An integrated transcriptomic and proteomic approach to identify the main *Torymus sinensis* venom components

Carmen Scieuzo1,2,8, Rosanna Salvia1,2,8, Antonio Franco1,2, Marco Pezzi3, Flora Cozzolino4,5, Milvia Chicca2, Chiara Scapoli3, Heiko Vogel6, Maria Monti4,5, Chiara Ferracini7, Pietro Pucci4,5, Alberto Alma7 & Patrizia Falabella1,2,*

During oviposition, ectoparasitoid wasps not only inject their eggs but also a complex mixture of proteins and peptides (venom) in order to regulate the host physiology to benefit their progeny. Although several endoparasitoid venom proteins have been identified, little is known about the components of ectoparasitoid venom. To characterize the protein composition of *Torymus sinensis* Kajino (Hymenoptera: Torymidae) venom, we used an integrated transcriptomic and proteomic approach and identified 143 venom proteins. Moreover, focusing on venom gland transcriptome, we selected additional 52 transcripts encoding putative venom proteins. As in other parasitoid venoms, hydrolases, including proteases, phosphatases, esterases, and nucleases, constitute the most abundant families in *T. sinensis* venom, followed by protease inhibitors. These proteins are potentially involved in the complex parasitic syndrome, with different effects on the immune system, physiological processes and development of the host, and contribute to provide nutrients to the parasitoid progeny. Although additional in vivo studies are needed, initial findings offer important information about venom factors and their putative host effects, which are essential to ensure the success of parasitism.

Insects are characterized by the highest level of biodiversity of all organisms, with around 1 million described species1,2. A molecular and functional understanding of major associations between insects and other organisms may result in the identification of useful genes and molecules for the development of new strategies to control harmful insects3,4, and also for biomedical and industrial applications through biotechnologies5,6. Phytophagous insects are responsible for massive negative impacts on agricultural production, inducing directly and indirectly plant damage, i.e. through the transmission of plant pathogens7. However, insects can be attacked by a wide range of natural antagonists, which are commonly known as entomophages and categorized as predators or parasitoids8. Predators attack and consume their prey directly and immediately8. Parasitoids, instead, preserve the victim for their own progeny by injecting secretions into the hosts that can regulate their physiology, transforming the host into a suitable environment for the development of the new-born larvae9. Hymenopteran parasitoids show the most sophisticated strategies. Whereas ectoparasitoid wasps lay their eggs on the surface of hosts or in the environment close to the host, and their larvae develop outside the host body, endoparasitoid wasps oviposit and their progeny develop inside the host hemocoel10.

The main alterations observed in the parasitized host are induced by secretions (venom and ovarian calyx fluid) injected by the adult female during oviposition10. Venom of parasitoid insects is produced by glands attached to the female reproductive system and consists of a complex mixture of proteins, peptides, and some

1Department of Sciences, University of Basilicata, Via dell’Ateneo Lucano 10, 85100 Potenza, Italy. 2Spinoff XFlies S.R.L., University of Basilicata, Via dell’Ateneo Lucano 10, 85100 Potenza, Italy. 3Department of Life Sciences and Biotechnology, University of Ferrara, Via L. Borsari 46, 44121 Ferrara, Italy. 4Department of Chemical Sciences, University Federico II of Napoli, Via Cintia 6, 80126 Naples, Italy. 5CEINGE Advanced Biotechnology, Via Gaetano Salvatore 486, 80126 Naples, Italy. 6Department of Entomology, Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, 07745 Jena, Germany. 7Department of Agricultural, Forest and Food Sciences, University of Torino, Largo Paolo Braccini 2, 10095 Grugliasco, Italy. 8These authors contributed equally: Carmen Scieuzo and Rosanna Salvia. *Email: hvogel@ice.mpg.de; patrizia.falabella@unibas.it
Bioactive components in venom are responsible for alterations of host development and metabolism, in order to optimize nutrient supply for parasitoid offspring. Unlike endoparasitoids whose venom can induce various effects ranging from the regulation/alteration of host physiology to transiently paralyzing even lethal effects, ectoparasitoid venoms paralyze or rapidly kill the host. To better understand the pathological syndrome observed in parasitized hosts, including alterations of host physiology, development, and reproduction, it is essential to identify and characterize the components of venom and ovarian fluid.

To date, several venom proteins have been identified both in endo- and ectoparasitoid wasps using different approaches. Only after candidate venom protein identification, they can be functionally characterized to understand how they alter the host physiological processes. These alterations, in combination with other parasitic factors, induce changes in the immune system, both humoral (suppression of melanization processes) and cellular (inhibition of the encapsulation of foreign bodies by the hemocytes), in reproductive processes and in the digestive system (host tissues provide suitable nutrients for the parasitoid offspring). Moreover, some of the venom protein components can also play a paralyzing role, preserving the host tissues in response to the nutritional needs of the parasitoid progeny.

Although several venom proteins from endoparasitoids have been identified, proteins from ectoparasitoid venom are still largely unknown. The most reliable strategy for the identification of these proteins consists of an integrated bottom-up proteomic and transcriptomic approach, using a combination of high-throughput next-generation transcriptome sequencing and mass spectrometry.

D. kuriphilus, a univoltine ectoparasitoid, is considered as the main biocontrol agent of D. kuriphilus, and its biological cycle is perfectly synchronized with its host. The adult female inserts its ovipositor in the newly formed galls of D. kuriphilus and lays eggs in the inner wall of the gall or on the surface of the D. kuriphilus larva (Supplementary Fig. S1). Adults of T. sinensis emerge from the gall in early spring and mate, starting the biological cycle again; the lack of a host may cause up to a 12-month diapause. For these reasons, T. sinensis, native to China, has been introduced into several countries of Asia, North America and Europe to control populations of Asian chestnut gall wasps.

Here, we employed an effective approach that combined next-generation transcriptome sequencing and proteomics to identify the major protein components of T. sinensis venom. The transcriptome of the T. sinensis venom gland was built by using a high-throughput nucleic acid sequencing method. Transcriptomic information provided an overall picture of the putative proteins of the venom gland and on their molecular functions, biological processes, and putative cellular compartments. Proteomic analysis was carried out on the components of the venom, fractionated by SDS-PAGE electrophoresis, and analyzed by mass spectrometry (nanoLC-MS/MS). The comparison between translated transcriptomic and proteomic data allowed us to identify the expressed venom proteins. Based on similarities in databases, we obtained a number of functional annotated proteins and a group of novel proteins (without any similarities in databases).

By understanding the role of venom in parasitized hosts, we hope to apply these molecules as bioinsecticides in integrated pest control.

Results
Transcriptome assembly and functional analysis by gene ontology. Next-generation sequencing (RNAseq) performed with RNA isolated from the venom glands of T. sinensis (Fig. 1a) allowed us to generate a de novo transcriptome assembly, which contained 22,875 contigs, with a maximum contig length of 19,306 bp. The six reading frames of the 22,875 nucleotide sequences were translated into their corresponding amino acid sequences, resulting in 137,250 predicted proteins ("T. sinensis protein database").

To define similarities with annotated proteins, the contig sequences of the de novo transcriptome of T. sinensis venom glands were searched using the BLASTx algorithm against a non-redundant (nr) NCBI protein database with an E-value cut-off of 10^-5 identifying 7,466 contigs (33%) matching entries. The species distribution of the top BLAST hit against the nr database for the T. sinensis venom gland transcriptome showed that the majority of obtained top hits matched N. vitripennis (Fig. 2).

For functional annotation, all sequences were subjected to gene ontology (GO) analysis in Blast2GO, with 12,714 (56%) of the 22,875 contigs showing significant similarity to proteins with assigned molecular functions in the GO database, whereas 44% of the total transcripts did not match any annotated sequences in the nr database indicating a consistent group of noncoding transcript and species-specific putative proteins or more probably transcript not annotated yet.

The 12,714 annotated contigs were classified into three GO categories: biological processes, cellular components and molecular functions. The most prominent GO Cellular Component categories (Level 3) were organelles (51%) (Fig. 3a). The most prominent Molecular Function (Level 3) were different kinds of protein with binding activity (74%), followed by hydrolases (10%) and transferases (9%) (Fig. 3b). The most prominent category of GO Biological Processes (Level 2-3-4) was composed of proteins involved in metabolic and cellular processes (Fig. 3c-e). This result was linked to the very large number of general GO terms, terms that include basic processes for a living organism.

A GO analysis was performed on the identified 195 venom proteins (Fig. 4). The most abundant categories of Biological Processes (Level 4) were macromolecules, proteins and organonitrogen metabolic processes (Fig. 4a). Four macro-categories were identified through Molecular Function (Level 4) analysis: peptidases, serine proteinases, hydrolases and cation binding activity proteins (Fig. 4b). A further “Enrichment Analysis” highlighted that proteins with proteolytic and serine-type endopeptidase activity are the most abundant in T. sinensis venom, comparing venom protein components and total T. sinensis transcripts (Fig. 4c).
As previously reported, contigs not matching any known sequences in the nr database are relevant. The relatively high number of contigs without significant BLAST matches is quite common in the case of de novo transcriptome assemblies obtained from RNA-Seq data, and could be due to fragmented assembly, resulting in larger numbers of contigs corresponding to untranslated (UTR) regions of cDNAs, noncoding transcripts or species-specific un-annotated orphan genes. In *T. sinensis* venom transcriptome, the enzyme code distribution shows that the most abundant families of enzymes were transferases and hydrolases (Fig. 5).

A quantitative RNA-Seq analysis of the transcripts encoding for putative venom proteins based on the Reads Per Kilobase of transcript method, per Million mapped reads (RPKM) showed large differences in expression levels; the most expressed transcripts in the venom glands are reported in Table 1.

**Gene expression levels in *Torymus sinensis* venom gland.** To confirm the specific expression of genes in *T. sinensis* venom gland, a quantitative real time PCR (qPCR) experiment was performed on RNA extracted from venom glands and females deprived of the venom gland, focusing on a sample set of 10 genes selected among transcripts reported in Table 1. Female body without venom glands was used as calibrator. A significantly higher transcription level in venom glands was observed for all the selected genes (Fig. 6).

**SDS-PAGE and mass spectrometry.** The one-dimensional SDS-PAGE electrophoretic profile showed that the *T. sinensis* venom consists of a complex protein mixture whose molecular weights range from 10 to 250 KDa (Fig. 1b). The sample lane containing the venom proteins was cut into thin slices, and the protein bands
were excised, destained and subjected to in situ proteolytic digestion with trypsin. Protein identification was performed by direct nanoLC–MS/MS analysis of the corresponding peptide mixtures. Mass spectral data were processed with the Andromeda search engine and then compared with the "T. sinensis protein database", leading to the identification of the proteins expressed in venom. Mass spectrometry analysis provided a list of 15,381 peptide sequences.

Proteomic and transcriptomic data analysis. The comparison between proteomic and transcriptomic data sets led to the identification of 1322 contig. The surprisingly large number of proteins identified by mass spectrometry is common\(^\text{20,27,28,45}\). Further analyses were performed with the software Signal P 5.0 (http://www.cbs.dtu.dk/services/SignalP/) in order to select only proteins bearing the signal peptide and therefore are likely to be secreted. This in silico prediction allowed us to select a total of 195 secretory proteins.

To further confirm the venomous nature of these 195 proteins, their sequences were aligned and compared to venom proteins from \(N.\) \(vitripennis\), described in de Graaf et al.\(^\text{30}\), obtaining 74 matches reported in Supplementary Tables S1, S2 and S3. Among the 195 \(T.\) \(sinensis\) venom proteins, 121 proteins with signal peptides did not match any \(N.\) \(vitripennis\) venom protein. Although two contigs (12 and 2618) are not annotated, as they have no similarity with other known proteins, 69 proteins were found to show significant similarity with venom protein components of other parasitoid wasps or venomous animals or were uncharacterized proteins.

The 143 putative venom protein identified with the proteomic approach are listed in Supplementary Table S4 (Fig. 1c). The remaining 50 proteins without any similarity to known components of insect or other animal venom were considered either ‘venom trace elements’ with a limited function in the venom duct or in the reservoir, or contaminants of the venom gland released during the dissection of the gland tissue\(^\text{30,46}\). In addition to the proteomic approach, a key-word approach was used to identify a further group of putative venom protein in the venom gland transcriptome: all putative proteins annotated with the word "venom" were selected. 52 contigs annotated with "venom" keyword were identified (Supplementary Tables S5, S6; Fig. 1c). In addition, 2 contigs annotated with the "toxin" keyword were also identified (Supplementary Tables S5, S6). 18 of these contigs showed the signal peptide while this signature was missing in 34 contigs (Supplementary Table S5). Among the contigs missing the signal peptide, 18 had an incomplete sequence at 5’ end, 14 had a complete sequence and 2 had an open reading frame (ORF) too small to give information about signal peptide (Supplementary Table S6).

To further confirm that the 52 putative proteins annotated using "venom" as a keyword belonged to the mixture of \(T.\) \(sinensis\) venom protein components, the nucleotide sequences of these proteins were translated using Expasy-translate tool software (https://web.expasy.org/translate/), and the corresponding protein sequences were
compared to those of N. vitripennis, reported in de Graaf et al.\textsuperscript{30}, obtaining 46 matches reported in Supplementary Tables S7, S8, S9. 9 T. sinensis contigs did not match any of the N. vitripennis venom proteins but were similar to other venom protein components, including carboxylesterases, Kazal-type serine protease inhibitors, acid phosphatases and carboxypeptidases (Supplementary Table S10).

All T. sinensis venom proteins, identified by the proteomic approach or exclusively by using the transcriptomic approach, can be classified in different groups, according to their functional annotation (Table 2). Among the identified proteins, hydrolases constituted the most abundant family followed by protease inhibitors, recognition and binding proteins, isomerases and dehydrogenases. These findings are in agreement with data reported for the venom of other parasitoids\textsuperscript{18,28,29,47}. Among the hydrolases in the T. sinensis venom glands, we identified proteases and, in particular, serine proteases, metalloproteases, different kinds of peptidases, proteins of carbohydrate metabolism, lipases, phosphatases and nucleases.

