Journal of insect physiology* **58**, 941–948, <https://doi.org/10.1016/j.jinsphys.2012.04.010> (2012).
55. Pedda Kasim, D., Srideepthi, R., Suneetha, P., Krishna, M. S. R. & Lakshmisahitya, U. Identification of pheromone-binding proteins of the maize stem borer, *Chilo partellus* (Swinhoe, 1885) (Lepidoptera: Crambidae). *Egyptian Journal of Biological Pest Control* **28**, 6, <https://doi.org/10.1186/s41938-017-0007-y> (2018).
56. Li, S. *et al.* Multiple functions of an odorant-binding protein in the mosquito *Aedes aegypti*. *Biochemical and biophysical research communications* **372**, 464–468, <https://doi.org/10.1016/j.bbrc.2008.05.064> (2008).
57. Sarov-Blat, L., So, W. V., Liu, L. & Rosbash, M. The *Drosophila* takeout Gene Is a Novel Molecular Link between Circadian Rhythms and Feeding Behavior. *Cell* **101**, 647–656, [https://doi.org/10.1016/S0092-8674\(00\)80876-4](https://doi.org/10.1016/S0092-8674(00)80876-4) (2000).
58. Fujikawa, K., Seno, K. & Ozaki, M. A novel Takeout-like protein expressed in the taste and olfactory organs of the blowfly, *Phormia regina*. *The FEBS journal* **273**, 4311–4321, <https://doi.org/10.1111/j.1742-4658.2006.05422.x> (2006).
59. Saito, K. *et al.* Cloning and expression analysis of takeout/JHBP family genes of silkworm, *Bombyx mori*. *Insect molecular biology* **15**, 245–251, <https://doi.org/10.1111/j.1365-2583.2006.00612.x> (2006).
60. Jones, W. D., Nguyen, T. A., Kloss, B., Lee, K. J. & Vosshall, L. B. Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Current biology: CB* **15**, R119–121, <https://doi.org/10.1016/j.cub.2005.02.007> (2005).
61. Krieger, J., Klink, O., Mohl, C., Raming, K. & Breer, H. A candidate olfactory receptor subtype highly conserved across different insect orders. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology* **189**, 519–526, <https://doi.org/10.1007/s00359-003-0427-x> (2003).
62. Chang, J. C. & Ramasamy, S. Transcriptome analysis in the beet webworm, *Spoladea recurvalis* (Lepidoptera: Crambidae). *Insect science* **25**, 33–44, <https://doi.org/10.1111/1744-7917.12375> (2018).
63. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* **25**, 402–408, <https://doi.org/10.1006/meth.2001.1262> (2001).
64. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673–4680 (1994).
65. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* **30**, 2725–2729, <https://doi.org/10.1093/molbev/mst197> (2013).
66. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785–786, <https://doi.org/10.1038/nmeth.1701> (2011).

## Acknowledgements

The authors also gratefully acknowledge the core donors to World Vegetable Center: Republic of China (Taiwan), UK aid from the UK government, United States Agency for International Development (USAID), Australian Centre for International Agricultural Research (ACIAR), Germany, Thailand, Philippines, Korea, and Japan. This research was funded by the Federal Ministry for Economic Cooperation and Development (BMZ), Germany through *icipe* (Project number: 13.1432.7-001.00; Contract number: 81170265).

## Author Contributions

R.S. developed the main conceptual ideas and the project. R.S.K. carried out the experiments, analyzed the data and wrote the manuscript. R.S. critically reviewed and revised the manuscript. All authors approved the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-39822-x>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019