

# Melt With This Kiss: Paralyzing and Liquefying Venom of The Assassin Bug *Pristhesancus plagipennis* (Hemiptera: Reduviidae)\*<sup>§</sup>

Andrew A. Walker<sup>‡§</sup>, Bruno Madio<sup>‡</sup>, Jiayi Jin<sup>‡</sup>, Eivind A. B. Undheim<sup>¶</sup>, Bryan G. Fry<sup>||</sup>, and Glenn F. King<sup>‡§</sup>

Assassin bugs (Hemiptera: Heteroptera: Reduviidae) are venomous insects, most of which prey on invertebrates. Assassin bug venom has features in common with venoms from other animals, such as paralyzing and lethal activity when injected, and a molecular composition that includes disulfide-rich peptide neurotoxins. Uniquely, this venom also has strong liquefying activity that has been hypothesized to facilitate feeding through the narrow channel of the proboscis—a structure inherited from sap- and phloem-feeding phytophagous hemipterans and adapted during the evolution of Heteroptera into a fang and feeding structure. However, further understanding of the function of assassin bug venom is impeded by the lack of proteomic studies detailing its molecular composition.

By using a combined transcriptomic/proteomic approach, we show that the venom proteome of the harpactorine assassin bug *Pristhesancus plagipennis* includes a complex suite of >100 proteins comprising disulfide-rich peptides, CUB domain proteins, cystatins, putative cytolytic toxins, triabin-like protein, odorant-binding protein, S1 proteases, catabolic enzymes, putative nutrient-binding proteins, plus eight families of proteins without homology to characterized proteins. S1 proteases, CUB domain proteins, putative cytolytic toxins, and other novel proteins in the 10–16-kDa mass range, were the most abundant venom components. Thus, in addition to putative neurotoxins, assassin bug venom includes a high proportion of enzymatic and cytolytic venom components likely to be well suited to tissue liquefaction. Our results also provide insight into the trophic switch to blood-feeding by the kissing bugs (Reduviidae: Triatominae). Although some protein families such as triabins occur in the venoms of both predaceous and blood-feeding reduviids, the composition of venoms produced by these two groups is revealed to differ markedly. These results pro-

vide insights into the venom evolution in the insect sub-order Heteroptera. *Molecular & Cellular Proteomics* 16: 10.1074/mcp.M116.063321, 552–566, 2017.

Venoms are chemical arsenals injected by one animal into another to disrupt the homeostasis of the injected animal in ways that assist predation, defense, or feeding by the injecting animal (1). Typically, venoms are composed of multiple toxins, including peptides, enzymes, and small molecules, such as polyamines, that bind to and affect the function of multiple molecular targets in the injected animal. Because of their key role governing life-or-death interactions between animals, venom toxins are subject to selection pressures that have resulted in unique evolutionary patterns such as massive duplication and accelerated evolution of toxin-encoding genes (2–5). In addition, the properties that ensure that toxins confer a fitness advantage to the animals that produce them, including high stability and potency, make them well suited for use as insecticides, therapeutics, and pharmacological tools (6–11). However, our understanding of the factors shaping venom evolution, and our ability to repurpose venom toxins for biotechnological use, is limited by the current focus of research on a small number of prominent groups of venomous animals: the scorpions, spiders, snakes, and cone snails. Studies on neglected taxa (12–16) are essential to gain a more general understanding of how venom systems evolve and how venom evolution is influenced by factors such as geographical, trophic, and morphological constraints.

Assassin bugs (family Reduviidae) are a large and diverse group of insects consisting of ~6800 species in 25 subfamilies distributed over all continents except Antarctica (17). Like other hemipterans such as cicadas and aphids, which feed on plants, reduviids have mouthparts that are extensively elongated and modified to form a proboscis that is specialized for piercing and sucking. Reduviids however (together with the majority of other heteropteran families), use their piercing-sucking mouthparts to inject venom into, and feed from, prey. Exceptions are the blood-feeding kissing bugs (Reduviidae: Triatominae) that use their venom to facilitate acquisition of blood meals from vertebrates, including humans (18). The venom apparatus of reduviids includes morphologically com-

From the <sup>‡</sup>Institute for Molecular Bioscience, <sup>¶</sup>Centre for Advanced Imaging, and <sup>||</sup>School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia

Received August 17, 2016, and in revised form, January 4, 2017

Published, MCP Papers in Press, January 27, 2017, DOI 10.1074/mcp.M116.063321

Author contributions: A.A.W., E.A.U., and G.F.K. designed the research; A.A.W., B.M., and J.J. performed the research; B.G.F. contributed new reagents or analytic tools; A.A.W. analyzed the data; and A.A.W., B.M., J.J., E.A.U., B.G.F., and G.F.K. wrote the paper.

plex paired secretory glands within the thorax/abdomen, a muscle-driven pump within the head, and a devoted venom channel formed by interlocking maxillary stylets through which venom can be injected into the prey (16). Assassin bug venoms are paralytic and lethal to invertebrates and small vertebrates (19–22), hyperalgesic to vertebrates (23), cytolytic (19, 24), and antibacterial (25). Numerous enzymatic activities are present, including protease, phospholipase, and hyaluronidase (16, 19, 26). The ability of assassin bug venom to liquefy tissue (19) is usually interpreted as a form of extra-oral digestion (EOD)<sup>1</sup> (27) that allows feeding through the proboscis.

The presence of neurotoxins in assassin bug venom has been demonstrated by the ability of venom to reversibly block nerve conduction (19, 22). Corzo *et al.* (28) determined the primary structure of three disulfide-rich peptides in assassin bug venom and showed that one of these, Ptu1, was neurotoxic by virtue of its ability to inhibit the voltage-gated calcium channel Ca<sub>v</sub>2.2. Ptu1 was shown to have the inhibitor cystine knot (ICK) fold (29) that is common among toxins from numerous other venomous animals, including spiders, cone snails, scorpions, and sea anemones (30).

Despite the natural abundance of assassin bugs and their importance as predators of invertebrates in agricultural and other ecosystems (31–33), the detailed composition of assassin bug venom remains unknown. Here, we elucidate the protein composition of venom from the harpactorine assassin bug *Pristhesancus plagipennis* using a combined transcriptomic/proteomic approach. We show that this venom contains more than a hundred individual components, including putative neurotoxins, cytolytic toxins, digestive enzymes, and members of eight novel protein families. Our data reveal convergent evolution between assassin bugs and other venomous animals, as well as unique differences in reduviid venom that may be related to the dual requirement of assassin bug venom to both paralyze and liquefy prey.

#### EXPERIMENTAL PROCEDURES

**Insects and Venom Collection**—Assassin bugs (*P. plagipennis*) were collected in Brisbane, Australia, fed on crickets (*Acheta domesticus*; Pisces Live Food, Brisbane, Australia), and housed in individual containers to avoid cannibalism. Venom was harvested from adults of both sexes by electrostimulation using a non-lethal protocol. Bugs were restrained on a foam platform with a rubber band over the thorax, and the proboscis was gently inserted into a P200 pipette tip, and electrostimulation (30 V, 5 ms pulses, 5 Hz) was applied to the thorax using an S48 square pulse stimulator (Grass Technologies, Warwick, RI) with electrodes installed on tweezers (supplemental Video 1). Venom was immediately transferred to a tube on dry ice and stored at –80 °C until analysis.

<sup>1</sup> The abbreviations used are: EOD, extra-oral digestion; ICK, inhibitor cystine knot; BLAST, basic local alignment search tool; OCLP,  $\omega$ -conotoxin-like peptide; CUB domain, complement C1r/C1s, Uegf, Bmp1 domain; BAPP, biogenic amine binding protein; mya, million years ago; PLA<sub>2</sub>, phospholipase A<sub>2</sub>; RNA-Seq, RNA sequencing; rcf, relative centrifugal force; AMP, antimicrobial peptide.

**Experimental Design and Statistical Rationale**—The purpose of this study was to determine the venom proteome of *P. plagipennis*. Therefore, we employed methods to maximize the number and diversity of proteins identified, including strategies to identify peptides (short minimum contig length for RNA-Seq assemblies; LC-MS/MS of HPLC fractions) and low abundance proteins (combining multiple RNA-Seq assemblies, and LC-MS/MS of HPLC fractions and 2D gel electrophoresis spots). Overall, our LC-MS/MS analysis included 62 HPLC fractions, 54 1D SDS-PAGE bands, and 156 2D SDS-PAGE spots. The eventual dataset includes 17 technical replicates of selected HPLC fractions, two technical replicates of reduced and alkylated but undigested venom, and two replicates of reduced, alkylated, and digested venom. The remaining samples are neither biological nor technical replicates but are subsets of venom components fractionated to maximize the potential of LC-MS/MS for protein identification. Initially, spectra were compared by Paragon searches (34) against a database containing all ORFs with a length of >30 amino acids in our three venom gland transcriptomes. Data from MS samples were pooled for a single search wherever possible to facilitate identification of the optimal data set of proteins with minimum redundancy using the Paragon/ProtGroup algorithms in ProteinPilot. In practice, it was necessary to run a separate Paragon search for 2D gel spots, as they were alkylated with iodoacetamide (all other samples were alkylated with iodoethanol). The proteins and peptides identified by these two searches were reviewed manually, and poor-quality identifications were excluded to yield a draft dataset of 130 proteins. To check this manual process, we re-validated protein identifications using further Paragon searches against a database containing just these 130 sequences, leading us to further discard three sequences with ProteinPilot Unused values below the threshold of 1.3 (corresponding to >95% confidence at protein level). After establishing the proteome, we performed an extra Paragon search of each of our 277 experimental samples individually against the venom proteome, providing insights into assignment of individual gel spots and HPLC peaks.

**Transcriptomics**—For RNA extraction, venom glands were harvested from two adult female and two final instar bugs after anesthesia with CO<sub>2</sub> for ~5 min. The main gland posterior lobe, main gland anterior lobe, and accessory gland were removed and stored separately in >10× the glandular volume of RNAlater (Ambion, Austin, TX). Total RNA was extracted using a DNeasy kit (Qiagen, Mississauga, Canada), and mRNA was isolated using a Dynabeads mRNA Direct kit (Ambion) according to the manufacturer's instructions. This process yielded 3120, 680, and 119 ng of mRNA from the main gland posterior lobe, main gland anterior lobe, and accessory gland, respectively. RNA sequencing (RNA-Seq) was performed on 340 ng from each lobe of the main gland and 119 ng from the accessory gland on the Illumina, San Diego, CA NextSeq instrument at the IMB Sequencing Facility. After TruSeq library preparation, each sample was run on four lanes of a 150 cycle mid-output run to generate 150-bp paired-end reads (main gland posterior lobe, 95,700,337 reads; main gland anterior lobe, 78,592,255 reads; accessory gland, 54,480,493 reads).

For each gland region, eight assemblies were constructed using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) and Trinity (35). For CLC assemblies, reads with *q*-scores below 30 were trimmed and the reads assembled using minimum contig length of 150 bp, minimum similarity to join contig 0.95, and word (*k*-mer) sizes of 21, 24, 29, 34, 44, 54, and 64. For Trinity, which employs a fixed *k*-mer method, reads were assembled using the default trimming parameters and minimum contig length of 150 bp. For each gland compartment, contigs from the Trinity and CLC assemblies were pooled and clustered using CD-HIT (36; threshold 95%) and then re-imported into CLC Genomics Workbench where trimmed reads were re-mapped and used to update the final contigs. Contigs from all compartments were then pooled, and a total of 149,776 open reading

frames (ORFs) of 90 bp and greater were extracted using GetORF (37), to which 155 common LC-MS/MS contaminant sequences were added to produce the final sequence database for searching.