Discussion
In order to identify the venom proteins of the ectoparasitoid Torymus sinensis, a combined proteomic and transcriptomic approach was used (hereafter defined as the proteomic approach). The transcriptomic analysis provided general information about putative proteins of the venom gland, focusing on their molecular functions, biological processes, and cellular compartments. This first level of analysis allowed us to select putative candidate proteins, annotated as “venom/toxin”. The proteomic analysis was performed on the venom extract.
using mass spectrometry, and the expressed proteins were identified when their sequences matched with the translated transcripts (contigs) of the venom gland transcriptome. Proteins containing a putative signal peptide for extracellular localization and predicted cleavage sites were reported as venom-expressed components. All putative (predicted from the transcriptome) or expressed (confirmed by the proteomic approach) proteins identified as venom protein components were further compared to *Nasonia vitripennis* venom proteins described by de

**Figure 4.** Gene Ontology sequence annotation of *T. sinensis* 195 venom proteins. Gene Ontology (GO) assignments for the *T. sinensis* venom proteins. GO assignments as predicted for their involvement in (a) biological processes and (b) molecular functions. All data are presented at level 4 GO categorization. Classified gene objects are depicted as absolute numbers and percentages (in brackets) of the total number of gene objects with GO assignments. In (c) enriched distribution of Gene Ontology (GO) terms in *T. sinensis* venom proteins were identified. Bar charts show the GO terms that were significantly (false discovery rate (FDR) < 0.05) enriched in the group of venom proteins compared to the complete *T. sinensis* gland transcriptome. The GO terms are sorted in an ascending order according to their FDR value, starting with the most significantly enriched. Only the most specific GO terms are displayed. Differences are shown as the percentage of sequences associated with a specific GO category in the test set (venom protein-encoding contigs) versus the reference set (transcriptome backbone assembly) using Fisher’s exact test in OmicsBox.

**Figure 5.** Enzyme Code (EC) Classes of the *T. sinensis* contigs encoding enzymes. The most abundant families of enzymes found in the transcriptome of *T. sinensis* venom glands are shown.
Graaf et al., as both ectoparasitoids belong to the Chalcidoidea superfamily. The comparison with *N. vitripennis* venom proteins further confirmed the venomous nature of *T. sinensis* predicted and expressed proteins. We also identified putative or expressed *T. sinensis* venom proteins showing similarity to venom protein components of other parasitoids or animals. Lastly, a group of *T. sinensis* venom proteins was identified with sequence similarities to unknown, hypothetical or uncharacterized proteins. All the identified proteins were grouped according to their functions, and their possible role in the complex parasitic syndrome after envenomation was discussed.

Hydrolases. Hydrolases have been found in several endo- and ectoparasitoid venoms. The hydrolases identified in *T. sinensis* venom can be grouped in different classes: proteases (serine proteases, metalloproteases, dipeptidases, amino- and carboxypeptidases), phosphatases, esterases, lipases, glucosidases, galactosidases, amylases, trehalases, nucleases. Some of these proteins belong to carbohydrate metabolism, including glucosidases, galactosidases, amylases and trehalases.

Proteases have been reported as abundant components in numerous parasitoid venoms. Among proteases, serine proteases are the most abundant family, including trypsins, chymotrypsins and serine protease homologues. Serine proteases and serine protease homologues were found in different parasitoid species, such as the endoparasitoids *Aphidius ervi* (49), *Microplitis mediator* (47), *Pimpla hypochondriaca* (24), *Pteromalus puparum* (29), *Toxoneuron nigriceps* (18) and *Cotesia rubecula* (50). Serine proteases can play a crucial role in regulating

| Contig | Protein name | RPKM |
|--------|--------------|------|
| 445    | Alpha-amylase-like | 9,7718 |
| 4      | Serine protease 33 isoform X2 | 9,71077 |
| 270    | Trypsin-like | 9,66998 |
| 1650   | Peptidyl-prolyl cis-trans isomerase 5 | 9,54057 |
| 157    | Venom protein U precursor | 9,49706 |
| 1105   | Serine protease inhibitor 3/4 isoform X16 | 9,17066 |
| 293    | Cathepsin L | 8,64858 |
| 2095   | Carboxylesterase clade B, member 2 precursor | 8,59989 |
| 289    | Low-density lipoprotein receptor-related protein 2-like | 8,58940 |
| 428    | Lyosomal aspartic protease-like | 8,45276 |
| 940    | Heat shock protein 70 | 8,39405 |
| 1194   | Ferritin precursor | 8,33117 |
| 1001   | Calreticulin precursor | 8,31852 |
| 1747   | Endoplasmin | 8,09155 |
| 1971   | Chymotrypsin-1-like | 7,94740 |
| 392    | Lipase 3-like | 7,90375 |
| 181    | Protein disulfide-isomerase A3 | 7,84263 |
| 495    | Pancreatic triacylglycerol lipase | 7,75945 |
| 4156   | Ribonuclease Oy-like | 7,67279 |
| 3120   | Venom allergen 5-like | 7,52323 |
| 874    | Adipocyte plasma membrane-associated protein-like | 7,17361 |
| 2736   | Chitotriosidase-1 isoform X1 | 7,14344 |
| 128    | Pancreatic triacylglycerol lipase | 7,03636 |
| 1330   | Alpha-glucosidase-like | 6,95432 |
| 1835   | Serine protease homolog 21 precursor | 6,76623 |
| 1890   | Ovomucoid-like | 6,61221 |
| 3708   | Trypsin beta-like | 6,55134 |
| 2538   | Apolipoprotein D-like | 6,40265 |
| 575    | Mesencephalic astrocyte-derived neurotrophic factor homolog | 6,31104 |
| 2298   | Dipeptidase 1 isoform X1 | 6,29161 |
| 69     | Trypsin-3-like | 6,28669 |
| 18     | Serine protease inhibitor 28Dc isoform X2 | 6,21690 |
| 1034   | Vitellogenin | 6,21393 |
| 2025   | Serpin 5 precursor | 6,17429 |
| 5412   | Calnexin isoform X2 | 6,02715 |
| 3815   | Inosine-uridine preferring nucleoside hydrolase-like precursor | 6,02358 |

Table 1. Most expressed *T. sinensis* venom proteins. Proteins identified by SDS-PAGE and LC–MS/MS of venom extract, ordered by descending RPKM values. Expression levels are reported as log2-transformed RPKM values.
the immune system by inhibiting melanization in the host hemolymph blocking the phenoloxidase cascade\textsuperscript{51,52}. Generally, serine protease homologues can act as co-enzymes for prophenoloxidase-activating proteinases and are important for the activation of prophenoloxidase and melanization\textsuperscript{53}. In \textit{C. rubecula}, a particular protease named Vn50 was structurally similar to other serine protease homologs but showed a different function. Vn50 inhibits melanization presumably by competing with host serine protease homologs for binding the prophenoloxidase that remains un-cleaved in the hemolymph\textsuperscript{50,51}. Besides the action of serine protease homologues, serine protease may also play a role in down-regulating prophenoloxidase\textsuperscript{20,54}. A different model for the involvement of venom serine protease in the melanization response was proposed by Choo et al.\textsuperscript{55} In the venom of \textit{Apis mellifera} and \textit{Bombus} spp., serine proteases seemed to be involved in the hyperactivation of prophenoloxidases, resulting in an excessive melanization response leading the target insects (\textit{Bombyx mori}, \textit{Spodoptera exigua} and \textit{Pieris rapae}) to death. A similar poisoning effect may be induced by injecting ectoparasitoid venoms into the host. In addition to its involvement in host immunity disruption, a serine protease in \textit{N. vitripennis} venom showed a putative cytotoxic function in assays with a \textit{Spodoptera frugiperda} cell line\textsuperscript{56}, whereas a trypsin-like enzyme found in \textit{salivary secretions of the ectoparasitoid Euplectrus separatae} larva seemed to be able to digest host tissues\textsuperscript{57}, prompting the hypothesis that serine proteases are extensively involved in the parasitic syndrome. In \textit{T. sinensis} venom, many serine proteases were identified: trypsins (peptides identified by the proteomic approach matching to 25 contigs), chymotrypsins (peptides identified by the proteomic approach matching to 6 contigs), general serine proteases (peptides identified by the proteomic approach matching to 7 contigs, and 2 contigs identified by the transcriptomic approach) and a serine protease homologue (identified by the proteomic approach). Seven trypsins, and two general serine proteases identified in \textit{T. sinensis} venom did not show the classic serine protease catalytic site, composed of serine, histidine and aspartic acid.

In the predicted serine proteases identified from the \textit{T. sinensis} data, the relative positions of the amino acids of catalytic triad are mostly conserved: histidine at position 70, aspartic acid at positions 125 and 130, and serine at positions 240 and 245. A few anomalies have to be mentioned: in three contigs, the catalytic triad showed slight differences; in a trypsin, the serine residue is substituted by arginine, and in the serine protease homologue, the histidine residue is substituted by glycine. For some contig-derived amino acid sequences, the absence of the catalytic triad can be explained by the incompleteness of the contig sequence.

The high number of serine proteases identified as part of the complex \textit{T. sinensis} venom and the absence of clear one-to-one orthology to other insect proteases (Supplementary Fig. S2), including \textit{N. vitripennis} venom proteins, indicates species-specific gene duplication events in \textit{T. sinensis}. Such large-scale gene duplication events of the serine protease gene family is a frequent phenomenon\textsuperscript{29,58} and we suspect that a greatly expanded set of proteases is associated with complex venom function.

Metalloproteases are involved in several biological and disease-related processes, such as intracellular signaling, matrix degradation, inflammation, and coagulation disorders\textsuperscript{59,60}. In insect, metalloproteases are related to the immune response; indeed zinc-dependent proteases were highly expressed in \textit{Manduca sexta} larvae infected with bacteria, and one of these proteases was quite similar to human neutral endopeptidase NEP 24.11, which is involved in the immune response\textsuperscript{61}. In addition, metalloproteases were found in the venom of \textit{P. hypochondriaca}\textsuperscript{24}.

---

**Figure 6.** Relative expression level of ten selected genes of \textit{T. sinensis} venom gland and \textit{T. sinensis} female body deprived of venom gland. Gene expression levels were quantified by quantitative real time PCR (qPCR). Data represent the mean of three independent replicates ± SEM. Samples were compared by the Unpaired t-test and statistically significant differences between samples are indicated with asterisk (*p = 0.015, **p = 0.011, ***p = 0.0009, ****p < 0.0001). Reference genes: GAPDH and beta-tubulin. Calibrator sample: female body deprived of venom gland.
| Protein category         | Contig number                                                                 | Approach     |
|--------------------------|-------------------------------------------------------------------------------|--------------|
| **Hydrolases**           |                                                                               |              |
| Serine proteases         | 1, 4, 25, 69, 70, 87, 125, 235, 270, 368, 378, 706, 753, 775, 789, 991, 1058, 1059, 1143, 1233, 1310, 1617, 1636, 1794, 1835, 1855, 1907, 1971, 2101, 2283, 2399, 3012, 3708, 5058, 6126, 6382, 11,004, 11,592 | Proteomic   |
|                         | 7239, 8948                                                                   | Transcriptic |
| Metalloproteases         | 8199                                                                         | Proteomic    |
| Dipeptidases             | 2298, 6303, 11,199                                                           | Proteomic    |
|                         | 7070                                                                         | Transcriptic |
| Aminopeptidases          | 5399, 7646, 18,029                                                           | Proteomic    |
| Lysosomal proteases      | 293                                                                          | Proteomic    |
| Carboxypeptidases        | 1049, 12,519                                                                 | Proteomic    |
|                         | 7893, 15,076                                                                 | Transcriptic |
| Venom acid phosphatases  | 3591, 3878                                                                   | Proteomic    |
|                         | 3670, 7321, 7343, 7836, 8187, 8691, 8857, 10,928, 13,041, 13,453, 14,872, 14,873, 14,873, 16,876, 16,269, 18,015, 18,733, 18,844, 20,332 | Transcriptic |
| Carboxylesterases        | 106, 1651, 1652, 2095, 2872, 3239, 9844                                     | Proteomic    |
| Lipases                  | 4090, 7684, 11,671, 12,240, 15,429, 19,895, 21,975                         | Transcriptic |
| Glucosidases             | 128, 392, 495                                                                | Proteomic    |
| Galactosidases           | 1330, 2201, 5291, 8029, 10,216, 10,716                                     | Proteomic    |
| Amylases                 | 7966, 4813                                                                   | Proteomic    |
| Trehalase                | 445, 5249                                                                    | Proteomic    |
| Chitinases               | 4724                                                                        | Proteomic    |
| Nucleases                | 2736                                                                        | Proteomic    |
|                         | 2219, 3815, 4156, 4479, 6082, 1587, 1953                                   | Proteomic    |
| **Protease inhibitors**  |                                                                               |              |
| Serine protease inhibitors| 18, 1105, 2025                                                               | Proteomic    |
| Kazal type serine protease inhibition-like venom proteins | 1890 | Proteomic |
|                         | 1535, 11,921                                                                 | Transcriptic |
| Cysteine-rich proteins   | 5346, 21,875                                                                 | Transcriptic |
| **Immune related proteins** |                                                                              |              |
| Calreticulin             | 1001                                                                        | Proteomic    |
| Calnexin                 | 5412                                                                        | Proteomic    |
| Nucleobindin             | 4917                                                                        | Proteomic    |
| C1q-like venom protein   | 9506                                                                        | Transcriptic |
| **Recognition/binding proteins** |                                                                              |              |
| Odorant binding proteins | 281, 590, 1303, 1372, 4531, 4663, 7398, 8693, 9106, 10,799, 14,598          | Proteomic    |
| Chemosensory proteins    | 5573                                                                        | Proteomic    |
| Low-density lipoprotein receptors | 289, 2622, 1256                             | Proteomic    |
| Apolipoporphines         | 784, 2538                                                                   | Proteomic    |
| Insulin-like growth factor-binding protein | 2057 | Proteomic |
| **Glutathione metabolism** |                                                                              |              |
| γ-Glutamyl cyclotransferase | 5910                          | Transcriptic |
| **Oxidase**              |                                                                               |              |
| Laccase                  | 17,081                                                                      | Transcriptic |
| **Dehydrogenases**       |                                                                               |              |
| Alcohol dehydrogenase    | 2176                                                                        | Proteomic    |
| **Isomerases**           |                                                                               |              |
| Peptidyl-prolyl cis/trans isomerase | 1650                        | Proteomic    |
| FK506-binding protein    | 3257                                                                        | Proteomic    |
| Protein disulfide-isomerase A3 | 181                          | Proteomic    |
| **Heat shock proteins**  |                                                                               |              |
| Endoplasmic              | 1747                                                                        | Proteomic    |
| Heat shock 70 kDa        | 940                                                                         | Proteomic    |
| **Other proteins**       |                                                                               |              |
| Continued                |                                                                               |              |
and *M. mediator* endoparasitoids, and *Eulophus pennicornis* ectoparasitoid. In the latter cases, metalloproteases were shown to be responsible for the alteration of host development, reducing and even blocking host larval growth and metamorphosis, and promoting parasitoid development. Some of *M. mediator* metalloproteases belong to the M12B subfamily, a member of which was also identified in the venom of endoparasitoid *T. nigriceps*. According to the MEROPS database (http://merops.sanger.ac.uk), the *T. sinensis* venom metalloprotease, identified by the proteomic approach, belongs to the M12B subfamily, whose components are able to regulate processes related to neoplastic progression in mammals, such as immune response evasion, matrix degradation, metastasis and inflammation processes.

