Protein Discovery: Combined Transcriptomic and Proteomic Analyses of Venom from the Endoparasitoid *Cotesia chilonis* (Hymenoptera: Braconidae)

Zi-Wen Teng [1], Shi-Jiao Xiong [1], Gang Xu [1], Shi-Yu Gan [1], Xuan Chen [1], David Stanley [2], Zhi-Chao Yan [1], Gong-Yin Ye [1] and Qi Fang [1,*]

[1] State Key Laboratory of Rice Biology & Key Laboratory of Agricultural Entomology of Ministry of Agriculture, Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China; tzwbat@126.com (Z.-W.T.); xiongshijiao@zju.edu.cn (S.-J.X.); xugang0503@126.com (G.X.); 3120100103@zju.edu.cn (S.-Y.G.); 3120100105@zju.edu.cn (X.C.); yan_zc@126.com (Z.-C.Y.); chu@zju.edu.cn (G.-Y.Y.)

[2] USDA/Agricultural Research Service Biological Control of Insects Research Laboratory, Columbia, MO 65203, USA; stanleyd@missouri.edu

* Correspondence: angqi@zju.edu.cn; Tel.: +86-571-8898-2696; Fax: +86-571-8898-2988

Abstract: Many species of endoparasitoid wasps provide biological control services in agroecosystems. Although there is a great deal of information on the ecology and physiology of host/parasitoid interactions, relatively little is known about the protein composition of venom and how specific venom proteins influence physiological systems within host insects. This is a crucial gap in our knowledge because venom proteins act in modulating host physiology in ways that favor parasitoid development. Here, we identified 37 possible venom proteins from the polydnavirus-carrying endoparasitoid *Cotesia chilonis* by combining transcriptomic and proteomic analyses. The most abundant proteins were hydrolases, such as proteases, peptidases, esterases, glycosyl hydrolase, and endonucleases. Some components are classical parasitoid venom proteins with known functions, including extracellular superoxide dismutase 3, serine protease inhibitor and calreticulin. The venom contains novel proteins, not recorded from any other parasitoid species, including tolloid-like proteins, chitooligosaccharidolytic β-*N*-acetylglucosaminidase, FK506-binding protein 14, corticotropin-releasing factor-binding protein and vascular endothelial growth factor receptor 2. These new data generate hypotheses and provide a platform for functional analysis of venom components.

Keywords: parasitoid wasps; hosts; venom proteins; venom gland; transcriptomics; proteomics

1. Introduction

Most hymenopteran parasitoids are very small insects that produce minute quantities of venom, which helps explain why research in this area is limited to about 70 of the approximately 300,000 venom-producing species, mainly stinging wasps, bees, and ants [1,2]. Beyond understanding the mechanisms of host/parasitoid relationships, research into parasitoid venom has the potential to uncover a wealth of biomolecules of value in agriculture and pharmacology [1,3,4].

Parasitoid venoms are highly diverse among species [2]. Generally, venoms from ectoparasitoids result in short- to long-term paralysis/lethargy in hosts [5,6]. Compounds identified from ectoparasitoid venom include gamma-aminobutyric acid, β alanine and taurine from *Ampulex compressa* (Hymenoptera, Ampulicidae) [5] and philanthotoxins from *Philanthus triangulum* (Hymenoptera:

Crabronidae) [6]. Ectoparasitoid venoms alter host biology, including immunosuppression, developmental arrest, apoptosis, stress response, and regulating metabolism [7,8]. Endoparasitoid venoms also induce transient paralysis, as seen in two ichneumonids, *Pimpla hypochondriaca* (Hymenoptera: Ichneumonidae) [9], and *Pimpla turionellae* (Hymenoptera: Ichneumonidae) [10], as well as in four braconid wasps, including *Asobara tabida* (Hymenoptera: Braconidae) [11], two *Binodoxys* species (Hymenoptera: Braconidae) [12], and *Chelonus inanitus* (Hymenoptera: Braconidae) [13]. One paralytic factor, pimplin, has been purified from *Pimpla hypochondriaca* venom [9]. Transient host paralysis may be an adaptation to reduce self-superparasitism [12]. Endoparasitoid venom can lead to permanent paralysis, for example, *Asobara japonica* (Hymenoptera: Braconidae) venom induces permanent paralysis that leads to the death of *Drosophila melanogaster* (Diptera: Drosophilidae) larvae. The paralysis can be experimentally reversed by injecting unfractionated ovarian extracts from female wasps, although this biological significance of the observation is unclear [14]. Endoparasitoid venom proteins can also influence host cellular immunity, thought to protect their progeny from host immune responses [1,3,15]. Known immunosuppressive venom components include P4 (containing a RhoGAP domain) from *Leptopilina boulardi* (Hymenoptera: Eucoilidae) [16,17], calreticulin (CRT) from *Cotesia rubecula* (Hymenoptera: Braconidae) [18], VPr1, VPr3 from *Pimpla hypochondriaca* [19], Vn.11 from *Pteromalus puparum* (Hymenoptera: Pteromalidae) [20], and sarco/endoplasmic reticulum calcium ATPase (SERCA) from *Ganaspis* sp.1 (Hymenoptera: Figitidae) [21]. Other venom proteins and peptides, including Vn4.6 [22] and Vn50 [23] from *Cotesia rubecula*, LbSPNy [24] and extracellular Cu, Zn-superoxide dismutase [25] from *Leptopilina* species inhibit host melanization. Endoparasitoid venoms also influence host development. *Glyptapanteles liparidis* (Hymenoptera: Ichneumonidae) venom results in prolonged larval development and higher pupal mortality of *Lymantria dispar* (Lepidoptera: Lymantriidae) [26], and *Pteromalus puparum* venom is an endocrine disrupter of *Pieris rapae* (Lepidoptera: Pieridae) [27]. A venom γ-glutamyl transpeptidase of the braconid *Aphidius ervi* (Hymenoptera: Braconidae) castrates the host *Acyrthosiphon pisum* (Homoptera: Aphididae) by inducing apoptosis in host ovarioles. This may re-direct nutrient flow from ovaries to the hemolymph, possibly advantaging developing parasitoids [1,28,29]. On the other hand, in some polydnavirus (PDV)-carrying endoparasitoids, venoms synergize the effect of PDVs in immunosuppression [30,31] and developmental arrest [13,32], probably by promoting persistence and expression of PDVs [31,33].

Combined transcriptomic and proteomic analyses were first used to characterize *Chelonus inanitus* venom proteins, which discriminated between proteins in the venom gland cells and those in the venom [4,34]. This approach has led to insights into venoms from *Chelonus inanitus*, *Hyposoter didymator* (Hymenoptera: Ichneumonidae), *Leptopilina boulardi*, *Leptopilina heterotoma*, *Microplitis demolitor* (Hymenoptera: Braconidae), *Aphidius ervi*, and *Toxoneuron nigriceps* (Hymenoptera: Braconidae) [2,4,35–38]. Nonetheless, in view of the very large number of parasitoid species, very little is known about the diversity, compositions, and functions of their many venoms. Analysis of venom proteins from a phylogenetically-broad range of parasitoid species will lead to the discovery of new proteins, new chemistries, and valuable information on host/parasitoid relationships.

Cotesia chilonis (Hymenoptera: Braconidae) is an obligate larval parasitoid that effectively regulates *Chilo suppressalis* populations (Lepidoptera: Crambidae), an important rice pest in Asia, Northern Africa, and Southern Europe [39]. It is widely distributed in China, Japan, Indonesia, and North Korea where the natural parasitism rates range from 10% to 30% and may be as high as 90%, for example, in Jiangsu province, China [40,41]. The *Cotesia chilonis*/*Chilo suppressalis* relationship is our model host/parasitoid system, from which we reported that *Cotesia chilonis* parasitization influences the expression of host fat body and hemocyte genes [39]. Transcriptome analysis of adult parasitoids revealed chemoreception-related and bracovirus-related genes may be involved in successful parasitism [41]. We found that *Cotesia chilonis* venom inhibits host humoral immunity and it synergizes the immunosuppressive effects of calyx fluid [42]. Here we report the identification of venom proteins consistent with the biology of known parasitoid venoms and several novel proteins, not previously reported from parasitoid wasps. These findings generate new hypotheses and provide a platform for functional analysis of parasitoid venom components.

2. Results and Discussion

2.1. Transcriptome Analysis of Cotesia chilonis Venom Glands

Transcriptome sequence data were generated from two cDNA libraries, a venom gland library (VG) and an adult female carcass (without venom gland) library (FAC). We acquired 67,097,344 bp raw reads from the VG and 63,686,328 bp raw reads from the FAC. After removing reads containing adapter sequences, poly-N sequences and low quality reads, 65,515,114 (VG) and 62,221,866 bp (FAC) clean reads remained (Table 1). The clean reads were assembled using Trinity software [43]. Total clean base pairs yielded 115,957 transcripts with an N50 length of 2393 bp and a mean length of 1254 bp and 71,617 unigenes with an N50 length of 1309 bp and a mean length of 759 bp (Table 1). The *q*-value < 0.005 and ｜log2(fold change)｜ > 1 was set as the threshold for significant differential expression. By these criteria, 479 VG unigenes were upregulated compared to FAC unigenes, defined here as upregulated VG unigenes (UVG) (Supplementary Table S1).

Table 1. Summary statistics of the analysis of the *Cotesia chilonis* venom gland and female adult carcass reads.

Transcripts/Unigenes	Raw Reads		Clean Reads		No. of Transcripts/Unigenes	N50	Meanlength
	Venom Gland	Female Adult Carcass	Venom Gland	Female Adult Carcass	-	-	-
Transcripts	67,097,344	63,686,328	65,515,114	62,221,866	115,957	2393	1254
Unigenes					71,617	1309	759

We annotated 23,613 unigenes (32.97%) in GO. At level 2, UVG were classified into nine molecular functional categories (Table 2), among which "structural molecule activity" (GO:0003735) and "structural constituent of ribosome" (GO:0003735) categories were highly represented (42 and 37 unigenes respectively). This may relate to gene transcription and translation. Table 2 shows proteins functionally annotated as "peptidase activity" (GO:0008233), "metallopeptidase activity" (GO:0008237), "hydrolase activity, hydrolyzing *O*-glycosyl compounds" (GO:0004553), "hydrolase activity, acting on glycosyl bonds" (GO:0016798), "phospholipase activity" (GO:0004620), and "phospholipase C activity" (GO:0004629). Venom proteins containing conserved enzymatic domains constitute a large proportion of venom components, although their biological roles are not clear in most cases [4,34].

Table 2. Functional characterization of up regulated unigenes from *Cotesia chilonis* venom gland.

GO_Accession	Description	# Unigenes	Percentages *
GO:0003735	Structural constituent of ribosome	37	13.3
GO:0004553	Hydrolase activity, hydrolyzing *O*-glycosyl compounds	14	5.0
GO:0004620	Phospholipase activity	7	2.5
GO:0004629	Phospholipase C activity	6	2.2
GO:0005198	Structural molecule activity	42	15.1
GO:0008233	Peptidase activity	30	10.8
GO:0008237	Metallopeptidase activity	15	5.4
GO:0016798	Hydrolase activity, acting on glycosyl bonds	14	5.0
GO:0046982	Protein heterodimerization activity	3	1.1

* Percent values show the percent of upregulated VG unigenes (UVG).

2.2. Identification of Secreted Proteins

The venom proteins were separated on 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, with bands ranging from 10 to more than 250 kDa (apparent molecular mass, Figure 1), which we resolved into 922 proteins by LC-MS/MS. The peptide sequences of these proteins matched the VG and FAC unigene databases (Supplementary Table S2), among

which 55 unigenes (Figure 2) were included in the UVG database (Supplementary Table S3). Some of these unigenes correspond to cellular proteins such as ribosomal proteins without predicted signal peptides (e.g., comp44503_c0, comp44417_c0, comp11049_c0, comp11512_c0, comp44398_c0, and comp44381_c0), which were likely contaminants from cell leakage during venom collection. Limiting the venom components in silico to the putative proteins with predicted signal peptides, or with venom homologs in other parasitoid species, totally resulted in 37 candidate venom proteins (Figure 2 and Table 3). Abundant venom proteins are usually more important [1]. The LC-MS/MS system (LTQ-VELOS; Thermo Finnigan, San Jose, CA, USA) and Sequest search algorithm we used could not provide us quantitative information exactly, so here we divided the venom components into three relative transcript levels according to RPKM values: (1) low abundance, RPKM < 100; (2) medium abundance, RPKM from 100 to 1000; and (3) high abundance, RPKM > 1000 (listed in Table 3). Nine components belong to the high-abundance group, including a carboxypeptidase and three esterases implicating the important roles of enzymes played in the venom cocktail. Thirteen venom components fall into low-abundance group, including a mesencephalic astrocyte-derived neurotrophic factor (MANF), two DnaJ homolog subfamily members and two endoplasmic reticulum proteins (ERp). These proteins may be mainly expressed in secretory cells and regulate venom protein folding and secretion. Whether these proteins are actual venom components or tissue contamination is still unclear. We validated the venom proteins by determining the tissue specificity of these venom gene transcription because these venom genes would be expressed exclusively, or nearly so, in the venom gland. Although some possible venom genes, such as FKBP14 (Figure 3E), IEP-2 (Figure 3H), CRT (Figure 3J) and CRF-BP (Figure 3K) were not specifically expressed in the venom gland, we did not exclude their possible roles as virulence factors when expressed in the venom gland, even at low concentration [36,44]. These data indicate specific venom tissue expression of most of the venom genes (Figure 3).

