MOLECULAR DIVERSITY OF PBAN FAMILY PEPTIDES FROM FIRE ANTS

Man-Yeon Choi, Robert K. Vander Meer, and Steven M. Valles
United States Department of Agriculture-Agricultural Research Service, Center of Medical, Agricultural and Veterinary Entomology (CMAVE), Gainesville, Florida

The PBAN/Pyrokinin peptide family is a major neuropeptide family characterized with a common FXPRLamide in the C-termini. These peptides are ubiquitously distributed in the Insecta and are involved in many essential endocrinial functions, e.g., pheromone production. Previous work demonstrated the localization of PBAN in the fire ant central nervous system, and identified a new family of PBAN from the red imported fire ant, Solenopsis invicta. In this study, we identified five more PBAN/Pyrokinin genes from S. geminata, S. richteri, S. pergandii, S. carolinensis, and a hybrid of S. invicta and S. richteri. The gene sequences were used to determine the phylogenetic relationships of these species and hybrid, which compared well to the morphologically defined fire ant subgroup complexes. The putative PBAN and other peptides were determined from the amino acid sequences of the PBAN/Pyrokinin genes. We summarized all known insect PBAN family neuropeptides, and for the first time constructed a phylogenetic tree based on the full amino acid sequences translated from representative PBAN cDNAs. The PBAN/Pyrokinin gene is well conserved in Insecta and probably extends into the Arthropod phylum; however, translated pre-propeptides may vary and functional diversity may be retained, lost, or modified during the evolutionary process. Published 2010 Wiley Periodicals, Inc.

Keywords: fire ant; PBAN; neuropeptide; phylogenetic tree; Solenopsis

The cDNA sequences reported in this paper have been deposited in the GenBank database (GenBank Accession numbers: GQ872197; GQ872198; GQ872199; GQ872200; GQ872201).
Correspondence to: Man-Yeon Choi or Robert K. Vander Meer, United States Department of Agriculture-Agricultural Research Service, Center of Medical, Agricultural and Veterinary Entomology (CMAVE), 1600 SW 23rd Dr. Gainesville, FL 32608. E-mail: mychoi@ars.usda.gov; bob.vandermeer@ars.usda.gov

This article is a U.S. government work and, as such, is in the public domain in the United States of America.

Published online in Wiley InterScience (www.interscience.wiley.com).
Published 2010 Wiley Periodicals, Inc. DOI: 10.1002/arch.20396
INTRODUCTION

Insect neurohormones function as intercellular communicators that regulate a variety of physiological and behavioral events during development and reproduction. The Pyrokinin/PBAN (Pheromone Biosynthesis Activating Neuropeptide) represents a major neurohormone family, characterized by a 5-amino-acid C-terminal sequence, FXPRamide. A pyrokinin (leucopyrokinin) member of this family from the cockroach (Leucophaea maderae) was the first to be isolated and characterized, as a myotropin (Holman et al., 1986). Subsequently, other myotropic peptides were identified from various insect orders (Nachman et al., 1986). PBAN has been the subject of a great deal of interest since the first PBAN molecule was identified from Helicoverpa zea adults two decades ago (Raina et al., 1989). The PBAN/pyrokinin peptide family exhibiting the FXPRamide functional epitope is expected to be widely distributed in the Insecta and influence a variety of physiological functions, e.g., (1) stimulation of pheromone biosynthesis in female moths (Raina et al., 1989); (2) induction of melanization in moth larvae (Matsumoto et al., 1990); (3) induction of embryonic diapause in the silk moth (Suwan et al., 1994); (4) stimulation of visceral muscle contraction in cockroaches (Predel and Nachman, 2001); (5) acceleration of puparium formation in the flesh fly (Zdarek et al., 1997); and (6) termination of development of pupal diapause in heliothine moths (Xu and Denlinger, 2003). The PBAN/pyrokinin peptides are cross-reactive in that each peptide can activate all physiological functions noted above in experimental models.