**Dipeptidases** including Dipeptidyl peptidase IV, dipeptidase 1 and the Angiotensin-converting enzyme, are exopeptidases that catalyze hydrolysis of dipeptides. Dipeptidyl peptidase IV is a very common enzyme in snake venom and in the venom of some Hymenoptera, such as *A. mellifera*, *Vespa basalis* and *Polistes dominula*. Although its function is not completely known, it seems to be involved in the processing of precursors of venom protein components. This enzyme is a serine protease that cleaves dipeptides from the N-terminus of peptides with proline or alanine in the penultimate position, and its function could be related to the maturation of toxic peptides, as proposed for mastoparan B, the major toxin in the venom of *V. basalis*. Mastoparan B is indeed synthesized as pro-peptide and then activated through an enzymatic cleavage by a dipeptidyl peptidase IV able to generate a consecutive release of dipeptides. Dipeptidyl peptidase IV cleavage activity was also related to other functions, such as the regulation of inflammatory and immunological responses, signal transduction and apoptosis by degrading physiological substances. The angiotensin-converting enzyme is a peptidyl dipeptidase that removes dipeptides from the C-terminus of short oligopeptides. In mammals, angiotensin I is converted into angiotensin II and bradykinin is activated, thus regulating blood pressure and electrolyte homeostasis. This enzyme was identified in the venom of the endoparasitic wasp *P. hypochondriaca*, in *Thalassophryne matteri* and *Scorpaena plumieri*, venomous fishes typical of the north-eastern coast of Brazil. Although its specific function remains unknown, it is probably involved in processing peptide precursors.

**Aminopeptidases** remove one or more specific N-terminal residues from target proteins or peptides, and are common in venom of snakes, in the venom of the predatory ant *Fuchyzondyla striata*, in the venom of the velvet spider *Stegodyphus mimosarum* and in the venom of the genus *Thoracobombus* (Hymenoptera), even if its function has not been understood yet. Mammalian aminopeptidases, which cleave brain angiotensin II to angiotensin III, are implicated in the control of arterial blood pressure. This suggests that aminopeptidases might be involved, in synergy with angiotensin-converting enzymes, in angiogenic mechanisms, such as in regulating blood vessel formation and blood pressure. However, the most probable functions in venoms are

**Table 2. T. sinensis venom proteins, identified by the transcriptomic or proteomic approach, categorized according to their role. The table contains the contig number and the approach by which each protein was identified.**

| Protein category | Contig number | Approach |
|------------------|---------------|----------|
| Allergens 3 and 5| 3120, 8578    | Proteomic|
| Major royal jelly protein | 1427, 7940, 12,687, 12,783 | Transcriptomic |
| Adipocyte plasma membrane-associated protein | 11,185 | Proteomic |
| Ferritin | 874 | Proteomic |
| Lachesin | 1194, 1432 | Proteomic |
| Mesencephalic astrocyte-derived neurotrophic factor | 3847 | Proteomic |
| Vitellogenin | 575 | Proteomic |
| Agatoxin | 1034 | Proteomic |
| Venom protein D | 12,568 | Transcriptomic |
| Venom protein F | 3227 | Transcriptomic |
| Venom protein L | 1029 | Transcriptomic |
| Venom protein N | 2092 | Transcriptomic |
| Venom protein O | 10,580 | Transcriptomic |
| Venom protein Q | 19,579 | Transcriptomic |
| Venom protein R | 472 | Transcriptomic |
| Venom protein T | 2910 | Transcriptomic |
| Venom protein U | 2487 | Transcriptomic |
| Venom protein V | 157, 1301 | Transcriptomic |
| Uncharacterized proteins | 4647 | Transcriptomic |
| Hypothetical proteins | 529, 990, 1109 | Proteomic |
1) to help degrade the host tissues extracellular matrix in order to increase its permeability to venom protein components and 2) to contribute in transforming host tissues into nutrients for parasitoid progeny. In *T. sinensis* venom, three aminopeptidases were identified using the proteomic approach. Lysosomal proteases belong to the aspartic, cysteine, or serine endopeptidase family; despite the adjective “lysosomal,” they are usually detected within all vesicles of the endocytic pathway. They are also known as cathepsins. Lysosomal aspartic proteases are enzymes whose catalytic sites consist of two aspartate residues. A similar protein, a cathepsin D, was found in the venom of the endoparasitoids *Leptopilina heterotoma* and *Chouioia cunca*. A putative lysosomal aspartic protease was also upregulated in the pupal transcriptome of *Sarcophaga crassipalpis* 25 h after envenomation by *N. vitripennis*, supporting the hypothesis that this protease plays a key role in the success of parasitism. Lysosomal proteases, such as cathepsins, are activated in apoptotic and necrotic cells and during autophagy phenomena. The regulation of autophagy, which is associated with starvation, nutrient recycling and cell cycle arrest, can be a strategy used by parasitoids to manipulate host development and metabolism for its progeny. Although the specific role of lysosomal proteases in Hymenoptera parasitoids is still unclear, they may be involved in the production of venom protein components or in blocking host immunity, contributing to their offspring development. In *T. sinensis* venom, one lysosomal aspartic protease was identified using the proteomic approach. Cathepsin L is a lysosomal cysteine protease with a catalytic dyad consisting of cysteine and histidine. Lysosomal cysteine proteases are involved in extra- and intracellular protein degradation, antigen presentation and cellular development, and in various orders of insects are considered important digestive enzymes. In parasitoid venom, cathepsin L was first found in *T. nigriceps* venom. A large quantity of this protein was found in fat body/hemocytes complex of a *Spodoptera littoralis* larva parasitized by the ectoparasitoid *Bracon nigricans*; in this latter case it was hypothesized that cathepsin L induced the formation of these complexes, because of the rapid degradation of fat body required to mobilize the stored nutrients in the parasitoid offspring. In *T. sinensis* venom, one cathepsin L was identified using the proteomic approach. Moreover, a protein annotated as “uncharacterized,” identified by the proteomic approach, might also be included in the Lysosomal protease category as it contains a cysteine protease domain and a region similar to cathepsin L.

Carboxypeptidases cleave peptide bonds at the C-terminal of a protein. Putative venom serine carboxypeptidases were found in the venom of the ants *Odontomachus monticola* and *Tetramorium bicornutum*, the snake *Crotalus durissus terrificus*, the Hymenoptera of the Apidae family and in the venom gland of the ectoparasitoid *Anisopteromalus calandrae*. A serine carboxypeptidase was also found in the venom of the endoparasitoids *Psytalla concolor* and *T. nigriceps*. Although the specific role of carboxypeptidases in parasitoid venom is not clear yet, this enzyme could be involved in the degradation of the host tissues, most likely as the aminopeptidase enzymes. In *T. sinensis* venom, four carboxypeptidases were identified: three using the proteomic approach, and one using the transcriptomic one.

Venom acid phosphatases have been identified in the venom of some hymenopteran species, such as *A. melifera* and *Apis cerana*, the endoparasitoids *P. hypochondriaca*, *Pimpla turionella*, *P. puparum* and the ectoparasitoids *A. calandrae* and *Bracon hebetor*. Venom acid phosphatases were also identified in the venom gland of the endoparasitoid *M. pulchricornis*. They are characterized by a conserved catalytic core containing a histidine residue which is phosphorylated during catalysis. In venom, they have a neurotoxic, myotoxic, anticoagulant and inflammatory effect. In *T. sinensis* venom, two acid phosphatases were identified by the proteomic approach, and eighteen acid phosphatases were identified by the transcriptomic approach.

Carboxylesterases are hydrolases containing a catalytic triad of three residues, serine, glutamate or aspartate and histidine, and the mechanism involves a nucleophilic attack on a carbonyl carbon atom. Although their function in venom has not been identified yet, they could be allergen proteins; indeed, carboxylesterase-6 is one of the main allergens in honeybee venom. In *N. vitripennis* venom, two types of esterases have been found: an arylsulfatases B isoform X1 and a carboxylesterase clade B member 2 precursor. In *T. sinensis* venom, seven carboxylesterases were identified by the proteomic approach and seven by the transcriptomic one.

Lipases, which act on carboxylic esters, have the same catalytic triad as esterases. In general, lipases have essential roles in the digestion, transport, and processing of dietary lipids in most living organisms. It has been essential to identify N. vitripennis venom induces alterations in the host lipid metabolism, although the specific role of lipases has not been clarified yet. Recently, lipase activity has been found in the venom of the endoparasitoid wasps *P. hypochondriaca*, *Psytalla lounsburyi* and *P. concolor*, *Microctonus aethiopes*, *Ooencyrtus telenomicida* and the ectoparasitoid *B. nigricans*. An interesting example is reported in the endoparasitoid *Cotesia kariyai*: several days after parasitization, the total amount of lipid from the fat body of the parasitized hosts decreased as the lipase activity of parasitoid larvae increased. Although in this case lipase was not annotated as a component of the endoparasitoid venom, this observation could support the hypothesis that, in general, lipases can digest host lipids and provide nutrients to the parasitoid larvae. Lysis of fat body cells, with the increase in the hemolymphatic lipid content, was also observed in the lepidoptera *Pseudactia separata* after it was parasitized by ectoparasitoid wasps of the genus *Euctetes*. Lipases can then be considered involved, at least partially, in increasing the suitability of the host environment in favor of parasitoid progeny. In *T. sinensis* venom three lipases were identified by the proteomic approach.

Glucosidases hydrolyze glycosidic bonds from glycosides and oligosaccharides and remove non-reducing terminal glucosyl residues releasing glucose as product. Among parasitoids, β-glucosidases were detected in the venom of the endoparasitoids *P. hypochondriaca*, *Microplitis demolitor*, *P. lounsburyi* and *P. concolor*. The release of glucose, deriving from host hemolymph carbohydrates, may increase the amount of energetic nutrients available for the developing parasitoid larvae, suggesting that glucosidases are involved in modifying host metabolic pathways in favor of parasitoid development. In *T. sinensis* venom six glucosidases were identified by the proteomic approach.
Galactosidasases catalyze the hydrolysis of galactoside molecules by breaking glycosidic bonds\(^{211}\). Among parasitoids, galactosidasases were identified in the venom of the endoparasitoid \(P.\) hypochondriaci\(^{28}\). Like glucosidasases, galactosidasases could also release carbohydrates into the host hemolymph for supply of parasitoid larvae. In \(T.\) sinensis venom, two galactosidasases were identified by the proteomic approach.

Amylases are enzymes catalyzing the hydrolysis of alpha-1,4 glycosidic bonds of starch and glycogen into sugars, such as maltose, maltotriose and residual branched maltodextrins\(^{212}\). Most amylases have been identified in insect salivary glands or digestive tracts; few examples have been reported in the venom of parasitoid wasps, such as \(Nasonia\) species and \(P.\) puparum\(^{113,114}\). As carbohydrates are essential for metabolism, it is proposed that amylases expressed in the venom of \(P.\) puparum and secreted in the host (\(P.\) rapae) hemolymph are involved in the degradation of host polysaccharides for energy intake and parasitoid larva development\(^{114}\). In \(T.\) sinensis venom, two amylases were identified by the proteomic approach.

Trehalases, enzymes which catalyze the conversion of trehalose to glucose, have been found in the venom of wasps such as \(Cerceris\) rybyensis\(^{155}\), the endoparasitoids \(Ampulex\) compressa\(^{116}\), \(H.\) hypochondriaca\(^{121}\) and the ectoparasitoid \(N.\) vitripennis\(^{116}\). Trehalases can modify the high concentration of trehalose in the host hemolymph to glucose in order to provide a source of energy for the parasitoid larva development\(^{118,119}\). In \(T.\) sinensis venom, one trehalase was identified by the proteomic approach.

Collectively glucosidasases, galactosidasases, amylases and trehalases may play a key role in metabolic pathways, providing nutrients to the parasitoid offspring.

Chitinases are able to disrupt and digest chitin, one of the components of the exoskeleton of arthropods\(^{210}\). This enzyme is vital for animals during ecysis and metamorphosis, as well as for animals that feed on organisms whose structures are composed of chitin. Although this enzyme has been detected in the venom of spiders\(^{213}\) and scorpions\(^{210,214}\), it is also important in the physiology of different parasitoids, like \(T.\) nigriceps, where teratocysts release chitinases to support the emergence of larvae\(^{112}\). Chitinases may be performed by chitinase in the venom of the endoparasitoid wasp \(Chelonus\) inanus, helping the parasitoid larva to hydrolyze the host embryonic cuticle to reach host embryo\(^{124}\). In \(T.\) sinensis venom, one chitinotidase was identified by the proteomic approach. Moreover, a protein annotated as “uncharacterized protein”, found by the proteomic approach was included in this group as it contains chitin-binding domains in its structure.