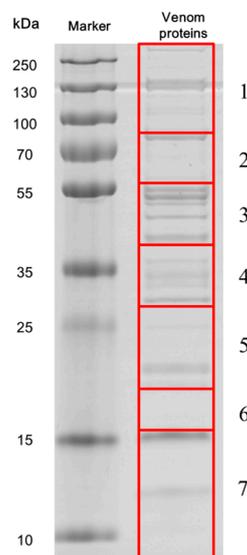In the following sections we discuss the major groups of venom proteins.

Figure 1. SDS-PAGE analysis of *Cotesia chilonis* venom proteins. Proteins were separated on a 12% SDS-PAGE gel and stained with Coomassie Brilliant Blue R-250. The left lane shows molecular weight markers and the right lane shows the venom proteins. The gel was cut into seven large sections, indicated by the red boxes. Tryptic peptides extracted from each section were excised and analysed on LC-MS/MS.

Table 3. Proteins recorded from *Cotesia chilonis* venom.

Protein Name	RPKM/UniquePepCount	Putative Function	Sequence Length (Signal Peptide)	Blast Information (E-Value; Genbank No.; Species)	Gene ID	Accession No.
Hydrolases						
Proteases						
Cc-Ven1	106.52 (M)/2	Serine protease	390 (Y)	$2e^{-130}$; NP_001155078.1; *Nasonia vitripennis*	comp32062_c0	KU663618
Cc-Ven2	47.99 (L)/9	Disintegrin and Metalloproteinase with thrombospondin motifs	509 (Y)	$1e^{-58}$; EGI57486.1; *Acromyrmex echinatior*	comp35977_c0	KU663619
Cc-Ven3	212.13 (M)/11	Tolloid-like protein	478 (Y)	$1.00e^{-94}$; XP_003695260.2; *Apis florea*	comp35892_c0	KU663620
Cc-Ven4	219.32 (M)/30	Tolloid-like protein	477 (Y)	$7.00e^{-124}$; P_003695260.2; *Apis florea*	comp39295_c0	KU663621
Peptidases						
Cc-Ven5	1665.96 (H)/18	Retinoid-inducible serine carboxypeptidase-like	415 (Y)	$5.00e^{-71}$; XP_001605442.1; *Nasonia vitripennis*	comp11050_c0	KU663622
Cc-Ven6	261.11 (M)/1	Aminopeptidase N-like	148 (NC)	$3.00e^{-41}$; EFN65598.1; *Camponotus floridanus*	comp43465_c0	KU663623
Esterases						
Cc-Ven7	4876.97 (H)/9	1-phosphatidyl inositol phosphodiesterase precursor	230 (NC)	$6.00e^{-10}$; XP 007809166.1; *Metarhizium acridum*	comp44319_c0	KU663624
Cc-Ven8	1010.49 (H)/30	Phosphatidylinositol-specific phospholipase	323 (Y)	$4.00e^{-52}$; EZA44899.1; *Microplitis demolitor*	comp43453_c0	KU663625
Cc-Ven9	7010.08 (H)/29	Phosphatidylinositol-specific phospholipase	323 (Y)	$2.00e^{-55}$; EZA44899.1; *Microplitis demolitor*	comp43453_c1	KU663626
Cc-Ven10	194.85 (M)/9	Phospholipase A1	295 (Y)	$1.00e^{-54}$; EGI61859.1; *Acromyrmex echinatior*	comp11611_c0	KU663627
Cc-Ven11	427.00 (M)/14	Phospholipase A1	313 (Y)	$6.00e^{-89}$; XP 011258702.1; *Camponotus floridanus*	comp35865_c0	KU663628
Glycosyl hydrolase						
Cc-Ven12	437.68 (M)/35	Chitooligosaccharidolytic β-*N*-acetylglucosaminidase	603 (Y)	0; XP_008213962.1; *Nasonia vitripennis*	comp35842_c0	KU663629
Endonuclease						
Cc-Ven13	24.39 (L)/1	Enzymatic polyprotein Endonuclease; Reverse transcriptase	1660 (NC)	$1e^{-126}$; XP 002431503.1; *Pediculus humanus corporis*	comp43476_c1	KU663630
Isomerase						
Cc-Ven14	202.99 (M)/12	FK506-binding protein 14	239 (Y)	$6e^{-105}$; NP_476973.1; *Drosophila melanogaster*	comp12229_c0	KU663631
Cc-Ven15	203.89 (M)/14	Protein disulfide-isomerase	2002 (Y)	0; XP_001602967.1; *Nasonia vitripennis*	comp32146_c0	-
Cc-Ven16	412.07 (M)/1	Protein disulfide-isomerase	3554 (Y)	$8e^{-167}$; XP_011138070.1; *Harpegnathos saltator*	comp40525_c0	-
Extracellular superoxide dismutase 3						
Cc-Ven17	10.97 (L)/1	Extracellular superoxide dismutase 3	172 (Y)	$2.00e^{-56}$; AET83769.1/AET83767.1; *Leptopilina boulardi*	comp40765_c0	KU663632
Serine protease inhibitors						
Cc-Ven18	466.16 (M)/1	Kazal-type proteinase inhibitor	120 (Y)	$3.00e^{-22}$; BAO48212.1; *Camponotus japonicus*	comp23198_c0	KU663633
Cc-Ven19	4970.74 (H)/16	Serine protease inhibitor (Serpin)	406 (Y)	$3.00e^{-71}$; XP_008201843.1; *Nasonia vitripennis*	comp23147_c0	KU663634

Table 3. *Cont.*

Protein Name	RPKM/UniquePepCount	Putative Function	Sequence Length (Signal Peptide)	Blast Information (E-Value; Genbank No.; Species)	Gene ID	Accession No.
Other venom proteins (Non enzymatic and inhibitor activities)						
Cc-Ven20	89.09 (L)/10	Immunoevasive protein-2	320 (NC)	$1.00e^{-62}$; BAB72015.1; *Cotesia kariyai*	comp39249_c0	KU663635
Cc-Ven21	103.23 (M)/7	Immunoevasive protein-2	180 (NC)	$5.00e^{-13}$; BAB72015.1; *Cotesia kariyai*	comp42303_c0	KU663636
Cc-Ven22	377.40 (M)/21	Calreticulin	403 (Y)	0; AAN73309.1; *Cotesia rubecula*	comp11089_c0	KU663637
Cc-Ven23	11187.79 (H)/34	Venom protein Ci-48a	384 (NC)	$7.00e^{-14}$; CBM69271.1; *Chelonus inanitus*	comp39158_c0	KU663638
Cc-Ven24	1862.61 (H)/5	Icarapin-like precursor	227 (Y)	$2.00e^{-20}$; NP 001012431.1; *Apis mellifera*	comp44327_c0	KU663639
Cc-Ven25	24.06 (L)/5	Corticotropin-releasing factor-binding protein	327 (Y)	$3.00e^{-95}$; XP_003692566.1; *Apis florea*	comp25125_c0	KU663640
Cc-Ven26	36.45 (L)/8	Vascular endothelial growth factor receptor 2	231 (Y)	e^{-13}; EFN76191.1; *Harpegnathos saltator*	comp23754_c0	KU663641
Cc-Ven27	36.81 (L)/11	Mesencephalic astrocyte-derived neurotrophic factor	842 (Y)	$3.82e^{-66}$; BAM18078.1; *Papilio xuthus*	comp10689_c0	-
Cc-Ven28	59.91 (L)/1	DnaJ homolog subfamily C member 10-like	2650 (Y)	0; XP_001606269.2; *Nasonia vitripennis*	comp32038_c0	-
Cc-Ven29	64.12 (L)/3	DnaJ homolog subfamily B member 11-like	1665 (Y)	0; XP_624603.2; *Apis mellifera*	comp45528_c0	-
Cc-Ven30	67.36 (L)/6	Endoplasmic reticulum protein ERp29	987 (Y)	$3.36e^{-103}$; EGI59372.1; *Acromyrmex echinatior*	comp32086_c0	-
Cc-Ven31	92.63 (L)/1	Endoplasmic reticulum resident protein 44	1350 (Y)	0; XP_624571.2; *Apis mellifera*	comp11316_c0	-
Cc-Ven32	101.69 (M)/52	Heat shock 70 kDa protein	2840 (Y)	0; EFN61604.1; *Camponotus floridanus*	comp35875_c0	-
Cc-Ven33	72.01 (L)/4	Unknown	357 (Y)	-	comp38220_c0	KU663642
Cc-Ven34	84.75 (L)/1	Unknown	132 (Y)	-	comp44992_c0	KU663643
Cc-Ven35	270.62 (M)/4	Unknown	516 (Y)	-	comp23465_c0	KU663644
Cc-Ven36	1465.76 (H)/20	Unknown	291 (Y)	-	comp39041_c2	KU663645
Cc-Ven37	1659.20 (H)/43	Unknown	1033 (Y)		comp39175_c0	KU663646

NC means that prediction of secretion could not be performed due to the incompleteness of the sequence and Y represents the deduced amino acid sequences with predicted signal peptides. L, M, and H represent low, medium and high abundance of unigene expression levels, respectively, according to the RPKM values.
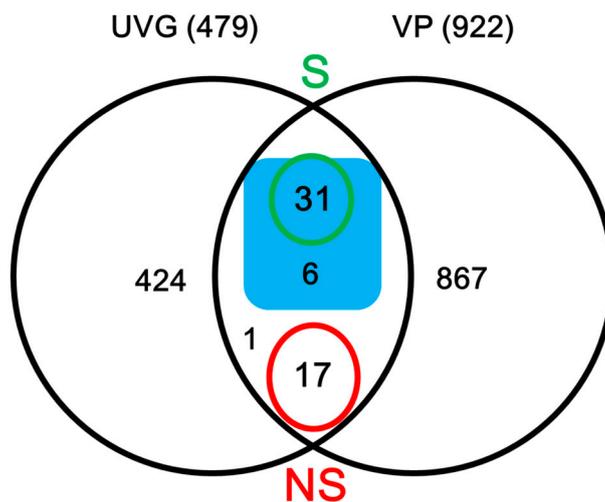
Figure 2. Venn diagram illustrating a likely set of venom proteins. The diagram shows the intersection of UVG unigenes and those matched to the proteomic databases (VP). The numbers in the green and red ellipses represent the numbers of unigenes with complete sequences encoding secreted (green; S) or not secreted proteins (red; NS). The numbers in blue rectangle correspond to the numbers of proteins considered "possible venom proteins" [1,45].
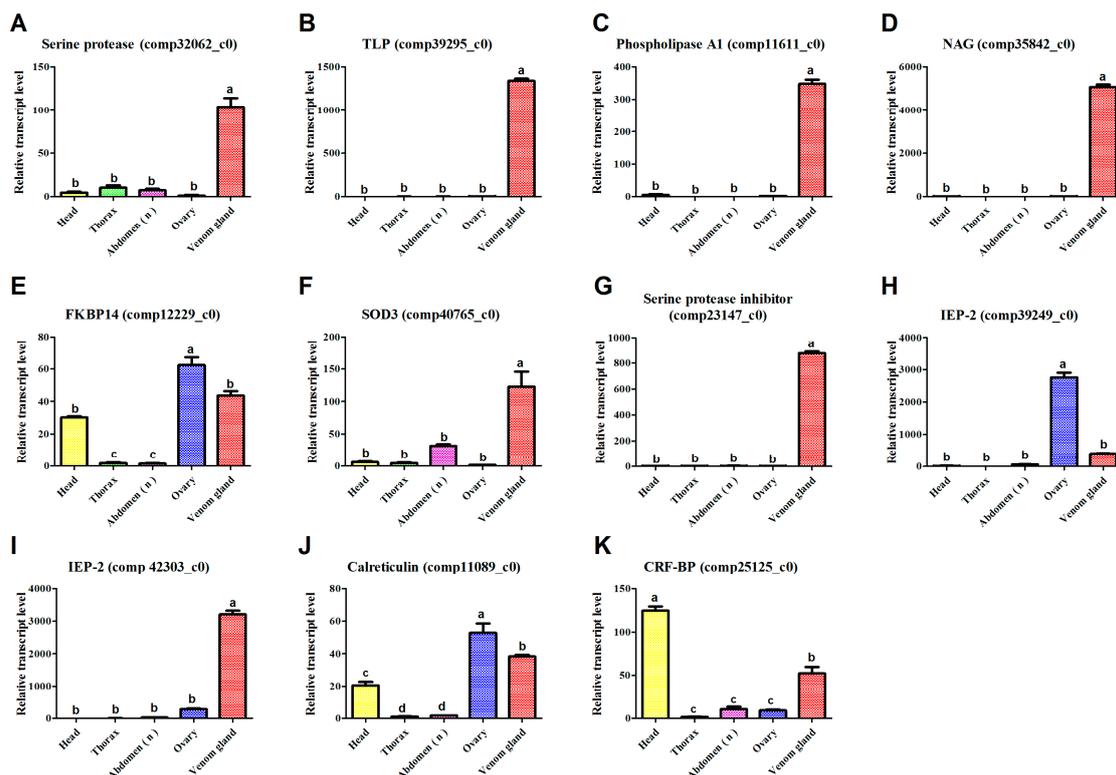

Figure 3. qPCR results showing the mRNA abundance levels of selected unigenes in indicated female tissues. (**A**) Cc-Ven1, Serine protease; (**B**) Cc-Ven4, Tolloid-like protein; (**C**) Cc-Ven10, Phospholipase A1; (**D**) Cc-Ven12, Chitooligosaccharidolytic β-*N*-acetylglucosaminidase; (**E**) Cc-Ven14, FK506-binding protein 14; (**F**) Cc-Ven17, Extracellular superoxide dismutase 3; (**G**) Cc-Ven19, Serine protease inhibitor (Serpin); (**H**) Cc-Ven20, Immunoevasive protein-2; (**I**) Cc-Ven21, Immunoevasive protein-2; (**J**) Cc-Ven22, Calreticulin; (**K**) Cc-Ven25, Corticotropin-releasing factor-binding protein.

