

## Article

# Identification of Novel Toxin Genes from the Stinging Nettle Caterpillar *Parasa lepida* (Cramer, 1799): Insights into the Evolution of Lepidoptera Toxins

Natrada Mitpuangchon <sup>1</sup>, Kwan Nualcharoen <sup>2</sup>, Singtoe Boonrotpong <sup>2</sup> and Patamarerk Engsontia <sup>2,\*</sup> 

<sup>1</sup> Princess Maha Chakri Sirindhorn Natural History Museum, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand; natrada.p@psu.ac.th

<sup>2</sup> Division of Biological Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand; kwan.n@psu.ac.th (K.N.); singtoe.b@psu.ac.th (S.B.)

\* Correspondence: patamarerk.e@psu.ac.th

**Simple Summary:** Many caterpillar species can produce toxins that cause harmful reactions to humans, varying from mild irritation to death. Currently, there is very limited knowledge about caterpillar toxin diversity, because only a few species have been investigated. We used the transcriptome technique to identify candidate toxin genes from the nettle caterpillar *Parasa lepida* (Cramer, 1799). It is a common pest of oil palm, coconut, and mango in South and South-East Asia, which can cause severe pain and allergic responses to those in contact with them. We reported 168 candidate toxin genes. Most of them are members of the toxin genes families commonly recruited in animal venoms such as serine protease and serine protease inhibitors. However, we identified 21 novel genes encoding knottin-like peptides expressed at a high level in the transcriptome. Their predicted 3D structures are similar to neurotoxins in scorpion and tarantula. Our study suggests that *P. lepida* venom contains diverse toxin proteins that potentially cause allergic reactions and pain. This study sheds light on the hidden diversity of toxin proteins in caterpillar lineage, which could be future fruitful new drug sources.

**Abstract:** Many animal species can produce venom for defense, predation, and competition. The venom usually contains diverse peptide and protein toxins, including neurotoxins, proteolytic enzymes, protease inhibitors, and allergens. Some drugs for cancer, neurological disorders, and analgesics were developed based on animal toxin structures and functions. Several caterpillar species possess venoms that cause varying effects on humans both locally and systemically. However, toxins from only a few species have been investigated, limiting the full understanding of the Lepidoptera toxin diversity and evolution. We used the RNA-seq technique to identify toxin genes from the stinging nettle caterpillar, *Parasa lepida* (Cramer, 1799). We constructed a transcriptome from caterpillar urticating hairs and reported 34,968 unique transcripts. Using our toxin gene annotation pipeline, we identified 168 candidate toxin genes, including protease inhibitors, proteolytic enzymes, and allergens. The 21 *P. lepida* novel Knottin-like peptides, which do not show sequence similarity to any known peptide, have predicted 3D structures similar to tarantula, scorpion, and cone snail neurotoxins. We highlighted the importance of convergent evolution in the Lepidoptera toxin evolution and the possible mechanisms. This study opens a new path to understanding the hidden diversity of Lepidoptera toxins, which could be a fruitful source for developing new drugs.

**Keywords:** caterpillar envenomation; urticating dermatitis; venom evolution; venomics



**Citation:** Mitpuangchon, N.; Nualcharoen, K.; Boonrotpong, S.; Engsontia, P. Identification of Novel Toxin Genes from the Stinging Nettle Caterpillar *Parasa lepida* (Cramer, 1799): Insights into the Evolution of Lepidoptera Toxins. *Insects* **2021**, *12*, 396. <https://doi.org/10.3390/insects12050396>

Academic Editor: Michael Kristensen

Received: 6 April 2021

Accepted: 26 April 2021

Published: 29 April 2021

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



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Various animal species have evolved venoms used primarily for defense, predation, and competition [1,2]. Venoms are usually a cocktail of diverse peptides, proteins, and other organic compounds, which have multiple physiological effects on the recipients'

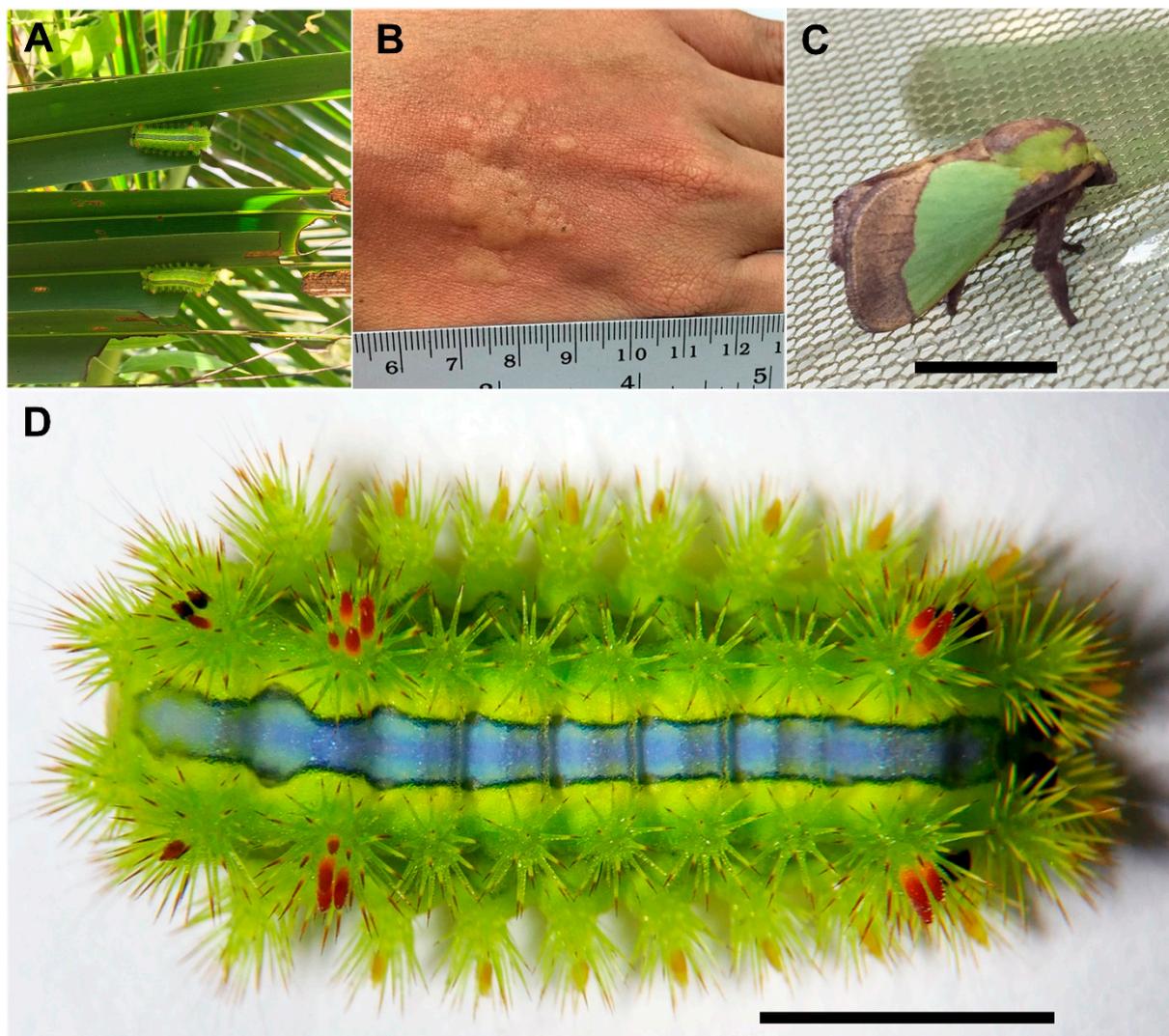
body as a result of millions of years of natural selection [3]. Venom toxins have long been a subject of interest due to their effects on human health, such as causing severe pain, allergic reactions, organ system failure, and death [4,5]. In contrast, many of them have potential pharmaceutical (e.g., anti-cancer, antimicrobial, neurodegenerative disorder treatment, and analgesic drugs) and pest control applications (e.g., insecticide for agriculture and mosquito-killing) [6–11].

Lepidoptera (moth and butterfly), with more than 150,000 described species, represent another lineage of animal that can produce toxins [12], of which more than 70 species from 15 families are medically important, which cause varying adverse effects to humans [13]. Unlike other venomous insects, the caterpillar does not have a specialized venom gland [14]. The secretory epithelial cells synthesize venoms and accumulate them inside the hollow canal of the injecting organs, such as urticating hairs, spines, and bristles [15–17]. These organs usually have sharp tips that can pierce the predators' skin, breaking and releasing venom into the contact area [2,18]. Caterpillar venom varies in its effect from mild to severe. Local effects include contact dermatitis, e.g., allergic reaction, edema, erythema, and local tissue damage; systemic effects include hemostatic disturbance, acute kidney injury, intracerebral bleeding, and paraneoplastic-associated phalangeal peri-arthritis [15,18–20].

Currently, knowledge about the diversity and evolution of Lepidoptera toxins is still under-represented. The giant silkworm moth, *Lonomia obliqua* (Saturniidae), has been comprehensively studied most due to its venom's lethal effects. The essential toxin proteins are Lopap (prothrombin activator) and Losac (factor X activator) proteins, which interfere with blood coagulation, leading to hemorrhagic syndrome [21–23]. Proteomic approaches and enzymatic assays have been applied to study toxins from other caterpillars from family taxa, including Saturniidae (*Lonomia achelous*, *Leucanella memusae*) [22,23], Limacodidae (*Lotoia consocia*) [16], Lymantriidae (*Euproctis chryssorrhoea*, *E. subflara*) [24], Notodontidae (*Thaumetopoea pityocampa*) [25], Erebidae (*Premolis semirufa*) [26,27], and Megalopygidae (*Podalia ca. fuscescens*) [23]. Various toxin classes, including serine proteases, C-type lectin, phospholipase A2, hyaluronidase, and protease inhibitors, were reported. In *T. pityocampa*, the major toxin proteins are IgE-mediated allergens, including unrelated proteins Tha p 1 and Tha p 2 [25,28]. Further information on the Lepidoptera toxin can be found in the previous excellent review [18].

We noted that the enzymatic assay used in most previous studies could not reveal the true diversity of toxins in each caterpillar species, because this method selects only a few candidate enzymes for the analysis. Recently, high-throughput sequencing technology, i.e., RNA sequencing, has been applied to study toxin genes in *Premolis semirufa* and identify more than 400 putative toxin genes [27]. This technique also provides the DNA and protein sequence and expression level of the candidate toxins genes. Thus, RNA sequencing is a promising tool to study Lepidoptera toxin diversity.

The stinging nettle caterpillar, *Parasa lepida*, is a common pest in South-East Asia, South Asia, and Southern China that feeds on leaves of oil palm, coconut, and mango trees [29,30]. The caterpillar possesses urticating hairs protruding from the dorsal side along its body's longitudinal axis (Figure 1). It is recognized as one of the most harmful caterpillars in Thailand, causing intense pain, burning and itching sensations, and severe dermatitis [18]. Despite its adverse effects on the health of many farmers, its toxin has not been investigated. Similarly, little is known about toxins from other genera of Limacodidae, e.g., *Acharia*, *Adoneta*, *Darna*, *Euclaea*, *Isa*, *Thosea*, *Phobetron*, and *Sibine* [18].



**Figure 1.** The stinging nettle caterpillar, *Parasa lepida*. (A) Caterpillars feeding on coconut leaves, (B) dermatitis symptom on a hand after exposure to caterpillar, (C) adult stage, and (D) final instar caterpillar with urticating hairs on its back (scale bar = 1 cm).

This study aimed to identify putative toxin genes from the urticating hair transcriptome of the *P. lepida* using the RNA-sequencing technique and to investigate some aspects of the gene family evolution. Our findings serve as primers to understand the hidden diversity and the evolution of lepidopteran toxin genes.

## 2. Materials and Methods

### 2.1. Caterpillars and RNA Isolation

The *Parasa lepida* caterpillars were collected from a local coconut farm in the Tapa district, Songkhla province, Thailand. They were brought to the laboratory and fed with fresh coconut leaves until used. To isolate the total RNA from the venom glands, final instar caterpillars were anesthetized in dry ice and placed under a stereomicroscope. The urticating hairs on the dorsal side (all clusters) (Figure 1D) were dissected at the bases. The total RNA was isolated from these tissues (~100 mg) using GENEzol™ reagent (Geneaid) following the manufacturer's protocol. RNA quantity and its integrity were assessed using NanoDrop™ and gel electrophoresis, respectively.