Nucleases are capable of cleaving the phosphodiester bonds among nucleotides. Nucleases are very common in snake\(^{222}\) and Cnidaria venoms\(^{226}\). It was predicted that these enzymes play a central role in strategies of prey immobilization, as free adenosine molecules may induce the inhibition of neurotransmitter release. Moreover, these enzymes may induce renal failure and cardiac arrest and increase vascular permeability, thereby helping the spread of toxins in host tissue\(^{223}\). Endonuclease-like venom proteins are characterized by a DNA/RNA non-specific endonuclease conserved domain that gives the protein the ability to cut double-stranded and single-stranded nucleic acids\(^{227}\). They were found in the venom of the ectoparasitoid \(A.\) calandrae\(^{211}\) and the endoparasitoid \(Cotesia\) chilonis\(^{218}\). A ribonuclease Oy-like found in the venom of the ectoparasitoid \(Pachycerdequeus\) videmmiae may be related to the cleavage of host RNA to face host defensive reactions\(^{228}\). Inosine uridine-prefering nucleoside hydrolase-like is an enzyme capable of hydrolyzing nucleotides in nucleosides, with a preference for inosine and uridine\(^{218}\). It was found in the venom of the seed-parasitic wasp, \(Megascyclus\) spermocephal\(^{219}\) and in the venom of the endoparasitoids \(Leptopilina\) bouardi and \(L.\) heterotoma\(^{220}\). However, its role in venom is still unknown. Another kind of deoxyribonuclease with predicted function in DNA degradation found in \(T.\) sinensis venom is planctinotoxin. Planctinotoxin, first detected in the venom of the starfish commonly known as crown-of-thorns, \(Acanthaster\) planici, is a DNAase II able to reduce cellular antioxidant level in response to high oxidative stress and induce hepatotoxic damage\(^{229}\). It was found in the venom gland of the endoparasitoids \(M.\) demolitor\(^{130}\) and \(Lysiphlebus\) fabarum\(^{131}\), although its role in parasitoid venom remains unknown. In \(T.\) sinensis, several nucleases were found by the proteomic approach: one ribonuclease Oy-like, two endonuclease-like venom protein precursors, one inosine uridine-prefering nucleoside hydrolase-like precursor, one poly(U)-specific endonuclease homolog. Moreover, one endonuclease-like venom protein precursor and one plancitoxin were identified by the transcriptomic approach. Finally, one protein annotated as “uncharacterized protein” found by the proteomic approach can also be included in this group, as in its structure it contains a DDE_Tnp_4 domain, belonging to the endonuclease family.

Protease inhibitors. Serine protease inhibitors, known also as serpin proteins, are common in the venom of different species, such as snakes, scorpions and wasps. Serpin proteins were found in the venom of endoparasitoids \(L.\) bouardi\(^{155}\), \(A.\) ervi\(^{159}\), \(A.\) calandrae\(^{210}\), \(T.\) nigriceps\(^{198}\) and \(P.\) puparum\(^{149}\) and in the venom gland of the endoparasitoid \(M.\) pulchricornis\(^{29}\). Serpins can form permanent covalent complexes with target serine proteases\(^{134}\) and are involved in regulating the phenoloxidase cascade as well as blocking melanization process\(^{231,135}\). In \(T.\) sinensis venom, one serpin 5, a serine protease inhibitor \(28\)Dc isoform X2 and a serine protease inhibitor 3/4 isoform X16 were identified using the proteomic approach. Moreover, a further protein annotated as “uncharacterized protein” found by the transcriptomic approach. Finally, one protein annotated as “uncharacterized protein” found by the proteomic approach can also be included in this group, as in its structure it contains a DDE_Tnp_4 domain, belonging to the Serpin superfamily.

Kazal-type serine protease inhibition-like venom proteins are involved in the inhibition of serine proteases, such as trypsin, chymotrypsin, and elastases. The inhibitory domain contains a specific peptide bond, which may serve as a substrate for the cognate enzyme. The reactive site peptide bond is located within a loop whose conformation is identical in all Kazal inhibitors and all enzyme/inhibitor complexes. Similar domains, which are also present in follistatin and follistatin-like family members, play an important role in regulating specific tissues\(^{136}\). Four Kazal proteins may be involved in the immune response. Indeed, in \(B.\) mori, Kazal inhibitor proteins showed
antimicrobial activity playing a putative protective role against invading pathogenic microorganisms. Moreover, Kazal inhibitor proteins were also found in the saliva of blood-sucking insects, indicating a putative anticoagulant role. Few studies on these proteins have been carried out in parasitoid wasps, unlike in other insects. Studies on Kazal-type serine protease inhibition-like proteins of N. vitripennis venom showed an inhibition of prophenoloxidase 2 activation in parasitized Musca domestica hemolymph. The same activity was observed in the endoparasitoid wasps Venturia canescens. The prophenoloxidase cascade, one of the main components of the immune system in arthropods, activates specific humoral immune responses to non-self organisms through the melanization and damages their tissues. This process is mediated by the enzyme phenoloxidase, which is synthesized as the zymogen prophenoloxidase. The activation of prophenoloxidase in phenoloxidase is tightly regulated by the serine proteases cascade and by serpins. Accordingly, it was proposed that Kazal-type serine protease inhibition-like venom proteins may play a role in suppressing host melanization. One protein belonging to Kazal-type serine protease inhibition-like venom protein-1 and one protein belonging to Kazal-type serine protease inhibition-like venom protein-2 were identified in T. sinensis venom, using the transcriptomic approach.

Finally, an ovomucoid-like protein was identified by the proteomic approach. Ovomucoids, proteins found in the whites of eggs, are composed of three Kazal-type domains. This protein, found in the venom of the snake Bothricichis schlegeli, has a putative function of a serine protease inhibitor, according to Kazal-type serine protease inhibitor role.

Cysteine-rich proteins are protease inhibitors showing a distribution of cysteine residues similar to toxin proteins and serine protease inhibitors of insects and crustaceans. Specifically, the Kunitz (KU) type motif was found in toxins from amphibians, spiders, cones, snails and sea anemones. In addition to the classical function of serine proteases inhibition, these cysteine-rich/KU venom proteins can block ion channels, which are essential for regulating various physiological processes such as blood coagulation, fibrinolysis and host defense, favoring the spread of parasitization events. Moreover, cysteine-rich protease inhibitors could be involved in disrupting host immunity by inactivating the prophenoloxidase cascade, as the Kazal-type serine protease inhibition-like venom protein of the endoparasitoids P. hypochondriaca and P. turionellae and the ectoparasitoids A. calandrae and N. vitripennis, using the transcriptomic approach, one cysteine-rich/KU venom protein and one cysteine-rich/pacifastin venom protein-2 were identified.

**Immune-related proteins.** Several studies report that N. vitripennis venom inhibits host cellular immune response. Venom proteins involved in the suppression of host defense are collectively named “Immune-related proteins”.

Calreticulin is a Ca\(^{2+}\)-binding chaperone that was found in endoparasitoid C. rubecola venom, where it may compete with host calreticulin on the surface of hemocytes, acting as an antigen of hemocyte activation in early encapsulation reactions. Calreticulin, indeed, seems to play an important role in encapsulation and phagocytosis, inhibiting hemocytes diffusion and suppressing host immune reaction. Calreticulin has also been identified in the venom of other endoparasitoids such as P. paparum and T. nigriceps, suggesting a role similar to that of C. rubecola calreticulin in hemocyte encapsulation. Beside that, by altering the intracellular calcium balance, calreticulin might affect the biological processes in which Ca\(^{2+}\) is involved, such as apoptosis, inflammation and the activation of hydrolytic enzymes. In T. sinensis venom, a calreticulin was identified by the proteomic approach.

Calnexin is another Ca\(^{2+}\)-binding protein found in the venom of snake Bothrops colombiensis; together with calreticulin, calnexin may be involved in the process of toxin secretion. In T. sinensis venom, calnexin was identified by the proteomic approach.

Nucleobindin-2 is another protein that can be included in the Ca\(^{2+}\)-binding protein group. In mammals, nucleobindin-2 is the precursor of a DNA- and calcium-binding protein, nesfatin-1, that seems to be associated with brain changes in stress situations. Nucleobindin-2 was reported in the venom of A. mellifera, in the venom of Ornithorhyncus anatinus platypus, a venomous monotreme, and in the venom of the snake Bothrops jararaca, in which it could induce excitotoxicity, that is the alteration of nerve cells by excessive neurotransmitter stimulation. In T. sinensis venom, Nucleobindin-2 was identified by the proteomic approach.

Complement component 1q (C1q) is a key protein in the classical complement pathway and represents the joining link between the acquired and the innate immune response. As the human C1q domain directly interacts with lipopolysaccharides from gram-negative bacteria, it is possible that this protein plays a role in opsonizing molecules. This protein was also found in A. mellifera venom. In T. sinensis venom, one C1q-like venom protein was identified through the transcriptomic approach.

**Recognition/binding proteins.** Odorant binding proteins (OBPs) are a group of small globular and soluble polypeptides highly concentrated in olfactory organs as nasal mucus and tears in vertebrates and sensillar lymph in insect sensilla. In invertebrates, they are characterized by a pattern of six conserved cysteine residues, paired in three disulphide bridges. Beside the general OBP-like (GOBP-like) proteins in N. vitripennis venom, OBPs were found in the venom of the endoparasitoids P. paparum, L. heterotoma, C. inanitus, in the fire ant Dinoponera quadriceps and in the giant ant Dinoponera quaedvirens. In the giant ant, OBP-like proteins induced IgE antibody production in prey (insects, small birds, and mammals) and acted as a powerful allergy-inducing molecule. Their possible role in Hymenoptera venom and in host-parasitoid interactions is not known yet, but OBPs were supposed to be involved in host selection and in the search for an appropriate substrate to oviposit. In addition to this classical chemosensory function, OBP-like proteins in venom may play a role in transporting hydrophobic molecules, such as components of the nourishment, ensuring nutrients to the parasitoid rather than the host. In T. sinensis venom, eleven general OBP-like proteins were identified by the proteomic approach.
**Chemosensory proteins**, similarly to OBPs, are small soluble proteins mediating sensory perception in insects. Their function in parasitoid venom could be very similar to OBPs mediating host selection and carrying hydrophobic feed molecules for the parasitoid offspring. Chemosensory proteins were found in the venom of the endoparasitoid *C. inanitus* and the ectoparasitoid *A. calandrae*. In *T. sinensis* venom, one chemosensory protein was identified using the proteomic approach.

**Low-density lipoprotein receptors** have a central role in cholesterol and other lipoprotein metabolism, even though they are uncommon among venom proteins and their role in envenomation is still unknown. Low-density lipoprotein receptors were described in the venom of parasitoids, such as in *N. vitripennis* and *P. puparum* and the spider *Latrodectus hesperus*. An example of this protein is PH-4 (neuropeptide prohormone-4) of *Profundicoris* cone snail genus. This protein, found also in the venom of other sea snails, has a mature sequence and a precursor-related peptide containing low-density lipoprotein receptor A domain. Another example is PS1 (peptidase S1) from the crustacean *Xibalbanus tulumensis*, which also contains a low-density lipoprotein receptor A domain. This domain allows PS1 to bind lipoproteins, while the PS1 domain facilitates their digestion. These proteins are hypothesized to facilitate the interaction with lipoproteins of the prey/host to create a substrate for predators and parasitoids. In *T. sinensis* venom, two low-density lipoprotein receptors were identified using the proteomic approach and one using the transcriptomic one.

**Apolipophorins**, which belong to apolipoprotein family and are involved in lipid transport processes in the insect hemolymph, play also an important role in immunity, programmed cell death and the detoxification approach. Moreover, this group of apolipophorins may include one protein annotated as "hypothetical protein", which belong to apolipoprotein family and are involved in lipid transport processes in the insect hemolymph. They were found in venom of *A. mellifera* and *B. pascuorum* wasps and in predatory ant *O. monticola*. In *T. sinensis* venom, two apolipophorins were identified using the proteomic approach. Moreover, this group of apolipophorins may include one protein annotated as "hypothetical protein", as it contains an apolipophorin domain in its structure.

**Glutathione metabolism.** γ-Glutamyl cyclotransferase proteins are involved in glutathione metabolism. This enzyme was found in transcripts of the venom gland in the ectoparasitoid *A. calandrae* and in the venom of *M. spermotrophus*. The function of this protein in the parasitoid venoms is still unknown. However, glutathione is fundamental in regulating homeostasis in the cell, and its alteration could result in oxidative stress and apoptosis, as reported for the γ-Glutamyl transpeptidase like venom protein identified in *A. ervi* venom. In *T. sinensis* venom, one γ-glutamyl cyclotransferase-like venom protein was identified using the transcriptomic approach.

**Oxidases.** Laccases, which belong to a group of proteins collectively known as multicopper oxidases, were supposed to play an important role in insect cuticle sclerotization. During this extracellular process, cuticular proteins are cross-linked into a matrix as result of oxidative and nucleophilic reactions of catechols to their corresponding quinones. This enzyme was found also in the venom of both the ectoparasitoid *N. vitripennis* and the endoparasitoid *P. hypochondriaca*. In *T. sinensis* venom, one laccase was identified by the transcriptomic approach.

**Dehydrogenases.** Alcohol dehydrogenases are oxidoreductase enzymes that catabolize otherwise toxic alcohols. They are not common in venom; to the best of our knowledge, this enzyme was previously identified exclusively in the venom of *A. mellifera* and in the venom of scorpion species *Leiurus abdullahbayrami*. In *T. sinensis* venom, one alcohol dehydrogenase was identified by the proteomic approach.

**Isomerases.** Peptidyl-prolyl cis/trans isomerases are enzymes that catalyze the cis-to-trans isomerization around proline, allowing proteins to fold into their correct conformation. In nature, proline is the only amino acid existing in cis and trans isomerization form; correct protein folding is often not possible when a proline peptide bond is in the incorrect configuration, and proper isomerization is necessary. Peptidyl-prolyl cis/trans isomerases were identified in the venom gland of predatory marine cone snails *Conus novaezollandiae*, where they facilitate the in vitro folding of conotoxins. This enzyme was also identified in the venom of the endoparasitoids *C. chilonis* and *T. nigriceps*, in the venom of the ectoparasitoid *B. nigricans* and in the venom of the predatory ant *O. monticola*. Although the specific function of these enzymes in parasitoid venom remains to be determined, they may be involved in the folding of toxin peptides, as it occurs in other venomous organisms. In *T. sinensis* venom, one isomerase, annotated as peptidyl-prolyl cis–trans isomerase 5-like, was identified by the proteomic approach.

**FK506-binding proteins** are immunophilins that bind immunosuppressive drugs such as FK506, rapamycin and cyclosporin A. They often have peptidyl-prolyl cis–trans isomerase activity. This enzyme was found in venom of the endoparasitoid *C. chilonis*. In *T. sinensis* venom, one isomerase, annotated as FK506-binding protein, was identified by the proteomic approach.