2.2.1. Hydrolases

Proteases

Serine protease (SP) Cc-Ven1 (comp32062_c0) shares high sequence similarity (BlastP, 50% identity, E-value = 2e^{-130}) to SP homolog 42 isoform 1 precursor (GenBank: NP_001155078.1) identified from the venom of an ectoparasitoid wasp, *Nasonia vitripennis* (Hymenoptera: Pteromalidae). There were 15 members of the SP family screened from *Nasonia vitripennis* venom using bioinformatic and/or proteomic approaches. The SP family is the best represented of all *Nasonia vitripennis* venom constituents [46]. The functions of all *Nasonia vitripennis* venom SPs have not been thoroughly investigated. Bee venom SP kills insects via melanization and exhibits fibrin(ogen)olytic activity [47]. Serine protease homologs (SPH), formed by an amino acid substitution in the three catalytic triads (His-Asp-Ser), share similar amino acid sequences with active SPs, and function in several physiological processes of insects [9]. For instance, a venom SPH (Vn50) from *Cotesia rubecula* without proteolytic activity (because the conserved serine of the catalytic triad is glycine) significantly inhibited melanization of host hemolymph via suppressing prophenoloxidase activation. The deduced amino acid sequence of Cc-Ven1 is composed by two domains, including a CUB domain (cd00041, interval: 31–112, E-value: 1.93e^{-04}) at N-terminus and a catalytic domain (cd00190, interval: 148–379, E-value: 2.20e^{-53}) at C-terminus. The CUB domains are usually present in the proteins related to development [48]. This domain is commonly characterized by four conserved cysteine residues, forming two pairs of disulfide bonds (Figure 4A). Compared to the CUB-domain SPs (CUBs), the identifications and functions of Clip-domain SPs (CLIPs) are studied very well. The sequences of CUBs and CLIPs are similar with each other in their catalytic domains. According to this point, we did the multiple sequence alignment of Cc-Ven1 with other arthropod SPs, including CLIPs and CUBs, based on their catalytic domains [49]. In the catalytic domain, Cc-Ven1 also contains the three catalytic triads and three disulfide linkages, like the condition in the catalytic domains of other SPs. We found that Cc-Ven1 is more similar to Group I SPs. A difference between Group I and II SPs is that the cleavage sites in all Group II members are followed with Lys or Arg, while the sites in most enzymes of Group I are followed by Leu or Ile. Additionally, the catalytic domains of Group II members contain a short insertion, including two specific Cys residues, present between the catalytic triads His and Asp (Figure 4B) [49]. Group I enzyme contains coagulation factor B and proclotting enzyme of *Tachypleus tridentatus* (Xiphosurida Limulidae), which indicates to us it may act in regulating host hemolymph clotting. However, Cc-Ven1 differs from non-venom SPs by the replacement of the determinant Gly by Ala in the specificity pocket, thus indicating that Cc-Ven1 has a distinct protease target specificity.
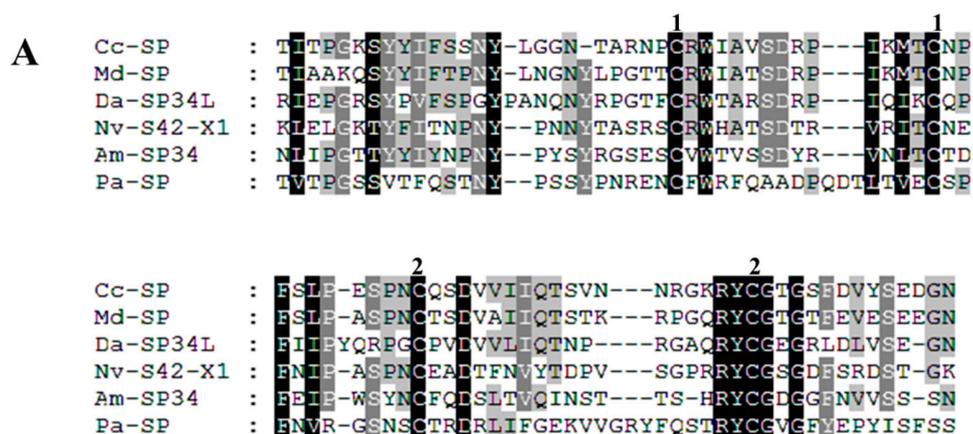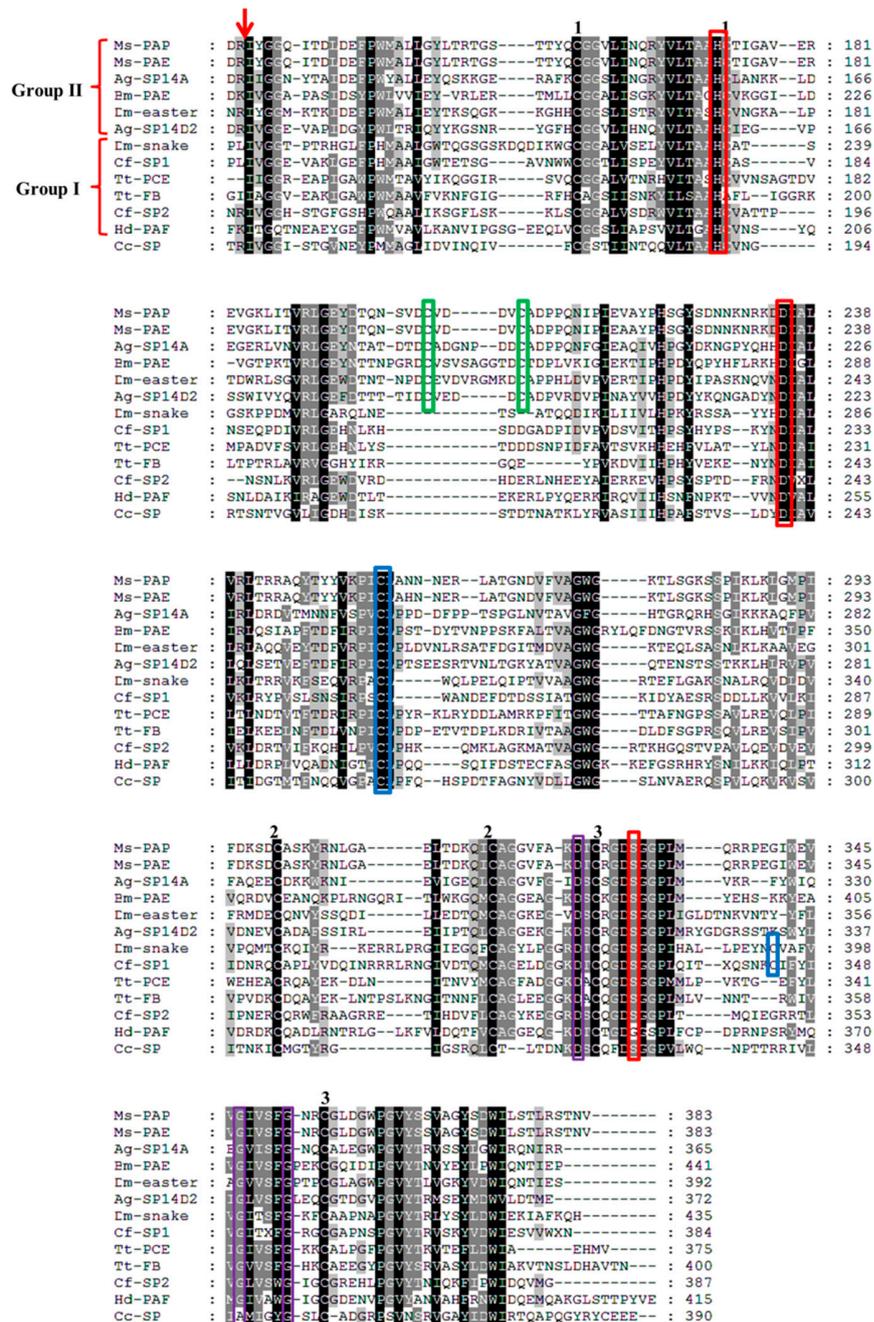

Figure 4. *Cont.*

Figure 4. (**A**) Multiple sequence alignment of CUB domains of serine proteases. The paired numbers above the conserved Cys residues indicate two pairs of the disulfide linkages (**B**) Multiple alignment of the catalytic domains of arthropod serine proteases. The amino acids of the catalytic triad are in the red boxes and the determinants of the specificity pocket are in the purple boxes. The paired numbers above conserved Cys residues indicate the disulfide linkage in the catalytic domain. Cysteines involved in interdomain disulfide bonds are in the blue boxes. The two unique Cys residues in most group II proteases are shown in the green boxes. The amino acid prior to the proteolytic activation site (arrow) is included at the beginning of each sequence. The sequence analyses of catalytic domains were mainly based on [50]. Multiple alignment was performed using ClustalX2. Protein full names and sequence accession numbers are provided in Supplementary Table S4.

For the phylogenetic analysis, we chose a total of 57 of CLIPs and 13 of CUBs, respectively according to recently reported literature [51–54]. We also chose venom CLIPs and CUBs from two

parasitoid wasps: *Nasonia vitripennis* and *Pteromalus puparum* [46,55]. According to [52,56], CLIPs are divided into four subfamilies. Subfamily A (CLIPA) is composed by SPHs solely, whereas subfamilies B (CLIPB), C (CLIPC), and D (CLIPD) comprise SPs, mainly. All of CUBs including the venom CUB, Cc-Ven1 (CUB-V-CcSP) are clustered into the same clade. Two venom CUBs from *Nasonia vitripennis* are clustered into the CUB clade. Three venom CLIPs from *Nasonia vitripennis* are clustered into CLIPB, while two venom CLIPs from *Pteromalus puparum* are clustered into CLIPB and CLIPD, respectively (Figure 5). Four CLIPs may represent lineages of SP derived from ancient evolutionary events since these subfamilies already exist in *Anopheles gambiae* (Diptera: Culicidae), *Drosophila melanogaster*, and *Helicoverpa armigera* (Lepidoptera: Noctuidae) [51,52,56]. Moreover, expansion of individual groups occurred several times to account for the gene clusters observed in *Tribolium castaneum* (Coleoptera: Tenebrionidae), *Anopheles* gambiae, and *Drosophila melanogaster* genomes [51,56,57]. The three venom CLIPs from *Nasonia vitripennis* might originate from SPs in CLIPB, whereas two venom CLIPs of *Pteromalus puparum* might originate from two different subfamilies (CLIPB and CLIPD) and were recruited to venom respectively. Cc-ven1 and two venom CUBs are clustered together in CUB clade, which indicates a similar function they probably achieve. Our phylogenetic analysis implies that the evolutionary processes of these venom SPs from parasitoid wasps may be strongly diverse. However, the essential functions of Cc-ven1 and other venom SPs are still unknown. We infer that venom SPs probably play different roles in the successful parasitization, including increasing the efficiency of venom, suppressing host immunity, and regulating the host development.
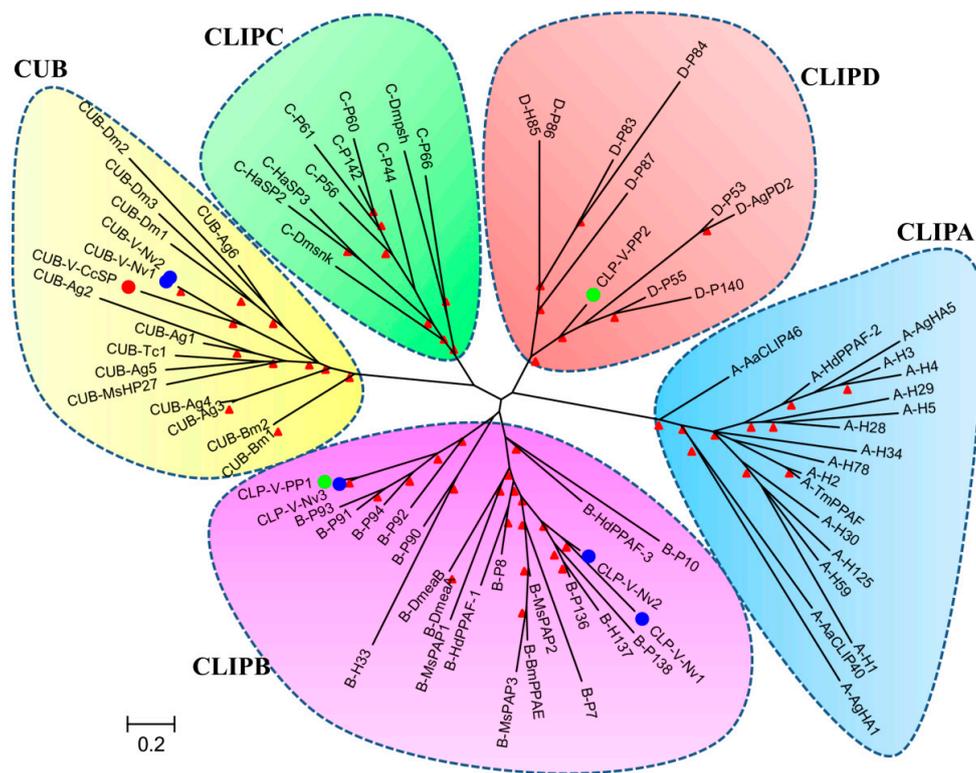

Figure 5. Phylogenetic tree built based on an alignment of the catalytic and protease-like domains of serine proteases and serine protease homologs. Maximum likelihood unrooted tree is based on deduced amino acid sequences of different catalytic and protease-like domains and constructed using PhylML with Smart Model Selection [58]. The best model is LG+G+I. Red arrowheads at nodes indicate bootstrap values higher than 50%. Circles with different colors represent the venom serine proteases from different parasitoid wasps (red: *Cotesia chilonis*; blue: *Nasonia vitripennis*; green: *Pteromalus puparum*). The full names of the proteins and accession numbers of the sequences are listed in Supplementary Table S4.