## 2.2. RNA Sequencing and Transcriptome Analysis

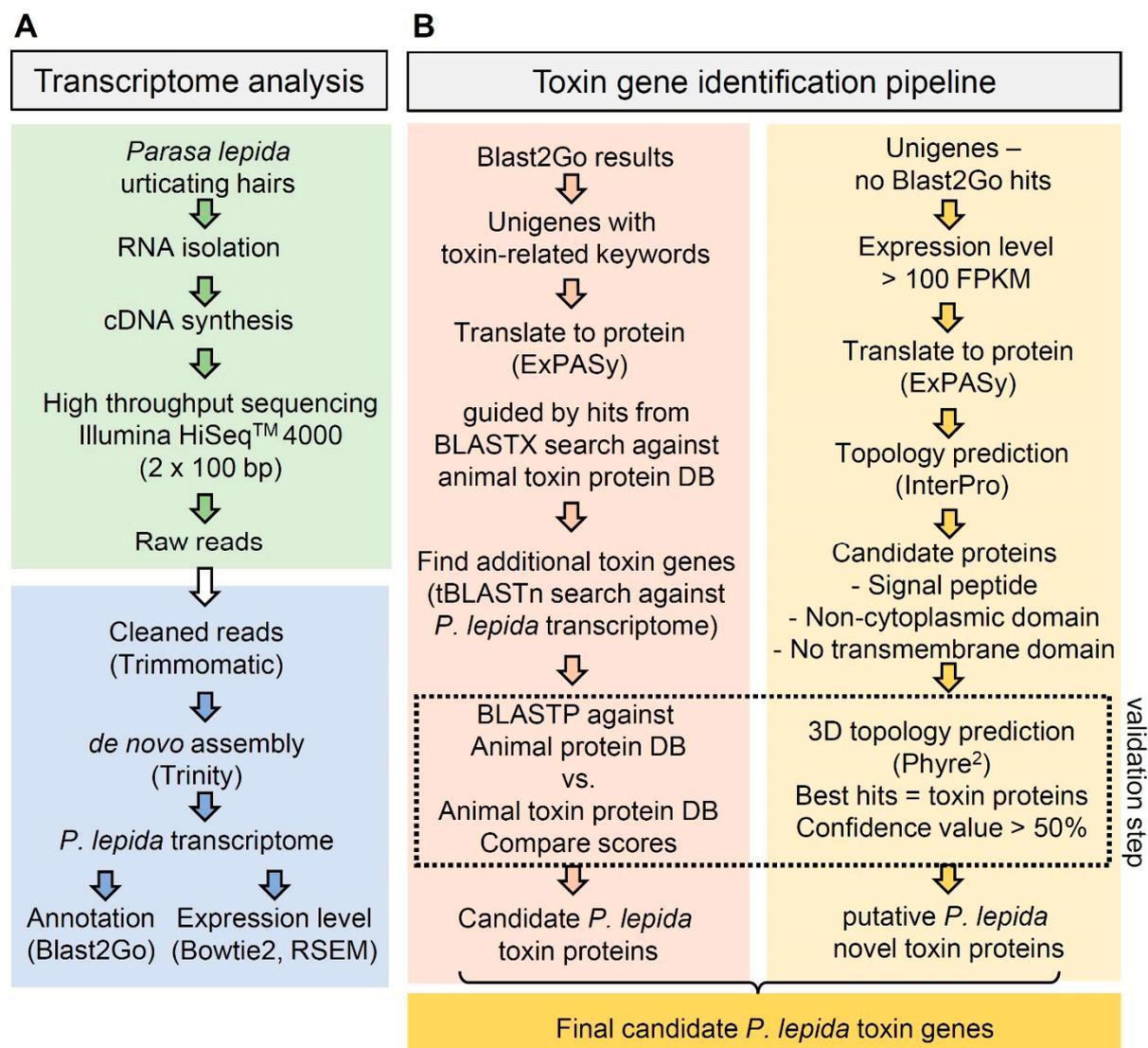
The RNA sample was sent to BGI TECH SOLUTIONS (Hongkong) for the cDNA library preparation and high-throughput sequencing. In brief, RNA quality was re-evaluated using an Agilent 2100 Bioanalyzer (concentration ~ 800 ng/ $\mu$ L; RIN = 7.3). The mRNA was selected using magnetic beads with Oligo (dT) and fragmented using a fragmentation buffer. The cDNA was synthesized using the RNA fragments as templates and was then ligated with adapters. Sequencing was carried out on an Illumina HiSeq<sup>TM</sup> 4000 (paired-ended reads,  $2 \times 100$  bp).

Low-quality reads (more than 20% of the base's qualities are lower than 10, N bases more than 5%, reads with adapters) were removed using Trimmomatic v.0.36 [31]. Cleaned reads were used for de novo assembly using Trinity v.2.0.6 [32]. To remove redundant sequences, assembled transcripts were clustered to a final set of Unigenes using TGICL v.2.0.6 [33]. The Unigenes are divided into two types; one type is a cluster with the prefix "CL", followed by the cluster-id (each cluster may have several Unigenes with a similarity between them being more than 70%). Another type, the singleton, has the prefix "Unigene".

Gene annotation of the whole transcriptome was conducted using Blast2GO v.2.5.0 [34] based on BLAST search against nucleotide and protein databases (NT, NR, KOG, KEGG, and SwissProt), gene ontology annotation, and InterPro annotation. To assess expression levels of each Unigene, cleaned reads were mapped to Unigene using Bowtie2 v.2.2.5 [35]. Expression levels were calculated into FPKM units using RSEM v.1.2.12 [36]. The full parameters used to run each program are summarized in Supplementary Data S1.

## 2.3. Identification of Toxin Genes

From the Blast2GO results, Unigenes with toxin-related keywords in the BLAST hits or InterProt prediction, e.g., lopap, serine protease, protease inhibitor, serpin, venom, toxin, allergen, CAP, and phospholipase A2 [37], were selected as candidate *P. lepidata* toxin genes. These Unigenes were used as queries for the BLASTX search against our custom toxin protein database (13,124 proteins, in total), containing animal toxin proteins from the UniProt database (Animal toxin annotation project; both reviewed and nonreviewed—downloaded on the 1 May 2018), recently reported toxin proteins from Crustacea [38], and toxin proteins from *Lonomia obliqua* [21,39]. BLAST results were used to identify the translation frame of *P. lepidata* Unigenes. DNA sequences were translated to proteins using the ExPASy translation tool (<https://www.expasy.org> (accessed on 5 May 2018)). Proteins from the correct translation frame were selected. The completeness of proteins was noted using a suffix in the gene name: Full = full gene model, NTE = N terminus missing, CTE = C terminus missing, and NC = N and C terminus missing. These candidate toxin proteins were used as queries for the tBLASTn search against the *P. lepidata* transcriptome to find additional candidate toxin genes that have not been identified by Blast2GO. Additional proteins were added to the candidate toxin dataset of *P. lepidata*. To remove potentially nontoxic orthologs, all candidate *P. lepidata* toxin proteins were used as queries for the local-BLASTP search (e-value =  $1 \times 10^{-5}$ ) against two databases: (1) animal-reviewed proteins in the UniprotDB (558,125 proteins—downloaded on the 8 October 2018) and (2) our custom toxin protein database, as explained previously (Figure 2). Unigenes were filtered out from the candidate toxin gene dataset if the BLAST score of the best hit from the animal protein database was higher than the animal toxin database results. However, if the best hits were from venomous or poisonous animals, the sequences were still kept in the dataset. This guarantees that the best BLAST hits for all candidate *P. lepidata* toxin proteins were known toxin proteins or proteins from venomous and poisonous animals.



**Figure 2.** Flow chart of the *P. lepidata* toxin gene identification pipeline: (A) transcriptome analysis—green box = RNA isolation and sequencing, blue box = transcriptome analysis; (B) toxin gene identification pipeline: identification based on orthologous relationships (left) and identification of novel toxin genes (right).

All candidate toxin proteins of *P. lepidata* were then separately evaluated by the protein family. They were aligned with proteins in the same family from other arthropod species (Supplementary Data S2) using MAFFT v.7. Proteins from *P. lepidata*, without a conserved region, were discarded. Some proteins from two Unigenes were merged if they showed overlapping regions in the alignment, and the suffix ‘Join’ was added to the gene name. The important motifs and signatures of each protein were confirmed by InterProScan [40]. Candidate toxin genes were finally confirmed by phylogenetic relationships with previously reported toxin proteins.

To further identify putative novel toxins in *P. lepidata*, which do not show significant similarity to any known protein, we selected 217 Unigenes from the *P. lepidata* transcriptome that do not have hits to any database (nonredundant, nt, SwissProt, KEGG, InterPro, and GO) but have relatively high expression levels (within the top 2.5% of the highest expressed genes; expression level ranging from 100 to 24,515 FPKM) for the screening. Each sequence was inspected manually. The DNA sequences were translated using the ExPASy Translate tool, and protein domains were predicted using InterProScan. Proteins with venom protein signatures, including signal peptides, the noncytoplasmic domain, and

without a transmembrane domain, were kept. These proteins were submitted to Phire<sup>2</sup> (Protein Homology/analogy Recognition Engine V. 2.0) for the 3D structure prediction. Proteins were considered as putative novel toxins of *P. lepidida* if their predicted 3D structures were most similar to the reported toxin proteins with a confidence value higher than 50%.

#### 2.4. Gene Orthology and Phylogenetic Analysis

We analyzed whether some families of *P. lepidida* candidate toxin genes have undergone expansion or contraction by comparing the number of orthologous genes from other Lepidoptera species. To identify orthologous genes, we used OrthoVenn2 [41] to search against available lepidopteran genomes (*Danaus plexippus*, *Heliconious melpomene*, *Malitaea cinxia*, and *Bombyx mori*) and a toxin gene dataset from *Lonomia obliqua* [21,39] using *P. lepidida* genes as queries (E-value =  $1 \times 10^{-5}$ ; inflation value = 1.5). Then, we used the chi-square test to examine whether the number of genes in some toxin gene families of *P. lepidida* deviates from a 1:1 ratio, which indicates gene family expansion or contraction.

Phylogenetic analyses of individual toxin families were conducted to reconstruct the evolutionary relationship of these toxin proteins within the arthropod lineage. *P. lepidida* proteins were aligned with toxin proteins from other arthropods, including Chelicerata, Myriapoda, Crustacea, and other insects retrieved from the NCBI UniProt database (Supplementary Data S2). Sequence alignment was carried out using MAFFT version 7.0 followed by removing gappy regions using TrimAl (<http://trimal.cgenomics.org> (accessed on 5 May 2018)). Trees were made using PhyML with Akaike Information Criterion automatic model selection [42]. Support values for branches were analyzed using 1000 replications bootstrapping. Trees were illustrated by FigTree version 1.4.4 (<http://tree.bio.edu.ac.uk/software/figtree> (accessed on 5 May 2018)).

### 3. Results

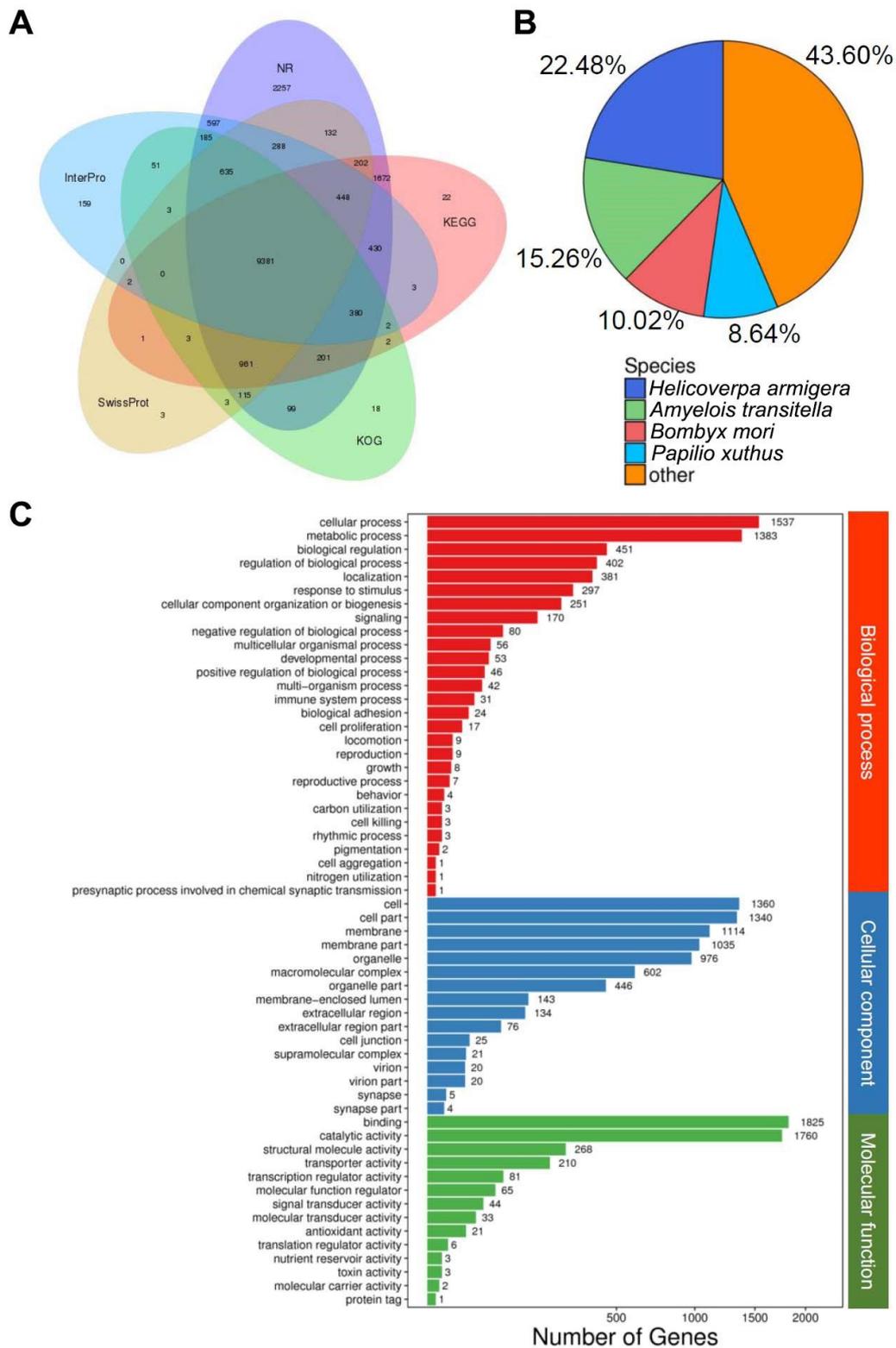
#### 3.1. Transcriptome Assembly and Gene Annotation

Illumina HiSeq4000 sequencing generated about 68 million reads and 6.8 billion bases in total (Table 1). After de novo assembly and filtering the redundant sequences, we received 34,968 Unigenes for the downstream analyses. The total length, average length, N50, and GC content of the analyzed Unigenes were 27,807,171 bp, 795 bp, 1315 bp, and 36.73%.