**Protein disulfide-isomerase A3** is a molecular chaperone, supporting folding and processing of glycoprotein after their synthesis in the endoplasmic reticulum. It was identified in the venom of the endoparasitoid *P. puparum*, although its role in venom is still unknown. In *T. sinensis* venom, one isomerase, annotated as protein disulfide-isomerase A3, was identified by the proteomic approach.
Heat shock proteins (HSPs). *Endoplasmins* are molecular chaperones belonging to the HSP 90 family and involved in the final processing and export of secreted proteins. They may also play a role in the stabilization of other proteins. They were found in the venom of the endoparasitoid *A. ervi*, and, according to HSP functions, they may help protect parasitoid proteins during secretion and transport in host cells. The same protein was identified in venom of the endoparasitoids *P. lounsburyi*, *P. concolor* and *C. cunea* and snake *Crotalus adamanteus*. In *T. sinensis* venom, a heat shock protein, annotated as endoplasmalin, was identified by the proteomic approach.

Heat shock 70 kDa, molecular chaperones involved in the folding of other proteins, were found in venom of the Hymenoptera wasps *Apis mellifera* and *Polybia paulista* and in venom of the ant *Neoponera villosa*. These proteins also occur in the venom glands of *A. cerana* and in venom of the endoparasitoid *T. nigriceps*. In *T. sinensis* venom a heat shock protein, annotated as generic heat shock 70 kDa, was identified by the proteomic approach.

Other proteins. In addition to the previously described categories, other proteins were identified, difficult to categorize but similar to known proteins.

**Allergens 3 and 5** are allergen proteins. Mostly studied and characterized in humans, they are also found in several wasp and ant species. Venom allergen 3-like and allergen 5 were found in the venom of the ectoparasitoid *A. calandrae*, allergen 3 in the venom of the endoparasitoid *P. puparum*, allergen 5 in the venom of the endoparasitoids *T. nigriceps* and *C. inanitus* and the ectoparasitoid *B. nigriceps*, venom allergen 5-protein contains a sperm-coating protein (SCP)-like extracellular protein domain, that may function as endopeptidase. This protein might be involved in protein proteolysis and tissue degradation by the parasitoid. In *T. sinensis* venom, six allergens proteins annotated as allergens 3 and 5 have been identified, two of which by the proteomic approach and four by the transcriptomic one. These proteins are very similar to *N. vitripennis* antigen 5-like proteins, which is also one of three major allergenic proteins found in the venom of *Vespula* and *Dolichovespula*.

**Major royal jelly proteins** are proteins involved in the development of bee larvae. MRJP7 is highly expressed in nurse bees and bees that feed the worker and the queen with jelly secreted from specific glands. Moreover, proteins very similar to MRJP8 and 9 were also identified as components of honeybee venom and of the venom of parasitoids, such as the endoparasitoid *P. puparum* and the ectoparasitoid *P. vindemmiae*, although their function in parasitoid venom is still unknown. Recently it was hypothesized that they could be allergens or proteins related to storage of nutrients. In *T. sinensis* venom, one protein annotated as major royal jelly protein was identified by the proteomic approach. Moreover, a protein annotated as “Uncharacterized protein” found by the proteomic approach could also be included in this group, as it contains a major royal jelly protein domain in its structure. The same protein was found in the ectoparasitoid *P. vindemmiae* venom.

**Adipocyte plasma membrane-associated protein** was previously identified in the endoparasitoid *Tetrastichus brentispae*. In this study, the protein was found on the surface of the parasitoid egg and it was supposed to be involved in evading the host immunity response and protecting the egg during the early parasitoid stages, in association with a lipophorin protein. The adipocyte plasma membrane-associated protein was very similar to hemomucin, an O-glycosylated surface mucin found on the extraembryonic membrane of many parasitoid eggs that may allow the embryo to evade the host encapsulation reaction. Therefore, this protein could have protective properties in host-parasitoid systems. In *T. sinensis* venom, one protein annotated as adipocyte plasma membrane-associated protein was identified by the proteomic approach.

**Ferritin** is an intracellular protein that carries and stores iron. This protein is found in the venom of scorpion *Centruroides vittatus* and the endoparasitoids *M. aethiopoides* and *Microtonus hyperodae* and in venom gland of the endoparasitoid *M. pulchricornis*, but its role is still unknown. In *T. sinensis* venom two proteins annotated as ferritin were identified by the proteomic approach.

**Lachesins** belong to the disintegrin family, typically of viper venoms, and act as potent inhibitors of platelet aggregation and integrin-dependent cell adhesion. *Lachesin* was characterized in snake *Lachesis muta* venom. Although this protein is not found in parasitoid venom, it could be involved in defense against host immunity responses. In *T. sinensis* venom one protein annotated as lachesin was identified by the proteomic approach.

**Mesencephalic astrocyte-derived neurotrophic factor** is an endoplasmic reticulum stress-inducible protein, originally identified as protein protecting rat dopaminergic neurons in vitro and prevents neuron degeneration in Parkinson’s disease. It was found in venom of endoparasitoid *C. chilonis*, in which probably it is involved in cell protection from endoplasmic reticulum stress. In *T. sinensis* venom one protein, annotated as mesencephalic astrocyte-derived neurotrophic factor, was identified by the proteomic approach.

**Vitellogenin** is a protein involved in lipid transport from ovarian follicle cells to oocytes, providing nutrition during embryogenesis and playing a role as egg yolk protein precursor in the ovaries. Despite its main role, it is also component of venom of Hymenoptera, such as *A. mellifera* and *Vespula vulgaris* in which it represents one of the allergens, with its IgE-reactive allergenic properties, and *A. cerana* in which it was hypothesized that could be involved in response to microbial infection and oxidative stress, ensuring protection to DNA against ROS. Indeed, it could be also considered an antimicrobial and antioxidant agent. It was also found in venom of the ant *O. monticola* and in venom of the ectoparasitoid *A. calandrae*. In *T. sinensis* venom a protein annotated as vitellogenin was identified by the proteomic approach.

**Agatoxins** are polyamine and peptide toxins isolated from spider and scorpion venoms. Their mechanism of action led to the inactivation of several ion channels, causing neurotoxic effects. *Agatoxin* was found also in the venom of the giant ant *D. quadriceps* and in the venom gland transcriptome of solitary and social wasps, such as *Vespula velutina*. In *T. sinensis* one u5-agatoxin-a01-like isoform 1 was identified exclusively by...
the transcriptomic approach. To the best of our knowledge, the *T. sinensis* venom is the only one containing a putative agatoxin among parasitoid venoms. Because agatoxins are strongly involved in blocking ion channels, their action in parasitic syndrome could be strictly related to prey paralysis, that is one of the main effects of ectoparasitoid wasps attack12.

**Unknown and hypothetical proteins.** We found different proteins that are not associated with known proteins, but they are very similar to the same unknown *N. vitripennis* proteins; they are named from “venom protein A” to “venom protein Z”.

In *T. sinensis* we found:

- venom protein D (one protein) identified by the transcriptomic approach, also found in the venom of *P. puparum*27 and *A. calandrae*14;
- venom protein F (one protein) identified by the transcriptomic approach, also found in the venom of the wasps *M. spermotrophus*129, *Megaphragma amalphitanum*, *Ceratosolen solmsi*, the endoparasitoid Trichogramma pretiosum280, and *A. calandrae*, in this case with a putative role in actin polymerization and in transcription regulation of cholesterol and fatty acid homeostasis31;
- venom protein L (one protein) identified by the transcriptomic approach, also found in the venom of *P. puparum*27, *A. calandrae*31 and *P. vindemmiae*20;
- venom protein R (one protein) identified by the transcriptomic approach, also found in the venom of *P. puparum*27, *A. calandrae*31 and *P. vindemmiae*20;
- venom protein T (one protein) identified by the transcriptomic approach, also found in the venom of *A. calandrae*14 and *N. giraulti*114.

In addition to the above-mentioned proteins, a group of “uncharacterized proteins” and a group of “hypothetical proteins” were also found by the proteomic approach. A hypothetical protein is a protein whose existence has been predicted since it derives from an ORF, but there is no experimental evidence of translation208. The identification of these proteins in venom by the proteomic approach is an experimental evidence of their existence, but their characterization and their role remain unknown. In *T. sinensis* venom, we identified two venom protein protein A*27* and *A. calandrae*14; one protein identified by the proteomic approach, also found in the venom of *N. vitripennis*, also found in the venom of *T. sinensis*27, *N. giraulti*114, *M. amalphitanum*, *C. solmsi*, *T. pretiosum*280, *Tetrastichus brontispae*284,285 and *M. spermotrophus*280. RNA sequencing performed on abdomen tissue in *Ishimura elegans* also revealed a venom protein t-like, a toxin and hemolymph juvenile hormone binding protein, that regulates embryogenesis, larva development and reproductive maturation280.

- venom protein U (two proteins) identified by the proteomic approach, also found in the venom of *P. puparum*27, *A. calandrae*31 and *P. vindemmiae*20;
- venom protein V (one protein) identified by the transcriptomic approach, also found in the venom of *P. vindemmiae*20 and *A. calandrae*, in which a chaperone_ClpB domain was detected31. Chaperone ClpBs from several microorganisms are essential for survival under severe stress conditions207.

In addition to the above-mentioned proteins, a group of “uncharacterized proteins” and a group of “hypothetical proteins” were also found by the proteomic approach. A hypothetical protein is a protein whose existence has been predicted since it derives from an ORF, but there is no experimental evidence of translation208.

The identification of these proteins in venom by the proteomic approach is an experimental evidence of their expression, but their characterization and their role remain unknown. In *T. sinensis* venom, we identified two proteins similar to hypothetical protein LOC100679659 isoform 1 and uncharacterized protein LOC100118367 of *N. vitripennis*, also found in the venom of *P. puparum*27. The analysis of the structure of uncharacterized protein LOC100113619 led to the identification of a SWVC domain (single-domain von Willebrand factor type C). Proteins characterized by this domain were also identified in the venom of the spider *Pamphobeteus verdolaga*209 and in the venom gland transcriptome of the scorpions *Hadrurus spadix*209, *Centruroides hentzi*211 and *Paravaejovis schwenkmeyeri*212. In the structures of uncharacterized protein LOC106783674 isoform X2 and uncharacterized protein LOC108911535 isoform X1, a DUF protein families were still functionally uncharacterized. Overall, in *T. sinensis* venom, sixteen uncharacterized proteins and three hypothetical proteins were identified by the proteomic approach. According to the functional domains identified through the BLASTp software (https://blast.ncbi.nlm.nih.gov/Blast.cgi), some of the uncharacterized/hypothetical proteins can be included in the previous groups, as already described.

Further investigation will provide more information about these proteins.

**Conclusions**

The integrated transcriptomic and proteomic approach used to analyze *Torymus sinensis* venom and the accompanying analysis, using a transcriptomic approach, provides an overview of venom’s major protein components, in order to understand the mechanism underlying the complex host/parasitoid interaction. The study of endoparasitoid venoms, using these approaches, is the starting point for detailed knowledge of the molecular biology, evolution and effects of venom proteins in the host/parasitoid interactions. Indeed, although general physiological effects of Hymenoptera parasitoid venoms have been recorded, their exact composition is not completely known, also considering the high number of species and the differences among them. Specifically, in *T. sinensis* venom, a large number of proteins was identified, involved both in inhibition of the host immune system and in providing nutrients to the parasitoid progeny. Although additional in vivo studies are needed,
this first view of venom protein components confirms how venom, combined with other maternal factors, helps ensure the success of parasitism.

Materials and methods

Insect rearing.  
_Torymus sinensis_ adults were reared from galls of _Dryocosmus kuriphilus_ collected in 26 sites within Cuneo province (NW Italy), where 40% of the forestry area is covered by sweet chestnut, _Castanea sativa_ Mill. (Eudicots: Fagaceae) and nearly 10% by _Quercus_ spp.123. This area has been widely infested by _D. kuriphilus_ and includes the initial sites of _T. sinensis_ release in Europe, performed in 2005. Chestnut trees were sampled from both mixed forests and chestnut orchards. Galls were randomly collected by hand from low branches and with the aid of lopping shears from the medium–high canopy, according to a previously described method36,214. They were separated from any non-gall plant material to avoid contamination by other insects not associated with the galls and then stored in containers outdoors in cardboard rearing boxes. Up to 2000 galls were kept in a container. Rearing boxes were checked once per week until the emergence of the first parasitoid wasp, then parasitoids were collected daily and their date of emergence recorded. All _T. sinensis_ wasps were removed using an entomological pooter, then stored in 99% ethanol. All specimens were divided by sex by observing their morphological characters and then sent to the University of Basilicata’s laboratory for the subsequent analysis.

Venom gland collection.  
Before venom glands were collected, _T. sinensis_ adult females were briefly (a few minutes) anesthetized on ice and subsequently placed in a phosphate-buffered saline solution (1X PBS) (Sigma, St. Louis, MO, USA) in a Petri dish (Sigma, St. Louis, MO, USA). A stereo microscope (Nikon, Tokyo, Japan) was used for dissections. The entire reproductive apparatus was removed with a pair of forceps and placed in a drop of 1X PBS solution. Subsequently, the venom glands were isolated and placed in a centrifuge tube (Eppendorf, Hamburg, Germany) containing TRI Reagent (Sigma, St. Louis, MO, USA) for the following RNA extraction and on slides for the subsequent microscopic observations (Sigma, St. Louis, MO, USA).

Light microscopy.  
Venom glands were carefully transferred onto slides and fixed at room temperature for 10 min with 4% formaldehyde pH 6.9 (Carlo Erba, Milano, Italy). Then samples were observed using a Nikon Eclipse 80i (Instruments Europe, Amsterdam, The Netherlands) at 10X magnification. Images were recorded by a Nikon Digital Sight Ds-Fil camera (Nikon Instruments Europe).