We identified three metalloproteinases belonging to the metzincin protease superfamily [59]. Cc-Ven2 (comp35977_c0) displays 29% identity (BlastP, E-value = 1e^{-58}) to a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS) (GenBank: EGI57486.1) from the leaf-cutting ant, *Acromyrmex echinatior* (Hymenoptera: Formicidae). ADAMTS proteins are structurally organized into a proteinase and an ancillary domain and they are found in all metazoans, acting in the structure and function of the extracellular matrix [59]. ADAMTS occur in venoms from snakes, social stinging wasps and endoparasitoids, contributing to tissue damage [50,60]. Cc-Ven2 contains an ADAM cysteine-rich domain (ACR) (smart00608, E-value = 4.83e^{-03}), which regulates metalloprotease activity [61], but lacks an ancillary domain responsible for substrate-binding preferences [59]. Truncated single-domain proteins acting as virulence factors are for other venoms. Lacking selectivity domains could make them broadly toxic [62,63].

Cc-Ven3 and Cc-Ven4 share 42% and 44% identity (BlastP, E-value = 1.00e^{-94} and 7.00e^{-124}), respectively, with a Tolloid-like protein 1 (TLP 1) from *Apis florea* (GenBank: XP_003695260.2). They act in activation of growth factors, degradation of polypeptides, serve as venom toxins of brown spiders, digest molecules and spread other toxins through host bodies [64–66]. A typical Tolloid structure contains an N-terminal Astacin domain and C-terminal CUB domains [65], which determine specificity. Cc-Ven3 and Cc-Ven4 feature an Astacin domain (pfam01400) (Cc-Ven3: interval 45–245, E-value = 2.07e^{-65}; Cc-Ven4: interval 45–238, E-value = 2.01e^{-70}) (Figure 6A), and two CUB domains (pfam00431) (Cc-Ven3: interval 246–358, E-value = 3.11e^{-15} and interval 379–473, E-value = 1.42e^{-28}; Cc-Ven4: interval 241–354, E-value = 2.25e^{-22} and interval 358–469, E-value = 1.13e^{-28}). For phylogenetic analysis of proteins containing Astacin domains, we chose proteins according to the reference described by Moran et al. in 2013 [62]. Consistent with the previous results [62], two tolloid-like venom proteins of *Nematostella vectensis* (Actiniaria: Edwardsiidae), NEP-6 and NEP-14 without C-terminal CUB domains, which are different from typical Tolloid structure are clustered into the subgroup of Tolloid proteins, indicating evolved from a tolloid-like ancestor by loss of the CUB domains. By contrast, spider Astacin-like toxins are clustered out of the BMP1/Tolloid subgroup of the Astacin family, which indicates they are probably recruited into venom from other subgroups of Astacin family by losing other C-terminal domains, although the recruiting mechanism is not clear and need to be further studied in future. Here, the phylogenetic results show that Cc-Ven3 and Cc-Ven4 share a much closer evolutionary relationship with Cnidaria venom proteins than with spider venom proteins (Figure 6B). Although the reports of Astacin-like proteins as toxins are limited, the phylogenetic results imply that there may be a great diversity of evolutionary origins in arthropod venom proteins.

Peptidases

We identified two peptidases, CcVen5 and 6. Cc-Ven5 (comp11050_c0) shares 36% identity (BlastP, E-value = 5.00e^{-71}) with a retinoid-inducible serine carboxypeptidase-like protein (GenBank: XP_001605442.1) from *Nasonia vitripennis*. CcVen5 contains a serine carboxypeptidase domain (pfam00450, E-value=6.29e^{-27}). Cc-Ven6 (comp43465_c0; incomplete amino acid sequence) shares 41% identity (BlastP, E-value = 3.00e^{-41}) to an aminopeptidase N-like protein (GenBank: EFN65598.1) from *Camponotus floridanus* (Hymenoptera: Formicidae) CcVen6 contains a peptidase_M1 domain (pfam01433, E-value = 8.50e^{-09}). Peptidases are common components of parasitoid venoms [1]. Dipeptidyl peptidase IV (DDPIV) has been reported in the sting wasp, *Vespa basalis* (Hymenoptera: Vespidae) [67] and angiotensin-converting enzymes (ACEs) have been identified from the venoms of the parasitoid wasps *Chelonus inanitus* and *Nasonia vitripennis*, respectively [4].

Esterases

We identified five esterases. Cc-Ven7 (comp44319_c0) has 40% shared identity (BlastP, E-value = 1.00e^{-43}) to 1-phosphatidylinositol phosphodiesterase-like protein from *Microplitis demolitor* (GenBank: XP_008561140.1). Cc-Ven8 (comp43453_c0) and Cc-Ven9 (comp43453_c1) share about 39% shared identity (BlastP, E-value = 4.00e^{-52}; BlastP, E-value = 2.00e^{-55}) to phosphatidylinositol-specific

phospholipase C from the braconid *Microplitis demolitor* (GenBank: EZA44899.1). Cc-Ven7 and Cc-Ven9 contained a catalytic domain of *Bacillus cereus* (Bacillales: Bacillaceae) phosphatidylinositol-specific phospholipases C (PI-PLCc_BcPLC_like domain, cd08586, E-value = $5.60e^{-20}$ and $4.40e^{-23}$ respectively), and Cc-Ven8 has a catalytic domain of bacterial phosphatidylinositol-specific phospholipase C (PI-PLCc_bacteria_like domain, cd08557, E-value = $5.22e^{-23}$). Bacterial PI-PLCc acts in Ca^{2+}-independent lipid phosphatidylinositol (PI) metabolism [68,69].
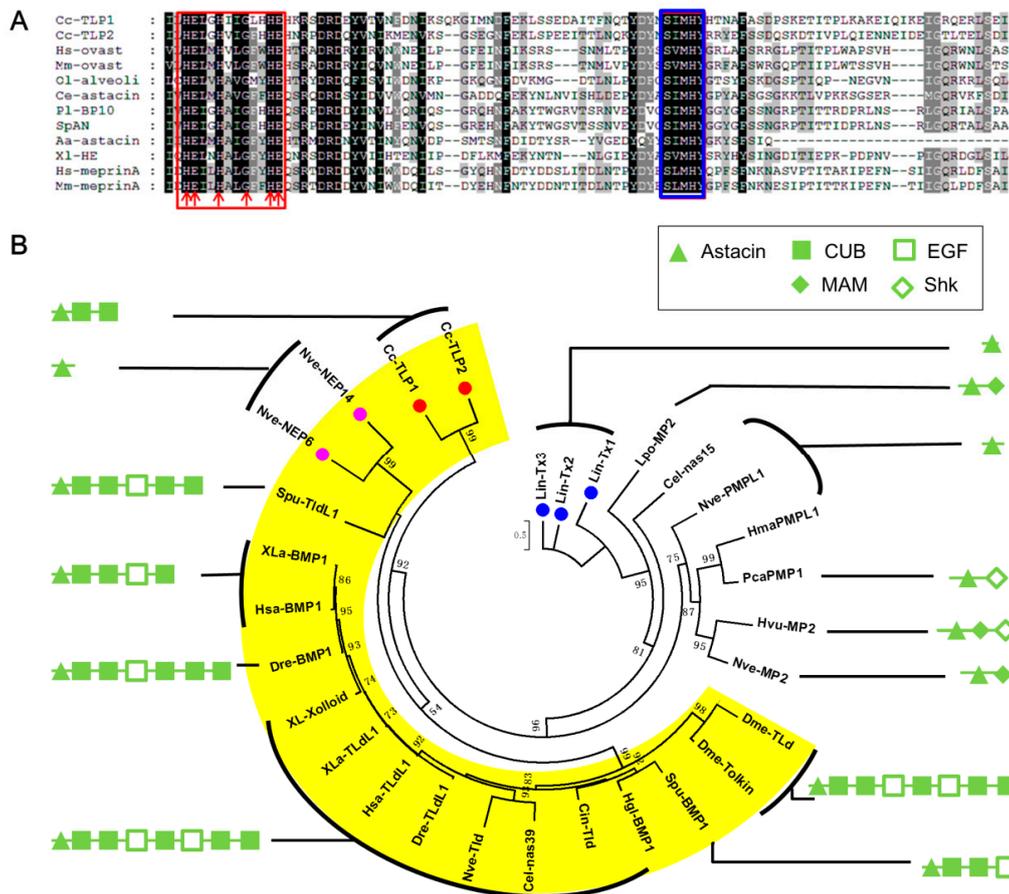
Figure 6. (**A**) Multiple alignment based on parts of catalytic domains of selected astacins. The alignment was performed using ClustalX2. Arrows in the red box point to the astacin signature sequence and blue box indicates the Met-turn; (**B**) Phylogeny of proteins containing Astacin domains. The maximum likelihood unrooted tree is based on deduced amino acid sequences of different Astacin domains and constructed using PhylML with Smart Model Selection [58]. The best model is LG+G+I. Bootstrap values higher than 50% are indicated at each corresponding node. Red circles indicate two astacin-like venom proteins from *Cotesia chilonis*. Purple and blue circles indicate astacin-like venom proteins identified from *Nematostella vectensis* and *Loxosceles intermedia* (Araneae: Sicariidae), respectively. BMP1/Tolloid subgroup of the Astacin family is shaded by yellow. The full names of the proteins and accession numbers of the sequences are listed in Supplementary Table S4.

Glycosyl Hydrolase

Cc-Ven12 (comp35842_c0) has 57% sequence similarity (BlastP, 57% identity, E-value = 0) to a chitooligosaccharidolytic-β-*N*-acetylglucosaminidase (GenBank: XP_008213962.1) from *Nasonia vitripennis*. β-*N*-acetylhexosaminidases (*N*-acetylglucosaminidases, NAGs) catalyze the removal of β-1,4-linked *N*-acetyl-D-hexosamine residues from the ends of *N*-acetyl-β-D-hexosaminides including *N*-acetylglucosides and *N*-acetylgalactosides [70]. NAG is primarily responsible for the production of monomers from chitooligosaccharides for recycling [71].

NAGs have not been found in parasitoid venoms, but a 52-kDa chitinase was identified in the venom of the egg parasitoid *Chelonus* near *curvimaculatus* (Hymenoptera: Braconidae) containing all the conserved residues for chitinolytic activity. A role in dissociation and degradation of host tissue cells for endoparasitoid larval feeding has been proposed [72,73]. comp36742_c0 shows sequence similarity (BlastP, 88% identity, E-value = 0) to a chitinase domain-containing protein 1 from *Microplitis demolitor* (GenBank: XP 008553138.1) containing a Glyco_hydro_18 domain (pfam00704), corresponding to the glycosyl hydrolases family 18 (E-value = $3.17e^{-21}$).

Endonuclease

Cc-Ven13 (comp43476_c1) displays 60% identity (BlastP, E-value = $1e^{-126}$) with the enzymatic ortholog (endonuclease; reverse transcriptase) from the body louse *Pediculus humanus corporis* (Phthiraptera: Pediculidae) (GenBank: XP 002431503.1). Endonucleases have been reported in the venom of some snakes [74], *Physalia physalis* (Cnidaria: Siphonophora) [75], *Chrysaora quinquecirrha* (Cnidaria, Scyphozoa) [76], and *Nasonia vitripennis* [46].

2.2.2. Isomerase

Cc-Ven14 (comp12229_c0) was originally identified as FK506-binding protein 2 (FKBP2) in the transcriptome. Its closest ortholog of six well characterized *Drosophila melanogaster* FKBPs is FKBP14 (GenBank: NP_476973.1), with 72% overall sequence identity (E-value = $6e^{-105}$) (Supplementary Table S5). For consistency with the literature, we rename Cc-Ven14 FKBP14 (Table 3). FKBPs belong to a family of highly-conserved proteins, immunophilins that bind to the immunosuppressive drugs FK506, rapamycin, and cyclosporin A. They often have peptidyl-prolyl cis-trans isomerase (PPIase) activity, which acts in a variety of functions, including protein folding and transcription [77,78]. FKBPs occur in insects, including *Spodoptera Frugiperda* (Lepidoptera: Noctuidae) [79], and *Manduca sexta* (Lepidoptera: Sphingidae) [80].