**Table 1.** Information of sequence reads generated by RNA-seq, and Trinity de novo-assembled contigs.

Sequence Analysis	Counts	Length (bp)
Raw sequence reads acquired	67,876,696	6,787,669,600
Sequence reads after QC	66,361,788	6,636,178,800
Contigs after trinity assembly	48,761	34,507,843
Unique contigs (Unigene)	34,968	27,807,171
GC content of Unigene	36.73%	
Average length		795
N50		1315

All Unigenes were annotated by searching against seven available databases using Blast2GO (BioBam<sup>®</sup>). In summary, 17,983 (51.43%) and 6175 Unigenes (17.66%) were annotated from nonredundant nucleotide and protein databases from NCBI, respectively. In addition, 12,177 (34.82%), 12,039 (34.43%), 13,710 (39.21%), 4107 (11.75%), and 12,564 (35.93%) were annotated from Swissprot, KOG, KEGG, and InterPro databases, respectively (Figure 3A). Species of the best matching BLAST hits were *Helicoverpa armigera* (22.48%), *Amyelois transitella* (15.26%), and *Bombyx mori* (10.02%) (Figure 3B).



**Figure 3.** Overview of the BLAST2GO results: (A) intersection of different databases and numbers of contigs, (B) species distribution of best BLAST hits, and (C) genes classified by different categories (biological process, cellular component, and molecular function) and number of genes.

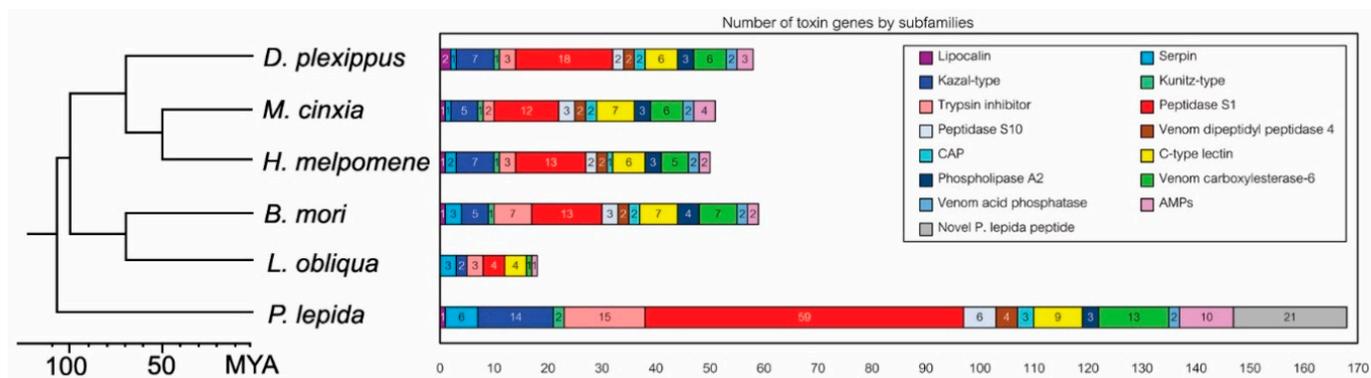
Gene ontology analysis indicated the expression of diverse proteins in the urticating hairs of the *P. lepida*. They were classified into three main categories: biological process, cellular component, and molecular function. The top three subgroups for each category

in terms of the number of genes were as follows: (1) cellular process, metabolic process, biological regulation (biological process) (1537, 1383, and 451 Unigenes, respectively); (2) cell, cell part, and membrane (cellular component) (1360, 1340, and 1114 Unigenes, respectively); and (3) binding, catalytic activity, and structural molecular activity (molecular function) (1825, 1760, and 268 Unigenes, respectively) (Figure 3C). The full annotation results of the *P. lepida* urticating hair transcriptomes are reported in Supplementary Data S3.

### 3.2. Candidate Toxin Genes of *Parasa lepida*

Following our toxin gene identification pipeline, we reported 168 candidate toxin genes classified into four main groups and 15 subgroups. The first three main groups were conventional toxin classes that share homologies with known toxins from other venomous animals: (1) protease inhibitor (lipocalin, serpin, Kazal-type protease inhibitor, Kunitz-type protease inhibitor, and trypsin inhibitor), (2) proteolytic enzyme (peptidase S1, peptidase S10, and venom dipeptidyl peptidase-4), and (3) other toxins (CAP (CRISPs, Ag5, and Pr-1), C-type lectin, phospholipase A2, venom carboxylesterase6, venom acid phosphatase, and antimicrobial peptides). The fourth group involved the candidate novel toxin proteins from *P. lepida* that do not share homologies to any sequence in the NCBI database. All of these proteins contain a Knottin-like domain, thus referred to here as a Knottin-like peptide. Full data of these candidate toxin proteins are reported in the Supplementary Data S4.

The largest toxin group was a serine protease, which has 59 genes in total, accounting for 35% of all toxin genes identified in this study, followed by the *P. lepida* Knottin-like peptide (21 genes), trypsin inhibitor (15 genes), Kazal-type protease inhibitor (14 genes), venom carboxylesterase-6 (13 genes), antimicrobial peptide (10 genes), C-type lectin (9 genes), peptidase S10 (6 genes), serpin (6 genes), venom dipeptidyl peptidase 4 (4 genes), CAP (3 genes), phospholipase A2 (3 genes), venom acid phosphatase (2 genes), Kunitz-type protease inhibitor (2 genes), and lipocalin (1 gene) (Figure 4).



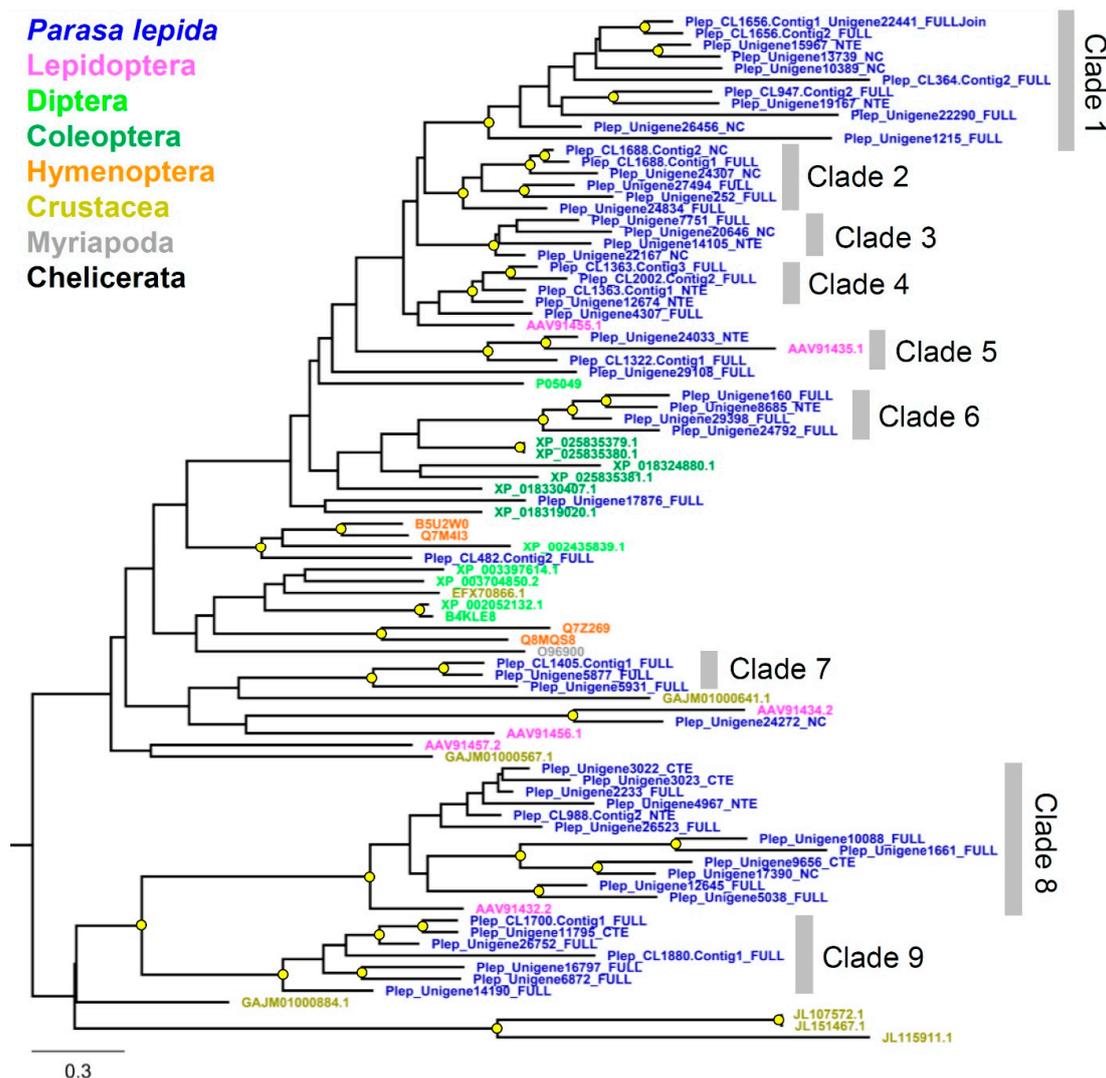
**Figure 4.** Number of toxin genes by subfamilies of *P. lepida* and the number of their orthologous genes presented in other Lepidopteran genomes (*Danaus plexippus*, *Melitaea cinxia*, *Heliconius melpomene*, and *Bombyx mori*) and cDNA library (*Lonomia obliqua*).

### 3.3. Gene Family Expansion

We identified orthologous genes of the *P. lepida* candidate toxin genes in other Lepidopteran species using OrthoVenn2 (E-value =  $1 \times 10^{-5}$ ; inflation value = 1.5) (Figure 4 and Supplementary Data S5). The number of genes for each toxin gene family varied between species; however, comparing the four species with genome data, the number of orthologous genes was in a ratio of about 1:1 for all gene families. The number of genes in *L. obliqua* was fewer than in other species, most possibly due to incomplete annotation. The number of genes in subfamily Peptidase S1 in *P. lepida* was much higher (59 genes) than in other species (12–18 genes), suggesting lineage-specific expansion. We performed a chi-square test on gene subfamilies with more than five genes in all species (Venom carboxylesterase-6, C-type lectin, and Peptidase S1). The number of genes in the peptidase S1 subfamily significantly deviated from the 1:1 ratio ( $\chi^2 = 71.391$ ,  $df = 4$ ,  $p$ -value < 0.0001),

but this was not the case for venom carboxylesterase-6 and C-type lectin ( $\chi^2 = 5.568$ ,  $df = 4$ ,  $p$ -value = 0.2339;  $\chi^2 = 0.857$ ,  $df = 4$ ,  $p$ -value = 0.9306, respectively).

Phylogenetic relationships of *P. lepida* Peptidase S1 genes revealed multiple lineage-specific expansions (Figure 5), explaining this subfamily’s large size. Genes are divided into nine clades. The expansions were observed in clades 1, 8, and 9 (11, 12, and 7 genes, respectively). These genes’ closest relatives are serine proteases from *L. oblique* (AAV91435.1, AAV91432.2, and AAV91434.2).