The neurohormonal action of PBAN for pheromone biosynthesis in lepidopteran moths has been well studied (Rafaeli, 2009; Rafaeli and Jurenka, 2003). PBAN synthesized in the subesophageal ganglion (SG) is released into the hemolymph via the corpora cardiaca (CC), a neurohemal organ. The first PBAN identified was the 33-amino-acid peptide from H. zea, Hez-PBAN (Helicoverpa zea PBAN) (Raina et al., 1989). Subsequently, PBAN amino acid sequences were determined from Bombyx mori (Kitamura et al., 1989) and Lymantria dispar (Masler et al., 1994), through direct isolation and purification of peptides. DNA cloning methods yielded additional PBAN encoding genes (= PBN or DH-PBAN) from 18 lepidopteran moths (summarized in Choi and Vander Meer, 2009). In addition to PBAN, a diapause hormone and three additional FXPR/KL neuropeptides (NFs: α, β, γ) were deduced from the same gene and are well conserved in moths. Non-lepidopteran PBAN/Pyrokinin peptides have been detected in the genome sequences of the mosquito, Anopheles gambiae (Riehle et al., 2002), the red flour beetle, Tribolium castaneum (Li et al., 2008), and the honeybee, Apis mellifera (Hummon et al., 2006); however, the physiological functions of these peptides are unknown.

The fire ant is probably the most studied ant species in the world and a great deal is known about the pheromone systems used to reduce reproductive competition, recruit to resources, and maintain colony social structure and territoriality (Vander Meer and Alonso, 1998, 2002; Vargo, 1998). In spite of decades of study on fire ant pheromones, virtually nothing is known about how pheromone production and release are regulated, nor whether protein hormones, especially neuropeptides, are involved in key physiological and endocrinical processes during development. Recently, we demonstrated the presence of PBAN/Pyrokinin immunoreactive neurons in the central nervous system of the fire ant, Solenopsis invicta (Choi et al., 2009), and identified and characterized PBAN and three additional F/PXRPL peptides from the Sol-PBAN cDNA (Choi and Vander Meer, 2009). In the present study, we have
extended this research by identifying five additional Solenopsis PBAN genes from S. geminata, S. richteri, S. pergandii, S. carolinensis, and a hybrid of S. invicta and S. richteri. Molecular sequences of putative PBANs and three neuropeptides deduced from these PBAN cDNAs were used to determine molecular diversity and phylogenetic relationships in the fire ant-group complexes. This is the first demonstration of social insect species classification or diversity measurements at the genus level using an insect neuropeptide family. We also summarize all identified PBANs and related neuropeptides, along with the basic PBAN gene sequences in insects.

MATERIALS AND METHODS

Fire Ants

Sample collection dates and locations were as follows: S. richteri, Moulton, Alabama (July 2005); S. invicta/richteri hybrid, Huntsville, Alabama (August 2005); S. geminata, Gainesville, Florida (August 2005); S. carolinensis, Starke, Florida (July 2005); S. pergandii, Gainesville, Florida (September 2005). Fire ant species were identified using taxonomic keys (Trager, 1991; Wojcik et al., 1976). Classification of Solenopsis richteri and the S. invicta/richteri hybrid were verified by venom alkaloid and cuticular hydrocarbon analyses (Ross et al., 1987; Vander Meer and Lofgren, 1990; Vander Meer et al., 1985). Samples of S. invicta, S. richteri, S. geminata, S. pergandii, and S. invicta/richteri hybrids were collected by plunging a 20-ml scintillation vial into the ant mound. S. carolinensis was collected by burying a 20-dram glass vial containing a small piece of a cookie as bait. The plastic cap of the vial was punched with multiple holes using a 26-gauge needle. The vial was buried 4 inches below the soil surface and left overnight. The vials with S. carolinensis were retrieved 24 h later.

RNA Isolation and cDNA Synthesis

One hundred brain-subesophageal ganglia (Br-SGs) of S. geminata were dissected from the female alates in an RNase-free cold hymenopteran saline (130 mM NaCl, 6 mM KCl, 4 mM MgCl₂, 5 mM CaCl₂, 160 mM sucrose, 25 mM glucose, and 10 mM HEPES, adjusted to pH 7.2 with NaOH) and stored at −80°C until use. Poly (A)⁺ RNA was isolated from the dissected Br-SGs by Micro Fast mRNA purification kit (Invitrogen, Carlsbad, CA), and used to synthesize cDNA with the GeneRacer cDNA synthesis kit (Invitrogen). The other species collected above were returned immediately to the laboratory and total RNA extracted from 10 to 20 worker ants with Trizol reagent according to the manufacturer's instructions (Invitrogen). The total RNA (~5 μg) isolated from each species and hybrid was used to synthesize cDNA using a 3'-RACE (Rapid Amplification of cDNA Ends) (Invitrogen).