**Mass Spectrometry**—For MS analysis, venom was centrifuged (12,000 rcf, 4 °C) to remove particulate matter and prepared for MS from crude venom or fractionated using 1D and 2D gel electrophoresis or HPLC. For 1D electrophoresis, 1, 2, 4, 8, or 16  $\mu\text{g}$  ( $A_{280}$  eq) of venom was run on a 12.5% Tris-glycine polyacrylamide gel after denaturation (5 min 95 °C) in sample buffer with or without the reducing agent dithiothreitol (DTT). Gels were stained with Coomassie Brilliant Blue R-250. For 2D electrophoresis, 400  $\mu\text{g}$  ( $A_{280}$  eq) of venom was lyophilized and re-suspended in 125  $\mu\text{l}$  of resuspension buffer (8 M urea, 4% CHAPS) supplemented with 10 mM DTT and 1% ampholyte solution (Life Sciences catalog no. 17-6000-88; GE Healthcare, Little Chalfont, UK) and then absorbed into ReadyStrip immobilized pH gradient strip (pH 3–10, nonlinear, Life Sciences catalog no. 163-2005) for 16 h. Isoelectric focusing was then conducted under mineral oil on an Ettan IPGphor3 isoelectric focuser (GE Healthcare) using a program of 5900 total V-h. The strip was then incubated in reducing reagent (50 mM Tris-HCl, pH 8.8, 6 M urea, 2% SDS, 30% glycerol, and 1.5% DTT) for 10 min and then alkylation reagent (50 mM Tris-HCl, pH 8.8, 6 M urea, 2% SDS, 30% glycerol, and 2% iodoacetamide) for 20 min. Second dimension SDS-PAGE was then run using a 15% Tris-glycine polyacrylamide gel alongside a Precision Plus ladder (catalog no. 161037, Bio-Rad, Hercules, CA) and the gel stained with colloidal Coomassie stain (38). Protein bands and spots were then excised with a scalpel, incubated for 2 h at 37 °C in reduction/alkylation buffer (50 mM ammonium carbonate, pH 11.0, 1% iodoethanol, 0.025% triethylphosphine in 48.5% acetonitrile), and then incubated for 4–16 h at 37 °C in 5–20  $\mu\text{l}$  of digestion reagent (20 ng/ $\mu\text{l}$  trypsin (catalog no. 7575, Sigma-Aldrich, St. Louis, MO; cleaves on carboxylic side of Arg and Lys residues) in 40 mM ammonium bicarbonate, pH 8.0, 5% acetonitrile). Samples were then incubated for 30 min at room temperature in extraction reagent (50% acetonitrile, 5% formic acid) with occasional vortexing.

For HPLC,  $\sim 1$  mg of venom ( $A_{280}$  eq) was diluted  $>5$ -fold in loading buffer consisting of 95% solvent A (0.05% TFA) and 5% solvent B (0.043% TFA, 90% acetonitrile). After centrifugation (10 min, 17,000 rcf, 4 °C), the supernatant was fractionated using a Jupiter 250  $\times$  10-mm Jupiter C4 column (10- $\mu\text{m}$  particle size, 300 Å pore size, catalog no. 00G-4168-NO, Phenomenex, Torrance, CA) using a linear gradient from 5 to 75% solvent B in solvent A over 56 min and a flow rate of 3 ml/min, yielding 64 fractions. After lyophilization and resuspension in 30  $\mu\text{l}$  of milliQ water, each fraction was analyzed by MALDI-TOF MS and LC-MS/MS. For MALDI-TOF MS, each fraction was diluted in MALDI solvent (70% acetonitrile, 1% formic acid), spotted together with the same volume of  $\alpha$ -cyano-4-hydroxycinnamic acid (5 mg/ml in MALDI solvent), and analyzed on a 4700 MALDI TOF/TOF Proteomics Analyzer (AB SCIEX, Washington, D.C.) operated in reflectron mode with a laser power of 3400–3800 V. For LC-MS/MS preparation, 3  $\mu\text{l}$  of each resuspended HPLC fraction was incubated with 7  $\mu\text{l}$  of reduction/alkylation buffer (2 h, 37 °C), lyophilized, and incubated in 5  $\mu\text{l}$  of digestion reagent (16 h, 37 °C) before the reaction was terminated by addition of 20  $\mu\text{l}$  of 5% formic acid. For crude venom samples, 5  $\mu\text{g}$  of protein ( $A_{280}$  eq) was processed using the same protocol as for HPLC fractions, except terminated at the appropriate stage to produce samples that were either native, reduced, alkylated but undigested or reduced, alkylated, and digested.

Peptide digests from gel spots and bands, HPLC fractions, and crude venom preparations were resuspended in 1% formic acid, 2.5% acetonitrile and analyzed by LC-MS/MS. Liquid chromatography was performed using either a Nexera X2 LC system (Shimadzu, Kyoto, Japan) with a 100  $\times$  2.1-mm Zorbax 300SB-C18 column (1.8

$\mu\text{m}$  particle size, 300-Å pore size, catalog no. 858750–902, Agilent, Santa Clara, CA) or a Shimadzu Nano LC system coupled to a 150  $\times$  0.1-mm Zorbax 300SB-C18 column (3.5  $\mu\text{m}$  particle size, 300 Å pore size, Agilent catalog no. 5065-9910). The LC outflow was coupled to a 5600 Triple TOF mass spectrometer (AB SCIEX) equipped with a Turbo V ion source. Peptides were eluted over 14- or 25-min gradients of 1–40% solvent B (90% acetonitrile, 0.1% formic acid) in solvent A (0.1% formic acid) at a flow rate of 0.2 ml/min. MS<sup>1</sup> scans were collected between 350 and 1800  $m/z$ , and precursor ions in the range  $m/z$  350–1500 with charge +2 to +5 and signal  $>100$  counts/s selected for analysis, excluding isotopes within 2 Da. MS/MS scans were acquired at with an accumulation time of 250 ms and a cycle time of 4 s. The “Rolling collision energy” option was selected in Analyst, allowing collision energy to be varied dynamically based on  $m/z$  and  $z$  of the precursor ion. Up to 20 similar MS/MS spectra were pooled from precursor ions differing by less than 0.1 Da. The resulting mass spectra in WIFF format were then compared with a library of ORFs extracted from transcriptomes generated from RNA-Seq experiments (together with a list of common MS contaminants) using a Paragon 4.0.0.0 algorithm implemented in ProteinPilot 4.0.8085 software (AB SCIEX). A mass tolerance of 50 mDa, which is ProteinPilot default value for data obtained on an AB SCIEX 5600 mass spectrometer, was used for both precursor and MS/MS ions. MS contaminants considered are included as a FASTA file “MS Contaminants.fa” in the PRIDE submission associated with this study (PXD004804). For all searches, the options “Biological modifications,” “Amino acid substitutions,” and “Thorough ID” were selected. The Biological modifications option allows detection of 232 modifications in addition to the 59 default modifications considered by the Paragon algorithm in ProteinPilot, which are described in the file “Modifications Catalogue and Translations.xls” in the same PRIDE submission. The Amino acid substitutions option allows detection of substitutions of all standard amino acids. In addition to searching predicted cleavage products of protein sequences against precursor ions, the Thorough ID option runs an algorithm to match sequence tag information to amino acid sequences independently of expected cleavage sites (34). For the 2D gel spot search, the option “Gel based ID” was also selected, which computationally prioritizes oxidative modifications that are common artifacts of electrophoresis. The results of the pooled-sample Paragon searches were reviewed manually. Low quality identifications were excluded by only reporting proteins for which three or more peptides were observed ( $p > 95\%$ ) or one or more peptides with  $p > 95\%$  plus a secretion signal sequence with a  $D$ -score  $> 0.65$  according to SignalP 4.1 (39). FDR analyses were generated by ProteinPilot default method, which employs a decoy database (40).

**Sequence Analysis**—Proteins identified by MS were compiled in a spreadsheet and annotated using BlastP results with  $E < 0.05$  against GenBank™ nr database, HMMER domain predictions with  $E < 0.05$  against the Pfam database, signal sequence probability, and mass of predicted mature toxin (supplemental Table 1). For Ptu1 and lipocalin/triabin family proteins, further family members not identified by MS were recovered from transcriptomes using MS-identified Ptu1 or lipocalin/triabin family proteins as bait for BLAST searches with  $E < 0.05$  against the whole-gland database of possible protein sequences. For phylogenetic analysis of protein families, the most closely related proteins to those from *P. plagiipennis* transcriptomes were retrieved from the GenBank™ nr database using BlastP with  $E < 0.05$ . Further lipocalin/triabin protein family members from blood-sucking reduviids for which functional data has been published were retrieved manually. Proteins were aligned using MAFFT (41). Protein phylogenies were constructed using Bayesian inference in MrBayes 3.2.1 (42) using Iset rates = invgamma with prset aamodelpr = mixed.



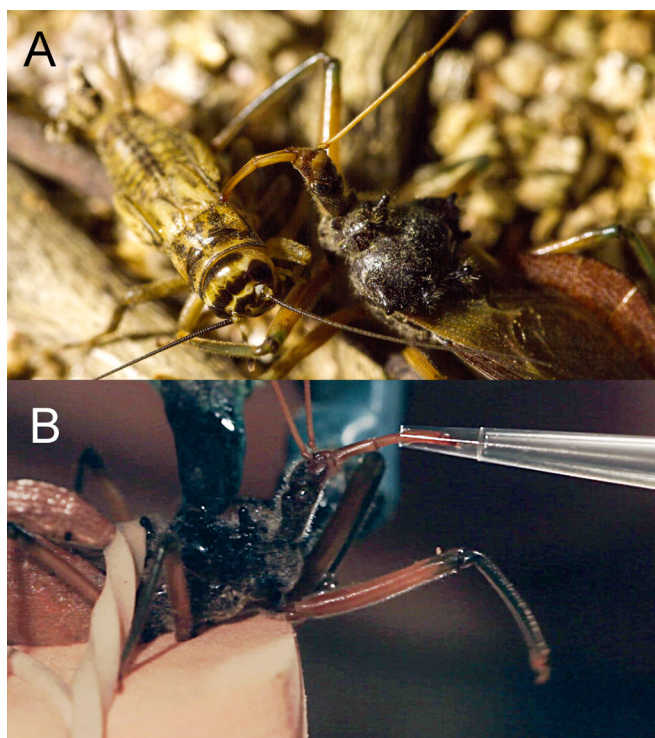


FIG. 1. A, assassin bug *P. plagipennis* envenomating feeder cricket (*A. domestica*). B, extraction of venom from *P. plagipennis* using electrostimulation.

## RESULTS

**Determination of Venom Proteome Using RNA Sequencing and Proteomics**—Assassin bugs (*P. plagipennis*) kept in captivity readily envenomated and killed house crickets (*A. domestica*; Fig. 1A; supplemental Video 1). We developed a method to harvest venom non-lethally from *P. plagipennis* by electrostimulation, which typically yielded 2–10  $\mu$ l of venom of ~50–250 mg/ml ( $A_{280}$  eq; Fig. 1B; supplemental Video 2).