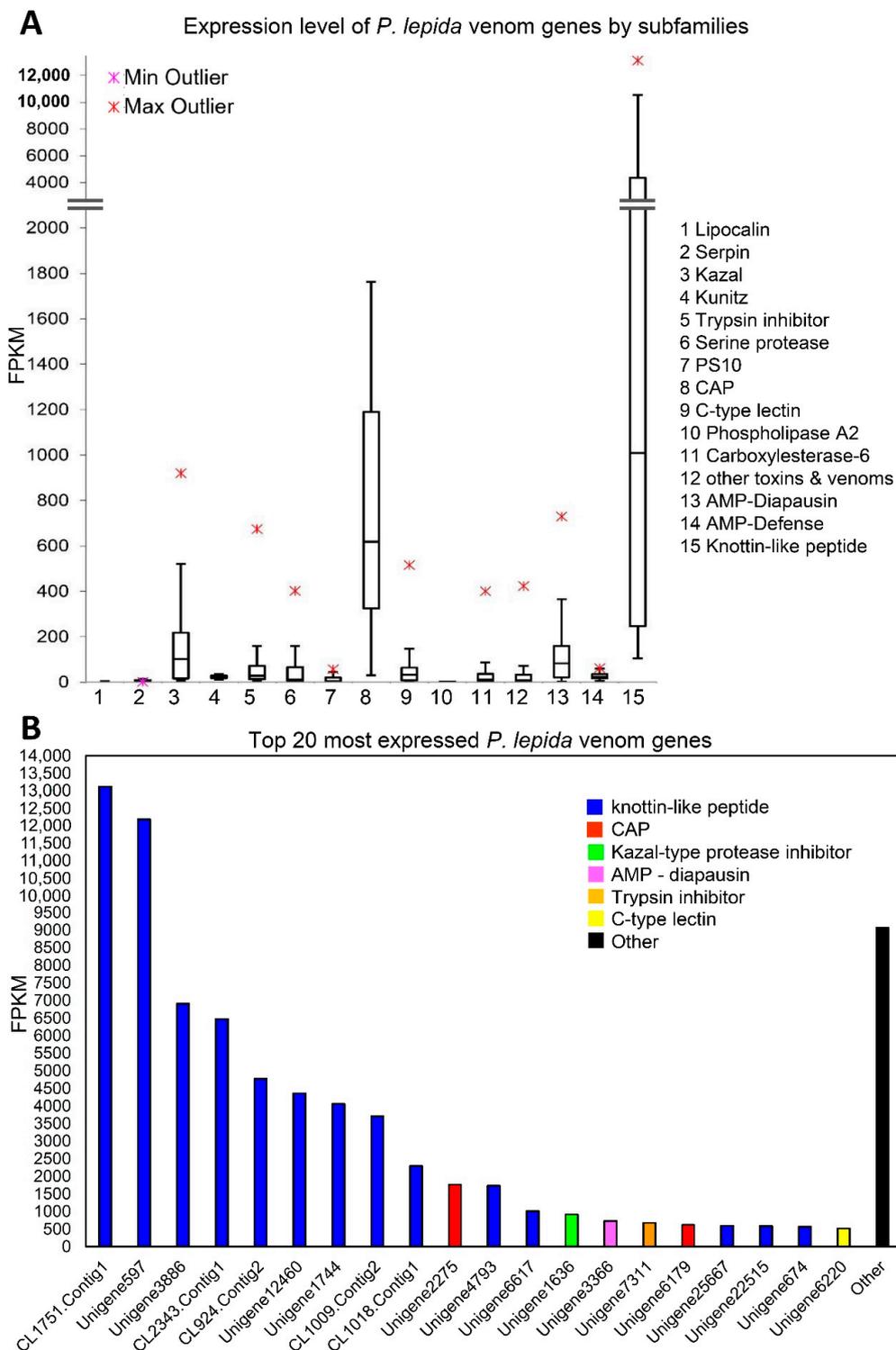


**Figure 5.** Phylogenetic tree of the serine protease gene family from *Parasa lepida* and other Arthropods. Clades were defined based on phylogenetic relationships with branch support. The tree was constructed using the maximum likelihood method by PhyML 3.0. A branch support value higher than 70% (1000 bootstrap replications) is shown with a yellow circle. The tree was rooted at the midpoint.

### 3.4. The Expression Level of the Candidate *Parasa lepida* Toxin Genes

Expression levels of genes were estimated by mapping reads to assembled contigs. The numbers of mapped reads (counts) was normalized by contig length and total mapped reads. The expression levels among candidate toxin genes varied between 0.38 and 13,119.23 FPKM (Fragments Per Kilobase of transcript per Million mapped reads). The top three most expressed families (according to median) were Knottin-like peptides, CAP, and Kazal (Figure 6A). Interestingly, nine out of ten most expressed toxin genes were novel Knottin-like peptides (Figure 6B). The expression levels of the 21 novel Knottin-like toxins

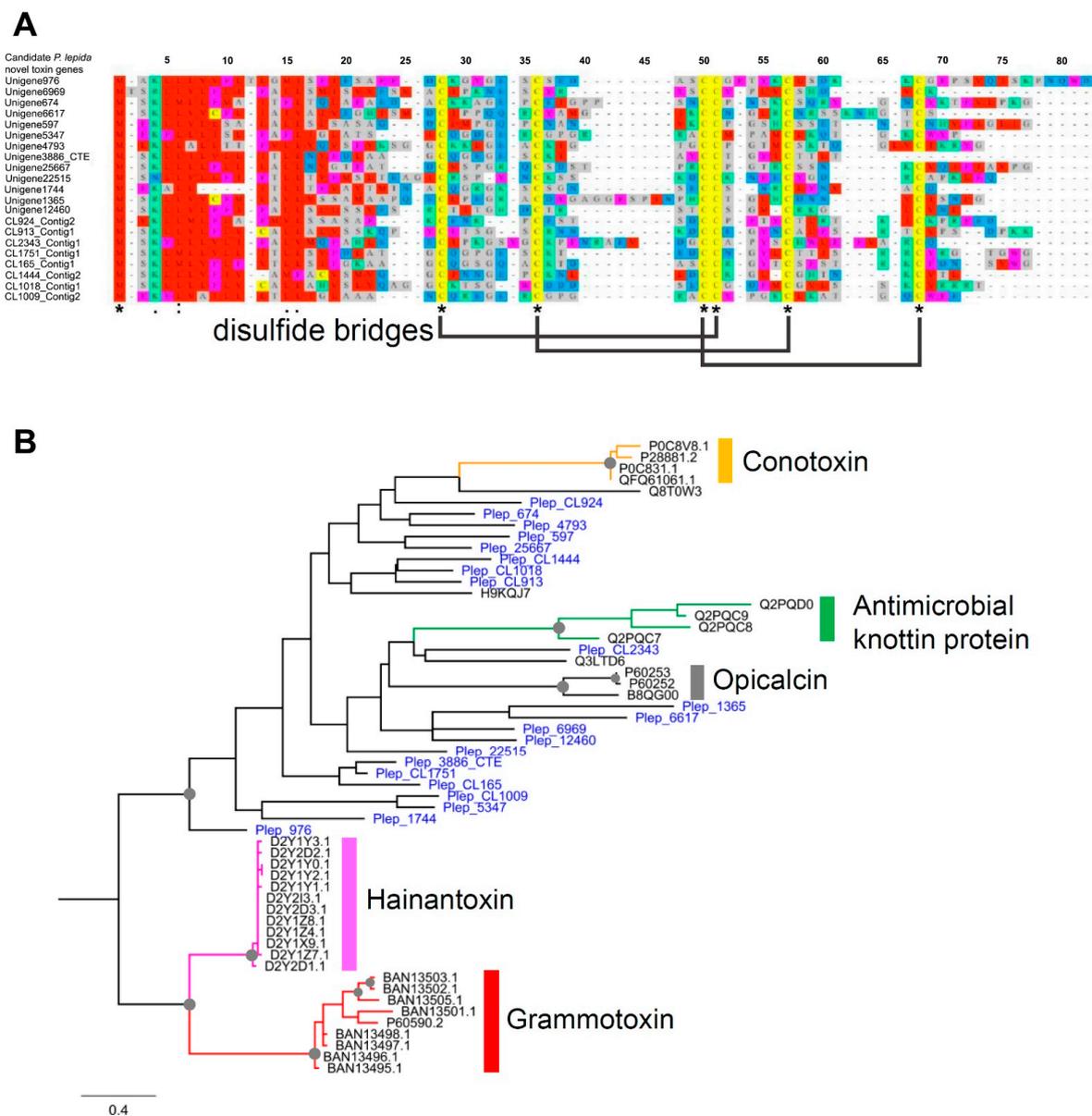
were also within the top 0.5% most expressed genes in the *P. lepida* urticating hair transcriptome. Other genes in the top 20 were CAP, Kazal-type protease inhibitor, AMP-diapausin, trypsin inhibitor, and C-type lectin (Figure 6B).



**Figure 6.** (A) Box-and-Whisker plot showing expression level (FPKM) of *P. lepida* venom genes by subfamily. Boxes show the upper quartile and lower quartile. The line across the box indicates the median. The maximum and minimum values (excluding outliers) are shown above and under the box, respectively. Upper and lower outliers are indicated using red and pink symbols, respectively. (B) The expression level of the top 20 most expressed *P. lepida* toxin genes.

### 3.5. Putative Novel Toxin Proteins from *P. lepida*

We identified 21 novel toxin peptides from *P. lepida*. Either DNA or protein sequence shares a similarity to known sequences in the NCBI database. Their sizes vary from 46 to 71 amino acids ( $\bar{x} \pm SE = 56 \pm 1.19$ ). These proteins' pairwise distances revealed a wide range of sequence diversity as the percent identity varies between 17.33 and 73.33%. These proteins have a signal peptide and a noncytoplasmic domain. Protein alignment revealed six conserved cysteine residues (Figure 7A). These conserved domains were predicted from the three disulfide bridges structure of the inhibitor cystine knot motif. Phire2 (Protein fold recognition server) predicted the 3D structures of these proteins similar to the neurotoxins from spiders, scorpions, and cone snails such as hptx2, ptu-1, hainantoxin-iii, actx-2, and omega-conotoxin (Supplementary Data S4).



**Figure 7.** (A) Protein alignment of novel Knottin-domain protein from *P. lepida* showing six conserved cysteine residues and putative three disulfide bridges; (B) phylogenetic tree of Knottin-like proteins from *P. lepida* and other animals: Conotoxin—*Conus achatinus* (Little frog cone), Antimicrobial Knottin protein—*Bemisia tabaci* (Silverleaf whitefly), Opicalcin—*Opisthophthalmus carinatus* (African yellow leg scorpion), Hainantoxin—*Cyriopagopus hainanus* (Chinese bird spider), Grammotoxin—*Grammostola rosea* (Chilean rose tarantula). Branch supports higher than 70% (1000 replications bootstrapping) are highlighted with grey circles.

Phylogenetic analysis of these 21 *P. lepida* novel toxins and other Knottin domain-containing toxins from other animals, including spiders, scorpions, insects, and cone snails, revealed nonmonophyletic relationships of *P. lepida* proteins (Figure 7B). These proteins are more likely to arise from convergent evolution based on sequence divergence and unclear relationships with other animals' toxins. Besides, the long branches in a phylogenetic tree suggested that they are not the product of recent gene duplication or lineage-specific expansion.

## 4. Discussion

### 4.1. Predicted Functions of *P. lepida* Toxins

#### 4.1.1. Knottin-Like Peptides

We predicted that Knottin-like peptides play significant roles in the *P. lepida* venom, and they are neurotoxins. This is because they showed the highest expression genes among *P. lepida* toxin gene families (Figure 6) and were unique to *P. lepida*, suggesting that they are the product of the lineage-specific adaptation in *P. lepida* venom evolution. Their predicted 3D structures are similar to Knottin-containing domain neurotoxins from distantly related animals (Supplementary Data S4). For example, conotoxins from cone snail and grammotoxins from tarantula block voltage-gated  $\text{Ca}^{2+}$  channels and cause paralysis [43,44]. Other neurotoxins are hainantoxins from the spider that can inhibit voltage-gated  $\text{Na}^{2+}$  channels, and opicalcins from the scorpion that can activate  $\text{Ca}^{2+}$  release channel/ryanodine receptors, which also cause paralysis [45–47]. Knottin proteins ( $\delta$ -theraphotoxin-Hm1a and -Hm1b) in *Heteroscodra maculate* tarantula venom activate the voltage-gated  $\text{Na}^{2+}$  channels nociceptor, which elicits the pain sensation [48]. A recent study demonstrated that short peptides from the *Parasa (Latoia) consocia* caterpillar play essential roles in the caterpillar venom by activating the nociceptive ion channel TRPV1, which causes the burning pain sensation [16]. However, the proteins have not yet been fully characterized, so we cannot conclude whether they are orthologous proteins. Further investigations are required to confirm whether the novel Knottin-like proteins play a crucial role in *P. lepida* venom by activating the recipients' pain receptors. Interestingly, a recent transcriptomic study in the Brazilian moth, *Premolis semirufa*, also identified putative knottin-like toxins (0.7% of all candidate toxins), which have protein sequences similar to neurotoxins from spiders [27]. This may reflect the convergent evolution of knottin-like toxins in venomous Lepidoptera, warranting further investigation.

#### 4.1.2. CAP Family

The second-highest expression belongs to the CAP family (Figure 6). This family has three subclasses: Cysteine-rich secretory proteins (CRISPs), Antigen 5 (Ag5), and Pathogenesis-related 1 proteins (Pr-1) [49]. All of the three genes identified from *P. lepida* are in the Ag5 subclass. In Hymenoptera, CAP proteins are the primary allergen that activates IgE, causing an extreme allergic reaction [50–52]. In snake venom, CRISPs inhibit smooth muscle contraction [53], and in cone snail venom, CRISPs act as protease [54].