Molecular Cloning and Characterization

The synthesized cDNA of S. geminata was amplified with degenerate primer set, 5'-GGNATGTTGTTYGGNCCNMGYNTAXGGNMG-3' and 5' -CKNCCNARNCKNNGNCC RAACGCACATNCC -3. PCR was performed with the following temperature program: 5 cycles at 95°C for 30 s, 67°C for 30 s, and 72°C for 1 min, 5 cycles at 95°C for 30 s, 62°C for 30 s, and 72°C for 1 min, and 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min. The PCR product was gel purified and cloned using PCR2.
vector from TOPO TA cloning kit (Invitrogen) and sequenced. Based on this partial sequence, the further identification of the full cDNA of Sog-PBAN was performed using gene-specific primers set: sense primer, 5'-ATCGCCAGCG GCAATGGTGTTGGGTC-3' for 5'-RACE and antisense primer, 5'-CTAACTCGGATAGTTGGCC TATAGTTTCG-3' for 5'-RACE with GeneRacer kits (Invitrogen). The RACE PCR products were purified and cloned using the cloning kit above and sequenced. The sequence data obtained were used to construct the full-length PBAN gene sequence. Sequences were aligned and compared with Genetyx DNA software (Genetyx Co., Tokyo).

After the PBAN gene sequence was elucidated, oligonucleotide primers upstream and downstream of the N-terminal and C-terminal transcription sites were used to design a common primer set: 5'-GGAGCGACGTTCCGATGCCTA TAG-3'. These primers were used to amplify each PBAN ORF region. In the first round PCR, 2μl of cDNA synthesized from RNA isolated from S. richteri, hybrid of S. richteri and S. invicta, S. carolinensis, or S. pergandii was used as a template and conducted using a DNA Taq polymerase (Bio-Rad, Hercules, CA) with the following temperature program, 5 cycles at 95°C for 30 s, 62°C for 30 s, and 72°C for 1 min, 5 cycles at 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min, and 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min. Then, 1μl of the PCR product was directly applied to the second PCR amplification using the same primer set using a temperature program of 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min. The PCR products were isolated by Agarose gel purification, cloned, and sequenced by the same procedures described above.

Molecular Analysis of PBAN cDNAs Identified From Six Species

Nucleotide and translated amino acid sequences of the fire ant PBAN cDNAs were determined using homology and similarity based on S. invicta and S. geminata PBAN cDNAs, then cross-compared among six PBAN cDNAs using Genetyx DNA software. The evolutionary trees for Phylogenetic relationships were determined with the nucleotide sequences for fire ant species and with amino acid sequences for all insect PBAN using Unweighted Pair Group Method with Arithmetic (UPGMA) and Neighboring-Joining bootstrap methods from the same DNA software.