We totally identified two protein disulfide isomerases (PDIs) from the venom. Cc-Ven15 (comp32146_c0) has 76% identity (BlastP, E-value = 0) to PDI A6 (GenBank: XP_001602967.1) from *Nasonia vitripennis*. Cc-Ven16 (comp40525_c0) displays 49% identity (BlastP, E-value = $8e^{-167}$) with PDI (GenBank: XP_011138070.1) from *Harpegnathos saltator* (Formicidae: Ponerinae). PDIs are involved in the folding of the proteins containing disulphide bonds, via catalytic activities on disulfide bond formation and isomerization [81]. In some venomous organisms, PDIs have functions in the folding of toxin peptides [82,83]. For parasitoids, PDIs have been identified in the venom of *Aphidius ervi* [37] and *Psyttalia* species (Hymenoptera: Braconidae) [84], whose function remains to be elucidated.

2.2.3. Extracellular Superoxide Dismutase 3

Cc-Ven17 (comp40765_c0) has 52% identity (BlastP, E-value = $2.00e^{-56}$) with the extracellular superoxide dismutase 3 (SOD3) from *Leptopilina boulardi* (GenBank: AET83769.1/AET83767.1). SODs catalyze dismutation of superoxide radicals generated via the electron transport chain or via dietary pro-oxidant compounds into molecular oxygen. The SOD3 gene is expressed specifically in venom tissue in *Leptopilina boulardi*, but not *Leptopilina heterotoma* [25,44].

2.2.4. Serine Protease Inhibitors

Cc-Ven18 (comp23198_c0) displays 40% identity (BlastP, E-value = $3.00e^{-22}$) with the kazal-type protease inhibitor from *Camponotus japonicus* (GenBank: BAO48212.1) and Cc-Ven19 (comp23147_c0) displays 36% identity (BlastP, E-value = $3.00e^{-71}$) with a serine protease inhibitor from *Nasonia vitripennis* (GenBank: XP_008201843.1). Serine proteinase inhibitors have been described in several parasitoids, such as *Pimpla hypochondriaca* [85] and *Nasonia vitripennis* [46]. A *Leptopilina boulardi* venom serpin (LbSPNy) (Is_y type) interferes with the PO activation in *Drosophila yakuba* (Diptera: Drosophilidae) host larva hemolymph [24].

2.2.5. Other Venom Proteins

Alongside the enzymatic proteins mentioned above, we identified other components with high identities to the venom components previously characterized from other parasitoid wasps. Those are discussed in the following sections.

Immunoevasive Protein-2

Cc-Ven20 (comp39249_c0) and Cc-Ven21 (comp42303_c0) show 43% identity (BlastP, E-value = $1.00e^{-62}$) and 35% identity (BlastP, E-value = $5.00e^{-13}$), respectively, to immunoevasive protein-2 (IEP-2) (GenBank: BAB72015.1) from *Cotesia kariyai* (Hymenoptera: Braconidae). IEPs protect parasitoid eggs and *Cotesia kariyai* polydnavirus (CkPDV) from encapsulation by host hemocytes but not from hemocytes of larval cutworms, *Spodoptera litura* (Lepidoptera: Noctuidae), an incompatible *Cotesia kariyai* host [86]. The IEP protection for CkPDV depended on a thin surface layer covering virion particle [45].

CRT

Cc-Ven22 (comp11089_c0) has high similarity (BlastP, 98% identity, E-value = 0) to CRT from *Cotesia rubecula* (GenBank: AAN73309.1). CRT has been identified in several parasitoid venoms [34,36,46,87,88] and may act in many conserved roles contributing to the biology of parasites [88]. CRT from *Cotesia rubecula* (CrCRT) inhibited hemocyte-spreading behavior [33]. The coiled-coil domain helps CRT from *Pteromalus puparum* (PpCRT) enter *Pieris rapae* hemocytes, where it inhibits expression of host cellular immunity genes [89]. Siebert et al. used RNA interference directed to the venom CRT gene (*v-crc*) in *Nasonia vitripennis*, coupled with RNA-Seq in the envenomated insect host (vRNAi/eRNA-Seq) to investigate the *Nasonia vitripennis* v-crc, finding it inhibited melanization and reduced bleeding of the host during adult and larval parasitoid feeding [90].

Vascular Endothelial Growth Factor Receptor 2

Cc-Ven26 (comp23754_c0) displays 27% identify (BlastP, E-value = $7.00e^{-13}$) to vascular endothelial growth factor receptor 2 (VEGFR2) from the Indian jumping ant, *Harpegnathos saltator* (GenBank: EFN76191.1). VEGFR occurs in teratocytes of *Microplitis demolitor* [35].

Proteins Involved in the Folding and Export of Secreted Venom Proteins

Additionally, of the PDIs mentioned above, we also identified other proteins probably involved in the folding and export of secreted venom proteins. Cc-Ven27 (comp10689_c0), MANF is able to protect cells against the endoplasmic reticulum stress [91]. Cc-Ven28 (comp32038_c0) and Cc-Ven29 (comp45528_c0), annotated as DnaJ proteins, probably function as cofactors of heat shock proteins 70 kDa (Hsp70) and may modulate protein assembly, disassembly and translocation [92]. Cc-Ven30 (comp32086_c0), annotated as ERp29, is structurally similar to PDI and its function may be for protein folding and secretion [93], Cc-Ven31 (comp11316_c0), annotated as ERp44, probably functions as a multifunctional chaperone of the PDI family [94]. Cc-Ven32 (comp35875_c0) is annotated as a member of HSP70 family, and is one of the common chaperone of ubiquitous class. It is able to control the aspects of cellular proteostasis [95]. Hsp70 proteins were also identified in the venom of the parasitoids *Pteromalus puparum* [87] and two *Psyttalia* species [84], however, their functions are unknown. Actually, molecular chaperones were identified as venom proteins in some other parasitoid wasps [37,84]. For example, HSP70, calreticulin and PDI, were found in two *Psyttalia* species as venom components. It is speculated that they may play roles in the stabilization of other venom proteins, and/or their transport and targeting to the host cells [84]. Endoplasmins, another kind of molecular chaperones, have been identified from both *Aphidius ervi* and *Psyttalia* species venoms. It is reported that this protein is associated with the secretion of pancreatic lipases and their further internalization by intestinal cells

in the rat [96]. It is speculated that it may help the secretion, stabilization, transport and host targeting of other venom proteins.

Other Unknown Proteins

Five *Cotesia chilonis* venom proteins were not similar to identified proteins, including Cc-Ven33: comp38220_c0; Cc-Ven34: comp44992_c0; Cc-Ven35: comp23465_c0; Cc-Ven36: comp39041_c2; Cc-Ven37: comp39175_c0. This is not surprising because many venom proteins have not been identified, as seen in *Nasonia vitripennis* [46], *Chelonus inanitus* [4], and *Microplitis demolitor* [35].

3. Materials and Methods

3.1. Insect Rearing

Laboratory cultures of *Chilo suppressalis* and *Cotesia chilonis* were reared under 28 ± 1 °C, 70%–80% relative humidity, and a 16 h light/8 h dark photoperiod as described [97]. After eclosion, adults were held in glass containers, fed ad libitum on 20% (v/v) honey solution.

3.2. Venom Gland Collection and Total RNA Isolation

Mated female wasps aged 1–3 days were anaesthetized at -70 °C for 5 min, swabbed with 75% ethanol (v/v), dried and then dissected in sterile Pringle's phosphate-buffered saline (PBS) with 3 µL RNase inhibitor at 1 unit/µL (TOYOBO, Osaka, Japan) on an ice plate under a Leica MZ 16A stereomicroscope (Leica, Wetzlar, Germany). Venom glands and carcasses without venom apparatus were collected into TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Total RNA was extracted using TRIzol reagent according to the manufacture's protocol. We used agarose gel electrophoresis, Qubit Fluorometer (Thermo Scientific, Wilmington, DE, USA), Agilent 2100 (Agilent Technologies, Santa Clara, CA USA), and nanodrop 2000 (Thermo Scientific, Wilmington, DE, USA) to determine the quality and quantity of the total RNA samples, respectively.

3.3. Construction and Sequencing of cDNA Library

The construction and sequencing of cDNA library were done by Novogene Bioinformatics Institute (Beijing, China). Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Under elevated temperature (94 °C), fragmentation was carried out using divalent cations in an Illumina proprietary fragmentation buffer. After first and second strand cDNA synthesis, the remaining overhangs were converted into blunt ends via exonuclease/polymerase and enzymes were removed. Illumina PE adapter oligonucleotides were ligated to prepare for hybridization following adenylation of 3′ ends of DNA fragments. The library fragments of preferentially 200 bp with ligated adaptors on both ends were selectively enriched using Illumina PCR Primer Cocktail in a 10 cycle PCR reaction. Products were purified (AMPure XP system) and quantified using the Agilent high sensitivity DNA assay on an Agilent Bioanalyzer 2100 system.

After cluster generation using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA), the library preparations were sequenced on an Illumina Hiseq 2000 platform and 90 bp paired-end reads were generated.

3.4. Data Analysis

Raw reads in fastq format were processed through an in-house Perl scripts and then deposited into the NCBI Short Read Archive (SRA) database [98] by Novogene Bioinformatics (Beijing, China). We obtained the clean reads by removing the reads containing adapters, reads including ploy-N and low quality reads from the raw data. The remaining clean reads were assembled using Trinity v2012-10-05, as described for de novo transcriptome assembly without reference genome using the parameter min_kmer_cov set to two by default, and all other parameters set to default [43]. After assembling, the longest transcript from each transcript cluster was chosen as the unigene, according

to the proposal recommended by Trinity [43]. All unigenes were annotated by blastx search with a cutoff of $1e^{-5}$ against the following databases: NCBI non-redundant protein sequences (Nr), NCBI non-redundant nucleotide sequences (Nt), Protein family (Pfam), Clusters of Orthologous Groups of proteins (KOG/COG), Swiss-Prot (a manually annotated and reviewed protein sequence database), KEGG Ortholog database (KO), and Gene Ontology (GO).

Gene expression levels were estimated with the software RSEM (rsem-1.2.0) for each sample. Clean data were mapped back onto the assembled transcriptome, and the readcount for each gene was calculated from the mapping results. Prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR (3.0.8) through one scaling normalized factor [99]. Differential expression analysis of two samples, including venom glands and carcasses, was performed using the R package DEGSeq v1.2.2 [100]. The *p*-values were adjusted to q-values to account for multiple testing [101]. The *q*-value < 0.005 and |log2(foldchange)| > 1 was set as the threshold for significantly-differential expression. GO enrichment analysis of the differentially-expressed genes was implemented by GOseq R based Wallenius non-central hyper-geometric distribution [102], which adjusts for gene length bias in DEGs. GO terms at the second level was used to perform GO annotation.

3.5. SDS-PAGE of Venom and Protein Identification

Mated female wasps, aged 1–3 days, were dissected in sterile Pringle's PBS with 1 mM phenylmethanesulfonyl fluoride (PMSF) (Sigma, St. Louis, MO, USA) as described above. The venom reservoirs were isolated, washed three times, and transferred into 1.5 mL Eppendorf tubes. After centrifugation at $8000 \times g$ for 10 min at 4 °C, the supernatant was filtered with a 0.22 μm Millipore filter and stored at −70 °C until use. The concentration of venom protein was determined using the Modified Bradford Protein Assay Kit (Sangon Biotech, Shanghai, China).

Venom samples, 78 μg proteins, were separated on 12% SDS-PAGE gels and stained with Coomassie Brilliant Blue R-250. The SDS-PAGE and protein identification were performed by Shanghai Applied Protein Technology Co., Ltd. (Shanghai, China). The gel was cut into seven large sections, depending on the molecular masses of protein bands. Each gel slice was digested with trypsin, lyophilized and analysed on a LC-MS/MS system (LTQ-VELOS; Thermo Finnigan, San Jose, CA, USA). Samples were desalted on Zorbax 300 SB-C18 columns (Agilent Technologies, Wilmington, DE, USA) and then separated on a RP-C18 column (150 μm i.d., 150 mm length) (Column technology Inc., Fremont, CA, USA). Buffer A was water with 0.1% formic acid, buffer B was 84% acetonitrile with 0.1% formic acid. The buffer B gradient was: 0–50 min, from 4% to 50%; 50–54 min, from 50% to 100%; 54–60 min, 100%. The resulting MS/MS spectra were searched against the translated *Cotesia chilonis* transcriptome using the Sequest search algorithm [103]. Carbamidomethyl of cysteine and oxidation of methionine were set as fixed and variable modifications, respectively. Delta CN (≥0.1) and cross-correlation scores (Charge = 1, XCorr ≥ 1.9; Charge = 2, XCorr ≥ 2.2; Charge = 3, XCorr ≥ 3.75) were used to filter the peptide identification.

3.6. qPCR

The total RNA was isolated, separately, from wasp head, thorax, ovary, venom gland, and abdomen without ovary and venom gland as described above. cDNA was synthesized from 1 μg RNA using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Transgen, Beijing, China) for RT-PCR. Specific primers were designed with Primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/) (Supplementary Table S6). qPCR was conducted using the CFX96™ Real-Time PCR Detection System (Bio-rad, Hercules, CA, USA). 28S rRNA, was used as a reference gene. qPCR was done in 25 μL reactions containing 12.5 μL SYBR® *Premix Ex Taq*™ II (Tli RNaseH Plus) (Takara, Otsu, Japan), 1 μL each primer (10 μM), 2 μL cDNA template, 8.5 μL sterile H_2O, programmed at 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 30 s, followed by melting curve analysis.