#### 4.1.3. Venom Dipeptidyl Peptidase IV

Venom dipeptidyl peptidase IV is an enzyme that removes N-terminal dipeptide from the polypeptide [55]. We identified four venom dipeptidyl peptidase IV genes from the *P. lepida*. Their roles include activating an immune response, initiating the synthesis of antimicrobial peptides, and activating immune cells' movement [55,56]. This protein is a significant allergen in *Vespa vulgaris* and *Polistes dominula* venom and it has been reported in other animal venoms, including snakes, bees, and ants [55,57,58].

#### 4.1.4. Venom Acid Phosphatase

Venom acid phosphatase hydrolyzes phosphomonoester at acidic pH. We identified two venom acid phosphatase genes in the *P. lepida* transcriptome. It is a significant allergen in bee venoms that can induce wheal and flare reactions in humans by releasing

histamine [59,60]. They can also activate cell histolysis and tissue degeneration [61]. They are found in various insects, particularly from the order Hymenoptera [62–64].

#### 4.1.5. Carboxylesterase-6

Carboxylesterase-6 is a hydrolyzing enzyme that cuts carboxylic esters. We identified 13 carboxyl esterase-6 genes in the *P. lepida* transcriptome. Carboxylesterase-6 is the significant allergen in hymenopteran insects' venoms [59,65–67]. They also have roles as an odorant degrading enzyme in the insect olfactory sensilla [68].

#### 4.1.6. Serine Proteases

Serine proteases may also play essential roles in *P. lepida* venom, as observed from a large number of genes (59 genes), which is about four times higher than the number of orthologous genes shared among lepidopteran species (14 genes) (Figure 4). Serine proteases are the most common class of proteins presented in venoms of various animals, including many arthropods (e.g., centipede, Remipedia, bee, wasp, and tick), mollusk, annelid, and vertebrates (e.g., reptile, platypus, and vampire bat) [37,38,69–73]. They have diverse physiological effects, including blood coagulation, blood vessel relaxation, contraction of smooth muscles, initiation of pain, suppression of immunity, inflammation, and melanization [74,75].

#### 4.1.7. Lipocalins

Lipocalins are small extracellular proteins that form a large protein family. We identified one candidate *P. lepida* lipocalin toxin gene. Lipocalins have diverse physiological roles, including transporter (retinol, odorant, and pheromone), color mimicry, prostaglandin synthesis, cell homeostasis, and immunity [76]. Some of them are allergenic proteins causing severe anaphylaxis in humans [77]. They also function as an anticoagulant in the toxins of snakes, bats, and the saliva of ticks, kissing bugs, and mosquitoes [3,37,78,79]. Lopap of *L. oblique* is a lipocalin with serine protease properties. It functions as a prothrombin activator, which causes a severe consumptive coagulopathy [80].

#### 4.1.8. Serine Protease Inhibitor (Serpins)

Serpins inhibit serine protease by permanently altering its structure, and they can inhibit other enzymes such as caspases and papain-like cysteine protease [81]. This study reported six serpin genes in the *P. lepida* transcriptome. Serpin is a large and diverse protein family and it is present in the venom of *Lonomia obliqua*, *Leucanella memusae*, and *Podalia ca. fuscescens* caterpillar [23]. Serpin regulates protein digestion in the crucial biological process, e.g., blood coagulation, inflammation, and immune response [37,81]. Serpins are present in many animal venoms, including platypus, snake, cnidarian, leech, Remipedia, centipede, caterpillar, wasp, tick, and mosquito [37,81–83].

#### 4.1.9. Kazal-Type Protease Inhibitor

The Kazal-type protease inhibitor is a serpin containing one or more Kazal domains. These proteins inhibit serine protease, e.g., thrombin, trypsin, factor XIIa, subtilisin A, elastase, chymotrypsin, and plasmin [84]. We identified 14 Kazal-type protease inhibitors from *P. lepida*. Kazal-type protease inhibitors are present in the venoms of snake, bat, leech, jellyfish, tick, and many insects, e.g., caterpillar, termite, bee, wasp, mosquito [37,74,83,85].

#### 4.1.10. Kunitz-Type Protease Inhibitor

The Kunitz-type protease inhibitor is a serpin containing the Kunitz domain. We identified two Kunitz-type protease inhibitor genes from *P. lepida*. These proteins can inhibit trypsin, chymotrypsin, as well as potassium and voltage-gated sodium channels [86,87], which interfere with biological activities, such as blood coagulation, inflammation, and pain suppression [86,88,89]. This protein is also present in the venom of snake, frog, leech,

sea anemone, cone snail, scorpion, spider, tick, and many insects, including bees and wasps [37,90,91].

#### 4.1.11. Trypsin Inhibitor

The trypsin inhibitor is another serpin that inhibits trypsin, which functions as an irreversible competitive substrate and can impede blood coagulation [92]. We identified 15 trypsin inhibitor genes from *P. lepida*. They occur in the venom of *L. obliqua* [21] and scorpions, e.g., BmKAPi in *Tityus bahiensis*, *T. obscurus*, and *T. serrulatus* [92,93].

#### 4.1.12. Serine Carboxypeptidase

This protein is another member of the peptidase family, but there are few reports of serine carboxypeptidase in the animal venom than the serine protease. We identified six serine carboxypeptidase genes from *P. lepida*. They have diverse physiological roles, e.g., releasing histamine, destroying neurotransmitters, neurotoxicity, immunity, and venom proteins' phosphorylation [94]. They are allergens in bee venoms that elicit an IgE-mediated allergic reaction [94]. Serine carboxypeptidases are found in bee, wasp, kissing bug, and Remipedia venoms [37,69–73].

#### 4.1.13. C-Type Lectin

C-type lectin is an extracellular carbohydrate-binding protein, including multifunctions such as cell adhesion, pathogen recognition, an inflammatory response, hemostasis, and hemagglutination [95,96]. We identified nine C-type lectin genes from the *P. lepida*. They are also present in many animal venoms, including *L. obliqua*, snake, lionfish, cnidarian, crustacean, insects, caterpillars, and leech [21,37,96–99].

#### 4.1.14. Phospholipase A2

Phospholipase A2 hydrolyze various phospholipids. They can induce cell death, inflammation, edema, anticoagulation, and inhibit transmitters at the neuromuscular junction [3,96,100]. There are three putative venom phospholipase A2 genes in the *P. lepida* transcriptome. In *L. obliqua*, phospholipase A2 has an indirect hemolytic activity in human and rat red blood cells [22]. Phospholipase A2 is a potent toxin in the snake venom and is also found in the venom of other species, including reptile, cnidarian, mollusk, hymenopteran insect, scorpion, spider, and annelid [3,21,101,102].

#### 4.1.15. Antimicrobial Peptides

We identified two types of antimicrobial peptides, diapausin and attacin, from the *P. lepida* transcriptome (six and four genes, respectively). Diapausin is a cysteine-rich peptide that has an antifungal property by inhibiting  $\alpha$ -1,3-glucan synthesis [103]. They also function as an N-type voltage-gated calcium channel blocker due to their similar structure to  $\omega$ -conotoxin GVIA in cone snail toxins [104]. Diapausin is present in various arthropods, including collembola, beetles, and butterflies [105–107]. Attacin is a glycine-rich peptide that inhibits the growth of Gram-negative bacteria by inhibiting cell wall production [108–110]. They are reported in Lepidoptera and Diptera. This protein's toxicity is unclear, but it is expressed in the venom of *L. obliqua* [21].

### 4.2. *P. lepida* Toxins in Comparison with Other Venomous Lepidoptera

In total, we reported 168 candidate toxin genes from the *P. lepida* urticating hair transcriptome, indicating that the toxicity of *P. lepida* venom is most likely the product of a complex toxin cocktail. Similarly, the diversity of toxins is also observed in *Premolis semirufa* caterpillar as 418 candidate toxin genes are identified from its transcriptome [27]. We note that different annotation methods partly explain the distinct difference between these numbers. The study in *Premolis semirufa* does not have a validation step to filter out some housekeeping genes from the sequences that match known toxins. For this reason, the number of genes in each toxin class between the two species cannot be compared

directly. Some toxin classes of *P. lepida* are similar to those of *L. obliqua* and *Premolis semirufa*, including protease inhibitor (serpin, Kazal-type protease inhibitor, and trypsin inhibitor), proteolytic enzyme (peptidase S1), and other toxins, including lipocalin, C-type lectin, and AMPs [21,22,27]. Serine protease genes, which are greatly expanded (59 genes) in *P. lepida*, are also present in the venom of *Leucanella memusae*, *Euproctis chrysorrhoea*, *Euproctis subflara*, *Premolis semirufa*, and *Podalia ca. fuscescens* [23,26,27]. These species are from different taxonomic lineages, suggesting that some toxin classes, particularly serine protease, are repeatedly recruited in the Lepidoptera venom evolution.

Our transcriptome analysis did not detect closely related genes of Lopap and Losac, which are vital toxins in the *L. obliqua* venom. These proteins are prothrombin activator and factor x activator, respectively. They can interfere with the blood coagulation process, causing hemorrhagic syndrome and acute kidney injury [111,112]. The closely related species, *Lonomia achelous*, also has proteins that cause blood coagulation, e.g., Lonomin (urokinase-like), Lonomin III (prothrombin activator), Lonomin IV (Factor Xa-like), and Lonomin V (FXIII inactivator) [22]. The differences in toxin gene composition explain why the venom of *P. lepida* and *Lonomia* spp. causes different physiological effects.

Some toxin classes of *P. lepida* have not been reported in other venomous Lepidoptera but are shared with *Premolis semirufa*. These include venom peptidyl peptidase 4, venom acid phosphatase, and carboxylesterase-6 [18,27]. The putative functions of these proteins fit the effects of *P. lepida* venom, e.g., contact dermatitis. As discussed previously, the novel Knottin-like peptides are expressed at a high level and may function by activating pain receptors, causing stinging pain. Simultaneously, other proteins that are allergens similar to those found in hymenopteran venoms can cause allergic reactions in the contact area, including erythema, edema, intense itch, and inflammation.

However, we note that several aspects limit the comparison and interpretation. First, only a few venomous lepidopteran species have been investigated. Secondly, previous research did not use a high-throughput sequencing technique, e.g., transcriptome and genome; thus, the diversity of toxin genes might not have entirely been revealed. For these two reasons, the assumption that particular toxin classes are unique to a specific lepidopteran species or lineage might be incorrect. Third, *P. lepida* toxins' functions are predicted based solely on orthologous relationships with previously reported toxins or the predicted 3D structure with known toxins. Further functional tests are required to confirm the venom activity of the *P. lepida* candidate toxins.

#### 4.3. Evolution of *P. lepida* Toxin Genes

Many lepidopteran species from different families can produce venoms, suggesting that toxin genes evolved independently many times in Lepidoptera evolution. Convergent evolution is a fundamental process in animal venom evolution [3,113]. Our findings indicate that some toxin classes have been repeatedly recruited into the venom of Lepidoptera, particularly serine protease, protease inhibitor, phospholipase A2, and lectin, as reported in *P. lepida* and other species. The convergent evolution can also occur across animal lineage; for example, *P. lepida* and hymenopteran venoms share similar toxin compositions, including CAP, Kunitz, carboxylesterase-6, and venom acid phosphatase, which are allergens that trigger the recipient's immune system [57].

Due to a significant expansion of serine protease genes (59 genes) and Knottin-like peptide genes (21 genes) in *P. lepida*, processes involving gene duplication, followed by positive selection, may be responsible for this. This process is one of the most critical mechanisms in venom gene evolution, as documented in snakes, cone snails, spiders, scorpions, and centipedes [82,114–119]. Tandem gene duplication may explain the gene family expansion reported in the genome of *Stegodyphus mimosarum* spiders, which has 51 Knottin-like toxin genes [120], and *Mesobuthus martensii* scorpion, which has more than 100 neurotoxin genes (NaTx, KTx, and ClTx) [121]. When the genome data of *P. lepida* are available, and genes' locations are determined, it will confirm whether the serine protease genes of *P. lepida* arose from tandem gene duplication.

Interestingly, putative toxin genes encoding 21 Knottin-like peptides in *P. lepida* venom have no significant similarity to any previously reported gene in the NCBI database, suggesting that they may arise specifically in the ancestral lineage of *P. lepida*. Due to the limitation of its related species' genetic data, it is impossible to fully understand the origin of these genes. Phylogenetic analysis indicates that these 21 genes are not monophyletic (Figure 7B). They also have long branches, suggesting an ancient origin, and proteins from different branches may serve different functions. We hypothesized that they might arise from the nontoxic Knottin proteins in the ancestor of Lepidoptera. Mutations of housekeeping genes may lead to toxic functions, possibly neurotoxins, favored by natural selection. This fits the theory that toxins for defense purposes should be fast-acting by interfering with nerve transmission [1]. Other reported roles of Knottin peptides are antimicrobial peptides. For example, tachystatins from horseshoe crab hemocytes provides activity against Gram-negative and Gram-positive bacteria, and fungi [122]. In whitefly *Bemisia tabaci*, Knottin-like proteins play a role in regulating the yellow leaf curl virus, and they are upregulated after infection with *Pseudomonas aeruginosa* [123–125].

Our results support previous observations [3] that animal toxins may arise convergently from ancestral proteins, which have the following properties: secretory protein, functionally versatile, stable molecular scaffolds, and extensive disulfide cross-links. Once recruited in animal venom, these proteins evolved under positive selection and were subject to neofunctionalization.