To investigate the protein composition of *P. plagipennis* venom, we constructed a library of possible protein sequences produced by the venom gland complex. Because the glandular source of assassin bug venom has not been characterized in detail, we sequenced mRNA from each of the three compartments of the labial gland complex, comprising the main gland posterior lobe, main gland anterior lobe, and accessory gland. The resulting reads were assembled and ORF translations extracted to produce a library of possible protein sequences (see “Experimental Procedures”).

Mass spectra for protein identification were obtained by LC-MS/MS analysis of 277 protein samples derived from venom, comprising crude venom (5), 1D SDS-polyacrylamide gel bands (54), 2D SDS-PAGE spots (156; Fig. 2A), and HPLC fractions (62; Fig. 2B). A full list of samples analyzed (file “Sample numbering.xlsx”), along with raw data files and protein and peptide summaries, is included in the PRIDE submission accompanying this study (PXD004804). To determine the venom proteome of *P. plagipennis*, spectra from

multiple samples were combined into two Paragon searches against our ORF database “Pp123x.fasta”. One search “all\_iodoethanol\_samples” included venom crudes, HPLC fractions, and 1D gel bands, while the other search “2Dgelspots” included all 2D SDS-PAGE spots. Although it would have been ideal to combine all samples into a single Paragon search, two searches were performed because of the differing alkylating agents used for these two sample sets. The All\_iodoethanol\_samples and 2Dgelspots searches resulted in 175 and 174 protein identifications, respectively, at a protein confidence level of >95%. Removal of low quality identifications, decoy database hits and contaminants, and manual comparison of the results from these two searches resulted in a draft proteome consisting of 130 protein sequences. In some cases, identified protein sequences that appeared to be incomplete were compared by BLAST search against individual RNA-Seq assemblies to retrieve additional sequences. Final identification statistics were produced by Paragon search of each of the All\_iodoethanol\_samples and 2Dgelspots against a database containing the draft proteome plus contaminants (“Pp\_electrostim.fasta”). Three protein sequences were further discarded at this stage due to low ProteinPilot Unused values (indicating they were superfluous), resulting in a final proteome of 127 sequences (supplemental Table S1). Of these 127, four were identified from a single detected peptide, spectra for which are shown in supplemental Fig. S1. Re-examination of FDR analyses indicated these 127 sequences are ranked within the region of <1% global FDR in searches against the ORF database.

Protein families represented in the venom included disulfide-rich peptides similar to Ptu1 from *Peirates turpis* (28), CUB domain containing proteins, cystatins, homologues of trialyisin (a pore-forming toxin previously characterized from triatomine reduviids (43)), trypsin-like proteases, various catabolic enzymes, serpins, a triabin-like protein, bacterial permeability increasing-like protein, and novel protein families (Table 1; Figure 3). Putative enzymes detected in the venom were protein kinase, inositol-phosphate phosphatase, M12A-like metalloproteases, cathepsin B, peptidase S10, hexosaminidase, and nuclease. Seventeen proteins without any inferred putative function were present and were classified into eight families (heteropteran venom protein families 1–8). Proteins in families 3 and 5–8 showed homology to uncharacterized and predicted protein sequences from hemipterans and other insects, whereas families 1, 2, and 4 were novel.

To gain additional information about protein abundance, we then re-analyzed mass spectra from individual 2D gel spots (Fig. 2A), HPLC fractions (Fig. 2B), and other samples, comparing them using the Paragon algorithm to our venom proteome database Pp\_electrostim. Many of the individual samples, including low molecular weight gel spots and HPLC fractions, contained peptides from multiple larger proteins, suggesting venom underwent partial autoproteolysis before analysis, consistent with the high protease content of the

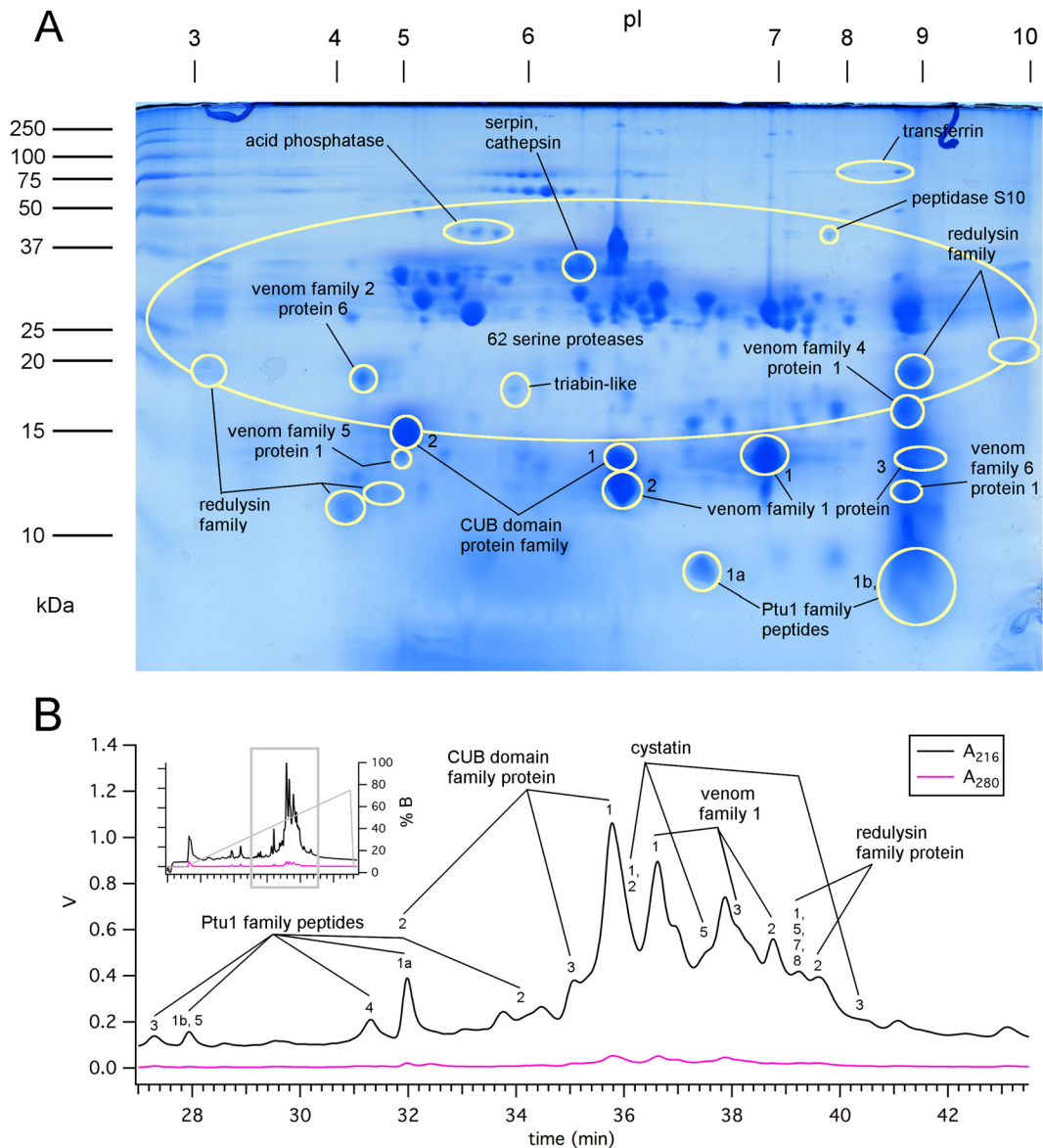


FIG. 2. Proteins detected by LC-MS/MS of 2D SDS-PAGE spots and HPLC fractions showing abundant proteases, CUB-domain proteins, and heteropteran venom family 1 proteins. A, 2D SDS-polyacrylamide gel of crude venom, showing protein families identified by LC-MS/MS of gel spots. B, HPLC trace of venom fractionation, showing protein families identified by LC-MS/MS of collected fractions.

venom. Nevertheless, it was possible in many cases to assign spots and fractions as particular proteins based on the number of peptides confidently detected (>95% confidence), the number of precursor signals counted as a proportion of total non-contaminant precursor counts, and (in the case of gel spots) the observed *versus* expected molecular weight. S1 proteases alone accounted for 65 of the 127 identified proteins, 62 of which were identified from spots in the 25–40 kDa range of a 2D SDS-polyacrylamide gel. Aside from proteases, the major venom proteins present are in the 10–16 kDa range and of unknown function. Intensely staining gel spot in this range (spot numbers 104, 62, 109, and 103) were attributed to CUB domain proteins 1 and 2 and venom family 1 proteins 1

and 2, respectively, each of which accounted for >90% of confidently identified peptides and >96% of precursor counts in each case (supplemental Table 2). In addition, each of these proteins is associated with a major peak in the HPLC trace (fractions 50, 44, 52, and 56, respectively). In each case, the identified protein accounts for >69% of precursor counts detected (with the exception of CUB domain protein 1, which accounted for 34% of precursor counts from the fraction collected at the highest UV peak). Further details of detected proteins are available as supplemental data (protein and nucleotide sequences, identification statistics, and annotation, see supplemental Table S1; alignments of selected protein families, see supplemental Fig. S2; files included in the PRIDE

TABLE I  
Venom proteome of *P. plagiennis*

Protein family/homology	No. of related sequences	Mass range	Putative function
		<i>kDa</i>	
Ptu1 family peptides	5	3–5	Neurotoxin
Cystatin	7	9–12	Protease inhibition/other
Venom families 4–8	6	9–17	Unknown
Venom family 3	2	12–13	Unknown
CUB domain protein	5	12–14	Unknown
Odorant-binding protein	1	13.2	Unknown
Venom family 1	3	14–15	Unknown
Triabin-like	1	19.9	Unknown
Venom family 2	7	22–29	Unknown
Redulyisin	9	24–50	Cytolysis
Protease M12A	2	27–30	Proteolysis
S1 protease	46	29–60	Proteolysis
Fibrillin	1	30	Unknown
Cathepsin B	1	35	Proteolysis
Inositol-phosphate phosphatase	1	35.1	Lipid catabolism
Lipase	2	36.8	Lipid catabolism
Histidine phosphatase	1	40.4	Phosphatase
Serpin	1	42.2	Protease inhibition
Nuclease	1	44	Nucleic acid catabolism
S1 protease + CUB	19	45–70	Proteolysis
Protein kinase	1	47.4	Unknown
Peptidase S10	1	50.6	Proteolysis
Hexosaminidase	1	67.8	Biopolymer catabolism
Transferrin	1	71.1	Nutrient dissemination
Bacterial permeability increasing	1	Not complete	Cytolysis
Total	127	3–72	Paralysis and liquefaction

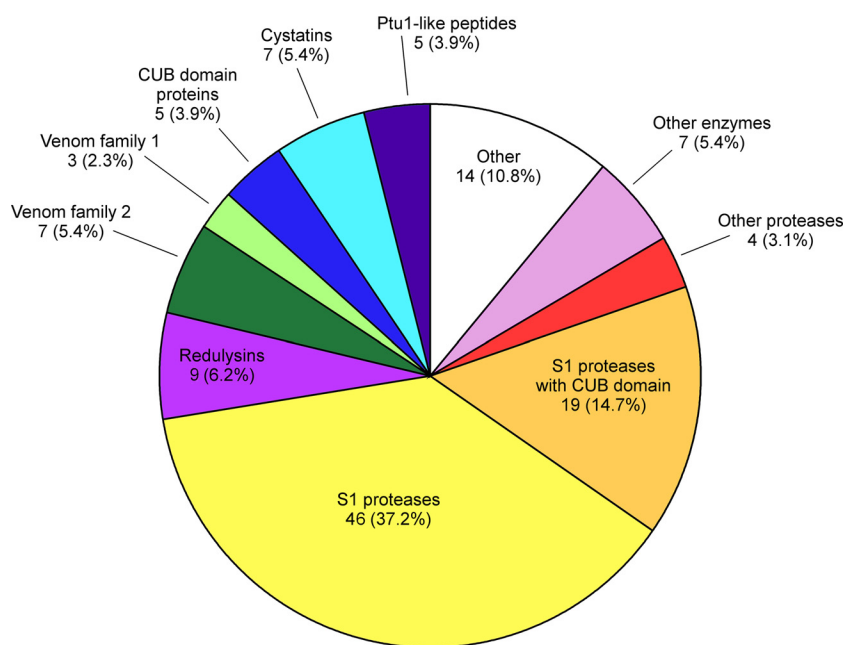


FIG. 3. Proportion of sequences belonging to each major protein class in the venom of *P. plagiennis*.

submission accompanying this study, see [supplemental Table S2](#); biological modifications, amino acid substitutions, and deviation from expected cleavage patterns detected may be

most conveniently viewed by opening ProteinPilot Peptide Summary files in the PRIDE submission using Microsoft Excel).