The relative accumulations of mRNAs in each tissue was calculated using the comparative $2^{-\Delta\Delta CT}$ method [104], following the guidelines of Bustin et al. [105]. We took the lowest mRNA abundance level

as the calibrator and assessed the relative mRNA abundances by comparing the abundances of each target gene in other tissues to the lowest one. The results are presented as mean mRNA abundances of three independent biological replicates. mRNA abundances among tissues were compared using one-way analysis of variance (ANOVA) and Tukey's test with statistical significance set at $p < 0.05$. All statistical analyses were performed using the Data Processing System (DPS) package (Version 9.5) [106].

3.7. Sequence Alignment and Phylogenetic Analysis

Prediction of signal peptide was performed by online software SignalP 4.1 [107]. Multiple sequence alignments of the amino acid sequences were performed with ClustalX2 [108] and edited with GeneDoc. For the phylogenetic analysis of CLIPs, we selected the sequences according to recent references [51–53], from different insect species including *Tribolium castaneum*, *Drosophila melanogaster*, *Anopheles gambiae*, *Bombyx mori* (Lepidoptera: Bombycidae), *Manduca sexta*, *Aedes aegypti* (Diptera: Culicidae), and *Helicoverpa armigera*. We selected CUBs from the same species described above. SPs of *Drosophila melanogaster*, *Anopheles gambiae*, *Bombyx mori* were obtained from NCBI RefSeq protein database, using Blastp with Cc-Ven1 as the query sequence. The sequences from database with the identities of E-value $< e^{-5}$ were chosen. Additionally, we searched the Cc-Ven1 against the BeetleBase [109] to identify the SP of *Tribolium castaneum* (E-value $< e^{-5}$). We used the library of Pfam HMM [110] to confirm the CUBs with a cutoff of 0.001. CUB of *Manduca sexta* we used was according to the reported literature [111]. Three venom CLIPs and one CUB (containing two isoforms) identified from *Nasonia vitripennis* [46], and two venom CLIPs from *Pteromalus puparum* [55] were also selected to construct the phylogenetic tree. For the phylogenetic analysis of proteins containing Astacin domains, the sequences we used were selected accordingly [62]. For the selected proteins used in analyses, not all proteins contained several domains (CUB, EGF, MAM, and Shk), but they all contained the common Astacin domain. The boundaries of catalytic domain (for SPs), proteinase-like domain (for SPHs), and Astacin domain were determined using Pfam. The multiple sequences containing the same domain were aligned using ClustalX2. For the catalytic and proteinase-like domains of SP and SPH, respectively, low quality alignment regions were removed by Gblocks Server [112]. PhyML with Smart Model Selection [58] was used to perform the phylogenetic analyses with 1000-fold bootstrap re-sampling. The full names of the proteins and accession numbers of the sequences used in this study are shown in Supplementary Table S4.

4. Conclusions

We offer several conclusions meant to provide a context for research into parasitoid venoms and their components. First, in their natural and in human-managed population dynamics, parasitoids provide a wide range of biological control services in agroecosystems. These services lead to reduced pest insect damage in cropping systems, reduced use of chemical insecticides, and they help slow the evolution of resistance in pest populations. It follows that research into all aspects of parasitoid biology, including biosystematics, ecology, physiology, biochemistry, and molecular biology contributes critically valuable new knowledge that can be applied to improve parasitoid efficacy. Recent research has led to new understandings and concepts in host-parasitoid relationships. Discovery of the immunosuppressive actions of parasitoid venoms and their active components guided the understanding that parasitoids have evolved specific mechanisms that operate to protect juvenile wasps from the dangerous host immunological reactions to overcome invaders. Looking forward, a large number of findings will drive research into the identities of specific venom components responsible for influencing various aspects of host biology.

Many parasitoid venom components, including proteins, have been reported [1]. Here, we report 37 putative venom proteins of *Cotesia chilonis* using the combination of transcriptomic and proteomic analyses. Our previous physiological research indicates that this parasitoid venom is able to partially inhibit host hemocyte-spreading behavior and encapsulation response to beads [42]. The results also imply that *Cotesia chilonis* uses a combination of passive and positive strategies to

keep its offspring alive in its host larvae [42]. Consistent with the physiological results, we identified not only the proteins involved in active immune suppression, but also the proteins including IEP-2 and Crp32 in contribution to "passive" strategy as the previous descriptions. A passive strategy may include deposition of proteins to protect parasitoid eggs and young larvae. Three egg surface proteins have been identified, including IEP from *Cotesia kariyai*, Crp32 from *Cotesia rubecula* and an *O*-glycosylated protein called hemomucin from *Macrocentrus cingulum* (Hymenoptera: Braconidae), that protect the offspring from encapsulation during the early stages of growth [45,113,114]. Here, we report more evidence on the active and passive mechanisms used by *Cotesia chilonis* to avoid host immune responses.

Changing of selective binding sites of proteases probably alter the enzymes to be the toxins. The replacement of the determinant Gly by Ala in the specific pocket structure of SP (Cc-Ven1) and the loss of an ancillary domain responsible for substrate-binding specification of disintegrin and metalloproteinase (Cc-Ven2) could make them better candidates as virulence factors in the venom of *Cotesia chilonis*, as the condition in other noxious animal species [62]. However, for two Astacin-domain containing metalloproteinases (Cc-Ven3 and Cc-Ven4), the present CUB domains may improve their target specificity. This is probably why two TLPs from *Cotesia chilonis* venom possess relatively low similarity (only 49% of identities).

Without belaboring the point, the idea here is identifying potential biological functions of venom proteins almost immediately leads to testable hypotheses about specific mechanisms of host-parasitoid interactions. Various organisms, including insects, are sources of commercially- and/or biomedically useful proteins and chemistries. Workers in insect science grasp the economic significance of the soil bacterium, *Bacillus thuringiensis*, whose toxins and cognate genes have been highly developed into potent insect management technologies [115]. Commercially valued insect products include honey and pollination services from honey bees, *Apis mellifera* (Hymenoptera: Apidae), silk from the silkworm, *Bombyx mori*, shellac from the lac scale insect, *Laccifer lacca* (Homoptera: Coccidae), indelible iron gall ink from Oak-produced Aleppo galls triggered by the cynipid wasp, *Cynips gallae-tinctoriae* (Hymenoptera: Cynipidae), and the red cochineal dye from the scale *Dactylopius coccus* (Hemiptera: Dactylopiidae), formerly used as a fabric dye and now used in cosmetics [116]. The discovery of insect antimicrobial peptides in the early 1980s (>500 now known) has been developed into a research platform for development of medical antibiotics [117], a still on-going enterprise. Here, the point is that current and future research into parasitoid venoms will lead to discoveries of new chemistries, new proteins and new genes, some of which may become products useful in agriculture and many other areas.

Overall, we foresee research into parasitoid venom components will yield new understanding of parasitology, help generate new concepts and hypotheses and possibly lead to new valuable products.

Author Contributions: Z.-W.T., Q.F., and G.-Y.Y. conceived and designed the experimental plan. Z.-W.T., S.-J.X., G.X., S.-Y.G., and X.C. performed the experiments. Z.-W.T., S.-J.X. and Z.-C.Y. analyzed and interpreted the sequence data and experimental data. Z.-W.T., Q.F., and G.-Y.Y. drafted the manuscript. D.S. revised the manuscript. All authors read and approved the final manuscript.

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

References

1. Asgari, S.; Rivers, D.B. Venom proteins from endoparasitoid wasps and their role in host-parasite interactions. *Annu. Rev. Entomol.* **2011**, *56*, 313–335. [CrossRef] [PubMed]

2. Pennacchio, F.; Strand, M.R. Evolution of developmental strategies in parasitic Hymenoptera. *Annu. Rev. Entomol.* **2006**, *51*, 233–258. [CrossRef] [PubMed]

3. Moreau, S.J.M.; Asgari, S. Venom proteins from parasitoid wasps and their biological functions. *Toxins* **2015**, *7*, 2385–2412. [CrossRef] [PubMed]

4. Vincent, B.; Kaeslin, M.; Roth, T.; Heller, M.; Poulain, J.; Cousserans, F.; Schaller, J.; Poirie, M.; Lanzrein, B.; Drezen, J.M.; et al. The venom composition of the parasitic wasp *Chelonus inanitus* resolved by combined expressed sequence tags analysis and proteomic approach. *BMC Genom.* **2010**, *11*, 693. [CrossRef] [PubMed]

5. Moore, E.L.; Haspel, G.; Libersat, F.; Adams, M.E. Parasitoid wasp sting: A cocktail of GABA, taurine, and β-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host. *J. Neurobiol.* **2006**, *66*, 811–820. [CrossRef] [PubMed]

6. Eldefrawi, A.T.; Eldefrawi, M.E.; Konno, K.; Mansour, N.A.; Nakanishi, K.; Oltz, E.; Usherwood, P.N.R. Structure and synthesis of a potent glutamate receptor antagonist in wasp venom. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4910–4913. [CrossRef]

7. Zhu, J.Y.; Wu, G.X.; Ze, S.Z.; Stanley, D.W.; Yang, B. Parasitization by *Scleroderma guani* influences protein expression in *Tenebrio molitor* pupae. *J. Insect Physiol.* **2014**, *66*, 37–44. [CrossRef]

8. Danneels, E.L.; Formesyn, E.M.; Hahn, D.A.; Denlinger, D.L.; Cardoen, D.; Wenseleers, T.; Schoofs, L.; de Graaf, D.C. Early changes in the pupal transcriptome of the flesh fly *Sarcophaga crassipalpis* to parasitization by the ectoparasitic wasp, *Nasonia vitripennis*. *Insect Biochem. Mol. Boil.* **2013**, *43*, 1189–1200. [CrossRef] [PubMed]

9. Parkinson, N.; Smith, I.; Audsley, N.; Edwards, J.P. Purification of pimplin, a paralytic heterodimeric polypeptide from venom of the parasitoid wasp *Pimpla hypochondriaca*, and cloning of the cDNA encoding one of the subunits. *Insect Biochem. Mol. Boil.* **2002**, *32*, 1769–1773. [CrossRef]

10. Ergin, E.; Uckan, F.; Rivers, D.B.; Sak, O. In vivo and in vitro activity of venom from the endoparasitic wasp *Pimpla turionellae* (L.) (Hymenoptera: Ichneumonidae). *Arch. Insect Biochem. Physiol.* **2006**, *61*, 87–97. [CrossRef] [PubMed]

11. Moreau, S.J.M.; Dingremont, A.; Doury, G.; Giordanengo, P. Effects of parasitism by *Asobara tabida* (Hymenoptera: Braconidae) on the development, survival and activity of *Drosophila melanogaster* larvae. *J. Insect Physiol.* **2002**, *48*, 337–347. [CrossRef]

12. Desneux, N.; Barta, R.J.; Delebecque, C.J.; Heimpel, G.E. Transient host paralysis as a means of reducing self-superparasitism in koinobiont endoparasitoids. *J. Insect Physiol.* **2009**, *55*, 321–327. [CrossRef] [PubMed]

13. Kaeslin, M.; Reinhard, M.; Buehler, D.; Roth, T.; Pfister-Wilhelm, R.; Lanzrein, B. Venom of the egg-larval parasitoid *Chelonus inanitus* is a complex mixture and has multiple biological effects. *J. Insect Physiol.* **2010**, *56*, 686–694. [CrossRef] [PubMed]

14. Mabiala-Moundoungou, A.D.N.; Doury, G.; Eslin, P.; Cherqui, A.; Prevost, G. Deadly venom of *Asobara japonica* parasitoid needs ovarian antidote to regulate host physiology. *J. Insect Physiol.* **2010**, *56*, 35–41. [CrossRef] [PubMed]

15. Asgari, S. Venom proteins from polydnavirus-producing endoparasitoids: Their role in host-parasite interactions. *Arch. Insect Biochem. Physiol.* **2006**, *61*, 146–156. [CrossRef] [PubMed]

16. Labrosse, C.; Eslin, P.; Doury, G.; Drezen, J.M.; Poirie, M. Haemocyte changes in *D. Melanogaster* in response to long gland components of the parasitoid wasp *Leptopilina boulardi*: A Rho-GAP protein as an important factor. *J. Insect Physiol.* **2005**, *51*, 161–170. [CrossRef] [PubMed]

17. Labrosse, C.; Staslak, K.; Lesobre, J.; Grangeia, A.; Huguet, E.; Drezen, J.M.; Poirie, M. A RhoGAP protein as a main immune suppressive factor in the *Leptopilina boulardi* (Hymenoptera, Figitidae)—*Drosophila melanogaster* interaction. *Insect Biochem. Mol. Boil.* **2005**, *35*, 93–103. [CrossRef] [PubMed]

18. Zhang, G.; Schmidt, O.; Asgari, S. A calreticulin-like protein from endoparasitoid venom fluid is involved in host hemocyte inactivation. *Dev. Comp. Immunol.* **2006**, *30*, 756–764. [CrossRef] [PubMed]

19. Richards, E.H.; Dani, M.P.; Bradish, H. Immunosuppressive properties of a protein (rVPr1) from the venom of the endoparasitic wasp, *Pimpla hypochondriaca*: Mechanism of action and potential use for improving biological control strategies. *J. Insect Physiol.* **2013**, *59*, 213–222. [CrossRef] [PubMed]