RNA extraction.  
Total RNA was extracted from 360 venom glands, collected as previously described, using TRI Reagent following the manufacturer’s instructions (Sigma, St. Louis, MO, USA). A Turbo DNase (Ambion Austin, TX, USA) treatment was carried out to eliminate any contaminating DNA. The DNase enzyme was then removed and the RNA was further purified using the RNeasy MinElut Clean up Kit (Qiagen, Venlo, The Netherlands) following the manufacturer’s protocol and eluted in 20 µl of RNAse free water (Ambion Austin, TX, USA). RNA integrity was verified on an Agilent 2100 Bioanalyzer using the RNA Nano chips (Agilent Technologies, Palo Alto, CA), and RNA quantity was determined by a Nanodrop ND1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

RNAseq data generation and de novo transcriptome assembly.  
Tissue-specific transcriptome sequencing of the RNA sample was performed with poly(A) + enriched mRNA (New England Biolabs, Ipswich, MA, USA) fragmented to an average of 240 nucleotides. Sequencing was carried out by the Max Planck Genome Center (http://mpgc.mpipz.mpg.de/home/) using standard TruSeq procedures on an Illumina HiSeq2500 sequencer (Illumina, San Diego, CA, USA), generating ~ 42 Mio paired-end (2 × 100 bp) reads for the venom tissue sample. Quality control measures, including the filtering of high-quality reads based on the score given in fastq files, removal of reads containing primer/adaptor sequences and the trimming of read length, were carried out using CLC Genomics Workbench v9 (http://www.clcbio.com). The de novo transcriptome assembly was carried out using CLC Genomics Workbench v9 with standard settings and two additional CLC-based assemblies with different parameters and then selecting the presumed optimal consensus transcriptome, as previously described215. The transcriptome was annotated using BLAST, Gene Ontology and InterProScan with Blast2GO Pro version 4.1216. For BLASTx searches against the non-redundant NCBI protein database (nr database), up to 20 best NR hits per transcript were retained, with an E-value cutoff of ≤ 1E-3 and a minimum match length of 15 amino acids. To optimize annotation of the obtained data, we used GO slim, a subset of GO terms that provides a high level of annotations and allows a global view of the result. The functions “Gene Ontology Graphs” and “Enrichment Analysis” (Fisher’s exact test) was used as part of OmicsBox (1.4.11) to identify the distribution of gene ontology (GO) terms as well as overrepresentation of GO terms among the _T. sinensis _venom protein database relative to the complete reference dataset (_T. sinensis_ transcriptome assembly). The GO-enriched bar charts were reduced to display only the most specific GO terms by removing parent terms representing existing child terms using the function “Reduce to most specific terms” implemented in OmicsBox. A GO term was considered significantly enriched if the p-value corrected for false discovery rate control (FDR) was less than 0.05. To assess transcriptome completeness, we performed a BUSCO (Benchmarking Universal Single-Copy Orthologs; http://busco.ezlab.org) analysis by comparing our assembled transcriptome against a set of highly conserved single-copy orthologs. This was accomplished using the BUSCO v3 pipeline217, comparing the predicted proteins of the _T. sinensis_ transcriptome to the predefined set of 1658 Insecta single-copy orthologs from the OrthoDB v9.1 database. This resulted in 91.9% complete and 5% missing BUSCO genes for the venom gland transcriptome assembly. The assembled and annotated venom gland transcriptome was used to generate a custom-made protein database. The six reading frames of the 22,874 nucleotide sequences were translated in their corresponding amino acid sequences by SEQtools software (http://www.seqtools.dk/).
Digital gene expression analysis. Digital gene expression analysis was carried out by using CLC Genomics workbench v9 (http://www.clcbio.com) to generate BAM (mapping) files and QSeq Software (DNAStar Inc., Madison, WI, USA) to remap the Illumina reads onto the reference transcriptome, and finally by counting the sequences to estimate expression levels, using previously described parameters for read-mapping and normalization. In particular, the expression abundance of each contig was calculated based on the reads per kilobase per million mapped reads (RPKM) method, using the formula: \( \text{RPKM} (A) = \frac{10,000 \times C}{N - L} \), where \( \text{RPKM} (A) \) is the abundance of gene \( A \), \( C \) is the number of reads that uniquely aligned to gene \( A \), \( N \) is the total number of reads that uniquely aligned to all genes, and \( L \) is the number of bases in gene \( A \).

Quantitative real time PCR (qPCR). The relative expression in *T. sinensis* venom gland and female body deprived of venom gland of 10 genes selected among those reported in Table 1 was evaluated by a quantitative real time PCR (qPCR) (ABI PRISM R 7500 Fast Real-Time PCR System Thermal Cycler—Applied Biosystems, Foster City, CA, USA). Glycerinaldehyde-3-phosphate dehydrogenase (GAPDH) and Beta-tubulin have been chosen as reference genes for normalization of qPCR data. Primers were designed using Primer3web (version 4.1.0) (Supplementary Table S1). PCR amplifications, with 30 ng of cDNA, were carried out using GoTaq qPCR Master Mix (Promega, Madison, WI, USA), following the manufacturer's instructions. Cycling programme was: 2 min at 95 °C, 40 cycles of 15 s at 95 °C and 1 min at 58 °C. Three technical replicates and three biological replicates were performed for each reaction. To evaluate qPCR expression levels, relative quantification was performed using equations described by Liu et al., based on PCR amplification efficiencies of reference and target genes. Amplification efficiency of each gene was calculated according to the equation \( E = 10^{-\frac{1}{S} - 1} \) (\( S \) is the slope of the curve derived from three serial tenfold cDNA dilutions). The efficiencies of the amplicons were approximately equal. Quantification analysis of amplification was performed using the comparative \( \Delta\Delta CT \) method. Data were expressed as mean ± SEM (standard error of mean) of independent biological replicates and were compared by the Unpaired t-test using GraphPad Prism 6.00 software for Windows (GraphPad Software, La Jolla, CA, USA).

Phylogenetic analyses of *T. sinensis* trypsins. We inferred the species-specific diversification patterns of putative trypsins identified in the *T. sinensis* venom gland transcriptome in phylogenetic analyses. We used all predicted serine protease sequences from *T. sinensis* as query to search for homologs in the NCBI nr protein database using Blastp (E-value threshold of 10−5), identified the top 50 best Blast hits and removed redundant entries. Next, we removed partial sequences with less than 50% of the typical protease length. The corresponding protein sequences were aligned using MAFFT implemented in Geneious (v11.0.4) with FFT-NS-i×1000 algorithm and BLOSUM62 scoring matrix. The alignments were trimmed manually. Maximum-likelihood phylogenetic trees were constructed in FastTree implemented in Geneious (v11.0.4) with 1000 ultrafast bootstrap replicates for the full dataset. The tree was visualized and processed in Figtree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree). Numbers next to the tree branches indicate the support values.

Collection of venom and SDS-PAGE electrophoresis. Wasps previously anesthetized on ice were submersed in 1 × PBS solution (Sigma, St. Louis, MO, USA) and their venom apparatus (venom glands and reservoir) was isolated. Each reservoir was gently opened with a dissecting needle in a drop of water (ratio 1 μl of water: 1 reservoir). The resulting crude extract was centrifuged at 5000 g for 5 min at 4 °C, and the supernatant was used for electrophoretic analysis. For the proteome analysis, the venom from 30 *T. sinensis* females was collected for a total of 150 μg of protein. Protein quantity was measured using the Bradford method, with bovine serum albumin as standard. An aliquot of venom proteins was loaded on a 12.5% polyacrylamide running gel on a Bio Rad Electrophoresis Cell Mini Protean II (Biorad, Hercules, CA, USA). After the run, the gel was stained with colloidal Coomassie Blue G-250 (Sigma, St. Louis, MO, USA) for 1 h and the excess dye was removed by washing in deionized water for 12 h.

In situ protein digestion. After electrophoresis and staining, whole lanes were cut in 35 bands. The bands were in situ hydrolysed by trypsin as reported in Medugno et al. Briefly, gel bands were destained by washes in acetonitrile (ACN) (Honeywell, Charlotte, NC, USA), and 50 mM ammonium bicarbonate (NH4HCO3). Cysteine residues involved in disulphide bridges were reduced by 10 mM of dithiothreitol (Sigma, St. Louis, MO, USA) in NH4HCO3 for 45 min at 56 °C and then alkylated in 55 mM iodoacetamide (Sigma, St. Louis, MO, USA) for 30 min. The excess reagents were finally removed by washing with ACN and 50 mM NH4HCO3 alternatively. The dehydrated gel bands were then treated overnight with trypsin and peptide mixtures extracted in 0.2% HCOOH and ACN. The obtained mixtures were vacuum dried by a Savant SpeedVac System (Thermo Fisher Scientific, Waltham, MA, USA).

LC–MS/MS and protein identification. Each peptide mixture was resuspended in 10 μl of 0.2% HCOOH (Sigma, St. Louis, Missouri, USA) and analyzed by nanoLC–MS/MS on a LTQ Orbitrap mass spectrometer equipped with a nano HPLC system (Thermo Fisher Scientific, Waltham, MA, USA). After loading, the peptide mixture was first concentrated and desalted in the precolumn (C18 Easy Column L = 2 cm, ID = 100 μm, Thermo Fisher Scientific Waltham, MA, USA). Each peptide sample was then fractionated on a C18 reverse-phase capillary column (C18 Easy Column L = 20 cm, ID = 7.5 μm, 3 μm, (Thermo Fisher Scientific Waltham, MA, USA) working at a flow rate of 250 nl/min. The gradient used for peptide elution ranged from 5% to 95% of buffer B. Buffers A and B have the following composition:—buffer A, 2% ACN LC–MS grade and 0.2% HCOOH;—buffer B, 95% ACN LC–MS grade and 0.2% HCOOH. The MS/MS method was set up in a data-dependent acquisition mode, with a full scan ranging from 400 to 1800 m/z range, followed by fragmentation in CID modality of the
top 10 ions (MS/MS scan) selected on the basis of intensity and charge state (+2, +3, +4 charges). An exclusion time of 40 s was applied to avoid the repetitive fragmentation of the same signals over a 40 s interval and to increase the number of fragmented peptides and, therefore, the number of available protein identification sequences. The raw files obtained from this analysis were used as inputs in the Andromeda search engine. The peak list generated was uploaded in Andromeda software and a research was performed using the “T. sinensis protein database.” These parameters were fixed: “trypsin” as an enzyme allowing up to 2 missed cleavages, carbamidomethyl as a fixed modification, oxidation of M, pyroGlu N-term Q, as variable modifications, 0.5 Da MS/MS tolerance, 10 ppm peptide tolerance. Scores used to evaluate the quality of matches for MS/MS data were higher than 10 for unmodified peptides, otherwise 40.

Transcriptomic and proteomic data analysis. Putative venom proteins in the venom gland transcriptome were identified using a key-word approach: all proteins annotated with the word “venom” or “toxin” were selected. Proteins identified with the proteomic and the transcriptomic approach were analyzed using the Signal P 5.0 software (http://www.cbs.dtu.dk/services/SignalP/) in order to pick out those with signal peptide. Then, a second filter was applied, consisting of the alignment of the amino acid sequences of these proteins with venom proteins of Nasonia vitripennis identified by de Graaf et al. 201080. The N. vitripennis database was used as reference because the top BLAST Hit Species Distribution showed a high level of matching with T. sinensis. The alignments were made using the software BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Most of these proteins have been updated in recent years, so the final amino acid sequences were identified in the NCBI protein database by using accession numbers or amino acid sequences.

Received: 10 September 2020; Accepted: 22 December 2020
Published online: 03 March 2021

References
1. Scudder, G.G.E. The Importance of Insects in Insect Biodiversity: Science and Society. (eds. Footitt, R.G. and Adler, P.H.) (Wiley Blackwell, 2017).
2. Stork, N. E. How many species of insects and other terrestrial arthropods are there on earth?. Annu. Rev. Entomol. 63, 31–45 (2018).
3. Vogel, E., Santos, D., Mingels, L., Verdonckt, T. W., & Broeck, J. V. RNA Interference in insects: Protecting beneficials and controlling pests. Front. Physiol. 9, 1912 (2019).
4. Wang, Z. Z., Liu, Y. Q., Shi, M., Huang, J. L., & Chen, X. X. Parasitoid wasps as effective biological control agents. J. Integr. Agr. 18, 705–715 (2019).
5. 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. 49, 945–954 (2003).
6. Bordon, K. C., Wiezel, G. A., Amorim, F. G. & Arantes, E. C. Arthropod venom Hyaluronidases: biochemical properties and potential applications in medicine and biotechnology. J. Venom. Anim. Toxins. Incl. Trop. Dis. 21, 43 (2015).
7. Desneux, N., Pimenta, D. M. & Basset, J. M. Strategies for enhanced crop resistance to insect pests. Annu. Rev. Plant. Biol. 69, 637–660 (2018).
8. Hoy, M. A. Natural enemies important in biological control in Encyclopedia of Entomology (ed. Capinera, J.L.) 2555–2564 (Springer, 2008).
9. Kalyanasundaram, M. & Merlin Kamala. Parasitoids in Ecofriendly Pest Management for Food Security. (ed. Omkar) 109–138 (Academic Press, Elsevier Inc 2016).
10. Pennacchio, F. & Strand, M. R. Evolution of developmental strategies in parasitic hymenoptera. Ann. Rev. Entomol. 51, 233–258 (2006).
11. Moreau, S. J. & Asgari, S. Venom proteins from parasitoid wasps and their biological functions. Toxins. 7, 2385–2412 (2015).
12. Nakamatsu, Y. & Tanaka, T. Venom of ectoparasitoid, Euplectrus sp. near plathypenae (Hymenoptera: Eulophidae) regulates the physiological state of Pseudaulidia separata (Lepidoptera: Noctuidae) host as a food resource. J. Insect Physiol. 49, 149–159 (2003).
13. Falabella, P. et al. A γ-glutamyl transpeptidase of Aphidius ervi (Hymenoptera: Braconidae) induces apoptosis in the ovaries of host aphids. Insect Biochem. Mol. Biol. 37, 453–465 (2007).
14. 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. 71, 45–53 (2009).
15. Desneux, N., Barta, R. I., Delebecque, C. I., & Heimpel, G. E. Transient host paralysis as a means of reducing self-superparasitism in koinobiont endoparasitoids. J. Insect Physiol. 55, 321–327 (2009).
16. Zhu, J. Y., Ye, G. Y., Dong, S. Z., & Hu, C. Venom of the endoparasitoid wasp pteromalus puparum: An overview. Psyche: J. Entomol. 2011, 520926 (2011).
17. Quistad, G. B., Leisy, D. J., & Bernasconi, P. & Leisy, D. J. Purification and characterization of insecticidal toxins from venom glands of the parasitic wasp Bracon hebetor. Insect Biochem. Mol. Biol. 24, 955–961 (1994).
18. Laurino, S. et al. Identification of major Toxoneuron nigriceps venom proteins using an integrated transcriptomic/proteomic approach. Insect Biochem. Mol. Biol. 76, 49–61 (2016).
19. Colinet, D., Schmitz, A., Depoix, D., Crochard, D. & Poirie, M. Convergent use of RhoGAP toxins by eukaryotic parasites and bacterial pathogens. PLoS Pathog. 3, 203 (2007).
20. Yang, L. et al. Identification and comparative analysis of venom proteins in a pupal endoparasitoid Pachycrepoideus vindemiatus. Insect Biochem. Mol. Biol. 37, 453–465 (2007).
21. Battaglia, D. et al. The effect of Leptomastix dactylopi parasitism and venom injection on host Planococcus citri. Invert. Surviv. J. 11, 273–285 (2014).
22. Labelle, C. et al. Identification of two Arginine Kinase forms of endoparasitoid Leptomastix dactylopi venom by bottom up sequencing tag approach. J. Mass Spectrom. 50, 756–765 (2015).
23. Falabella, P. The mechanism utilized by Toxoneuron nigriceps in inhibiting the host immune system. Invert. Surviv. J. 15, 240–255 (2018).
24. Parkinson, N., Conyers, C. & Smith, I. A venom protein from the endoparasitoid wasp Pimpla hypochondriaca is similar to snake venom repolysin-type metalloproteases. J. Invertebr. Pathol. 79, 129–131 (2002).
Moore, E. L., Haspel, G., Libersat, F. & Adams, M. E. Parasitoid wasp sting: A cocktail of GABA, taurine, and beta-alanine opens chloride channels for central synaptic block and transient paralysis of a cockroach host. J. Neurobiol. 66, 811–820 (2006).