**Ptu1 Family Proteins**—We detected five Ptu1-like peptides in *P. plagipennis* venom using LC-MS/MS (Pp1a, Pp1b, Pp2, Pp4, and Pp5). To determine if transcripts encoding further members of the Ptu1 family are produced in the labial gland complex, we performed a BlastP search against our library of possible protein sequences using *P. plagipennis* Ptu1 venom peptides and previously described assassin bug Ptu1 family venom peptides as queries. This strategy revealed further 10 Ptu1 family peptide sequences (KX752811 to KX752820). One of these, Pp3, was not identified in LC-MS/MS experiments but has a predicted mature mass (3581.4 Da) closely matching an observed venom component (3581.0 Da) that undergoes a mass shift of 270.7 Da upon alkylation with iodoethanol, close to the theoretical value expected for alkylation of the six Cys residues in the Pp3 sequence (270.4 Da). Pp3 was therefore classified as a putative venom peptide in further analyses. All identified Ptu1 family peptides have six conserved Cys residues that are homologous to those responsible for formation of the ICK fold of Ptu1 (Fig. 4A). Based on the  $A_{280}$  of crude venom and the quantity of peptides recovered after HPLC (judged by  $A_{280}$  calibrated with calculated extinction coefficients from identified peptide sequences), Ptu1-like peptides were estimated to account for ~1–3% of venom by dry weight.

To investigate the relationship between *P. plagipennis* Ptu1 family sequences and previously described proteins, we aligned them together with homologous sequences from the nr database retrieved by a BlastP search ( $E < 0.05$ ; Fig. 4A). We then performed a Bayesian inference of phylogeny using a mixed-evolution model and rooted the tree with a putative ICK-forming venom peptide from Remipedia, the sister group to Hexapoda (44). According to this analysis (Fig. 4B), Pp1a is monophyletic with the two previously described harpactorine assassin bug venom peptides, Ado1 and lob1 (posterior probability,  $p = 0.92$ ). Pp6a was more closely related to the bee peptide OCLP1 ( $p = 0.99$ ), a non-venom peptide expressed throughout the body, and to related sequences from ants (45). Another clade ( $p = 0.73$ ) is formed by Pp8 and Pp10–Pp13 and includes a putative antimicrobial peptide from a phytophagous hemipteran, the whitefly *Bemisia tabaci* (GenBank™ accession number O81338.1). The relationships between the remaining peptides, including all peptides that were detected in venom by MS apart from Pp1, were poorly resolved. Overall, these data are consistent with Ptu1/OCLP1 family peptides being widespread among insects where they perform a non-venom role, for example in the immune system or as phenol oxidase inhibitors, with independent recruitment as venom peptides in the orders Hemiptera and Hymenoptera.

Fractions from which Ptu1 family peptides were detected by LC-MS/MS were examined using MALDI-TOF MS to determine the masses of mature Ptu1 family peptides. For Pp1a and Pp3, the major mass detected in the venom fraction matched perfectly with that calculated for the mature toxin sequence predicted from the transcriptomic data, after re-

moval of the 20–23-residue signal peptides predicted by SignalP (supplemental Fig. S3). This suggests that post-translational processing of Ptu1-like venom peptides includes neither removal of a propeptide sequence nor enzymatic modification of mature toxin residues as is often the case for the venom peptides from cone snails and to a lesser extent spiders (46). Instead, their biosynthesis more closely resembles that of related peptides in hymenopterans (45).

**CUB Family Proteins**—CUB domains occur widely in multidomain proteins where they perform roles in protein recognition and as adaptor domains, and they are especially well represented in extracellular and developmentally regulated proteins (47). CUB domains also occur as one of several domains in some insect and crustacean S1 proteases, including the bee venom protease Api m 7 (48–50), and 19 of the 67 S1 proteases detected in *P. plagipennis* venom. In contrast to this pattern, we observed that two of the most abundant proteins (CUB domain proteins 1 and 2) in *P. plagipennis* venom consist of a solitary CUB domain; three further CUB domain proteins (3–5) were detected from HPLC fractions and 1D SDS-polyacrylamide gels. Each CUB domain protein contains 129–142 residues and is stabilized by two conserved disulfide bonds (Fig. 5A). To rule out the possibility that detection of CUB domain proteins is an artifact of proteolytic cleavage of CUB domains from S1 proteases and/or incomplete or wrongly assembled contigs, we examined each CUB domain protein coding sequence. Each contig encoded a CUB domain protein that included multiple stop codons in all reading frames downstream of the stop codon that terminates translation, without additional open reading frames homologous to S1 proteases or other proteins. To further rule out assembly errors, we examined >10 assemblies produced using either CLC Genomics Workbench or Trinity (51), and we found that all five contigs encoding CUB-only domains were present and complete in every assembly. In addition, CUB domain proteins 1 and 2 were very confidently identified from 2D gel spots with an apparent molecular mass that closely matched the predicted mature mass of these proteins (12.6–12.7 kDa). Thus, our data strongly suggest that proteins encoding single CUB domains are abundant components of the venom of *P. plagipennis*. A CUB domain only protein has previously been reported from a transcriptomic study of the venom glands of another cimicomorphan heteropteran, the minute pirate bug *Orius laevigatus* (Anthocoridae; Ref. 52), suggesting CUB domain proteins may be present in the venoms of phylogenetically diverse heteropterans.

According to BlastP searches, CUB domain sequences were most similar to N-terminal CUB domains of S1 proteases from insects and crustaceans. We examined the phylogenetic relationships between reduviid and anthocorid CUB domain proteins and arthropod CUB-S1 proteases using Bayesian inference of phylogeny, rooting the tree a CUB-S1 protease from the olfactory organ of the spiny lobster *Panulirus argus* (Fig. 5B). According to this analysis, heteropteran CUB do-

A

Remipede agatoxin-like	mkeanrrryiylclvvllsiiitte-----aeddrLFCIRHHKYGGHRP-----KRCCSGLF-CRCN--TFGTNCRQOGKALGKLISGK-----
Assassin bug Ado1	-----ADDD--CLPRGSKCLGEN-----KQCCGKTT-CM-----FYANRCVGV-----
Pp1a	mkifg---iifvlvtfialal-----aGDED--CLPRGSKCLGEN-----KQCCGKTT-CM-----FYANRCVGI-----
Assassin bug Iob1	-----aGDED--CLPRGSKCLGEN-----KQCCCKTT-CM-----FYANRCVGI-----
Pp2a	mk-----wifliviyllynvsaR---VFESRQ--CIPRLEKCVGGQ---AQCCPPSH-CL-----WYANKCI-----
Assassin bug Ptul	-----AEKD--CIAPGAPCGFTD---KPCCNPRAWCS-----SYANKCL-----
Pp3a	mkfsn---iislifivvllall-----dnatsDD--CIKRYGKCSKSN---SNCEPAQ-CY-----FSFNQCF-----
Pp9a	mkq-----lllaificflisc-cygVGFDFVDNDRM---CTPVGKQVMVD---FQCCAPYT-CD-----WLQNKQ-----
Pp6a	mk1----fgillfvfacial-vfa-----DTQG--CGRHGDPD-ND---AQCCGTVK-CH-----RYAKRCQVQLS---LPPRVD-----
Bee OCLP1	msk-----fillvcillltn-ivs-----aASK--CGRHGDSGV-SS---SDCCPGTW-CH-----TYANRCQVRIITEELMKQREKILGRKGKDY-----
Ant predicted	mak-----lmlyvfvvllaaslimg-----ATDK--CGRHGDPGV-SD---SQCCGTGIR-CH-----RYANRCQVRIITEELMAQREKILGRKGKDY-----
Ant conotoxin-like	mkf-----viftifvmiantm-----vasEEK--CSPLGQYCT-QY---SDCCRYTQ-CL-----TYAAKCVTKAGI--LVPGEDKRPLGPG-PYPPNMPE-----
Pp4a	mk1----ilflvfvialmaltvqs-----KEND--CIPGFQCGGAN---SRCCGLYV-CF-----NNRCIPTP-----
Pp5a	mk-----yllilflflacfdalyc-----AAPN--CIMPGNPGRGIN---ARCCSFNV-CV-----NGRCLAMPKR-----
Fly POI	mtamk---iitvllivalvsfaviss-----vesQ--CLANGSKCY-SH---DVCTTKR-CH-----NYAKKQVT-----
Pp8a	mavmk---tiavllivalvsfaviss-----vesQ--CLPLGASCNRLSTYP-KGCCSLTV-CS-----WDSSTCVPIKIN---IKN-----
Pp11a	mk-----tiavllivaivsftviss-----vesQ--CLPLGASCNRLDPIPRRRCSSMMV-CS-----WNTSTCAPPTYI---IK-----
Pp10a	matik---timlllvvalvlsfadada---yaRAQ--CLGAGQPCNRL--P-YQCCPMT-CD-----FFRGSTCQMN-----
Pp12a	mk-----y111flivtistvtfaA---VAND--CLPDGARGHHP---EQCCGGF-CT--NGMINGWGNHCEGVP-----
Pp7a	mr-----msikcfavivafafflg---qisraea--CIPDGGRCHESDP-G-PGCCSGF-CYRE--RNWKDGDCKRRP-----
Fly saliva AMP1	mh-----kylawfifavatlwmiag---smaQGN--CLPDGAPQADGS-M-GNCCFTF--CLQHEQPGGTPGHCTTRS-----
Hemiptera AMP	malpr---vilalfvvlvftfiefy-----saaQ--CYPENFPCNPGG--R-NYCCAPHV-CLLH-----RGIYVCRPES-----
Pp13a	