## 5. Conclusions

We identified 168 candidate toxin genes from the *P. lepida* transcriptome. Of these, 147 genes are from 13 toxin classes commonly found in venomous animals and 21 genes encoding novel peptides unique to *P. lepida*. We predicted that *P. lepida* toxins have diverse physiological roles, including neurotoxins that initiate intense pain (Knottin-like peptides), allergens activating the recipients' immune system (e.g., CAP, Kunitz, carboxylesterase-6, and venom acid phosphatase), and toxins causing inflammation (e.g., serine protease, phospholipase A2, and serpin). This prediction, however, requires future functional tests. We highlighted the importance of convergent evolution in the Lepidoptera venom and possible mechanisms, such as mutation of the house-keeping gene and gene duplication, followed by neofunctionalization. We believe this study opens further investigations into toxins in Lepidoptera, the understudied group that has potential benefits in pharmacy, medicine, biotechnology, and agriculture.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/insects12050396/s1>, Supplementary Data S1. Parameter settings used in transcriptomic analyses; Supplementary Data S2. Accession IDs of the arthropod toxin proteins; Supplementary Data S3. Gene annotation results from BLAST2GO; Supplementary Data S4. *Parasa lepida* candidate toxin genes; Supplementary Data S5 Lepidoptera orthologous genes identified by OrthoVenn2.

**Author Contributions:** Conceptualization, P.E., N.M. and S.B.; funding acquisition, P.E., N.M. and S.B.; methodology, P.E., N.M., K.N. and S.B.; investigation, P.E., N.M., K.N. and S.B.; project administration, N.M. and P.E.; writing—review and editing, P.E., N.M., K.N. and S.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the Plant Genetic Conservation Project under the Royal Initiative of Her Royal Highness Princess Maha Chakri Sirindhorn (Project ID: SCI600084S) and Prince of Songkla University.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and was approved by the Institutional Animal Care and Use Committee, Prince of Songkla University (2564-01-008: 29 March 2021).

**Data Availability Statement:** RNA sequencing reads have been deposited in the Sequenced Read Archive (SRA) under BioProject ID: PRJNA694339.

**Acknowledgments:** We would like to thank the funding sources and Anna Chatthong, International Relations Office, Faculty of Science, Prince of Songkla University for English language checking and editing.

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

## References

1. Arbuckle, K. Evolutionary context of venom in animals. In *Evolution of Venomous Animals and Their Toxins*; Malhotra, A., Ed.; Toxinology; Springer: Dordrecht, The Netherlands, 2017; pp. 3–31. ISBN 978-94-007-6458-3.
2. Schendel, V.; Rash, L.D.; Jenner, R.A.; Undheim, E.A.B. The diversity of venom: The importance of behavior and venom system morphology in understanding its ecology and evolution. *Toxins* **2019**, *11*, 666. [[CrossRef](#)]
3. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.A.; King, G.F.; Nevalainen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et al. The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genom. Human Genet.* **2009**, *10*, 483–511. [[CrossRef](#)]
4. Ericsson, C.D.; Hatz, C.; Junghanss, T.; Bodio, M. Medically important venomous animals: Biology, prevention, first aid, and clinical management. *Clin. Infect. Dis.* **2006**, *43*, 1309–1317. [[CrossRef](#)] [[PubMed](#)]
5. Ferraz, C.R.; Arrahman, A.; Xie, C.; Casewell, N.R.; Lewis, R.J.; Kool, J.; Cardoso, F.C. Multifunctional toxins in snake venoms and therapeutic implications: From pain to hemorrhage and necrosis. *Front. Ecol. Evol.* **2019**, *7*. [[CrossRef](#)]
6. Primon-Barros, M.; José Macedo, A. Animal venom peptides: Potential for new antimicrobial agents. *Curr. Top. Med. Chem.* **2017**, *17*, 1119–1156. [[CrossRef](#)]
7. Chatterjee, B. Animal venoms have potential to treat cancer. *Curr. Top. Med. Chem.* **2018**, *18*, 2555–2566. [[CrossRef](#)] [[PubMed](#)]
8. Pennington, M.W.; Czerwinski, A.; Norton, R.S. Peptide therapeutics from venom: Current status and potential. *Bioorganic Med. Chem.* **2018**, *26*, 2738–2758. [[CrossRef](#)]
9. King, G.F. Tying pest insects in knots: The deployment of spider-venom-derived knottins as bioinsecticides. *Pest Manag. Sci.* **2019**, *75*, 2437–2445. [[CrossRef](#)]
10. Maatuf, Y.; Geron, M.; Priel, A. The role of toxins in the pursuit for novel analgesics. *Toxins* **2019**, *11*, 131. [[CrossRef](#)]
11. Yang, X.; Wang, Y.; Wu, C.; Ling, E.-A. Animal venom peptides as a treasure trove for new therapeutics against neurodegenerative disorders. *Curr. Med. Chem.* **2019**, *26*, 4749–4774. [[CrossRef](#)] [[PubMed](#)]
12. Gullan, P.J.; Cranston, P.S. *The Insects: An Outline of Entomology*; John Wiley & Sons: Hoboken, NJ, USA, 2014; ISBN 978-1-118-84615-5.
13. Hossler, E.W. Caterpillars and moths. *Dermatol. Ther.* **2009**, *22*, 353–366. [[CrossRef](#)]
14. Veiga, A.B.; Blochtein, B.; Guimarães, J.A. Structures involved in production, secretion and injection of the venom produced by the caterpillar *Lonomia obliqua* (Lepidoptera, Saturniidae). *Toxicon* **2001**, *39*, 1343–1351. [[CrossRef](#)]
15. Seldeslachts, A.; Peigneur, S.; Tytgat, J. Caterpillar venom: A health hazard of the 21st century. *Biomedicines* **2020**, *8*, 143. [[CrossRef](#)]
16. Yao, Z.; Kamau, P.M.; Han, Y.; Hu, J.; Luo, A.; Luo, L.; Zheng, J.; Tian, Y.; Lai, R. The *Latoia consocia* caterpillar induces pain by targeting nociceptive ion channel TRPV1. *Toxins* **2019**, *11*, 695. [[CrossRef](#)] [[PubMed](#)]
17. Kawamoto, F.; Kumada, N. Biology and venoms of Lepidoptera. In *Handbook of Natural Toxins*; Marcel Dekker: New York, NY, USA, 1984.
18. Villas-Boas, I.M.; Bonfá, G.; Tambourgi, D.V. Venomous caterpillars: From inoculation apparatus to venom composition and envenomation. *Toxicon* **2018**, *153*, 39–52. [[CrossRef](#)]
19. Villas-Boas, I.M.; Pidde, G.; Lichtenstein, F.; Ching, A.T.C.; de Junqueira-de-Azevedo, I.L.M.; DeOcesano-Pereira, C.; Madureira Trufen, C.E.; Chudzinski-Tavassi, A.M.; Morais, K.L.P.; Tambourgi, D.V. Human chondrocyte activation by toxins from *Premolis semirufa*, an amazon rainforest moth caterpillar: Identifying an osteoarthritis signature. *Front. Immunol.* **2020**, *11*. [[CrossRef](#)] [[PubMed](#)]
20. Estrella, M.; Elston, D.M. What's eating you? Megalopyge opercularis. *Cutis* **2020**, *106*, 23–24; 27. [[CrossRef](#)]
21. Veiga, A.B.G.; Ribeiro, J.M.C.; Guimarães, J.A.; Francischetti, I.M.B. A catalog for the transcripts from the venomous structures of the caterpillar *Lonomia obliqua*: Identification of the proteins potentially involved in the coagulation disorder and hemorrhagic syndrome. *Gene* **2005**, *355*, 11–27. [[CrossRef](#)] [[PubMed](#)]
22. Carrijo-Carvalho, L.C.; Chudzinski-Tavassi, A.M. The Venom of the *Lonomia* caterpillar: An overview. *Toxicon* **2007**, *49*, 741–757. [[CrossRef](#)]
23. Quintana, M.A.; Sciani, J.M.; Auada, A.V.V.; Martínez, M.M.; Sánchez, M.N.; Santoro, M.L.; Fan, H.W.; Peichoto, M.E. Stinging caterpillars from the genera podalia, leucanella and lonomia in Misiones, Argentina: A preliminary comparative approach to understand their toxicity. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2017**, *202*, 55–62. [[CrossRef](#)] [[PubMed](#)]
24. de Jong, M.C.J.M.; Kawamoto, F.; Bleumink, E.; Kloosterhuis, A.J.; Meijer, G.T. A comparative study of the spicule venom of euproctis caterpillars. *Toxicon* **1982**, *20*, 477–485. [[CrossRef](#)]
25. Moneo, I.; Vega, J.M.; Caballero, M.L.; Vega, J.; Alday, E. Isolation and characterization of Tha p 1, a major allergen from the pine processionary caterpillar *Thaumetopoea pityocampa*. *Allergy* **2003**, *58*, 34–37. [[CrossRef](#)]
26. Villas-Boas, I.M.; Gonçalves-de-Andrade, R.M.; Pidde-Queiroz, G.; Assaf, S.L.M.R.; Portaro, F.C.V.; Sant'Anna, O.A.; van den Berg, C.W.; Tambourgi, D.V. *Premolis semirufa* (Walker, 1856) envenomation, disease affecting rubber tappers of the amazon: Searching for caterpillar-bristles toxic components. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1531. [[CrossRef](#)]

27. Pidde, G.; Nishiyama, M.Y.; de Oliveira, U.C.; Villas-Boas, I.M.; Paes-Leme, A.F.; Junqueira-de-Azevedo, I.L.; Marques-Porto, R.; Squaiella-Baptistão, C.C.; Tambourgi, D.V. Integrative multiomics analysis of *premolis semirufa* caterpillar venom in the search for molecules leading to a joint disease. *Sci. Rep.* **2021**, *11*, 1995. [[CrossRef](#)]
28. Rodriguez-Mahillo, A.I.; Gonzalez-Muñoz, M.; Vega, J.M.; López, J.A.; Yart, A.; Kerdelhué, C.; Camafeita, E.; Ortiz, J.C.G.; Vogel, H.; Toffolo, E.P.; et al. Setae from the pine processionary moth (*Thaumetopoea pityocampa*) contain several relevant allergens. *Contact Dermat.* **2012**, *67*, 367–374. [[CrossRef](#)]
29. Kapoor, K.; Deobhakta, S.R.; Dhamdhare, S. Bionomics of the Slug Caterpillar, *Latoia lepida* (Cramer) (Lepidoptera: Limacodidae) on Mango. *J. Entomol. Res.* **1985**, *9*, 235–236.
30. Arumugam, G.; Karuppiyah, H.; Sreeramulu, B.; Paulchamy, R.; Sundaram, J. Occurrence of natural lectin with bacterial agglutination property in the serum of lepidopteran pest, *Parasa lepida*. *Entomol. Sci.* **2019**, *22*, 239–249. [[CrossRef](#)]
31. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
32. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from rna-seq using the trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)]
33. Pertea, G.; Huang, X.; Liang, F.; Antonescu, V.; Sultana, R.; Karamycheva, S.; Lee, Y.; White, J.; Cheung, F.; Parvizi, B.; et al. TIGR Gene Indices Clustering Tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinformatics* **2003**, *19*, 651–652. [[CrossRef](#)]
34. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [[CrossRef](#)]
35. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
36. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)]
37. Verdes, A.; Simpson, D.; Holford, M. Are fireworms venomous? Evidence for the convergent evolution of toxin homologs in three species of fireworms (Annelida, Amphinomididae). *Genome. Biol. Evol.* **2017**, *10*, 249–268. [[CrossRef](#)]
38. von Reumont, B.M.; Undheim, E.A.B.; Jauss, R.-T.; Jenner, R.A. Venomics of remipede crustaceans reveals novel peptide diversity and illuminates the venom’s biological role. *Toxins* **2017**, *9*, 234. [[CrossRef](#)]
39. Chudzinski-Tavassi, A.M.; Alvarez Flores, M.P. Exploring new molecules and activities from *Lonomia obliqua* caterpillars. *Pathophysiol. Haemost. Thromb.* **2005**, *34*, 228–233. [[CrossRef](#)]
40. Quevillon, E.; Silventoinen, V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R. InterProScan: Protein domains identifier. *Nucleic Acids Res.* **2005**, *33*, W116–W120. [[CrossRef](#)]
41. Xu, L.; Dong, Z.; Fang, L.; Luo, Y.; Wei, Z.; Guo, H.; Zhang, G.; Gu, Y.Q.; Coleman-Derr, D.; Xia, Q.; et al. OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* **2019**, *47*, W52–W58. [[CrossRef](#)]
42. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)]
43. Ramírez, D.; Gonzalez, W.; Fissore, R.A.; Carvacho, I. Conotoxins as tools to understand the physiological function of voltage-gated calcium (CaV) channels. *Mar. Drugs* **2017**, *15*, 313. [[CrossRef](#)]
44. Takeuchi, K.; Park, E.; Lee, C.W.; Kim, J.I.; Takahashi, H.; Swartz, K.; Shimada, I. Solution structure of  $\omega$ -grammotoxin SIA, A gating modifier of P/Q and N-Type Ca<sup>2+</sup> Channel. *J. Mol. Biol.* **2002**, *321*, 517–526. [[CrossRef](#)]
45. Xiao, Y.-C.; Liang, S.-P. Purification and characterization of hainantoxin-V, a tetrodotoxin-sensitive sodium channel inhibitor from the venom of the spider *Selenocosmia hainana*. *Toxicon* **2003**, *41*, 643–650. [[CrossRef](#)]
46. Xiao, L.; Gurrola, G.B.; Zhang, J.; Valdivia, C.R.; SanMartin, M.; Zamudio, F.Z.; Zhang, L.; Possani, L.D.; Valdivia, H.H. Structure–function relationships of peptides forming the calxin family of ryanodine receptor ligands. *J. Gen. Physiol.* **2016**, *147*, 375–394. [[CrossRef](#)]
47. Fajloun, Z.; Kharrat, R.; Chen, L.; Lecomte, C.; Di Luccio, E.; Bichet, D.; El Ayeb, M.; Rochat, H.; Allen, P.D.; Pessah, I.N.; et al. Chemical synthesis and characterization of maurocalcine, a scorpion toxin that activates Ca(2+) release channel/ryanodine receptors. *FEBS Lett.* **2000**, *469*, 179–185. [[CrossRef](#)]
48. Osteen, J.D.; Herzig, V.; Gilchrist, J.; Emrick, J.J.; Zhang, C.; Wang, X.; Castro, J.; Garcia-Caraballo, S.; Grundy, L.; Rychkov, G.Y.; et al. Selective spider toxins reveal a role for Nav1.1 channel in mechanical pain. *Nature* **2016**, *534*, 494–499. [[CrossRef](#)] [[PubMed](#)]
49. Gibbs, G.M.; Roelants, K.; O’Bryan, M.K. The CAP superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—roles in reproduction, cancer, and immune defense. *Endocr. Rev.* **2008**, *29*, 865–897. [[CrossRef](#)] [[PubMed](#)]
50. Pantera, B.; Hoffman, D.R.; Carresi, L.; Cappugi, G.; Turillazzi, S.; Manao, G.; Severino, M.; Spadolini, I.; Orsomando, G.; Moneti, G.; et al. Characterization of the major allergens purified from the venom of the paper wasp *Polistes gallicus*. *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2003**, *1623*, 72–81. [[CrossRef](#)] [[PubMed](#)]
51. Müller, U.R.; Johansen, N.; Petersen, A.B.; Fromberg-Nielsen, J.; Haerberli, G. Hymenoptera venom allergy: Analysis of double positivity to honey bee and vespula venom by estimation of IgE antibodies to species-specific major allergens Api M1 and Ves V5. *Allergy* **2009**, *64*, 543–548. [[CrossRef](#)] [[PubMed](#)]