Goecks, J. et al. Integrative approach reveals composition of endoparasitoid wasp venoms. PLoS ONE 8, e64125 (2013).

Yan, Z. C. et al. Into the venom composition and evolution of an endoparasitoid wasp by combining proteomic and transcriptomic analyses. Sci. Rep. 6, 19604 (2016).

Teng, Z. W. et al. Protein discovery; Combined transcriptomic and proteomic analyses of venom from the endoparasitoid Cotesia chilonis (Hymenoptera: Braconidae). Toxins. 9, 135 (2017).

Yang, L. et al. Identification and characterization of serine protease inhibitors in a parasitic wasp Pteromalus puparum. Sci. Rep. 7, 15755 (2017).

de Graaf, D. C. et al. Insights into the venom composition of the ectoparasitoid wasp Nasonia vitripennis from bioinformatic and proteomic studies. Insect Mol. Biol. 19, 11–26 (2010).

Perkin, L. C., Friesen, K. S., Finn, P. W. & Oppert, B. Venom gland components of the ectoparasitoid wasp Anisopteromalus calandrae. J. Venom Res. 6, 19–37 (2015).

Becchimanzii, A. et al. Venomics of the ectoparasitoid wasp Bracon nigricans. BMC Genomics. 21, 34 (2020).

Duret, S. et al. Venoms-based drug discovery: proteomic and transcriptomic approaches in Venoms to Drugs (ed. Glenn, K.F) 80–96 (Royal Society of Chemistry, 2015).

Ferracini, C., Ferrari, E., Pontini, M., Saladini, M. A. & Alma, A. Effectiveness of Torymus sinensis: A successful long-term control of the Asian chestnut gall wasp in Italy. J. Pest. Sci. 92, 353–359 (2019).

Johnson, J. H., Kral, R. M. & Kraepho, K Jr. Insecticidal toxins from Bracon hebetor nucleic acid encoding said toxin and methods of use. US Patent No. 5874298 (1999).

Ullah, M. I. et al. Arthropods venom used as bio-pesticides: A new challenge to manage insect pests. Int. J. Agric. Appl. Sci. 9, 122–131 (2017).

Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids. Res. 25, 3389–3402 (1997).

Cox, J. et al. Andromeda: A peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10, 1794–1805 (2011).

Özbek, R. et al. Proteo-transcriptomic characterization of the venom from the endoparasitoid wasp Pimpla turionellae with aspects on its biology and evolution. Toxins. 11, 721 (2019).

de Oliveira, U. C., et al. Proteomic endorsed transcriptomic profiles of venom glands from Titus obscurus and T. serrulatus scorpions. PLoS One. 13, e0193739 (2018).

Lin, Z. et al. Insights into the venom protein components of Microplitis mediator, an endoparasitoid wasp. Insect Biochem. Mol. Biol. 105, 33–42 (2019).

Moreau, S. J. M. & Guillot, S. Advances and prospects on biosynthesis, structures and functions of venom proteins from parasitic wasps. Insect Biochem. Mol. Biol. 35, 1209–1223 (2005).

Colinet, D. et al. Identification of the main venom protein components of Aphidius ervi, a parasitoid wasp of the aphid model Acyrthosiphon pisum. BMC Genomics. 15, 342 (2014).

Asgari, S., Zhang, G., Zareie, R. & Schmidt, O. A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph. Insect Biochem. Mol. Biol. 33, 1017–1021 (2003).

Zhang, G., Lu, Z. Q., Jiang, H. & Asgari, S. Negative regulation of prophenoloxidase (proPO) activation by a clip-domain serine proteinase homolog (SPH) from endoparasitoid venom. Insect Biochem. Mol. Biol. 34, 477–483 (2004).

Thomas, P. & Asgari, S. Inhibition of melanization by a parasitoid serine protease homolog venom protein requires both the clip and the non-catalytic protease-like domains. Insects. 2, 509–514 (2011).

Yu, X., Jiang, H., Wang, Y. & Kanost, M. R. Nonproteolytic serine protease homologs are involved in prophenoloxidase activation in the tobacco hornworm Manduca sexta. Insect Biochem. Mol. Biol. 33, 197–208 (2003).

Danneels, E. L., Rivers, D. B. & de Graaf, D. C. Venom proteins of the parasitoid wasp Torymus sinensis. Biocontrol Sci. 59, 222–228 (2008).

Bao, Y. Y. et al. Genomic insights into the serine protease gene family and expression profile analysis in the planthopper Nilapavrita lugens. BMC Genomics. 21, 507 (2014).

Kini, R. M. & Koh, C. Y. Metalloproteases affecting blood coagulation, fibrinolysis and platelet aggregation from snake venoms: Definition and nomenclature of interaction sites. Toxins. 8, 284 (2016).

Fingleton, B. Matrix metalloproteinases as regulators of inflammatory processes. Biochem. Biophys. Acta. Mol. Cell. Res. 1864, 2036–2042 (2017).

Willett, E. & Tran, H. Q. Zinc and Manduca sexta hemocyte functions. J. Insect Sci. 2, 6 (2002).

Liu, Z. et al. A metalloproteinase homolog venom protein from a parasitoid wasp suppresses the toll pathway in host hemocytes. Front. Immunol. 9, 2301 (2018).

Price, D. R. G. et al. A venom metalloproteinase from the parasitic wasp Eulophus pensicornis is toxic towards its host, tomato moth (Lacanobia oleracea). Insect Mol. Biol. 18, 195–202 (2009).

Zadka, L., Kulus, M. J. & Pietek, K. ADAM protein family: Its role in tumorigenesis, mechanisms of chemoresistance and potential as diagnostic and prognostic factors. Neoplasma. 65, 823–839 (2018).

Rawlings, N.D. & Salvesen, G. Handbook of Proteolytic Enzymes. (eds. Rawlings, N.D. & Salvesen, G) e1-e39 (Academic Press, 2013).

Aird, S. D. Snake venom dipeptidyl peptidase IV: Taxonomic distribution and quantitative variation. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 150, 222–228 (2008).
1. Blank, S. et al. Identification, recombination expression, and characterization of the 100 kDa high molecular weight Hymenoptera venom allergens Api m 5 and Ves v 3. J. Immunol. 184, 5403–5413 (2010).

2. Hishe, S. K. et al. Functional expression and characterization of dipeptidyl peptidase IV from the black-bellied hornet Vespa basalis in SF21 insect cells. Biosci. Biotechnol. Biochem. 75, 2371–2375 (2011).

3. Schiener, M. et al. The high molecular weight dipeptidyl peptidase IV Pol d 3 is a major allergen of Polistes dominula venom. Sci. Rep. 8, 1318 (2018).

4. Kreil, L., Haiml, I. & Suchaneck, G. Stepwise cleavage of the pro part of promelittin by dipeptidylpeptidase IV – evidence for a new type of precursor-product conversion. Eur. J. Biochem. 111, 49–58 (1980).

5. Borloo, M. & De Meester, I. Dipeptidyl peptidase IV: development, design, synthesis and biological evaluation of inhibitors. Verh. K. Acad. Geneesk., Belg. 56, 57–88 (1994).

6. Dos Santos, D. M., de Souza, C. B. & Pereira, H. J. Angiotensin converting enzymes in fish venom. Toxicon 131, 63–67 (2017).

7. Ogawa, Y., Murayama, N., Fujita, Y. & Yanoshita, R. Characterization and cDNA cloning of aminopeptidase A from the venom of Gloydius blomhoffi brevicaudus. Toxicon 49, 1172–1181 (2007).

8. Vaiyapuri, S. et al. Purification and functional characterisation of rhinopeptidase A, a novel aminopeptidase from the venom of Bittis gabanica rhinoceras. PLoS Negl. Trop. Dis. 4, e796 (2010).

9. Santos, P. P. et al. Proteomic analysis of the venom of the predatory ant Pachycondyla striata (Hymenoptera: Formicidae). Arch. Insect Biochem. Physiol. 96, e21424. https://doi.org/10.1002/arch.21424 (2017).

10. Sanggaard, K. W. et al. Spider genomes provide insight into composition and evolution of venom and silk. Nat. Commun. 5, 3765 (2014).

11. Sarkar, N. P., Bayazit, M. B. & Ozel Demiralp, D. Proteomic Characterization of the Venom of Five Bombus (Thoracobombus) Species. Toxins. 9, 362 (2017).

12. Reaux, A. et al. Aminopeptidase A, which generates one of the main effectors peptides of the brain renin-angiotensin system, angiotensin III, has a key role in central control of arterial blood pressure. Biochem. Soc. Trans. 28, 435–440 (2000).

13. Marchiò, S. et al. Aminopeptidase A is a functional target in angiogenic blood vessels. Cancer Cell 5, 151–162 (2004).

14. Brix, K. Lysosomal Proteases: Revival of the Sleeping Beauty in Lysosomes (ed. Saftig, P.) (Springer, 2005).

15. Marchiò, S. et al. Early changes in the pupal transcriptome of the flesh fly Sarcophaga crassipalpis to parasitization by the ectoparasitoid wasp Nasonia vitripennis. Insect Biochem. Mol. Biol. 43, 1189–1200 (2013).

16. Stoka, V., Turk, B. & Turk, V. Lysosomal cysteine proteases: structural features and their role in apoptosis. JUBMB Life 57, 347–353 (2005).

17. Zhou, J., Zhang, Y. L., Li, Q. Y. & Cai, Z. H. Evolutionary History of Cathepsin D (L-like) Family Genes in Vertebrates. BMC Genomics. 15, 987 (2014).

18. Melani, R. D. et al. Seeing beyond the tip of the iceberg: a deep analysis of the venom of the Brazilian rattlesnake Crotalus durissus terrificus. EuPA Open Proteom. 8, 144–156 (2015).

19. Hubert, H. M. et al. Comparative venomics of Ptytiala lonshburyi and P. concor, two olive fruit fly parasitoids: a hypothetical role for a GH1 β-glucosidase. Sci. Rep. 6, 35873 (2016).

20. Kim, B. Y. & Jin, B. R. Molecular characterization of a venom acid phosphatase AcpH-1-like protein from the Asiatic honeybee Apis cerana. J. Asia-Pacific Entomol. 17, 695–700 (2014).

21. Hossei, M. S., Shapla, U. M., Gan, S. H. & Khalil, M. I. Impact of bee venom enzymes on diseases and immune responses. Molecules 22, 25 (2016).

22. Danneels, E. L. et al. Molecular cloning and characterization of a 100 kDa high molecular weight Hymenoptera venom allergen of Vespa velutina: Api m 5 and Ves v 3. J. Insect Physiol. 110, 2022–2030 (2017).

23. Tani, N. et al. Mass spectrometry analysis and biological characterization of the predatory ant odontomachus monticola venom components. Toxins. 11, 50 (2019).

24. Bouzid, W. et al. De Novo sequencing and transcriptome analysis for Tetramorium bicarinatum: A comprehensive venom gland transcriptome analysis from an ant species. BMC Genomics. 15, 987 (2014).

25. Musevo, V. et al. The major constituents of the venom gland of a braconid endoparasitoid, Bracon nigricans. Insect Biochem. Mol. Biol. 43, 373–381 (2005).

26. Zhau, J. Y., Ye, G. Y. & Hu, C. Molecular cloning and characterization of acid phosphatase in venom of the endoparasitoid wasp Pteromalus paparum (Hymenoptera: Pteromalidae). Toxicon 51, 1391–1399 (2008).

27. Manzoor, A., UAbdIn, Z., Webb, B. A., Arif, M. J. & Jamil, A. De novo sequencing and transcriptome analysis of female venom glands of ectoparasitoid Bracon helbetor (Say) (Hymenoptera: Braconidae). Comp. Biochem. Physiol. Part D Genomics Proteomics. 20, 101–110 (2016).

28. Yokoi, K. et al. The major constituents of the venom gland of a braconid endoparasitoid, Meteorus pulchricornis (Hymenoptera: Braconidae). J. Neuroendocrin. 25, 271–279 (2017).

29. Hoffman, D. R. Hymenoptera venom allergens. Clin. Rev. Allergy Immunol. 30, 109–128 (2006).

30. Aranda, J. et al. The catalytic mechanism of carboxylesterases: A computational study. Biochemistry 53, 5820–5829 (2014).

31. Matsiyak, I., Hajduck, J., Pietrzak, L., Schmelzer, C. E. & Kokot, Z. J. Shotgun proteome analysis of honeybee venom using targeted enrichment strategies. Toxicon 90, 255–264 (2014).

32. Saha, A. & Birner-Gruenberger, R. Lipases in Encyclopedia of Metalloproteins (eds. Kretzinger, R.H., Uversky, V.N. & Pernyakov, E.A.) (Springer, 2013).

33. Rivers, D. B. & Denlinger, D. L. Venom-induced alterations in fly lipid-metabolism and its impact on larval development of the ectoparasitoid Nasonia vitripennis (Walker) (Hymenoptera, Pteromalidae), J. Invertebr. Pathol. 66, 104–110 (1995).

34. Crawford, A. M. et al. The constituents of Microctonus sp. parasitoid venom. Insect. Mol. Biol. 17, 313–324 (2008).

35. Cusumano, A. et al. First extensive characterization of the venom gland from an egg parasitoid: Structure, transcriptome and functional role. J. Insect Physiol. 107, 68–80 (2018).