B

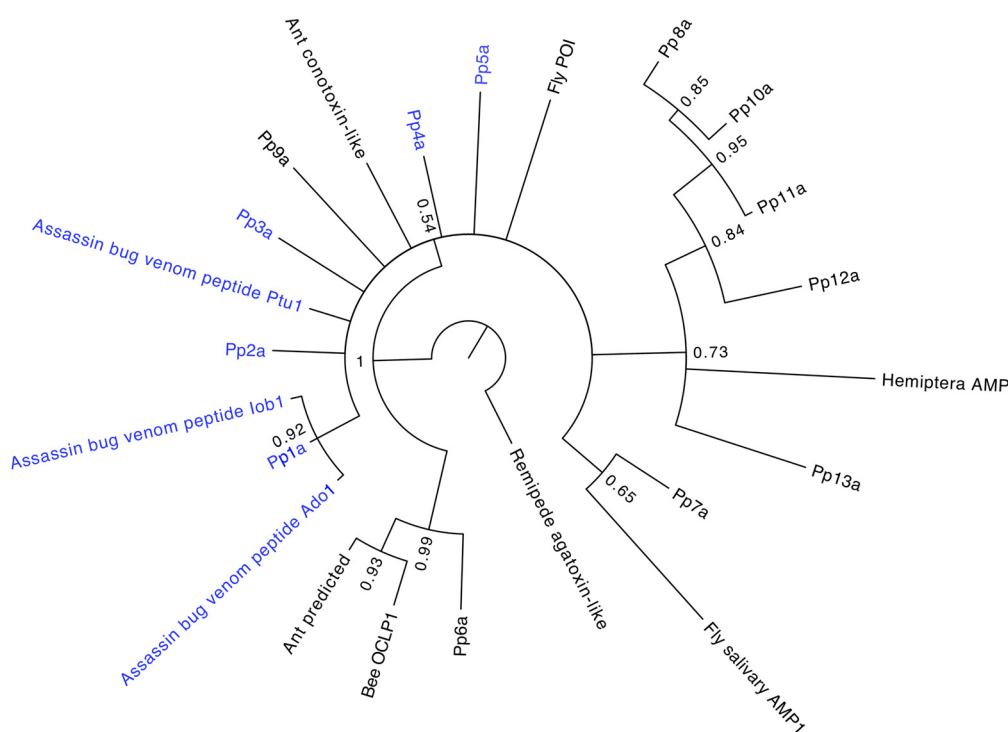


FIG. 4. Evolution of Ptul1-like peptides isolated from *P. plagiipennis* venom. Signal sequences predicted by SignalP4.1 are shown in lowercase; lines above the text show the disulfide connectivity for Ptul1. Sequence labels in blue indicate peptides detected in venom using proteomics. POI, phenoloxidase inhibitor; AMP, antimicrobial peptide. Node labels indicate posterior probabilities. Comparison sequences are remipede (*S. tulumensis*) agatoxin-like peptide (66), assassin bug venom peptides Ado1 (P58608.1), lob1 (P58609.1), and Ptul1 (P58606.1); bee (*Apis mellifera*) OCLP1 (H9KQJ7.1); ant (*Trachymyrmex cornetzi*) predicted protein (KYN16117.1); ant (*Wasmannia auropunctata*) conotoxin-like peptide (XP\_011706926.1); fly (*Bradysia hygida*) salivary putative AMP1 (ABA26621.1); and Hemiptera (*Bemisia tabaci*) putative AMP knottin peptide Btk-1 (ABC40569.1).

main venom proteins diverged anciently from a heteropteran CUB-S1 protease but are likely to be monophyletic ( $p = 0.75$ ). This result suggests that CUB domain proteins were recruited into venom prior to the split of Reduviidae and Anthocoridae ~190 mya (53). Their conservation over this extended period of time and the high abundance in the venom of *P. plagiipennis* suggest that they serve an important but as yet unknown functions.

**Redulysins**—Eight venom proteins showed homology to trialysin, a previously described Lys-rich cytolytic toxin isolated from venom of the blood-feeding triatomine reduviid *Triatoma infestans*. Because blood-feeding reduviids do not lyse red blood cells during feeding, it has been suggested that trialysin has an antimicrobial function that protects the venom gland from colonization by parasites (43). Regardless of its biological function, the cytolytic activity of trialysin is well



A

```

Anthocorid CUB protein m-lwlgvlllvqlal-gDAPVK-----KLEVPVGKEVVISSASYPTNSPPG
CUB domain protein 2 m-raiylclvlfiaa-vaadvq-----KLEVPPGKEVDIESENYPEDANAN
CUB domain protein 3 m-nsayllvvlvlgaa-taAITEH-----RITVKLSVPKGASVQLVSLDYPNIPAAAG
CUB domain protein 4 mklllavvllatlaial-----aEK-----RKVR-TSAFLGSMPI---NNFLRQWHPNGFPAAE
CUB domain protein 5 mkvllgl-iltctcalygsvnaNN---YEVHNATAIVRGLPL---RRFFDRWNPKGLLPPF
CUB-S1 protease 8 m-llfglfillalev-ancQQKSEdGVHINRRKKLkLYEPP---VVDIQNLNYPSEAPAA
* : : : * : * .

Anthocorid CUB protein THHVWELETDSdGQILAVCEVIQINSSRkCRDGFldVSD-NAMGKLCGGLRDFELTSQGN
CUB domain protein 2 SDLKWELSTTSdGHVINLKCSDIRLFEeKPCGPWVLRITdGDETEELCTTKFDYEkKFKTN
CUB domain protein 3 TTNTWYLTSEEGSKIELTCDdIWVAPSKLCERGWFkVAVGgKTDKfCEMRPGYTATSESN
CUB domain protein 1 NDFIYSLATEEGRKVELTcASIdI--EePCTENvFKLIIdGHSEtICGQkKDFVYkSKTh
CUB domain protein 4 TEFVYTVITdPDkTLEIHcPYISL--HQPCeEYSFDIvDGDkKkRICGkVKKFFYHSLITN
CUB domain protein 5 TEFVYTVITdPDkKIRIDcPLINL--HLPChQYSFDIIdTGPRIRICGkVVERVYYSNSE
CUB-S1 protease 8 LIQSWTITChYWLkLRFtCLIRIPENSPcEKYYVEVVKDGDkMKRYcGEVkdLLILSSS
: : : * * : * . : . *

Anthocorid CUB protein TMTVNFkLS-RV--PGHFkCTAKGIKkESDEE*-----144
CUB domain protein 2 KILVNFKTGPEA--RGFLQCTATAVKK*-----129
CUB domain protein 3 EMTVQIDLN-NT--TGRARCTAKAV*-----131
CUB domain protein 4 ILSLQLLsgANIi-KTTPeCSAKPID*-----130
CUB domain protein 1 KVDLMIRTGSKLV-KVAPrCTARPLKKN*-----137
CUB domain protein 5 KVQIIvRTGPIlV-PHPPTCSATVYKRARRY*-----141
CUB-S1 protease 8 QLEVTFkTG-KLYGGGLMNCNVeGFNPeFTTLAPPTTKSYLDID-->>(continues)
: : . *..
    
```

B

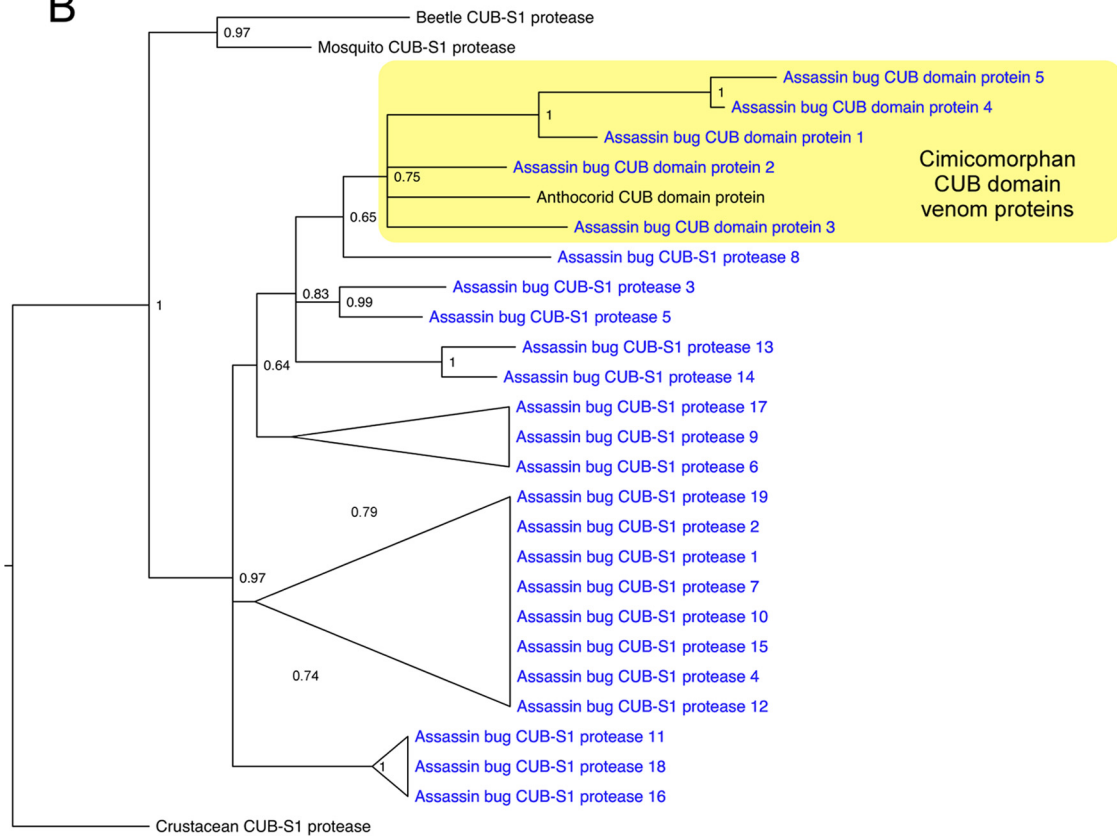


FIG. 5. Evolution of CUB domain proteins. Sequence names in blue indicate *P. plagiipennis* sequences determined in this study. A, alignment of CUB domain venom proteins from *P. plagiipennis* and the anthocorid *O. laevigatus* (52) with the N-terminal portion (including CUB domain) of *P. plagiipennis* venom S1 protease 8. B, phylogeny according to Bayesian inference showing support for monophyly of cimicomorphan venom CUB domain proteins (highlighted yellow). Anthocorid CUB domain protein, GBG01000009.1; crustacean CUB-S1 protease, AAK48894.1; beetle CUB-S1 protease, ENN74674.1; mosquito CUB-S1 protease, XP\_001657965.1.