RESULTS

Structure of PBAN cDNAs

S. geminata (accession number: GQ872197). A 513-bp-long amplicon contained the S. geminata PBAN (Sog-PBAN) and three additional neuropeptides (NPs), DH-homologue, β-NP, and γ-NP, and the entire 3' end of the transcript including the stop codon (TAG) and polyadenylation signal (AATAAA) (Fig. 1). 5'-RACE yielded a 599-bp-long cDNA that contained an entire open reading frame (ORF) of 534 nucleotides encoding 177 amino acids from the first initiation codon (ATG) to the termination codon (TAG) indicated in Figure 1. The TATA box was located 50 bp upstream of the transcription start site. The processing site for the signal peptide is predicted to be between the first 29 and 30 amino acids. The cDNA encoded for four putative peptides.
based on six possible endoproteolytic cleavage sites (Southey et al., 2008; Veenstra, 2000) K^{62}-R^{65}, G^{76}-R^{81}, K^{116}-R^{117}, G^{126}-R^{127}, G^{155}-R^{156}-R^{157}, and G^{167}-R^{168}, indicated in italics in Figure 1. The pro-peptides cleaved are predicted to have a C-terminal amide from (NH_{2}) provided from glycine (G). The PBAN encoding region of the cDNA of S. geminata showed ~98% similarity with the S. invicta PBAN cDNA. The third domain with a 27-amino-acid (AA) sequence, GSGEDLFSYGDAYEVDDEDDHPLFVPRLamide, is considered a putative PBAN homologue, Sog-PBAN, and contained an additional phenylalanine(F) compared with Soi-PBAN (Fig. 2). A 15-AA peptide, TSQDIAAGWMWFGPRLamide, was identified that is homologous to the DH domain in lepidopteran moths (see Fig. 4). To date, this type of peptide possesses a conserved WFGPRL sequence at the C-terminus, except in Adoxophyes moth species (Choi et al., 2004). The 8-AA peptide, QPFQFTPRL, positioned in the
S. invicta 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60
S. richteri 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60
S. inv.x ric. 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60
S. geminata 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60
S. pergandii 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60
S. carolinensis 1: MIVTRNSVRATIVCIHMLCQLSRASQGSYESREIGNSNGSSESERSFSNDFGSCIDGKC 60

S. invicta 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120
S. richteri 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120
S. inv.x ric. 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120
S. geminata 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120
S. pergandii 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120
S. carolinensis 61: IERTSQDIASGMFQPRLRKRSKEFQELSSEIEILANALDGVDEATVPASDPEFPRQ 120

Figure 2. Comparison of open reading frame (ORF) amino acid translated from six fire ant PBAN genes. Two groups of Solenopsis PBAN cDNAs translating 176-AA for S. invicta, S. richteri, and hybrid, and 177-AA for S. geminata, S. pergandii, and S. carolinensis, which is an additional residue, phenylalanine (F), in the PBAN domains. S. carolinensis PBAN gene indicates two different residues in DH and PBAN domains compared to the other species. Six predicted endoproteolytic cleavage sites and four putative pre-peptides are indicated by bold italics and underline. Identical amino acids in six and five sequences are indicated with asterisks (*) and period (.). Dashed lines indicate gaps.

second domain corresponds to a β-NP with an FTPRLaamidate at the C-terminus. The 9-AA peptide, LPWIPSPRLaamidate, in the fourth domain corresponds to the γ-NP homologue in moths (Fig. 1; see also Fig. 4).

S. richteri and the hybrid (accession numbers: GQ872200 and GQ872201). Three PBAN cDNAs sequences for S. invicta, S. richteri, and their hybrid are not shown because they are nearly identical to the S. invicta PBAN sequence. The PBAN cDNAs of S. richteri and the hybrid (S. invicta × S. richteri) are identical and differ from the PBAN cDNA of S. invicta at only one DNA codon located 120bp downstream; however, the nucleotide difference does not alter the amino acid coding (Table 1 and Fig. 2). Therefore, the three PBAN cDNAs of S. invicta, S. richteri, and the hybrid are identical at the amino acid sequence level.

S. pergandii (Accession number: GQ872198). S. pergandii PBAN (Sop-PBAN) cDNA was found to be 534 bp long (data not shown), encoding a 177-AA protein. This was identical in size to the AA sequence from the S. geminata (Sog-PBAN) cDNA except for an asparagine (N) found at the 51st residue instead of aspartate (D). The ORF contains four putative peptides, DH (15-AA), β-NP (8-AA), PBAN (Sop-PBAN: 27-AA), and γ-NP (9-AA) (Fig. 2; see also Fig. 4). Comparison of nucleotide sequences of S. pergandii PBAN genes with the other Solenopsis PBAN genes shows S. pergandii closest to S. geminata followed by S. richteri, S. invicta, and then S. carolinensis (Table 1, Fig. 3).