20. Wu, M.L.; Ye, G.Y.; Zhu, J.Y.; Chen, X.X.; Hu, C. Isolation and characterization of an immunosuppressive protein from venom of the pupa-specific endoparasitoid *Pteromalus puparum*. *J. Invertebr. Pathol.* **2008**, *99*, 186–191. [CrossRef] [PubMed]

21. Mortimer, N.T.; Goecks, J.; Kacsoh, B.Z.; Mobley, J.A.; Bowersock, G.J.; Taylor, J.; Schlenke, T.A. Parasitoid wasp venom SERCA regulates *Drosophila* calcium levels and inhibits cellular immunity. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9427–9432. [CrossRef] [PubMed]

22. Asgari, S.; Zareie, R.; Zhang, G.M.; Schmidt, O. Isolation and characterization of a novel venom protein from an endoparasitoid, *Cotesia rubecula* (Hym: Braconidae). *Arch. Insect Biochem. Physiol.* **2003**, *53*, 92–100. [CrossRef] [PubMed]

23. Asgari, S.; Zhang, G.M.; Zareie, R.; Schmidt, O. A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph. *Insect Biochem. Mol. Boil.* **2003**, *33*, 1017–1024. [CrossRef]

24. Colinet, D.; Dubuffet, A.; Cazes, D.; Moreau, S.; Drezen, J.-M.; Poirie, M. A serpin from the parasitoid wasp *Leptopilina boulardi* targets the *Drosophila* phenoloxidase cascade. *Dev. Comp. Immunol.* **2009**, *33*, 681–689. [CrossRef]

25. Colinet, D.; Cazes, D.; Belghazi, M.; Gatti, J.-L.; Poirie, M. Extracellular superoxide dismutase in Insects characterization, function, and interspecific variation in parasitoid wasp venom. *J. Biol. Chem.* **2011**, *286*, 40110–40121. [CrossRef] [PubMed]

26. Hoch, G.; Schopf, A. Effects of *Glyptapanteles liparidis* (Hym.: Braconidae) parasitism, polydnavirus, and venom on development of microsporidia-infected and uninfected *Lymantria dispar* (Lep.: Lymantriidae) larvae. *J. Invertebr. Pathol.* **2001**, *77*, 37–43. [CrossRef] [PubMed]

27. Zhu, J.Y.; Ye, G.Y.; Dong, S.Z.; Fang, Q.; Hu, C. Venom of *Pteromalus puparum* (Hymenoptera: Pteromalidae) induced endocrine changes in the hemolymph of its host, *Pieris rapae* (Lepidoptera: Pieridae). *Arch. Insect Biochem. Physiol.* **2009**, *71*, 45–53. [CrossRef] [PubMed]

28. Digilio, M.C.; Isidoro, N.; Tremblay, E.; Pennacchio, F. Host castration by *Aphidius ervi* venom proteins. *J. Insect Physiol.* **2000**, *46*, 1041–1050. [CrossRef]

29. Falabella, P.; Riviello, L.; Caccialupi, P.; Rossodivita, T.; Valente, M.T.; De Stradis, M.L.; Tranfaglia, A.; Varricchio, P.; Gigliotti, S.; Graziani, F.; et al. A γ-glutamyl transpeptidase of *Aphidius ervi* venom induces apoptosis in the ovaries of host aphids. *Insect Biochem. Mol. Boil.* **2007**, *37*, 453–465. [CrossRef] [PubMed]

30. Tanaka, T. Effect of the venom of the endoparasitoid, *Apanteles kariyai* Watanabe, on the cellular defence reaction of the host, *Pseudaletia separata* Walker. *J. Insect Physiol.* **1987**, *33*, 413–420. [CrossRef]

31. Stoltz, D.B.; Guzo, D.; Belland, E.R.; Lucarotti, C.J.; Mackinnon, E.A. Venom promotes uncoating in vitro and persistence in vivo of DNA from a braconid polydnavirus. *J. Gen. Virol.* **1988**, *69*, 903–907. [CrossRef]

32. Gupta, P.; Ferkovich, S.M. Interaction of calyx fluid and venom from *Microplitis croceipes* (Braconidae) on developmental disruption of the natural host, *Heliocoverpa zea*, and two atypical hosts, *Galleria mellonella* and *Spodoptera exigua*. *J. Insect Physiol.* **1998**, *44*, 713–719. [CrossRef]

33. Zhang, G.M.; Schmidt, O.; Asgari, S. A novel venom peptide from an endoparasitoid wasp is required for expression of polydnavirus genes in host hemocytes. *J. Biol. Chem.* **2004**, *279*, 41580–41585. [CrossRef] [PubMed]

34. Dorémus, T.; Urbach, S.; Jouan, V.; Cousserans, F.; Ravallec, M.; Demettre, E.; Wajnberg, E.; Poulain, J.; Azema-Dossat, C.; Darboux, I.; et al. Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins. *Insect Biochem. Mol. Boil.* **2013**, *43*, 292–307. [CrossRef] [PubMed]

35. Burke, G.R.; Strand, M.R. Systematic analysis of a wasp parasitism arsenal. *Mol. Ecol.* **2014**, *23*, 890–901. [CrossRef] [PubMed]

36. Goecks, J.; Mortimer, N.T.; Mobley, J.A.; Bowersock, G.J.; Taylor, J.; Schlenke, T.A. Integrative approach reveals composition of endoparasitoid wasp venoms. *PLoS ONE* **2013**, *8*, e64125. [CrossRef] [PubMed]

37. Colinet, D.; Anselme, C.; Deleury, E.; Mancini, D.; Poulain, J.; Azema-Dossat, C.; Belghazi, M.; Tares, S.; Pennacchio, F.; Poirie, M.; et al. Identification of the main venom protein components of *Aphidius ervi*, a parasitoid wasp of the aphid model *Acyrthosiphon pisum*. *BMC Genom.* **2014**, *15*, 342. [CrossRef] [PubMed]

38. Laurino, S.; Grossi, G.; Pucci, P.; Flagiello, A.; Bufo, S.A.; Bianco, G.; Salvia, R.; Vinson, S.B.; Vogel, H.; Falabella, P. Identification of major *Toxoneuron nigriceps* venom proteins using an integrated transcriptomic/proteomic approach. *Insect Biochem. Mol. Boil.* **2016**, *76*, 49–61. [CrossRef] [PubMed]

39. Wu, S.F.; Sun, F.D.; Qi, Y.X.; Yao, Y.; Fang, Q.; Huang, J.; Stanley, D.; Ye, G.Y. Parasitization by *Cotesia chilonis* Influences gene expression in fatbody and hemocytes of *Chilo suppressalis*. *PLoS ONE* **2013**, *8*, e74309. [CrossRef] [PubMed]

40. Lou, Y.G.; Zhang, G.R.; Zhang, W.Q.; Hu, Y.; Zhang, J. Reprint of: biological control of rice insect pests in China. *Biol. Control* **2014**, *68*, 103–116. [CrossRef]

41. Qi, Y.X.; Teng, Z.W.; Gao, L.F.; Wu, S.F.; Huang, J.; Ye, G.Y.; Fang, Q. Transcriptome analysis of an endoparasitoid wasp *Cotesia chilonis* (Hymenoptera: Braconidae) reveals genes involved in successful parasitism. *Arch. Insect Biochem. Physiol.* **2014**, *1*, 1–19.

42. Teng, Z.W.; Xu, G.; Gan, S.Y.; Chen, X.; Fang, Q.; Ye, G.Y. Effects of the endoparasitoid *Cotesia chilonis* (Hymenoptera: Braconidae) parasitism, venom, and calyx fluid on cellular and humoral immunity of its host *Chilo suppressalis* (Lepidoptera: Crambidae) larvae. *J. Insect Physiol.* **2016**, *85*, 46–56. [CrossRef] [PubMed]

43. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef] [PubMed]

44. Colinet, D.; Deleury, E.; Anselme, C.; Cazes, D.; Poulain, J.; Azema-Dossat, C.; Belghazi, M.; Gatti, J.L.; Poirie, M. Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: The case of Leptopilina parasitoids of Drosophila. *Insect Biochem. Mol. Boil.* **2013**, *43*, 601–611. [CrossRef] [PubMed]

45. Furihata, S.; Tanaka, K.; Ryuda, M.; Ochiai, M.; Matsumoto, H.; Csikos, G.; Hayakawa, Y. Immunoevasive protein (IEP)-containing surface layer covering polydnavirus particles is essential for viral infection. *J. Invertebr. Pathol.* **2014**, *115*, 26–32. [CrossRef] [PubMed]

46. De Graaf, D.C.; Aerts, M.; Brunain, M.; Desjardins, C.A.; Jacobs, F.J.; Werren, J.H.; Devreese, B. Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. *Insect Mol. Biol.* **2010**, *19*, 11–26. [CrossRef] [PubMed]

47. 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] [PubMed]

48. Yamada, K.; Takabatake, T.; Takeshima, K. Isolation and characterization of three novel serine protease genes from *Xenopus laevis*. *Gene* **2000**, *252*, 209–216. [CrossRef]

49. Jiang, H.; Kanost, M.R. The clip-domain family of serine proteinases in arthropods. *Insect Biochem. Mol. Boil.* **2000**, *30*, 95–105. [CrossRef]

50. Dos Santos, L.D.; Santos, K.S.; Aparecido Pinto, J.R.; Dias, N.B.; de Souza, B.M.; dos Santos, M.F.; Perales, J.; Domont, G.B.; Castro, F.M.; Kalil, J.E.; et al. Profiling the proteome of the venom from the social wasp *Polybia paulista*: A clue to understand the envenoming mechanism. *J. Proteome Res.* **2010**, *9*, 3867–3877. [CrossRef] [PubMed]

51. Zou, Z.; Evans, J.D.; Lu, Z.; Zhao, P.; Williams, M.; Sumathipala, N.; Hetru, C.; Hultmark, D.; Jiang, H. Comparative genomic analysis of the Tribolium immune system. *Genome Biol.* **2007**, *8*, R177. [CrossRef] [PubMed]

52. Xiong, G.-H.; Xing, L.-S.; Lin, Z.; Saha, T.T.; Wang, C.; Jiang, H.; Zou, Z. High throughput profiling of the cotton bollworm *Helicoverpa armigera* immunotranscriptome during the fungal and bacterial infections. *BMC Genom.* **2015**, *16*, 321. [CrossRef] [PubMed]

53. Waterhouse, R.M.; Kriventseva, E.V.; Meister, S.; Xi, Z.; Alvarez, K.S.; Bartholomay, L.C.; Barillas-Mury, C.; Bian, G.; Blandin, S.; Christensen, B.M.; et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* **2007**, *316*, 1738–1743. [CrossRef] [PubMed]

54. Wang, Y.-H.; Hu, Y.; Xing, L.-S.; Jiang, H.; Hu, S.-N.; Raikhel, A.S.; Zou, Z. A critical role for CLSP2 in the modulation of antifungal immune response in Mosquitoes. *PLoS Pathog.* **2015**, *11*. [CrossRef] [PubMed]

55. Yan, Z.; Fang, Q.; Wang, L.; Liu, J.; Zhu, Y.; Wang, F.; Li, F.; Werren, J.H.; Ye, G. Insights into the venom composition and evolution of an endoparasitoid wasp by combining proteomic and transcriptomic analyses. *Sci. Rep.* **2016**, *6*, 19604. [CrossRef] [PubMed]

56. Christophides, G.K.; Zdobnov, E.; Barillas-Mury, C.; Birney, E.; Blandin, S.; Blass, C.; Brey, P.T.; Collins, F.H.; Danielli, A.; Dimopoulos, G.; et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* **2002**, *298*, 159–165. [CrossRef] [PubMed]

57. Ross, J.; Jiang, H.; Kanost, M.R.; Wang, Y. Serine proteases and their homologs in the *Drosophila melanogaster* genome: An initial analysis of sequence conservation and phylogenetic relationships. *Gene* **2003**, *304*, 117–131. [CrossRef]

58. PhylML with Smart Model Selection. Available online: http://www.atgc-montpellier.fr/phyml-sms/ (accessed on 17 February 2017).

59. Kelwick, R.; Desanlis, I.; Wheeler, G.N.; Edwards, D.R. The ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family. *Genome Biol.* **2015**, *16*, 113. [CrossRef] [PubMed]

60. Takeda, S.; Takeya, H.; Iwanaga, S. Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochim. Biophys. Acta* **2012**, *1824*, 164–176. [CrossRef] [PubMed]

61. Smith, K.M.; Gaultier, A.; Cousin, H.; Alfandari, D.; White, J.M.; DeSimone, D.W. The cysteine-rich domain regulates ADAM protease function in vivo. *J. Cell Biol.* **2002**, *159*, 893–902. [CrossRef] [PubMed]

62. Moran, Y.; Praher, D.; Schlesinger, A.; Ayalon, A.; Tal, Y.; Technau, U. Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Mar. Biotechnol.* **2013**, *15*, 329–339. [CrossRef] [PubMed]

63. 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. Hum. G* **2009**, *10*, 483–511. [CrossRef] [PubMed]

64. Bond, J.S.; Beynon, R.J. The astacin family of metalloendopeptidases. *Protein Sci.* **1995**, *4*, 1247–1261. [CrossRef] [PubMed]

65. Xavier Gomis-Rueth, F.; Trillo-Muyo, S.; Stoecker, W. Functional and structural insights into astacin metallopeptidases. *Biol. Chem.* **2012**, *393*, 1027–1041.