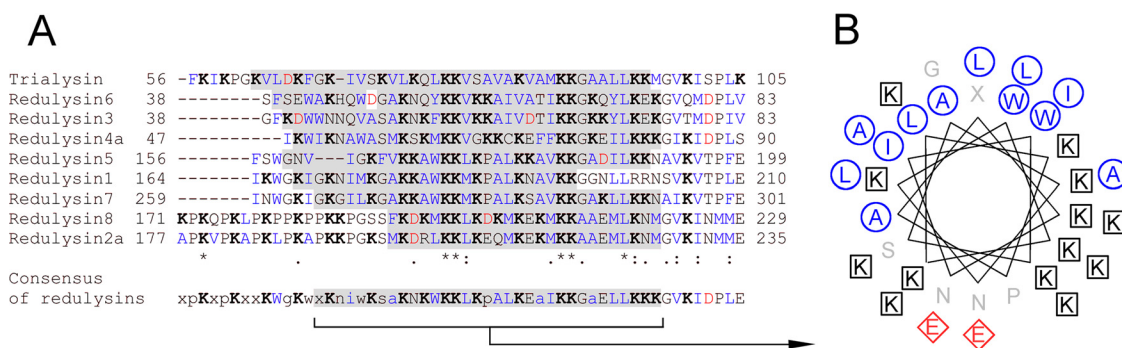


Fig. 6. **Conserved sequence features in redulyisin proteins.** A, amino acid sequence alignment of redulyisins with the cytolytic domain of trialysin (AAL82381.1). The consensus sequence of the redulyisin proteins is shown below. Lys residues are shown in *bold black*, positively charged residues in *red*, and hydrophobic residues in *blue*. Residues predicted by PSIPRED to form  $\alpha$ -helices are highlighted in *gray*. B, helical wheel diagram of the predicted helical portion of the consensus of redulyisin sequences showing clustering of Lys residues and hydrophobic residues on opposite sides of the helix.

established experimentally. Purified trialysin forms voltage-dependent channels in lipid bilayers (43), whereas synthetic peptides corresponding to an internal helical domain of trialysin are capable of lysing mammalian and bacterial cells. Martins *et al.* (54) used fluorophore quenching to demonstrate that trialysin is activated by proteolytic removal of a negatively charged N-terminal motif that exposes the positively charged helical domain. The same researchers used NMR to show that cytolytic peptides derived from the helical domain of trialysin form amphipathic  $\alpha$ -helices with hydrophobic and positively charged Lys residues on opposite sides of the helix, a pattern thought to underlie the ability of trialysin to bind negatively charged phospholipid headgroups and hydrophobic tails and therefore disrupt biological membranes (55). Because all of the eight trialysin homologs in *P. plagiipennis* venom feature a conserved motif homologous to the cytolytic motif of trialysin (Fig. 6), we have classified them as putative cytolytic or pore-forming toxins and named them redulyisins 1–8. The redulyisins are 233–458 residue proteins that are rich in Lys (14–17%) and consist of an N-terminal negatively charged motif, cytolytic motif, and a C-terminal domain stabilized by a pattern of eight conserved Cys residues. The putative cytolytic motif consists of 33 residues predicted to form an  $\alpha$ -helix by PSIPRED, ~40% of which are Lys residues arranged with conserved periodicity (Fig. 6). In contrast to the case for triatomine bugs, there exists a clear biological reason for predaceous reduviids to possess cytolytic toxins, as they might contribute to prey capture, pain induction in defensive envenomation, and/or liquefaction. Moreover, cytolytic and liquefying activities have previously been demonstrated experimentally in reduviid venoms (19, 24). Taken together, the combined data suggest that redulyisins are a cytolytic toxin family present in reduviid venoms, at least one member of which was retained by some triatomine bugs despite their shift to a blood diet 25–30 mya (17).

**Lipocalin/Triabins**—In blood-feeding reduviids, the lipocalin/triabin family has radiated to produce proteins with a wide variety of functions, including inhibition of coagulation factors,

inhibition of platelet aggregation through sequestration of ADP, sequestration of biological amines, and carriage of nitrous oxide (18, 56). In many triatomine bugs, lipocalin/triabin family proteins account for the majority of venom proteins both in number and weight; they also account for the majority of the functionally characterized components of triatomine venom. We identified one lipocalin/triabin family protein in the venom of *P. plagiipennis* by LC-MS/MS, a 178-residue protein with a predicted mature mass of 19.8 kDa. To further recover *P. plagiipennis* lipocalin/triabin family proteins for phylogenetic analysis, we performed a BlastP search of our library for possible protein sequences using a range of lipocalin/triabin family proteins from blood-feeding reduviids as queries. This strategy recovered a further 11 lipocalin/triabin family proteins (KX752800 to KX752810).

The relationships between lipocalin/triabin family proteins of triatomine reduviids have recently been examined in the context of understanding how triabins evolved to occupy their current roles as functionally diverse facilitators of blood-feeding (56). Because blood-feeding triatomines evolved from predaceous assassin bugs, an understanding of how triatomine triabins are related to those of predaceous assassin bugs is desirable but has not been previously possible. To investigate such relationships, we performed Bayesian inference of phylogeny using the 14 *P. plagiipennis* lipocalin/triabins, their closest homologs according to a BlastP search against the GenBank™ nr database and a range of functionally characterized triatomine triabin family proteins (Fig. 7). All triabin family proteins identified from triatomine venom and a subset of *P. plagiipennis* sequences (including triabin-like venom protein 1 identified in venom by LC-MS/MS) clustered together with high probability (Fig. 7, clade labeled “triabin-like proteins”;  $p = 0.92$ ). Another subset of lipocalin/triabin family proteins from *P. plagiipennis*, sequences from heteropterans such as the pentatomomorph *Halyomorpha halys*, the cimicomorph bedbug *Cimex lectularius*, and a representative of the “cockroach triabin” Bla g 4 (56, 57) were found not to be members of the triabin-like protein clade. We suggest that

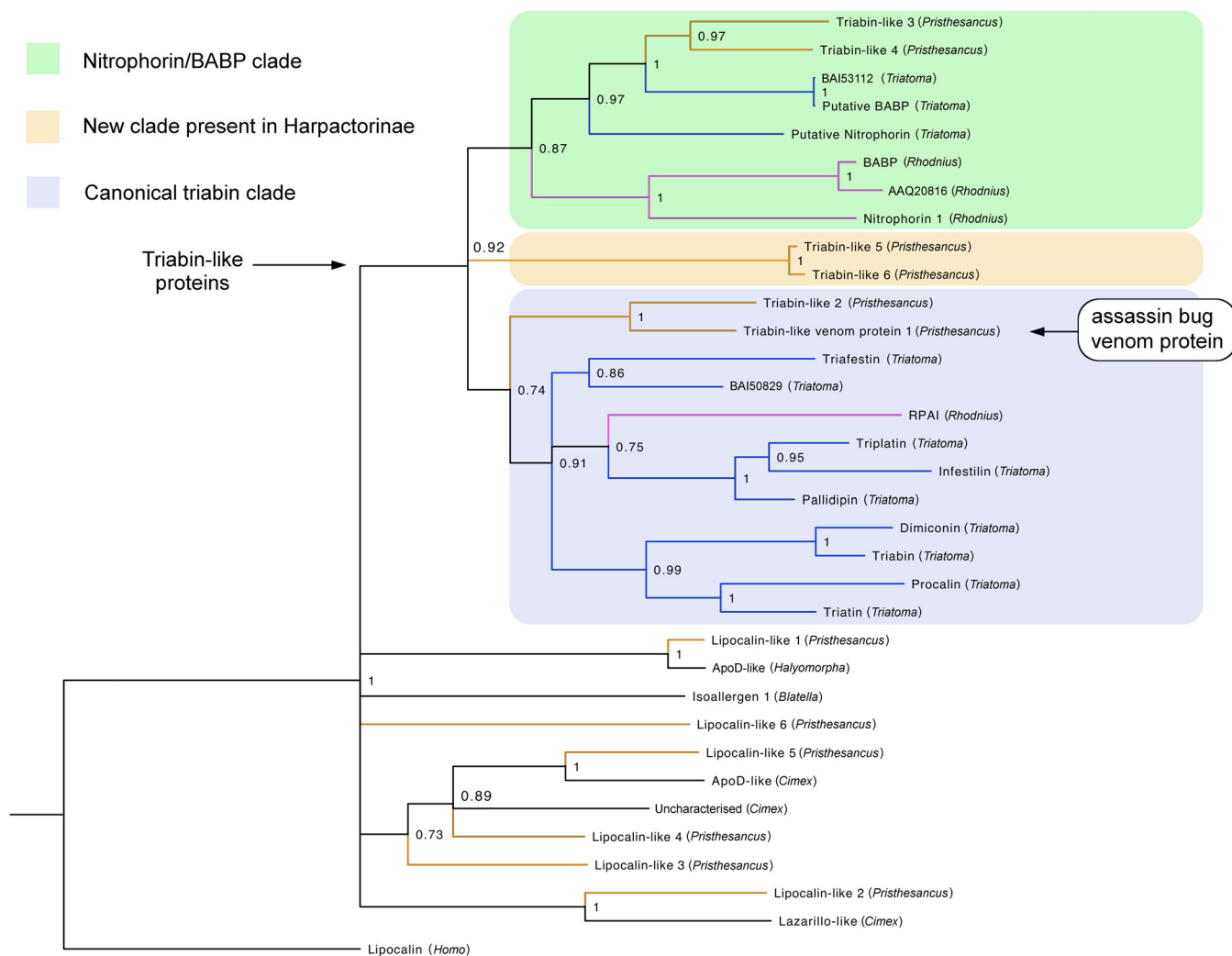


FIG. 7. Radiation of triabin/lipocalin superfamily in Reduviidae and allied insects calculated by Bayesian inference. Branches of *Pristhesancus*, *Triatoma*, and *Rhodnius* proteins are colored orange, blue, and pink, respectively; node labels indicate posterior probabilities. Accessions of sequences used in this analysis are as follows: putative BABP (*Triatoma*), D1MX91; putative nitrophorin (*Triatoma*), A0A023F6B8; Triafestin, AB292809.1; RPAI, JA76747.1; Triplatin, BAE96121.1; Infestilin, AAZ38958.1; Pallidipin, AAA30329.1; Dimiconin, BAI50848.1; Triabin, CAA56540.1; Procalin, AEM97970.1; Triatin, AAZ38956.1; ApoD-like (*Halymorpha*), XP\_014291570.1; Isoallergen 1 (*Blattella*), C3RWZ4; ApoD-like (*Cimex*), XP\_014256007.1; Uncharacterized (*Cimex*), XP\_014250080.1; Lazarillo-like (*Cimex*), XP\_014249890; Lipocalin (*Homo*), CAA47889.1.

membership of this clade, which includes venom proteins from both predaceous and blood-feeding reduviids but not any protein known to have a non-venom function and no non-reduviid proteins, constitutes a natural delineation for the “triabin” family.

According to our analysis, currently known triabins are divided into three clades. The clade highlighted in green (Fig. 7;  $p = 0.87$ ) contains members of the nitrophorin/BABP family, which are most prominent in venom from triatomine bugs from tribe Rhodniini (18) but are also expressed in venom glands of bugs from tribe Triatomi (58) and the harpactorine *P. plagiipennis*. Surprisingly, the nitrophorin/BABP family proteins from the triatomine tribe Triatomi are more closely related ( $p = 0.97$ ) to *P. plagiipennis* proteins rather than nitrophorin/BABP family members from the triatomine tribe Rho-

dniini ( $p = 0.97$ ), suggesting that at least two gene loci encoding members of this family existed in the common ancestor of Harpactorinae and Triatomiinae before they diverged  $\sim 75$  mya (17). In the blue clade (Fig. 7), *P. plagiipennis* triabin-like venom protein 1 and a related *P. plagiipennis* protein form the sister group to all remaining triabin family proteins (excluding the nitrophorin/BABP family) that have so far been documented in triatomine venom. The final clade highlighted in tan (Fig. 7) contains only *P. plagiipennis* proteins. These results illuminate the evolutionary history of the triabin family, suggesting that at least four genes encoding members of the triabin family existed in the common ancestor of Harpactorinae and Triatomiinae and that triabins subsequently radiated in both lineages.



## DISCUSSION

In this study we have provided the first holistic analysis of a venom proteome of a predaceous assassin bug through the combination of transcriptomic and proteomic approaches. Our analysis revealed that *P. plagipennis* produces highly complex venom containing at least 127 peptide and protein components. Many of these venom components have structural and inferred functional similarity to peptides and proteins involved in neurotoxicity (Ptu1 family), membrane disruption and cytolysis (redulysins, bacterial-permeability-increasing peptide), enzymatic catabolism (proteases, phosphatases, nucleases), and nutrient dissemination (transferrin). Thus, the suite of proteins present in the venom of *P. plagipennis* appears well suited to facilitating the dual activities of paralysis and tissue liquefaction previously observed to result from injection of reduviid venom (19).