36. Nakamatsu, Y., Fujii, S. & Tanaka, T. Larvae of an endoparasitoid Cotesia kariyai (Hymenoptera: Braconidae), feed on the host fat body directly in the second stadium with the help of teratocytes. J. Insect Physiol. 48, 1041–1052 (2002).

37. Nakamatsu, Y. & Tanaka, T. Venom of Euplectrus separatus causes hyperlipidemia by lysis of host fat body cells. J. Insect Physiol. 50, 267–275 (2004).

38. Kretudt Cairns, J. R. & Esen, A. β-glucosidases. Cell. Mol. Life Sci. 67, 3389–3405 (2010).

39. Burke, G. R. & Strand, M. R. Systematic analysis of a wasp parasitism arsenal. Mol. Ecol. 23, 890–901 (2014).

40. Juers, D. H., Mathews, B. W. & Huber, R. E. Lacc 2 β-galactosidase: structure and function of an enzyme of historical and molecular biological importance. Protein Sci. 21, 1792–1807 (2012).
121. Liberato, T., Troncone, L. R., Yamashiro, E. T., Serrano, S. M. T. & Zelanis, A. High-resolution proteomic profiling of spider venom. *Mol. Cell. Proteomics.*

122. Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C. & Schramm, V. L. Inosine-uridine nucleoside hydrolase from *Serratia.* *J. Insect Physiol.* 53, 5963–5970 (1996).

123. Consoli, F. L., Lewis, D., Keeley, L. & Vinson, S. B. Characterization of a cDNA encoding a putative chitinase from teratocytes of the endoparasitoid *Toxoneuron nigriceps.* *Entomol. Exp. App.* 122, 271–278 (2007).

124. Vincent, B. et al. The venom composition of the parasitic wasp *Chelonus inanitus* resolved by combined expressed sequence tags analysis and proteomic approach. *BMC Genomics.* 11, 693 (2010).

125. Dhananjaya, B. L. & D’Souza, C. J. An overview on nucleases (DNase, RNase, and phosphodiesterase) in snake venoms. *Biochemistry (Mosc).* 75, 1–6 (2010).

126. Neeman, L., Calton, G. J. & Burnett, J. W. Purification and characterization of the endonuclease present in *Physisila physalis venom.* *Comp. Biochem. Phys. B.* 67, 155–158 (1986).

127. Miller, M. D., Tanner, J., Alpaugh, M., Benedik, M. J. & Krause, K. L. A structure of *Serratia* endonuclease suggests a mechanism for binding to double-stranded DNA. *Nat. Struct. Biol.* 1, 461–468 (1994).

128. Gopal, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C. & Schramm, V. L. Inosine-uridine nucleoside hydrolase from *Critidia fasciiculata.* Genetic characterization, crystallization, and identification of histidine 241 as a catalytic site residue. *Biochemistry* 35, 5963–5970 (1996).

129. Paulson, A. R. et al. Transcriptome analysis provides insight into venom evolution in a sea-parasitic wasp *Megaistegus spermatophorus.* *Insect Mol. Biol.* 25, 604–616 (2016).

130. Lee, C. C., Hsieh, H. J., Hsieh, C. H. & Hwang, D. F. Plancitoxin I from the venom of crown-of-thorns starfish (*Acanthaster planci*) induces oxidative and endoplasmic reticulum stress associated cytotoxicity in *A375.S2* cells. *Exp. Mol. Pathol.* 99, 7–15 (2015).

131. Dennis, A. B., Patel, V., Oliver, K. M. & Vorburger, C. Parasitoid gene expression changes after adaptation to symbiont-protected hosts. *Evolution* 71, 2599–2617 (2017).

132. Colinet, D. et al. A serpin from the parasitic wasp *Leptopilina boulardi* targets the *Drosophila* phospholipase cascade. *Dev. Comp. Immunol.* 33, 681–689 (2009).

133. Van, Z. et al. A venom serpin splicing isoform of the endoparasitoid wasp *Pteromalus puparum* suppresses host phospholipase cascade by forming complexes with host hemolymph proteinases. *J. Biol. Chem.* 292, 1038–1051 (2017).

134. Gettins, P. G. Serpin structure, mechanism, and function. *Chem. Rev.* 102, 4751–4804 (2002).

135. Kanost, M.R. & Gorman, M.J. Phenoloxidases in *Insect immunology* (ed. Beckage, N.E.) 69–96 (Academic Press, 2008).

136. Niimi, T., Yokoyama, H., Goto, A., Beck, K. & Kitagawa, Y. A *Drosophila-*gene encoding multiple splice variants of Kazal-type serine protease inhibitor-like proteins with potential destinations of mitochondria, cytosol and the secretory pathway. *Eur. J. Biochem.* 266, 282–292 (1999).

137. Zheng, Q. L. et al. Expression, purification and characterization of a three-domain Kazal-type inhibitor from silkworm pupae (*Bombyx mori*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 146, 234–240 (2007).

138. Campos, I. T., Tanaka-Azevedo, A. M. & Tanaka, A. S. Identification and characterization of a novel factor XIIa inhibitor in the hematophagous insect, *Triatoma infestans* (Hemiptera: Reduviidae). *FEBS Lett.* 527, 512–516 (2004).

139. Qian, C., Fang, Q., Wang, L. & Ye, G. Y. Molecular cloning and functional studies of two Kazal-type serine protease inhibitors specifically expressed by *Nasonia vitripennis* venom apparatus. *Toxins* 7, 2888–2905 (2015).

140. Beck, M., Theopold, U. & Schmidt, O. Evidence for serine protease inhibitor activity in the ovarian calyx fluid of the endoparasitoid *Venturia canescens.* *J. Insect Physiol.* 46, 1275–1283 (2000).

141. Lu, A. et al. Insect phospholipaseA2: the view beyond immunity. *Front. Physiol.* 5, 252 (2014).

142. Fernández, J., Gutiérrez, J. M., Calvet, J. I., Sanz, L. & Lomonte, B. Characterization of a novel snake venom component: Kazal-type inhibitor-like protein from the arboreal pitviper Bothriechis schlegelii. *Biochimie* 125, 83–90 (2016).

143. Morjen, M. et al. PIVL, a snake venom Kunitz-type serine protease inhibitor, inhibits *in vitro* and *in vivo* angiogenesis. *Microvasc. Res.* 95, 149–156 (2014).

144. Parkinson, N. M. et al. Towards a comprehensive view of the primary structure of venom proteins from the parasitoid wasp *Pimpla hypothoracica.* *Insect Biochem. Mol. Biol.* 34, 565–571 (2004).

145. Rivers, D. B., Ruggiero, L. & Hayes, M. The ectoparasitic wasp *Nasonia vitripennis* (Walker) (Hymenoptera: *Pteromalidae*) differentially affects cells mediating the immune response of its flesh fly host, *Sarcophaga bullata* Parker (Diptera: Sarcophagidae). *J. Insect Physiol.* 48, 1053–1064 (2002).

146. Chot, J. V. et al. Calcreticulin enriched as an early-stage encapsulation protein in wax moth *Galleria mellonella* larvae. *Dev. Comp. Immunol.* 26, 335–343 (2002).

147. Asgari, S. & Schmidt, O. Is cell surface calreticulin involved in phagocytosis by insect hemocytes?. *J. Insect. Physiol.* 49, 545–550 (2003).

148. 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.* 75, 28–44 (2010).

149. Suntravat, M. et al. Gene expression profiling of the venom gland from the Venezuelan mapanere (*Bothrops colombianus*) using expressed sequence tags (ESTs). *BMC Mol. Biol.* 17, 7 (2016).

150. Wong, E. S. et al. Proteomics and deep sequencing comparison of seasonally active venom glands in the flatworm reveals novel venom peptides and distinct expression profiles. *Mol. Cell. Proteomics.* 11, 1354–1364 (2012).

151. Nicolau, C. A. An in-depth snake venom proteopeptide characterization: Benchmarking *Bothrops jararaca.* *J. Proteomics.* 151, 214–231 (2017).

152. de Graaf, D. C. et al. Two novel proteins expressed by the venom glands of *Apis mellifera* and *Nasonia vitripennis* share an ancient C1q-like domain. *Insect Mol. Biol.* 19, 1–10 (2010).
194. Liu, H., Tang, X. & Gong, L. Mesencephalic astrocyte-derived neurotrophic factor and cerebral dopamine neurotrophic factor: New endoplasmic reticulum stress response stress. *Eur. J. Pharmacol.* 759, 118–122 (2015).

195. Carducci, F., Biscotti, M. A. & Canapa, C. Vitellogenin gene family in vertebrates: Evolution and functions. *Eur. Zool. J.* 86, 233–240 (2019).

196. Blank, S. et al. Vitellogenins are new high molecular weight components and allergens (Api m 12 and Ves v 6) of *Apis mellifera* and *Vespula vulgaris* venom. *PLoS ONE* 8, e62009 (2013).

197. Park, H. G. et al. Honeybee (*Apis cerana*) vitellogenin acts as an antimicrobial and antioxidant agent in the body and venom. *Dev. Comp. Immunol.* 85, 51–60 (2018).

198. Luna-Ramírez, K., Quintero-Hernández, V., Juárez-González, V. R. & Possani, L. D. Whole transcriptome of the venom gland from *Urodacus yaschenkoi* scorpion. *PLoS ONE* 10, e0127883 (2015).

199. Adams, M. E. Agatoxins: ion channel specific toxins from the American funnel web spider *Agelenopsis aperta*. *Toxicol* 43, 509–525 (2004).

200. Durertre, S. & Lewis, R. J. Use of venom peptides to probe ion channel structure and function. *J. Biol. Chem.* 285, 13315–13320 (2010).

201. Torres, A. F. et al. Transcriptome analysis in venom gland of the predatory giant ant *Dinoponera quadriceps*: insights into the polypeptide toxin arsenal of hymenopterans. *PLoS ONE* 9, e87556 (2014).

202. Lee, S. H., Baek, J. H. & Yoon, K. Differential properties of venom peptides and proteins in solitary vs social hunting wasps. *Toxins* 8, 32 (2016).

203. Sharko, F. S. et al. A partial genome assembly of the miniature parasitoid wasp *Megaphragma ammonphilanthum*. *PLoS ONE* 14, e0226485 (2019).

204. Tang, B. Z. et al. Combination of label-free quantitative proteomics and transcriptomics reveals intraspecific venom variation between the two strains of *Tetrastichus brontispae*, a parasitoid of two invasive beetles. *J. Proteomics*. 192, 37–53 (2019).

205. Zhang, X. M. et al. Cloning and Immunosuppressive Properties of an Acyl-Activating Enzyme from the Venom Apparatus of *Tetrastichus brontispae* (Hymenoptera: Eupelminidae). *Toxins*. 11, 672 (2019).

206. Chauhan, P. et al. De novo transcriptome of *Ischnura elegans* provides insights into sensory biology, colour and vision genes. *BMC Genomics*. 15, 808 (2014).

207. Zhang, T. et al. Aggregate-reaction activity of the molecular chaperone CipB from *Ehrlichia chaffeensis*. *PLoS ONE* 8, e62454 (2013).

208. Lubé, G., Afshari-Sadat, L., Yang, J. W. & John, J. P. Searching for hypothetical proteins: Theory and practice based upon original data and literature. *Prog. Neurobiol.* 77, 90–127 (2005).

209. Estrada-Gomez, S. et al. Venomic, transcriptomic, and bioactivity analyses of pamphphobeteus verdolaga venom reveal complex disulphide-rich peptides that modulate calcium channels. *Toxins*. 11, 496 (2019).

210. Rokyta, D. R. & Ward, M. J. Venom-gland transcriptomics and venom proteomics of the black-back scorpion (*Hadrurus spadix*) reveal detectability challenges and an unexplored realm of animal toxin diversity. *Toxicon* 128, 23–37 (2017).

211. Ward, M. J., Ellsworth, S. A. & Rokyta, D. R. Venom-gland transcriptomics and venom proteomics of the Hentz striped scorpion (*Centruroides hentzi*; *Buthidae*) reveal high toxin diversity in a harmless member of a lethal family. *Toxicon* 142, 14–29 (2018).

212. Cid-Urbane, J. I. et al. The diversity of venom components of the scorpion species *Paravaejovis schwenkmeyeri* (Scorpiones: Vaejovidae) revealed by transcriptome and proteome analyses. *Toxins* 11, 47–62 (2018).

213. IPLA. Regione Piemonte. Collana Manuali Tecnico-divulgativi di Selvicoltura. Collana Manuali Tecnico-divulgativi di Selvicoltura. *Tipi forestali del Piemonte*. (Blu Edizioni, 2004).

214. Adams, M. E. Agatoxins: ion channel specific toxins from the American funnel web spider *Agelenopsis aperta*. *Toxicol* 43, 509–525 (2004).

215. Liu, H. et al. Venom gland transcriptomics and venom proteomics of the Hentz striped scorpion (*Centruroides hentzi*; *Buthidae*) reveal high toxin diversity in a harmless member of a lethal family. *Toxicon* 142, 14–29 (2018).

216. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by high-throughput functional annotation and data mining with the Blast2GO suite. *PLoS ONE* 3, 543–548 (2018).

217. Waterhouse, R. M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35, 543–548 (2018).

218. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*. 5, 621–628 (2008).

219. Liu, W. & Saint, D. A. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Anal. Biochem.* 302, 52–59 (2002).

220. Lee, C., Kim, J., Shin, S. G. & Hwang, S. Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *J. Biotechnol.* 123, 273–280 (2006).

221. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔC(T) method. *Methods* 285, 402–408 (2001).

222. Bradford, M. M. A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254 (1976).

223. Medugno, L. et al. A novel zinc finger transcriptional repressor, ZNF224, interacts with the negative regulatory element (AldA−) of the zinc-responsive transcription factor ZNF224. *FEBS Lett.* 534, 93–100 (2003).

Acknowledgements

We would like to thank Emily Wheeler for editorial assistance.

Author contributions

Conceptualization, P.F.; data curation, C.S. and R.S.; methodology, C.S., R.S., A.F., F.C., C.F., M.P., H.V. and P.F.; project administration, P.F.; supervision, P.F.; validation, H.V., P.P., A.A. and P.F.; writing—original draft, P.F.; writing—review and editing, C.S., R.S., A.F., M.P., F.C., M.C., Ch.Sc., H.V., M.M., C.F., P.P., A.A. and P.P.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-84385-5.