S. carolinensis (accession number: GQ872199). S. carolinensis PBAN (Soc-PBAN) cDNA was 534 bp long encoding a 177-AA protein containing four putative peptides, PBAN,
Table 1. Homology of Nucleotide Sequences of Entire PBAN mRNA in Fire Ants

<table>
<thead>
<tr>
<th>Species</th>
<th>S. invicta</th>
<th>S. richteri</th>
<th>S. inv x ric</th>
<th>S. geminata</th>
<th>S. pergandii</th>
<th>S. carolinensis</th>
<th>bp (PBAN)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. invicta</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td>41</td>
<td>531</td>
<td>(25)</td>
</tr>
<tr>
<td>S. richteri</td>
<td>99.81</td>
<td>100</td>
<td>7</td>
<td>9</td>
<td>42</td>
<td>531</td>
<td>(26)</td>
</tr>
<tr>
<td>S. inv x ric</td>
<td>99.81</td>
<td>100</td>
<td>7</td>
<td>9</td>
<td>42</td>
<td>531</td>
<td>(26)</td>
</tr>
<tr>
<td>S. geminata</td>
<td>98.69</td>
<td>98.69</td>
<td>98.69</td>
<td>2</td>
<td>39</td>
<td>534</td>
<td>(27)</td>
</tr>
<tr>
<td>S. pergandii</td>
<td>98.12</td>
<td>98.31</td>
<td>98.31</td>
<td>99.62</td>
<td>41</td>
<td>534</td>
<td>(27)</td>
</tr>
<tr>
<td>S. carolinensis</td>
<td>92.32</td>
<td>92.13</td>
<td>92.13</td>
<td>92.70</td>
<td>92.32</td>
<td>534</td>
<td>(27)</td>
</tr>
</tbody>
</table>

*Numbers indicate the identical difference (the upper diagonal) and percentage (the lower diagonal) in the total nucleotide base pair (bp) numbers shown in the "bp" column.

**Numbers in the parenthesis indicate the sequence residues of PBAN domains.

Figure 3. Phylogenetic tree based on entire nucleotide of fire ant PBAN genes identified from current and previous studies. The tree was made with a fixed distance scale using the UPGMA and NJ bootstraps of Genetyx 6.0 software and multiple alignments of nucleotide sequences. The numbers indicate phylogenetic distance values in the analysis.

DH, β-NP, and γ-NP (Fig. 2). The similarity of Soc-PBAN cDNA was distant from the other fire ant species and hybrid reported here and previously (~92% identity), considering either nucleotide or peptide sequences (Table 1, Fig. 3). The cleavage site for the signal peptide is predicted to be between the first 27 and 28 amino acids, whereas in the other four species and hybrid it is predicted to be between the 29th and 30th residue. The sequence of Soc-PBAN (27-AA) differs from Sog-PBAN in that three amino acids have been replaced with phenylalanine (F), alanine (A), and threonine (T). The corresponding DH domain of S. carolinensis has a threonine (T) substituted for a serine (S), which is found in N-termini of the other species and hybrid. These residues belong to the same polar and functional group. Overall, S. carolinensis appears to be phylogenetically distant from the other Solenopsis species examined based on PBAN gene homology.

Primary Structure and Phylogenetic Relationship of PBAN Family Peptides in Insects

Lepidopteran PBAN cDNAs translate PBAN and four additional peptides (DH, α-, β-, and γ-NP), but the domain of α-NP is absent in Hymenoptera, Diptera, and Coleoptera (Fig. 4). A gap of residues between β-NP and PBAN is found in the Coleoptera PBAN gene. The phylogenetic relationship of PBAN genes identified from 13 families is similar to insect phylogeny based on taxonomic characters (Fig. 5).
Figure 4. Comparison of PBAN and other neuropeptides in moths (top) and schematic diagram of PBAN-DH gene structures of Lepidoptera, Hymenoptera, Diptera, and Coleoptera. The full PBAN genes of *L. dispar* (Masler et al., 1994) and Pse-PTI of *P. separata* (Matsumoto et al., 1992) have not been identified. Compared full or partial amino acid sequences are from lepidopteran species, *H. zea* (Ma et al., 1994), *H. asulata* (Choi et al., 1998), *H. armigera* (Zhang et al., 2004), *H. virescens* (Xu and Denlinger, 2003), *A. ipsilon* (Duportet et al., 1999), *B. brassicae* (Jacquin-Joly et al., 1998), *S. littoralis* (Iglesias et al., 2002), *S. exigua* (Xu et al., 2007), *C. ananassae* (Jing et al., 2007), *O. mylitta* (Uchani et al., 2007), *A. peryyi* (Wei et al., 2008), *B. mori* (Sato et al., 1995), *B. mandarina* (Xu et al., 1999), *M. sexta* (Xu and Denlinger, 2004), *S. cynthiae* (Wei et al., 2004), *A. mylitta* sp. that has not been specified (Choi et al., 2004), *P. sylvestella* (Lee and Boo, 2005), *A. s. cretica* (Kawai et al., 2007), *S. invicta* (Choi and Vander Meer, 2009), *S. richteri* (this study), *S. geminata* (this study), *S. pergandii* (this study), *S. carolinensis* (this study), *A. mellifera* (Hummon et al., 2008), *A. gambiae* (Riehle et al., 2002), *A. aegypti* (GenBank accession no. Q16880), and *T. castaneum* (Li et al., 2008). DH, diapause hormone; NF, neuropeptide; PBAN: pheromone biosynthesis activating neuropeptide.