The venom used in this study was obtained by electrostimulation to avoid contamination by glandular tissue. Venom obtained from other species of the assassin bug by electrostimulation has been shown to have rapidly paralyzing and lethal effects on both vertebrates and invertebrates (22) and to contain neurotoxins (28). Thus, reduviid venom obtained by electrostimulation shows the biological activities expected of prey-capture venom. Although we argue that the venom analyzed in this study is likely to perform dual roles in paralysis and liquefaction, we note that it is possible that the labial gland complex produces more than one kind of secretion specialized for prey capture, defense, or feeding, as suggested previously (59). We are currently conducting studies to establish the glandular origin of venom obtained by electrostimulation and the biological roles of the different parts of the labial gland complex.

The biochemical composition of venom from predaceous reduviids is unique in comparison with venom from blood-feeding reduviids as well as predatory arthropods such as spiders, scorpions, and centipedes. The venoms of most marine and terrestrial invertebrate predators are dominated by highly diverse, often Cys-rich, peptides with a molecular mass less of than 10 kDa (9, 60–63). In contrast, we found that the putatively neurotoxic components in *P. plagipennis* venom, the Ptu1 family of small disulfide-rich peptides, make up a small proportion (1–3%) of venom, whereas the majority of venom proteins are proteases. Possibly, this is because some of the abundant proteins with unknown function, such as CUB domain proteins or venom family 1 proteins, confer the main neurotoxicity of the venom. The high proportion of enzymes compared with peptide neurotoxins may be related to the different feeding biology of assassin bugs compared with other venomous predators. Both spiders and scorpions, like assassin bugs, rely primarily on EOD for feeding (27). However, both spiders and scorpions practice “refluxing” EOD in which enzymes from the gut, and not from the stinger or fangs, are regurgitated onto or into prey to facilitate diges-

tion. In contrast, heteropterans are one of the few arthropods that practice non-refluxing EOD (27). According to the radiolabeling experiments of Cohen (64), the sole source of proteins injected into prey by assassin bugs is the labial glands, not the gut. Snakes and cone snails are similar to assassin bugs in the sense that their organs of envenomation have evolved from oral structures, but both these animals are capable of swallowing prey whole and digesting them internally. Although proteases and other enzymes in the venoms of snakes and spiders have often been proposed to have roles in the digestion of prey, evidence that they are essential for digestion is relatively weak (65). For assassin bugs, which can only ingest liquid food, and which use the same anatomical structure for injecting venom for both prey capture and EOD, the requirement for digestive activity in venom is likely to be much stronger. This fundamental difference in feeding physiology may explain the relative abundance of proteases compared with peptides in assassin bug venom. Supporting this notion, the molecular weight distribution of *P. plagipennis* venom toxins is similar to that of the remipede *Speleonectes tulumensis*, another arthropod whose venom is thought to have a role in EOD as well as prey capture (66).

The function of many *P. plagipennis* venom proteins is unknown, but our results provide some clues as to the evolutionary history and function of some protein families. For example, according to our phylogenetic analysis, the CUB domain protein family probably evolved from the CUB domains of S1 proteases through loss of the proteolytic domain prior to 190 mya. Thus, CUB domain proteins likely originated shortly after the divergence of Heteroptera and its sister group Coleorrhyncha (moss bugs) around 240 mya (67), coinciding with the switch by ancestral heteropterans from phytophagy to predation (68). We suggest that the most plausible account of the evolution of venoms in Heteroptera is that in between the divergence of Heteroptera and Coleorrhyncha and the last common ancestor of Heteroptera, the protease-rich salivary secretions used by both predaceous and phytophagous heteropterans for EOD (69, 70) evolved additional toxic and paralytic activities in response to selection for efficient prey capture. This might have occurred either through the recruitment of toxins from proteins expressed elsewhere in the body or via the evolution of salivary enzymes into venom toxins. The latter pattern is exemplified in another orally derived venom, the phospholipase A<sub>2</sub> (PLA<sub>2</sub>) family of snake venoms. PLA<sub>2</sub> enzymes are themselves powerful toxins that function by cleaving phospholipids into toxic signaling molecules, but one group (group II snake venom PLA<sub>2</sub> with <sup>49</sup>Lys) subsequently lost crucial active-site residues and acquired diverse non-enzymatic activities, including neurotoxicity, myotoxicity, and modulatory activity on coagulation and platelet aggregation pathways (71). CUB domain proteins apparently result from a similar evolutionary process, *i.e.* the neofunctionalization of a digestive enzyme present in the saliva of ancestral heteropterans into a non-enzymatic venom com-

ponent. Although the function of CUB domain proteins is unknown, their conservation over 190 million years and high abundance in *P. plagiipennis* venom suggest that they perform an important, non-enzymatic function in prey capture and/or feeding.

Functionally, the venoms of predaceous reduviids are very different to the comparatively well characterized venoms produced by blood-sucking triatomine reduviids. Instead of paralysis and liquefaction, triatomine venom works to disrupt host hemostatic systems (18). Reflecting these different actions, we found that the venom of *P. plagiipennis* was markedly different in protein composition compared with triatomine bug venom. Insofar as the venom proteome of *P. plagiipennis* can be generalized to other predaceous reduviids, our results indicate that the shift to blood-feeding by triatomines is likely to have been accompanied by a strong decrease in the expression of proteases, cytolytic toxins, and Ptu1 family peptides, with a concomitant increase in the number and expression of triabins accompanied by the recruitment of Kazal domain proteins and 5'-nucleotidase-type apyrases. Regardless of their major differences in composition, there are many protein families with representatives in the venom of both predaceous and blood-sucking reduviids, including proteases, redulysin/trialysins, inositol phosphate phosphatase, cystatin domain proteins, serpins, and triabins.

This study highlights the power of combining transcriptomic and proteomic approaches for providing a holistic overview of venom proteomes (72), and it further underscores the likelihood of finding novel protein families in poorly studied venomous taxa (73). Our results provide a solid foundation for understanding the role played by individual venom components in prey capture and liquefaction in predaceous assassin bugs and provide key insights into the different pathways of venom evolution in predaceous and hematophagous heteropterans.

Nucleic acid and amino acid sequences were deposited in GenBank™ with accession numbers KX459564–KX459693 and KX752800–KX752820.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (74) with the dataset identifier PXD004804 (available at <http://www.ebi.ac.uk/pride/>).

**Acknowledgments**—We thank Alun Jones at the Queensland Bioscience Precinct Mass Spectrometry Facility for help with proteomics experiments, Greg Baillie and Angelika Christ from the IMB Sequencing Facility, Max Rosenthal and Caleb Stewart for help with experimental procedures, and the many people who kindly donated assassin bugs to this project.

\* This work was supported by a University of Queensland Postdoctoral Fellowship (to A. A. W.), Australian Research Council Discovery Early Career Researcher Award DE160101142 (to E. U.), Linkage Grant LP140100832 (to G. F. K.), and Australian National Health and Medical Research Council Principal Research Fellowship (to G. F. K.).

§ This article contains [supplemental material](#).

§ To whom correspondence should be addressed. E-mail: a.walker@uq.edu.au or glenn.king@imb.uq.edu.au.

## REFERENCES

- Fry, B. G., Roelants, K., Champagne, D. E., Scheib, H., Tyndall, J. D., King, G. F., Nevalainen, T. J., Norman, J. A., Lewis, R. J., Norton, R. S., Renjifo, C., and de la Vega, R. C. (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **10**, 483–511
- Olivera, B. M., Walker, C., Cartier, G. E., Hooper, D., Santos, A. D., Schoenfeld, R., Shetty, R., Watkins, M., Bandyopadhyay, P., and Hillyard, D. R. (1999) Speciation of cone snails and interspecific hyperdivergence of their venom peptides. *Ann. N.Y. Acad. Sci.* **870**, 223–237
- Sollod, B. L., Wilson, D., Zhaxybayeva, O., Gogarten, J. P., Drinkwater, R., and King, G. F. (2005) Were arachnids the first to use combinatorial peptide libraries? *Peptides* **26**, 131–139
- Sunagar, K., Jackson, T. N., Undheim, E. A., Ali, S. A., Antunes, A., and Fry, B. G. (2013) Three-Fingered RAVERS: rapid accumulation of variations in exposed residues of snake venom toxins. *Toxins* **5**, 2172–2208
- Sunagar, K., and Moran, Y. (2015) The rise and fall of an evolutionary innovation: contrasting strategies of venom evolution in ancient and young animals. *PLoS Genet.* **11**, e1005596
- Dutertre, S., and Lewis, R. J. (2010) Use of venom peptides to probe ion channel structure and function. *J. Biol. Chem.* **285**, 13315–13320
- Kalia, J., Milescu, M., Salvatierra, J., Wagner, J., Klint, J. K., King, G. F., Olivera, B. M., and Bosmans, F. (2015) From foe to friend: using animal toxins to investigate ion channel function. *J. Mol. Biol.* **427**, 158–175
- King, G. F. (2011) Venoms as a platform for human drugs: translating toxins into therapeutics. *Expert Opin. Biol. Ther.* **11**, 1469–1484
- King, G. F., and Hardy, M. C. (2013) Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. *Annu. Rev. Entomol.* **58**, 475–496
- Lewis, R. J., and Garcia, M. L. (2003) Therapeutic potential of venom peptides. *Nat. Rev. Drug Discov.* **2**, 790–802
- Fox, J. W., and Serrano, S. M. (2007) Approaching the golden age of natural product pharmaceuticals from venom libraries: an overview of toxins and toxin-derivatives currently involved in therapeutic or diagnostic applications. *Curr. Pharm. Des.* **13**, 2927–2934
- Undheim, E. A., Fry, B. G., and King, G. F. (2015) Centipede venom: recent discoveries and current state of knowledge. *Toxins* **7**, 679–704
- von Reumont, B. M., Campbell, L. I., Richter, S., Hering, L., Sykes, D., Hetmank, J., Jenner, R. A., and Bleidorn, C. (2014) A polychaete's powerful punch: venom gland transcriptomics of *Glycera* reveals a complex cocktail of toxin homologs. *Genome Biol. Evol.* **6**, 2406–2423
- Wong, E. S., Morgenstern, D., Mofiz, E., Gombert, S., Morris, K. M., Temple-Smith, P., Renfree, M. B., Whittington, C. M., King, G. F., Warren, W. C., Papenfuss, A. T., and Belov, K. (2012) Proteomics and deep sequencing comparison of seasonally active venom glands in the platypus reveals novel venom peptides and distinct expression profiles. *Mol. Cell. Proteomics* **11**, 1354–1364
- Ziegman, R., and Alewood, P. (2015) Bioactive components in fish venoms. *Toxins* **7**, 1497–1531
- Walker, A. A., Weirauch, C., Fry, B. G., and King, G. F. (2016) Venoms of heteropteran insects: a treasure trove of diverse pharmacological toolkits. *Toxins* **8**, 43
- Hwang, W. S., and Weirauch, C. (2012) Evolutionary history of assassin bugs (Insecta: Hemiptera: Reduviidae): insights from divergence dating and ancestral state reconstruction. *PLoS ONE* **7**, e45523
- Ribeiro, J. M., Assumpção, T. C., and Francischetti, I. M. (2012) An insight into the sialomes of bloodsucking Heteroptera. *Psyche* **2012**, 1–16
- Edwards, J. S. (1961) The action and composition of the saliva of an assassin bug *Platymeris rhadamanthus* Gaerst. (Hemiptera, Reduviidae). *J. Exp. Biol.* **38**, 61–77
- Maran, S. P. M., and Ambrose, D. P. (2000) in *Biotechnological Applications for Integrated Pest Management* (Ignacimuth, A., Sen, A., and Janarthanan, S., eds) pp. 125–131, Oxford Publishing, New Delhi, India
- Sahayaraj, K., and Vinothkanna, A. (2011) Insecticidal activity of venomous saliva from *Rhynocoris fuscipes* (Reduviidae) against *Spodoptera litura* and *Helicoverpa armigera* by microinjection and oral administration. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **17**, 486–490