52. Torres, A.F.C.; Huang, C.; Chong, C.-M.; Leung, S.W.; Prieto-da-Silva, Á.R.B.; Havt, A.; Quinet, Y.P.; Martins, A.M.C.; Lee, S.M.Y.; Rádís-Baptista, G. Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: Insights into the polypeptide toxin arsenal of hymenopterans. *PLoS ONE* **2014**, *9*, e87556. [\[CrossRef\]](#)
53. Yamazaki, Y.; Morita, T. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* **2004**, *44*, 227–231. [\[CrossRef\]](#)
54. Milne, T.J.; Abbenante, G.; Tyndall, J.D.A.; Halliday, J.; Lewis, R.J. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J. Biol. Chem.* **2003**, *278*, 31105–31110. [\[CrossRef\]](#)
55. Blank, S.; Seismann, H.; Bockisch, B.; Braren, I.; Cifuentes, L.; McIntyre, M.; Rühl, D.; Ring, J.; Bredehorst, R.; Ollert, M.W.; et al. Identification, recombinant expression, and characterization of the 100 KDa high molecular weight hymenoptera venom allergens Api m 5 and Ves v 3. *J. Immunol.* **2010**, *184*, 5403–5413. [\[CrossRef\]](#)
56. Dani, M.P.; Richards, E.H.; Isaac, R.E.; Edwards, J.P. Antibacterial and proteolytic activity in venom from the endoparasitic wasp *Pimpla hypochondriaca* (Hymenoptera: Ichneumonidae). *J. Insect Physiol.* **2003**, *49*, 945–954. [\[CrossRef\]](#)
57. Robinson, S.D.; Mueller, A.; Clayton, D.; Starobova, H.; Hamilton, B.R.; Payne, R.J.; Vetter, I.; King, G.F.; Undheim, E.A.B. A comprehensive portrait of the venom of the giant red bull ant, *Myrmecia gulosa*, reveals a hyperdiverse hymenopteran toxin gene family. *Sci. Adv.* **2018**, *4*, eaau4640. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Schiener, M.; Hilger, C.; Eberlein, B.; Pascal, M.; Kuehn, A.; Revets, D.; Planchon, S.; Pietsch, G.; Serrano, P.; Moreno-Aguilar, C.; et al. The high molecular weight dipeptidyl peptidase IV Pol d 3 is a major allergen of *Polistes dominula* Venom. *Sci. Rep.* **2018**, *8*, 1318. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Matysiak, J.; Hajduk, J.; Pietrzak, Ł.; Schmelzer, C.E.H.; Kokot, Z.J. Shotgun proteome analysis of honeybee venom using targeted enrichment strategies. *Toxicon* **2014**, *90*, 255–264. [\[CrossRef\]](#)
60. Pushparani, D.S. High acid phosphatase level in the gingival tissues of periodontitis subjects. *J. Basic Clin. Pharm.* **2015**, *6*, 59–63. [\[CrossRef\]](#)
61. Zhu, J.; Ye, G.; Hu, C. Molecular cloning and characterization of acid phosphatase in venom of the endoparasitoid wasp *Pteromalus puparum* (Hymenoptera: Pteromalidae). *Toxicon* **2008**, *51*, 1391–1399. [\[CrossRef\]](#)
62. Bonasio, R.; Zhang, G.; Ye, C.; Mutti, N.S.; Fang, X.; Qin, N.; Donahue, G.; Yang, P.; Li, Q.; Li, C.; et al. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **2010**, *329*, 1068–1071. [\[CrossRef\]](#)
63. Dhananjaya, B.L.; D'Souza, C.J.M. The pharmacological role of phosphatases (Acid and Alkaline Phosphomonoesterases) in snake venoms related to release of purines—a Multitoxin. *Basic Clin. Pharmacol. Toxicol.* **2011**, *108*, 79–83. [\[CrossRef\]](#)
64. Heavner, M.E.; Gueguen, G.; Rajwani, R.; Pagan, P.E.; Small, C.; Govind, S. Partial venom gland transcriptome of a drosophila parasitoid wasp, *Leptopilina heterotoma*, reveals novel and shared bioactive profiles with stinging hymenoptera. *Gene* **2013**, *526*, 195–204. [\[CrossRef\]](#)
65. Spillner, E.; Blank, S.; Jakob, T. Hymenoptera allergens: From venom to “venome”. *Front. Immunol.* **2014**, *5*. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Perkin, L.C.; Friesen, K.S.; Flinn, P.W.; Oppert, B. Venom gland components of the ectoparasitoid wasp, *Anisopteromalus calandrae*. *J. Venom. Res.* **2015**, *6*, 19–37.
67. Chai, L.; Yang, X.; Liu, M.; Liu, C.; Han, L.; Guo, H.; Li, C.; Sun, Y.; Li, X.; Xiao, M.; et al. Biopanning of allergens from wasp sting patients. *Biosci. Rep.* **2018**, *38*, BSR20181113. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Chertemps, T.; François, A.; Durand, N.; Rosell, G.; Dekker, T.; Lucas, P.; Maïbèche-Coisne, M. A carboxylesterase, esterase-6, modulates sensory physiological and behavioral response dynamics to pheromone in *Drosophila*. *BMC Biol.* **2012**, *10*, 56. [\[CrossRef\]](#)
69. Whittington, C.M.; Papenfuss, A.T.; Locke, D.P.; Mardis, E.R.; Wilson, R.K.; Abubucker, S.; Mitreva, M.; Wong, E.S.; Hsu, A.L.; Kuchel, P.W.; et al. Novel venom gene discovery in the platypus. *Genome Biol.* **2010**, *11*, R95. [\[CrossRef\]](#)
70. Vaipayuri, S.; Wagstaff, S.C.; Harrison, R.A.; Gibbins, J.M.; Hutchinson, E.G. Evolutionary analysis of novel serine proteases in the venom gland transcriptome of *Bitis gabonica rhinoceros*. *PLoS ONE* **2011**, *6*, e21532. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Ruder, T.; Sunagar, K.; Undheim, E.A.B.; Ali, S.A.; Wai, T.-C.; Low, D.H.W.; Jackson, T.N.W.; King, G.F.; Antunes, A.; Fry, B.G. Molecular phylogeny and evolution of the proteins encoded by coleoid (Cuttlefish, Octopus, and Squid) posterior venom glands. *J. Mol. Evol.* **2013**, *76*, 192–204. [\[CrossRef\]](#)
72. Choo, Y.M.; Lee, K.S.; Yoon, H.J.; Kim, B.Y.; Sohn, M.R.; Roh, J.Y.; Je, Y.H.; Kim, N.J.; Kim, I.; Woo, S.D.; et al. Dual function of a bee venom serine protease: Prophenoloxidase-activating factor in arthropods and fibrin(Ogen)olytic enzyme in mammals. *PLoS ONE* **2010**, *5*, e10393. [\[CrossRef\]](#)
73. Ribeiro, J.M.C.; Genta, F.A.; Sorgine, M.H.F.; Logullo, R.; Mesquita, R.D.; Paiva-Silva, G.O.; Majerowicz, D.; Medeiros, M.; Koerich, L.; Terra, W.R.; et al. An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *PLoS Negl. Trop. Dis.* **2014**, *8*, e2594. [\[CrossRef\]](#)
74. von Reumont, B.M.; Blanke, A.; Richter, S.; Alvarez, F.; Bleidorn, C.; Jenner, R.A. The first venomous crustacean revealed by transcriptomics and functional morphology: Remipede venom glands express a unique toxin cocktail dominated by enzymes and a neurotoxin. *Mol. Biol. Evol.* **2014**, *31*, 48–58. [\[CrossRef\]](#)
75. Nakhleh, J.; El Moussawi, L.; Osta, M.A. Chapter Three-The melanization response in insect immunity. In *Advances in Insect Physiology*; Ligoxygakis, P., Ed.; Insect Immunity; Academic Press: Cambridge, MA, USA, 2017; Volume 52, pp. 83–109.
76. Flower, D.R. The lipocalin protein family: Structure and function. *Biochem. J.* **1996**, *318*, 1–14. [\[CrossRef\]](#)