**DISCUSSION**

As in other social insects, fire ants evolved sophisticated pheromonal communication to maintain colony cohesiveness and sociality. Several pheromone components from fire ants have been identified (Vander Meer and Alonso, 1998; Vander Meer et al., 2010). However, the mechanism used to control pheromone production and release is unknown, as is the role of neuropeptides in the developmental physiology of the ant. Since discovery of the first fire ant PBAN (Choi and Vander Meer, 2009), we wanted to determine the molecular diversity of PBAN genes from other fire ant species. Additionally, we were interested to know the phylogenetic relationships of PBANs and how these relationships corresponded with the classic morphologically based phylogeny of this group.
The Solenopsis group is a large genus covering approximately 185 species of fire ants (Pitts et al., 2005). The genus is a taxonomically difficult group due to the lack of reliable diagnostic characters. Recently, they have been re-classified into four groups: S. virtulens, S. tridens, S. geminata, and S. saevissima complexes (Pitts et al., 2005). Although a limited number of PBAN genes have now been identified, 5 fire ant species and S. invicta × S. richteri hybrid, it is possible to compare molecular diversity from nucleotide and peptide sequences in the fire ants. Basically, Solenopsis PBAN genes are divided into two groups of cDNAs translating 176- AA for S. invicta, S. richteri, and the hybrid, and 177-A for S. geminata, S. pergandii, and S. carolinensis. The difference is one additional residue, F, in the PBAN domains in the later group. PBAN domains of S. invicta and S. richteri consist of 26-AA, and S. geminata, S. pergandii, and S. carolinensis PBANs have 27-AA (Figs. 1, 2). In the current comparison among five fire ant PBAN genes, the S. carolinensis PBAN (Soc-PBAN) gene is the most different and distant from the other species based on nucleotide and peptide sequence homology, as well as having a different processing site for the signal peptide (Figs. 2, 3, Table 1). This phylogenetic classification by the neuropeptide sequence is consistent with the morphological cladistic analysis of Solenopsis genus (Pitts et al., 2005) that groups S. invicta and S. richteri in the S. saevissima complex and S. geminata in the S. geminata complex. Solenopsis pergandii and S. carolinensis belong to a very large group generally classified as thief ants that live primarily underground and steal brood and food from other ant species. This large group might be expected to exhibit correspondingly large variation; however, it is surprising that S. pergandii most closely resembles S. geminata, whereas S. carolinensis is distant from both the S. saevissima and S. geminata species groups, perhaps indicating that it is a more primitive species (Fig. 3).