22. Zerachia, T., Bergmann, F., and Shulov, A. (1973) in *Animal and Plant Toxins* (Kaiser, E., ed) pp. 143–146, Goldman, Munich, Germany
23. Readio, P. A. (1927) Studies on the biology of the Reduviidae of America North of Mexico. *Kans. Univ. Sci. Bull.* **17**, 5–249
24. Zerachia, T., Shulov, A., and Bergmann, F. (1973) in *Animal and Plant Toxins* (Kaiser, E., ed) pp. 147–150, Goldman, Munich, Germany
25. Sahayaraj, K., Borgio, J. F., Muthukumar, S., and Anandh, G. P. (2006) Antibacterial activity of *Rhynocoris marginatus* (Fab.) and *Catamirus brevipennis* (Servile) (Hemiptera: Reduviidae) venoms against human pathogens. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **12**, 487–496
26. Kitherian, S., Muthukumar, S., and Rivers, D. (2013) Biochemical and electrophoretic analyses of saliva from the predatory reduviid species *Rhynocoris marginatus* (Fab.). *Acta Biochim. Pol.* **60**, 91–97
27. Cohen, A. C. (1998) Solid-to-liquid feeding: the inside (s) out story of extra-oral digestion in predaceous arthropods. *Am. Entomol.* **44**, 103–116
28. Corzo, G., Adachi-Akahane, S., Nagao, T., Kusui, Y., and Nakajima, T. (2001) Novel peptides from assassin bugs (Hemiptera: Reduviidae): isolation, chemical and biological characterization. *FEBS Lett.* **499**, 256–261
29. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) A common structural motif incorporating a cystine knot and a triple-stranded  $\beta$ -sheet in toxic and inhibitory polypeptides. *Protein Sci.* **3**, 1833–1839
30. Undheim, E. A., Mobli, M., and King, G. F. (2016) Toxin structures as evolutionary tools: using conserved 3D folds to study the evolutionary trajectory of rapidly evolving peptides. *BioEssays* **38**, 539–548
31. Grundy, P. R. (2007) Utilizing the assassin bug, *Pristhesancus plagipennis* (Hemiptera: Reduviidae), as a biological control agent within an integrated pest management programme for *Helicoverpa* spp. (Lepidoptera: Noctuidae) and *Creontiades* spp. (Hemiptera: Miridae) in cotton. *Bull. Entomol. Res.* **97**, 281–290
32. Sahayaraj, K., and Ravi, C. (2007) Evaluation of reduviid predators and plant products against chosen groundnut pests. *Arch. Phytopathol. Plant Protect.* **40**, 281–290
33. Ambrose, D. P. (2000). In *Heteroptera of Economic Importance* (Schaeffer, C. W. and Panizzi, A. R., eds.) pp 695–712, CRC Press, Inc., Boca Raton, FL
34. Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M., and Schaeffer, D. A. (2007) The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **6**, 1638–1655
35. Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512
36. Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152
37. Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**, 276–277
38. Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G. M., Carnemolla, B., Orecchia, P., Zardi, L., and Righetti, P. G. (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**, 1327–1333
39. Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786
40. Tang, W. H., Shilov, I. V., and Seymour, S. L. (2008) Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J. Proteome Res.* **7**, 3661–3667
41. Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780
42. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542
43. Amino, R., Martins, R. M., Procopio, J., Hirata, I. Y., Juliano, M. A., and Schenkman, S. (2002) Trialysin, a novel pore-forming protein from saliva of hematophagous insects activated by limited proteolysis. *J. Biol. Chem.* **277**, 6207–6213
44. von Reumont, B. M., Jenner, R. A., Wills, M. A., Dell'ampio, E., Pass, G., Ebersberger, I., Meyer, B., Koenemann, S., Iliffe, T. M., Stamatakis, A., Niehuis, O., Meusemann, K., and Misof, B. (2012) Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol. Biol. Evol.* **29**, 1031–1045
45. Bloch, G., and Cohen, M. (2014) The expression and phylogenetics of the inhibitor cysteine knot peptide OCLP1 in the honey bee *Apis mellifera*. *J. Insect Physiol.* **65**, 1–8
46. Pineda, S. S., Sollod, B. L., Wilson, D., Darling, A., Sunagar, K., Undheim, E. A., Kely, L., Antunes, A., Fry, B. G., and King, G. F. (2014) Diversification of a single ancestral gene into a successful toxin superfamily in highly venomous Australian funnel-web spiders. *BMC Genomics* **15**, 177
47. Bork, P., and Beckmann, G. (1993) The CUB domain: a widespread module in developmentally regulated proteins. *J. Mol. Biol.* **231**, 539–545
48. Georgieva, D., Greunke, K., and Betzel, C. (2010) Three-dimensional model of the honeybee venom allergen Api m 7: structural and functional insights. *Mol. Biosyst.* **6**, 1056–1060
49. Levine, M. Z., Harrison, P. J., Walthall, W. W., Tai, P. C., and Derby, C. D. (2001) A CUB-serine protease in the olfactory organ of the spiny lobster *Panulirus argus*. *J. Neurobiol.* **49**, 277–302
50. Schmidt, M., Trulli, N. M., and Hoffman, D. R. (2006) Expression of the recombinant honeybee venom CUB protease allergen. *J. Allergy Clin. Immunol.* **117**, S309
51. Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652
52. Baek, J. H., and Lee, S. H. (2014) Differential gene expression profiles in the salivary gland of *Orius laevigatus*. *J. Asia-Pacif. Entomol.* **17**, 729–735
53. Li, M., Tian, Y., Zhao, Y., and Bu, W. (2012) Higher level phylogeny and the first divergence time estimation of Heteroptera (Insecta: Hemiptera) based on multiple genes. *PLoS ONE* **7**, e32152
54. Martins, R. M., Amino, R., Daghastanli, K. R., Cuccovia, I. M., Juliano, M. A., and Schenkman, S. (2008) A short proregion of trialysin, a pore-forming protein of *Triatoma infestans* salivary glands, controls activity by folding the N-terminal lytic motif. *FEBS J.* **275**, 994–1002
55. Martins, R. M., Sforça, M. L., Amino, R., Juliano, M. A., Oyama, S., Jr., Juliano, L., Pertinhez, T. A., Spisni, A., and Schenkman, S. (2006) Lytic activity and structural differences of amphipathic peptides derived from Trialysin. *Biochemistry* **45**, 1765–1774
56. Hernández-Vargas, M. J., Santibáñez-López, C. E., and Corzo, G. (2016) An insight into the triabin protein family of American hematophagous reduviids: functional, structural and phylogenetic analysis. *Toxins* **8**, 44
57. Jeong, K. Y., Yi, M. H., Jeong, K. J., Lee, H., Hong, C. S., and Yong, T. S. (2009) Sequence diversity of the Bla g 4 cockroach allergen, homologous to lipocalins, from *Blattella germanica*. *Int. Arch. Allergy Immunol.* **148**, 339–345
58. Kato, H., Jochim, R. C., Gomez, E. A., Sakoda, R., Iwata, H., Valenzuela, J. G., and Hashiguchi, Y. (2010) A repertoire of the dominant transcripts from the salivary glands of the blood-sucking bug, *Triatoma dimidiata*, a vector of Chagas disease. *Infect. Genet. Evol.* **10**, 184–191
59. Haridass, E. T., and Ananthkrishnan, T. N. (1981) Functional morphology of the salivary system in some Reduviidae (Insecta: Heteroptera). *Proc. Indian. Acad. Sci.* **90**, 145–160
60. Rodríguez de la Vega, R. C., Schwartz, E. F., and Possani, L. D. (2010) Mining on scorpion venom biodiversity. *Toxicon* **56**, 1155–1161
61. Dutertre, S., Jin, A. H., Kaas, Q., Jones, A., Alewood, P. F., and Lewis, R. J. (2013) Deep venomomics reveals the mechanism for expanded peptide diversity in cone snail venom. *Mol. Cell. Proteomics* **12**, 312–329
62. Undheim, E. A., Jones, A., Clauser, K. R., Holland, J. W., Pineda, S. S., King, G. F., and Fry, B. G. (2014) Clawing through evolution: toxin diversification and convergence in the ancient lineage Chilopoda (centipedes). *Mol. Biol. Evol.* **31**, 2124–2148



63. Jouiaei, M., Yanagihara, A. A., Madio, B., Nevalainen, T. J., Alewood, P. F., and Fry, B. G. (2015) Ancient venom systems: a review on Cnidaria toxins. *Toxins* **7**, 2251–2271
64. Cohen, A. C. (1990) Feeding adaptations in some predaceous Hemiptera. *Ann. Entomol. Soc. Am.* **83**, 1215–1223
65. Bottrall, J. L., Madaras, F., Biven, C. D., Venning, M. G., and Mirtschin, P. J. (2010) Proteolytic activity of elapid and viperid snake venoms and its implication to digestion. *J. Venom Res.* **1**, 18–28
66. von Reumont, B. M., Blanke, A., Richter, S., Alvarez, F., Bleidorn, C., and Jenner, R. A. (2014) The first venomous crustacean revealed by transcriptomics and functional morphology: Remipede venom glands express a unique toxin cocktail dominated by enzymes and a neurotoxin. *Mol. Biol. Evol.* **31**, 48–58
67. Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., Frandsen, P. B., Ware, J., Flouri, T., Beutel, R. G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., *et al.* (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767
68. Weirauch, C., and Schuh, R. T. (2011) Systematics and evolution of Hemiptera: 25 years of progress. *Annu. Rev. Entomol.* **56**, 487–510
69. Cohen, A. C. (1996) in *Zoophytophagous Heteroptera: Implication for Life History and Integrated Pest Management* (Alomar, O., ed) pp. 1–7, Thomas Say Publications in Entomology, Lanham, MD
70. Miles, P. W. (1972) in *The Saliva of Hemiptera* (Treherne, J. E., Berridge, M. J., and Wigglesworth, V. B., eds) pp. 183–255, Academic Press, New York
71. Sunagar, K., Tsai, I. H., Lomonte, B., Jackson, T. N. W., and Fry, B. G. (2015) In *Venomous Reptiles and their Toxins* (Fry, B. G., ed.) pp 335–340, Oxford University Press, Oxford, UK
72. Escoubas, P., Sollod, B., and King, G. F. (2006) Venom landscapes: mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon* **47**, 650–663
73. von Reumont, B. M., Campbell, L. I., and Jenner, R. A. (2014) Quo vadis venomomics? A roadmap to neglected venomous invertebrates. *Toxins* **6**, 3488–3551
74. Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456