66. Trevisan-Silva, D.; Gremski, L.H.; Chaim, O.M.; da Silveira, R.B.; Meissner, G.O.; Mangili, O.C.; Barbaro, K.C.; Gremski, W.; Veiga, S.S.; Senff-Ribeiro, A. Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the brown spider (genus Loxosceles). *Biochimie* **2010**, *92*, 21–32. [CrossRef] [PubMed]

67. Lee, V.S.Y.; Tu, W.C.; Jinn, T.R.; Peng, C.C.; Lin, L.J.; Tzen, J.T.C. Molecular cloning of the precursor polypeptide of mastoparan B and its putative processing enzyme, dipeptidyl peptidase IV, from the black-bellied hornet, *Vespa basalis*. *Insect Mol. Biol.* **2007**, *16*, 231–237. [CrossRef] [PubMed]

68. Shao, C.; Shi, X.; Wehbi, H.; Zambonelli, C.; Head, J.F.; Seaton, B.A.; Roberts, M.F. Dimer structure of an interfacially impaired phosphatidylinositol-specific phospholipase C. *J. Biol. Chem.* **2007**, *282*, 9228–9235. [CrossRef] [PubMed]

69. Chayakulkeeree, M.; Sorrel, T.C.; Siafakas, A.R.; Wilson, C.F.; Pantarat, N.; Gerik, K.J.; Boadle, R.; Djordievic, J.T. Role and mechanism of phosphatidylinositol-specific phospholipase C in survival and virulence of *Cryptococcus neoformans*. *Mol. Microbiol.* **2008**, *69*, 809–826. [PubMed]

70. Lemieux, M.J.; Mark, B.L.; Cherney, M.M.; Withers, S.G.; Mahuran, D.J.; James, M.N.G. Crystallographic structure of human β-hexosaminidase A: Interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *J. Mol. Biol.* **2006**, *359*, 913–929. [CrossRef] [PubMed]

71. Hogenkamp, D.G.; Arakane, Y.; Kramer, K.J.; Muthukrishnan, S.; Beeman, R.W. Characterization and expression of the β-*N*-acetylhexosaminidase gene family of *Tribolium castaneum*. *Insect Biochem. Mol. Boil.* **2008**, *38*, 478–489. [CrossRef] [PubMed]

72. Jones, D.; Wache, S.; Chhokar, V. Toxins produced by arthropod parasites: salivary gland proteins of human body lice and venom proteins of chelonine wasps. *Toxicon* **1996**, *34*, 1421–1429. [CrossRef]

73. Krishnan, A.; Nair, P.N.; Jones, D. Isolation, cloning, and characterization of new chitinase stored in active form in chitin-lined venom reservoir. *J. Biol. Chem.* **1994**, *269*, 20971–20976. [PubMed]

74. Danpaiboon, W.; Reamtong, O.; Sookrung, N.; Seesuay, W.; Sakolvaree, Y.; Thanongsaksrikul, J.; Dong-din-on, F.; Srimanote, P.; Thueng-in, K.; Chaicumpa, W. *Ophiophagus hannah* venom: Proteome, components bound by *Naja kaouthia* antivenin and neutralization by *N. kaouthia* neurotoxin-specific human ScFv. *Toxins* **2014**, *6*, 1526–1558. [CrossRef] [PubMed]

75. Neeman, I.; Calton, G.J.; Burnett, J.W. Purification and characterization of the endonuclease present in *Physalia physalis* venom. *Comp. Biochem. Phys. B* **1980**, *67*, 155–158. [CrossRef]

76. Neeman, I.; Calton, G.J.; Burnett, J.W. Purification of an endonuclease present in *Chrysaora quinquecirrha* venom. *Proc. Soc. Exp. Biol. Med.* **1981**, *166*, 374–382. [CrossRef] [PubMed]

77. Barik, S. Immunophilins: For the love of proteins. *Cell. Mol. Life Sci.* **2006**, *63*, 2889–2900. [CrossRef] [PubMed]

78. Kang, C.B.; Ye, H.; Dhe-Paganon, S.; Yoon, H.S. FKBP family proteins: Immunophilins with versatile biological functions. *Neurosignals* **2008**, *16*, 318–325. [CrossRef] [PubMed]

79. Alnemri, E.S.; Fernandesalnemri, T.; Pomerenke, K.; Robertson, N.M.; Dudley, K.; Dubois, G.C.; Litwack, G. FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a protein-kinase complex. *J. Biol. Chem.* **1994**, *269*, 30828–30834. [PubMed]

80. Song, Q.S.; Alnemri, E.S.; Litwack, G.; Gilbert, L.I. An immunophilin is a component of the insect ecdysone receptor (EcR) complex. *Insect Biochem. Mol. Boil.* **1997**, *27*, 973–982. [CrossRef]

81. Wilkinson, B.; Gilbert, H.F. Protein disulfide isomerase. *BBA Proteins Proteom.* **2004**, *1699*, 35–44. [CrossRef]

82. Safavi-Hemami, H.; Gorasia, D.G.; Steiner, A.M.; Williamson, N.A.; Karas, J.A.; Gajewiak, J.; Olivera, B.M.; Bulaj, G.; Purcell, A.W. Modulation of conotoxin structure and function is achieved through a multienzyme complex in the venom glands of cone snails. *J. Biol. Chem.* **2012**, *287*, 34288–34303. [CrossRef] [PubMed]

83. Di Luccio, E.; Azulay, D.O.; Regaya, I.; Fajloun, Z.; Sandoz, G.; Mansuelle, P.; Kharrat, R.; Fathallah, M.; Carrega, L.; Esteve, E.; et al. Parameters affecting in vitro oxidation/folding of maurotoxin, a four-disulphide-bridged scorpion toxin. *Biochem. J.* **2001**, *358*, 681–692. [CrossRef]

84. Mathe-Hubert, H.; Colinet, D.; Deleury, E.; Belghazi, M.; Ravallec, M.; Poulain, J.; Dossat, C.; Poirie, M.; Gatti, J.-L. Comparative venomics of *Psyttalia lounsburyi* and *P. concolor*, two olive fruit fly parasitoids: A hypothetical role for a GH1 β-glucosidase. *Sci. Rep.* **2016**, *6*, 35873. [CrossRef]

85. Parkinson, N.M.; Conyers, C.; Keen, J.; MacNicoll, A.; Smith, I.; Audsley, N.; Weaver, R. Towards a comprehensive view of the primary structure of venom proteins from the parasitoid wasp *Pimpla hypochondriaca*. *Insect Biochem. Mol. Boil.* **2004**, *34*, 565–571. [CrossRef] [PubMed]

86. Hayakawa, Y.; Yazaki, K. Envelope protein of parasitic wasp symbiont virus, polydnavirus, protects the wasp eggs from cellular immune reactions by the host insect. *Eur. J. Biochem.* **1997**, *246*, 820–826. [CrossRef]

87. Zhu, J.Y.; Fang, Q.; Wang, L.; Hu, C.; Ye, G.Y. Proteomic analysis of the venom from the endoparasitoid wasp *Pteromalus puparum* (Hymenoptera: Pteromalidae). *Arch. Insect Biochem. Physiol.* **2010**, *75*, 28–44. [CrossRef] [PubMed]

88. Crawford, A.M.; Brauning, R.; Smolenski, G.; Ferguson, C.; Barton, D.; Wheeler, T.T.; Mcculloch, A. The constituents of *Microctonus* sp. parasitoid venoms. *Insect Mol. Biol.* **2008**, *17*, 313–324. [CrossRef] [PubMed]

89. Ferreira, V.; Molina, M.C.; Valck, C.; Rojas, A.; Aguilar, L.; Ramirez, G.; Schwaeble, W.; Ferreira, A. Role of calreticulin from parasites in its interaction with vertebrate hosts. *Mol. Immunol.* **2004**, *40*, 1279–1291. [CrossRef] [PubMed]

90. Siebert, A.L.; Wheeler, D.; Werren, J.H. A new approach for investigating venom function applied to venom calreticulin in a parasitoid wasp. *Toxicon* **2015**, *107*, 304–316. [CrossRef] [PubMed]

91. Liu, H.; Tang, X.; Gong, L. Mesencephalic astrocyte-derived neurotrophic factor and cerebral dopamine neurotrophic factor: New endoplasmic reticulum stress response proteins. *Eur. J. Pharmacol.* **2015**, *750*, 118–122. [CrossRef] [PubMed]

92. Walsh, P.; Bursac, D.; Law, Y.C.; Cyr, D.; Lithgow, T. The J-protein family: Modulating protein assembly, disassembly and translocation. *EMBO Rep.* **2004**, *5*, 567–571. [CrossRef] [PubMed]

93. Baryshev, M.; Sargsyan, E.; Mkrtchian, S. ERp29 is an essential endoplasmic reticulum factor regulating secretion of thyroglobulin. *Biochem. Biophys. Res. Commun.* **2006**, *340*, 617–624. [CrossRef] [PubMed]

94. Sannino, S.; Anelli, T.; Cortini, M.; Masui, S.; Degano, M.; Fagioli, C.; Inaba, K.; Sitia, R. Progressive quality control of secretory proteins in the early secretory compartment by ERp44. *J. Cell Sci.* **2014**, *127*, 4260–4269. [CrossRef] [PubMed]

95. Radons, J. The human HSP70 family of chaperones: Where do we stand? *Cell Stress Chaperones* **2016**, *21*, 379–404. [CrossRef] [PubMed]

96. Bruneau, N.; Lombardo, D.; Levy, E.; Bendayan, M. Roles of molecular chaperones in pancreatic secretion and their involvement in intestinal absorption. *Microsc. Res. Techniq.* **2000**, *49*, 329–345. [CrossRef]

97. Han, L.Z.; Li, S.B.; Liu, P.L.; Peng, Y.F.; Hou, M.L. New artificial diet for continuous rearing of *Chilo suppressalis* (Lepidoptera: Crambidae). *Ann. Entomol. Soc. Am.* **2012**, *105*, 253–258. [CrossRef]

98. NCBI Short Read Archive Database. Available online: http://www.ncbi.nlm.nih.gov/Traces/sra/ (accessed on 12 October 2013).

99. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef] [PubMed]

100. Wang, L.K.; Feng, Z.X.; Wang, X.; Wang, X.W.; Zhang, X.G. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **2010**, *26*, 136–138. [CrossRef] [PubMed]

101. Storey, J.D. The positive false discovery rate: A Bayesian interpretation and the *q*-value. *Ann. Stat.* **2003**, *31*, 2013–2035. [CrossRef]

102. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **2010**, *11*, R14. [CrossRef] [PubMed]

103. Eng, J.K.; McCormack, A.L.; Yates, J.R. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989. [CrossRef]

104. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001**, *25*, 402–408. [CrossRef] [PubMed]

105. Bustin, S.A.; Benes, V.; Garson, J.A.; Hellemans, J.; Huggett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffl, M.W.; Shipley, G.L.; et al. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **2009**, *55*, 611–622. [CrossRef] [PubMed]

106. Tang, Q.Y.; Zhang, C.X. Data Processing System (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research. *Insect Sci.* **2013**, *20*, 254–260. [CrossRef] [PubMed]

107. SignalP 4.1 Server. Available online: http://www.cbs.dtu.dk/services/SignalP/ (accessed on 7 December 2016).

108. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef] [PubMed]

109. BeetleBase. Available online: http://www.bioinformatics.ksu.edu/BeetleBase/ (accessed on 17 February 2017).

110. Pfam HMM. Available online: http://www.ebi.ac.uk/Tools/pfa/pfamscan/ (accessed on 17 February 2017).

111. Cao, X.; He, Y.; Hu, Y.; Zhang, X.; Wang, Y.; Zou, Z.; Chen, Y.; Blissard, G.W.; Kanost, M.R.; Jiang, H. Sequence conservation, phylogenetic relationships, and expression profiles of nondigestive serine proteases and serine protease homologs in Manduca sexta. *Insect Biochem. Mol. Boil.* **2015**, *62*, 51–63. [CrossRef] [PubMed]

112. Gblocks Server. Available online: http://molevol.cmima.csic.es/castresana/Gblocks_server.html (accessed on 17 February 2017).

113. Hu, J.; Xu, Q.Y.; Hu, S.F.; Yu, X.Q.; Liang, Z.K.; Zhang, W.Q. Hemomucin, an *O*-glycosylated protein on embryos of the wasp *Macrocentrus cingulum* that protects It against encapsulation by hemocytes of the host *Ostrinia furnacalis*. *J. Innate Immun.* **2014**, *6*, 663–675. [CrossRef] [PubMed]

114. Asgari, S.; Theopold, U.; Wellby, C.; Schmidt, O. A protein with protective properties against the cellular defense reactions in insects. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3690–3695. [CrossRef] [PubMed]

115. Sanahuja, G.; Banakar, R.; Twyman, R.M.; Capell, T.; Christou, P. *Bacillus thuringiensis*: A century of research, development and commercial applications. *Plant Biotechnol. J.* **2011**, *9*, 283–300. [CrossRef] [PubMed]

116. Irwin, M.E.; Kampmeier, G.E. Commercial products from insects. In *Encyclopedia of Insects*; Resh, V.H., Cardé, R.T., Eds.; Academic Press: San Diego, CA, USA, 2003.

117. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **2002**, *415*, 389–395. [CrossRef] [PubMed]