PBAN and pyrokinin peptides have been found independently in various insect groups, but classified in the same peptide family because they were all characterized by a conserved penta-peptide (FXPRLamide) at the C-terminus. In spite of several different functions reported for these peptides during insect development, the mode of action for these peptides and their functions are still unclear. Thus far, only the regulation of pheromone biosynthesis by PBAN in lepidopteran moths has been elucidated (Rafaeli, 2009, Rafaeli and Jurenka, 2003). To date, PBAN from 29 species
in 13 families have been identified (Fig. 4): Lepidoptera (20 species in 9 families), Hymenoptera (6 species in 2 families), Coleoptera (1 species in 1 family), and Diptera (2 species in 1 family). Although the other PBAN family peptides beyond lepidopteran moths are only partially understood as to their physiological function, the pentapeptide epitopes in the C-termini are well conserved in the putative PBAN and additional four or three peptides. One of the most conserved motifs (WFGRPR) is recognized as DH homologues, which have been shown to induce embryonic diapause in the silkworm (Suwan et al., 1994) or accelerate pupal development in the moth (Xu and Denlinger, 2003). This common sequence has also been found in capability (CAPA) genes that encode multiple peptides containing PRV and WFGPRP epitopes at their C-termini (Predel and Wegener, 2006). The major difference in the structure of PBAN genes is that α-NP (VIIFPKL) is well conserved in all lepidopteran moths thus far studied, but absent in Hymenoptera, Coleoptera, and Diptera (Fig. 4). This shortest peptide has a similar sequence to leucopyrokinin peptides, known to stimulate contraction of hindgut muscles in cockroaches (Holman et al., 1986). However, it is not known if the peptide plays a similar role in lepidopteran moths. All α-NPs identified contain the FTPRP epitope at their C-termini, although it is slightly different in mosquitoes (Fig. 4). After the PBAN domains, γ-NPs are encoded for at the end of PBAN genes. They contain an FSPRL in C-termini, which is modified to PSPR in hymenopteran γ-NPs.

In case of the fire ant and honeybee, their PBANs do not have dibasic residues, K-K, K-R, and R-R, or single R basic amino acid residues in position -4, -6, or -8, which are known to act as potential cleavage sites during the pre-peptide process (Southey et al., 2008; Veenstra, 2000). The arginine(R) residue in two mosquito PBAN peptides also is not at a cleavage site, but is instead included in the PBAN amino acid sequence. The T. castaneum PBAN was predicted by peptidomics to have only 8-AA, the shortest PBAN domains sequence identified to date. As shown in Figure 4, there is a gap of amino acids between β-NP and PBAN of T. castaneum (Coleoptera) (HVVFQPRL-GRESGEEFVNNAPEDRVLQHETSGMLYQRSPFAPRL). This is because the basic residue, arginine(R) in Q8, is a predicted canonical cleavage site in the proteolytic processing (Li et al., 2008). When the entire sequence including a putative translation initiation site, indicated in italics, plus PBAN of T. castaneum is compared to hynmenopteran PBANs, they are approximately 40% identical, mostly in their C- and N-terminal regions. In Drosophila, the PK-2 (=hugin) gene has been suggested to correspond to the PBAN/pyrokinin genes of lepidopteran species and mosquitoes. There is an amino-acid-sequence gap between β-NP and γ-NP that does not contain FXPR-Like PBAN/pyrokinin family peptide (Bader et al., 2007; Choi et al., 2001). Diversity in neuropeptides, such as PBAN, between species as shown here could have evolved along with the functional diversity of animal endocellular systems.

Although the PBAN gene products, neuropeptides, are not characterized functionally except for lepidopteran moths, the phylogenetic relationships of all known PBAN genes were analyzed in the present study. The phylogenetic analysis of PBAN genes is similar to the taxonomic phylogeny based on the family level of classification and evolutionary tree of Insecta. This indicates that neuropeptide gene sequences can be used to infer insect phylogenetic relationships (Fig. 5). From the PBAN phylogenetic tree, the fire ant is most similar to the honeybee, another hymenopteran, followed by beetles (Coleoptera), mosquitoes (Diptera), and moths (Lepidoptera). Based on the amino acid sequences translated from PBAN genes from several insect families, the similarity is correlated with basic taxonomical or phylogenetic
classification of Insecta. The PBAN/pyrokinin gene is well conserved in Insecta, and probably extends to the entire Arthropod phylum; however, the translated prepropeptides might vary with functional diversity being retained or lost during the evolutionary process.

ACKNOWLEDGMENTS

We thank Drs. Clare Allen, Lee Cohnstaedt, and anonymous reviewers for helpful comments. We also thank D. Milne, M. Kwak, and C. Strong for technical support.

LITERATURE CITED


*Archives of Insect Biochemistry and Physiology*
silkworm *Bombyx mori*: characterization of the cDNA encoding the diapause hormone precursor and identification of additional peptides. Proc Natl Acad Sci USA 90:3251–3255.