77. Paddock, C.D.; McKerrow, J.H.; Hansell, E.; Foreman, K.W.; Hsieh, I.; Marshall, N. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. *J. Immunol.* **2001**, *167*, 2694–2699. [[CrossRef](#)]
78. Low, D.H.W.; Sunagar, K.; Undheim, E.A.B.; Ali, S.A.; Alagon, A.C.; Ruder, T.; Jackson, T.N.W.; Pineda Gonzalez, S.; King, G.F.; Jones, A.; et al. Dracula's children: Molecular evolution of vampire bat venom. *J. Proteom.* **2013**, *89*, 95–111. [[CrossRef](#)]
79. Junqueira-de-Azevedo, I.L.M.; Campos, P.F.; Ching, A.T.C.; Mackessy, S.P. Colubrid venom composition: An-omics perspective. *Toxins* **2016**, *8*, 230. [[CrossRef](#)] [[PubMed](#)]
80. Reis, C.V.; Portaro, F.C.; Andrade, S.A.; Fritzen, M.; Fernandes, B.L.; Sampaio, C.A.; Camargo, A.C.; Chudzinski-Tavassi, A.M. A prothrombin activator serine protease from the *Lonomia obliqua* caterpillar venom (Lopap) biochemical characterization. *Thromb. Res.* **2001**, *102*, 427–436. [[CrossRef](#)]
81. Law, R.H.; Zhang, Q.; McGowan, S.; Buckle, A.M.; Silverman, G.A.; Wong, W.; Rosado, C.J.; Langendorf, C.G.; Pike, R.N.; Bird, P.I.; et al. An overview of the serpin superfamily. *Genome Biol.* **2006**, *7*, 216. [[CrossRef](#)] [[PubMed](#)]
82. Wong, E.S.W.; Belov, K. Venom evolution through gene duplications. *Gene* **2012**, *496*, 1–7. [[CrossRef](#)] [[PubMed](#)]
83. Liu, D.; Wang, L.; Yang, L.; Qian, C.; Wei, G.; Dai, L.; Li, J.; Zhu, B.; Liu, C. Serpin-15 from *Bombyx mori* Inhibits Prophenoloxidase activation and expression of antimicrobial peptides. *Dev. Comp. Immunol.* **2015**, *51*, 22–28. [[CrossRef](#)]
84. Sigle, L.T.; Ramalho-Ortigão, M. Kazal-type serine proteinase inhibitors in the midgut of *Phlebotomus papatasi*. *Mem. Inst. Oswaldo Cruz.* **2013**, *108*, 671–678. [[CrossRef](#)]
85. Francischetti, I.M.B.; Assumpção, T.C.F.; Ma, D.; Li, Y.; Vicente, E.C.; Uieda, W.; Ribeiro, J.M.C. The “Vampirome”: Transcriptome and proteome analysis of the principal and accessory submaxillary glands of the vampire bat *Desmodus rotundus*, a Vector of human rabies. *J. Proteom.* **2013**, *82*, 288–319. [[CrossRef](#)] [[PubMed](#)]
86. Wan, H.; Lee, K.S.; Kim, B.Y.; Zou, F.M.; Yoon, H.J.; Je, Y.H.; Li, J.; Jin, B.R. A spider-derived kunitz-type serine protease inhibitor that acts as a plasmin inhibitor and an elastase inhibitor. *PLoS ONE* **2013**, *8*, e53343. [[CrossRef](#)] [[PubMed](#)]
87. Zhao, R.; Dai, H.; Qiu, S.; Li, T.; He, Y.; Ma, Y.; Chen, Z.; Wu, Y.; Li, W.; Cao, Z. SdPI, The first functionally characterized kunitz-type trypsin inhibitor from scorpion venom. *PLoS ONE* **2011**, *6*, e27548. [[CrossRef](#)] [[PubMed](#)]
88. Kozlov, S.A.; Andreev, Y.A.; Murashev, A.N.; Skobtsov, D.I.; D'yachenko, I.A.; Grishin, E.V. New polypeptide components from the *Heteractis crispa* sea anemone with analgesic activity. *Russ. J. Bioorg. Chem.* **2009**, *35*, 711. [[CrossRef](#)]
89. Mourão, C.B.F.; Schwartz, E.F. Protease inhibitors from marine venomous animals and their counterparts in terrestrial venomous animals. *Mar. Drugs* **2013**, *11*, 2069–2112. [[CrossRef](#)]
90. Yuan, C.-H.; He, Q.-Y.; Peng, K.; Diao, J.-B.; Jiang, L.-P.; Tang, X.; Liang, S.-P. Discovery of a distinct superfamily of kunitz-type toxin (KTT) from tarantulas. *PLoS ONE* **2008**, *3*, e3414. [[CrossRef](#)]
91. Wang, H.; Wang, L.; Zhou, M.; Yang, M.; Ma, C.; Chen, T.; Zhang, Y.; Zeller, M.; Hornshaw, M.; Shaw, C. Functional peptidomics of amphibian skin secretion: A novel kunitz-type chymotrypsin inhibitor from the african hyperoliid frog, *Kassina senegalensis*. *Biochimie* **2012**, *94*, 891–899. [[CrossRef](#)]
92. Zeng, X.-C.; Wang, S.-X.; Li, W.-X. Identification of BmKAPi, a novel type of scorpion venom peptide with peculiar disulfide bridge pattern from *Buthus martensii* Karsch. *Toxicon* **2002**, *40*, 1719–1722. [[CrossRef](#)]
93. de Oliveira, U.C.; Candido, D.M.; Coronado Dorce, V.A.; de Junqueira-de-Azevedo, I.L.M. The transcriptome recipe for the venom cocktail of *Tityus bahiensis* scorpion. *Toxicon* **2015**, *95*, 52–61. [[CrossRef](#)]
94. Li, R.; Zhang, L.; Fang, Y.; Han, B.; Lu, X.; Zhou, T.; Feng, M.; Li, J. Proteome and Phosphoproteome analysis of honeybee (*Apis mellifera*) venom collected from electrical stimulation and manual extraction of the venom gland. *BMC Genom.* **2013**, *14*, 766. [[CrossRef](#)]
95. Zelensky, A.N.; Gready, J.E. The C-type lectin-like domain superfamily. *FEBS J.* **2005**, *272*, 6179–6217. [[CrossRef](#)] [[PubMed](#)]
96. Huang, X.; Feng, J.-L.; Jin, M.; Ren, Q.; Wang, W. C-type lectin (MrCTL) from the giant freshwater prawn *Macrobrachium rosenbergii* participates in innate immunity. *Fish Shellfish Immunol.* **2016**, *58*, 136–144. [[CrossRef](#)] [[PubMed](#)]
97. Morita, T. Structures and functions of snake venom CLPs (C-Type Lectin-like Proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. *Toxicon* **2005**, *45*, 1099–1114. [[CrossRef](#)] [[PubMed](#)]
98. Magalhães, G.S.; Junqueira-de-Azevedo, I.L.M.; Lopes-Ferreira, M.; Lorenzini, D.M.; Ho, P.L.; Moura-da-Silva, A.M. Transcriptome analysis of expressed sequence tags from the venom glands of the fish *Thalassophryne nattereri*. *Biochimie* **2006**, *88*, 693–699. [[CrossRef](#)] [[PubMed](#)]
99. Kvist, S.; Ocegüera-Figueroa, A.; Tessler, M.; Jiménez-Armenta, J.; Freeman, R.M., Jr.; Giribet, G.; Siddall, M.E. When predator becomes prey: Investigating the salivary transcriptome of the shark-feeding leech *Pontobdella macrothela* (Hirudinea: Piscicolidae). *Zool. J. Linn. Soc.* **2017**, *179*, 725–737. [[CrossRef](#)]
100. Kini, R.M. Excitement ahead: Structure, function and mechanism of snake venom phospholipase A2 enzymes. *Toxicon* **2003**, *42*, 827–840. [[CrossRef](#)]
101. Nevalainen, T.J.; Graham, G.G.; Scott, K.F. Antibacterial actions of secreted phospholipases A2. review. *Biochim. Biophys. Acta* **2008**, *1781*, 1–9. [[CrossRef](#)]
102. Perez-Riverol, A.; Lasa, A.M.; dos Santos-Pinto, J.R.A.; Palma, M.S. Insect venom phospholipases A1 and A2: Roles in the envenoming process and allergy. *Insect Biochem. Mol. Biol.* **2019**, *105*, 10–24. [[CrossRef](#)]
103. Li, M.; Al Souhail, Q.; Veerapandian, R.; Vedyappan, G.; Kanost, M. Investigation of an antifungal peptide, diapausin, from *Manduca sexta*. *FASEB J.* **2019**, *33*, 800.2. [[CrossRef](#)]

104. Tanaka, H.; Sato, K.; Saito, Y.; Yamashita, T.; Agoh, M.; Okunishi, J.; Tachikawa, E.; Suzuki, K. Insect diapause-specific peptide from the leaf beetle has consensus with a putative iridovirus peptide. *Peptides* **2003**, *24*, 1327–1333. [[CrossRef](#)]
105. Kouno, T.; Mizuguchi, M.; Tanaka, H.; Yang, P.; Mori, Y.; Shinoda, H.; Unoki, K.; Aizawa, T.; Demura, M.; Suzuki, K.; et al. The structure of a novel insect peptide explains its Ca<sup>2+</sup> channel blocking and antifungal activities. *Biochemistry* **2007**, *46*, 13733–13741. [[CrossRef](#)] [[PubMed](#)]
106. Faddeeva-Vakhrusheva, A.; Derks, M.F.L.; Anvar, S.Y.; Agamennone, V.; Suring, W.; Smit, S.; van Straalen, N.M.; Roelofs, D. Gene family evolution reflects adaptation to soil environmental stressors in the genome of the collembolan *Orchesella cincta*. *Genome Biol. Evol.* **2016**, *8*, 2106–2117. [[CrossRef](#)] [[PubMed](#)]
107. Al Souhail, Q.; Hiromasa, Y.; Rahnamaeian, M.; Giraldo, M.C.; Takahashi, D.; Valent, B.; Vilcinskas, A.; Kanost, M.R. Characterization and regulation of expression of an antifungal peptide from hemolymph of an insect, *Manduca sexta*. *Dev. Comp. Immunol.* **2016**, *61*, 258–268. [[CrossRef](#)] [[PubMed](#)]
108. Hoffmann, J.A. Innate immunity of insects. *Curr. Opin. Immunol.* **1995**, *7*, 4–10. [[CrossRef](#)]
109. Carlsson, A.; Engström, P.; Palva, E.T.; Bennich, H. Attacin, an Antibacterial Protein from *Hyalophora cecropia*, Inhibits Synthesis of Outer Membrane Proteins in *Escherichia coli* by Interfering with Omp Gene Transcription. *Infect. Immun.* **1991**, *59*, 3040–3045. [[CrossRef](#)]
110. Wu, Q.; Patočka, J.; Kuča, K. Insect antimicrobial peptides, a mini review. *Toxins* **2018**, *10*, 461. [[CrossRef](#)]
111. Alvarez Flores, M.P.; Fritzen, M.; Reis, C.V.; Chudzinski-Tavassi, A.M. Losac, a factor X activator from *Lonomia obliqua* bristle extract: Its role in the pathophysiological mechanisms and cell survival. *Biochem. Biophys. Res. Commun.* **2006**, *343*, 1216–1223. [[CrossRef](#)]
112. Alvarez Flores, M.P.; Zannin, M.; Chudzinski-Tavassi, A.M. New insight into the mechanism of *Lonomia obliqua* envenoming: Toxin involvement and molecular approach. *Pathophysiol. Haemost. Thromb.* **2010**, *37*, 1–16. [[CrossRef](#)]
113. Suranse, V.; Srikanthan, A.; Sunagar, K. Animal venoms: Origin, diversity and evolution. In *ELS*; American Cancer Society: New York, NY, USA, 2018; pp. 1–20, ISBN 978-0-470-01590-2.
114. Chang, D.; Duda, T.F. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol.* **2012**, *29*, 2019–2029. [[CrossRef](#)]
115. Undheim, E.A.B.; Jones, A.; Clauser, K.R.; Holland, J.W.; Pineda, S.S.; King, G.F.; Fry, B.G. Clawing through evolution: Toxin diversification and convergence in the ancient lineage chilopoda (Centipedes). *Mol. Biol. Evol.* **2014**, *31*, 2124–2148. [[CrossRef](#)]
116. Hargreaves, A.D.; Swain, M.T.; Hegarty, M.J.; Logan, D.W.; Mulley, J.F. Restriction and recruitment—Gene duplication and the origin and evolution of snake venom toxins. *Genome Biol. Evol.* **2014**, *6*, 2088–2095. [[CrossRef](#)]
117. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Curr. Biol.* **2016**, *26*, 2434–2445. [[CrossRef](#)]
118. Sunagar, K.; Moran, Y. The rise and fall of an evolutionary innovation: Contrasting strategies of venom evolution in ancient and young animals. *PLoS Genet.* **2015**, *11*, e1005596. [[CrossRef](#)]
119. Malhotra, A. Mutation, duplication, and more in the evolution of venomous animals and their toxins. In *Evolution of Venomous Animals and Their Toxins*; Malhotra, A., Ed.; Toxinology; Springer: Dordrecht, The Netherlands, 2017; pp. 33–45. ISBN 978-94-007-6458-3.
120. Sanggaard, K.W.; Bechsgaard, J.S.; Fang, X.; Duan, J.; Dyrland, T.F.; Gupta, V.; Jiang, X.; Cheng, L.; Fan, D.; Feng, Y.; et al. Spider genomes provide insight into composition and evolution of venom and silk. *Nat. Commun.* **2014**, *5*, 3765. [[CrossRef](#)]
121. Cao, Z.; Yu, Y.; Wu, Y.; Hao, P.; Di, Z.; He, Y.; Chen, Z.; Yang, W.; Shen, Z.; He, X.; et al. The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods. *Nat. Commun.* **2013**, *4*. [[CrossRef](#)] [[PubMed](#)]
122. Osaki, T.; Omotezako, M.; Nagayama, R.; Hirata, M.; Iwanaga, S.; Kasahara, J.; Hattori, J.; Ito, I.; Sugiyama, H.; Kawabata, S. Horseshoe crab hemocyte-derived antimicrobial polypeptides, tachystatins, with sequence similarity to spider neurotoxins. *J. Biol. Chem.* **1999**, *274*, 26172–26178. [[CrossRef](#)] [[PubMed](#)]
123. Shatters, R.G., Jr.; McKenzie, C.L.; Boykin, L.M.; Gazit, S.; Sinisterra, X.; Weathersbee, A.A., III; Brown, J.K.; Czosnek, H. A knottin-like putative antimicrobial gene family in the whitefly *Bemisia tabaci* biotype B: Cloning and transcript regulation. *J. Insect Sci.* **2008**, *8*, 44–45.
124. Hariton Shalev, A.; Sobol, I.; Ghanim, M.; Liu, S.-S.; Czosnek, H. The whitefly *Bemisia tabaci* Knottin-1 Gene is implicated in regulating the quantity of tomato yellow leaf curl virus ingested and transmitted by the insect. *Viruses* **2016**, *8*, 205. [[CrossRef](#)] [[PubMed](#)]
125. Zhang, C.-R.; Zhang, S.; Xia, J.; Li, F.-F.; Xia, W.-Q.; Liu, S.-S.; Wang, X.-W. The immune strategy and stress response of the mediterranean species of the *Bemisia tabaci* complex to an orally delivered bacterial pathogen. *PLoS ONE* **2014**, *9*, e94477. [[CrossRef](#)]