Are Fireworms Venomous? Evidence for the Convergent Evolution of Toxin Homologs in Three Species of Fireworms (Annelida, Amphinomidae)

Aida Verdes[1,2,3,*], Danny Simpson[4], and Mandë Holford[1,2,5,*]

[1]Department of Chemistry, Hunter College Belfer Research Center, and The Graduate Center, Program in Biology, Chemistry and Biochemistry, City University of New York

[2]Department of Invertebrate Zoology, Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, New York

[3]Departamento de Biología (Zoología), Facultad de Ciencias, Universidad Autónoma de Madrid, Spain

[4]Department of Population Health, New York University School of Medicine

[5]Department of Biochemistry, Weill Cornell Medical College, Cornell University

*Corresponding authors: E-mails: averdes@gradcenter.cuny.edu; mholford@hunter.cuny.edu.

Abstract

Amphinomids, more commonly known as fireworms, are a basal lineage of marine annelids characterized by the presence of defensive dorsal calcareous chaetae, which break off upon contact. It has long been hypothesized that amphinomids are venomous and use the chaetae to inject a toxic substance. However, studies investigating fireworm venom from a morphological or molecular perspective are scarce and no venom gland has been identified to date, nor any toxin characterized at the molecular level. To investigate this question, we analyzed the transcriptomes of three species of fireworms—*Eurythoe complanata*, *Hermodice carunculata*, and *Paramphinome jeffreysii*—following a venomics approach to identify putative venom compounds. Our venomics pipeline involved de novo transcriptome assembly, open reading frame, and signal sequence prediction, followed by three different homology search strategies: BLAST, HMMER sequence, and HMMER domain. Following this pipeline, we identified 34 clusters of orthologous genes, representing 13 known toxin classes that have been repeatedly recruited into animal venoms. Specifically, the three species share a similar toxin profile with C-type lectins, peptidases, metalloproteinases, spider toxins, and CAP proteins found among the most highly expressed toxin homologs. Despite their great diversity, the putative toxins identified are predominantly involved in three major biological processes: hemostasis, inflammatory response, and allergic reactions, all of which are commonly disrupted after fireworm stings. Although the putative fireworm toxins identified here need to be further validated, our results strongly suggest that fireworms are venomous animals that use a complex mixture of toxins for defense against predators.

Key words: Annelida, Amphinomidae, fireworms, venom toxins, transcriptomics, venomics.

Introduction

Animal venoms, primarily defined as toxic biological secretions produced by one animal and delivered to another through the infliction of a wound, have originated independently in numerous lineages as adaptations for defense, predation, and intraspecific competition (Fry et al. 2009; Casewell et al. 2013). Venoms are among the most complex biochemical secretions known in nature, but despite this high complexity there is a remarkable degree of convergence in both the molecular scaffolds and the targets of venom components (Casewell et al. 2013; Gorson et al. 2015). A handful of easy to collect well-known venomous organisms such as snakes, spiders, or cone snails have been the focus of venom research for decades, and have shaped common hypotheses about venom evolution, such as gene duplication being the main driver of toxin gene evolution (Wong et al. 2012) or venom cocktails being mainly composed of neurotoxic peptides (von Reumont et al. 2014a, 2017). However, the

rise of -omics technologies, such as genomics, transcriptomics, and proteomics, has led to a revival of venom studies using an integrated approach referred to as venomics. Applying the venomics strategy to elucidate animal venom composition has revealed a high genetic and functional diversity of venom compounds across different taxa, and the molecular processes responsible for it (Vonk et al. 2013; Martinson et al. 2017; von Reumont et al. 2017). Additionally, venomics techniques have been a boon to the study of many taxa that were previously neglected due to their small size, difficulty in collection and/or limited venom quantities. This increase in taxon diversity in venom research is providing new perspectives and challenging traditional views about venom evolution (Whelan et al. 2014; Gorson et al. 2015; Modica et al. 2015; Perkin et al. 2015; Huang et al. 2016; Martinson et al. 2017; von Reumont et al. 2017). To understand the true diversity of animal venoms and to propose robust hypotheses regarding venom evolution, it is imperative to broaden venom taxon sampling and avoid biases derived from studying a limited pool of venomous organisms.

Annelids are a promising and often neglected group for venomics research. There are circa 17,000 described annelid species with an extraordinary diversity of body types, life strategies, feeding mechanisms, and striking adaptations (Weigert and Bleidorn 2016). Several annelid lineages use toxins for predation or defense (fig. 1), however apart from the widely studied salivary secretions of the hematophagous leech (fig. 1F) (Min et al. 2010; Kvist et al. 2013, 2014, 2016; Amorim et al. 2015; Siddall et al. 2016) and the potent neurotoxins of glycerid bloodworms (fig. 1E) (Michel and Keil 1975; Thieffry et al. 1982; Bon et al. 1985; von Reumont et al. 2014b; Richter et al. 2017), venom in segmented worms remains largely uncharacterized. Both leeches and bloodworms use venom to assist in feeding. Leeches use several anticoagulants and antihemostatic toxins to prevent coagulation in the host tissue during blood feeding (Kvist et al. 2016) while glycerid bloodworms use a mixture of toxins to capture and immobilize prey (von Reumont et al. 2014b). In addition to the predatory use of venom, there are also annelids that utilize toxins as a defensive mechanism (fig. 1A–D). For example, the earthworm *Eisenia foetida* (fig. 1D) expels its coelomic fluid when threatened, which contains Lysenin, a pore-forming hemolytic protein that is toxic to vertebrates (Kobayashi et al. 2004). Similarly, fireworms (Amphinomidae) (fig. 1A–C) which are common in shallow warm and temperate waters and include colorful reef-dwelling species, are well-known among divers and reef enthusiasts for their painful sting (Wiklund et al. 2008; Barroso et al. 2010; Borda et al. 2012, 2015; Ahrens et al. 2013).

Amphinomidae has been traditionally placed within Errantia (Rouse and Fauchald 1997) until recent molecular studies revealed their phylogenetic position as sister group, together with Sipuncula, to all Pleistoannelida (Sedentaria + Errantia) (Weigert et al. 2014; Andrade et al.

2015; Weigert and Bleidorn 2016). Their revised phylogenetic position suggests that amphinomids are crucial to clarify the basal relationships in the annelid tree of life, and to investigate the origin and evolution of key adaptations such as venom. There are approximately 200 described species and 25 genera of fireworms (Glasby et al. 2000; Rouse and Pleijel 2001) that share a series of morphological apomorphies such as nuchal organs on a sensory structure referred to as the caruncle, a ventral muscular eversible proboscis with thickened cuticle on circular lamellae, and calcareous chaetae (Rouse and Fauchald 1997; Wiklund et al. 2008; Borda et al. 2012). The calcareous chaetae have been hypothesized to function as hypodermic needles that deliver an active toxin, which was identified as a trimethylammonium compound that causes strong skin irritation from the species *Eurythoe complanata*, and was consequently named Complanine (Nakamura et al. 2008, 2010; Borda et al. 2012; von Reumont et al. 2014b). Gustafson (1930) attributed the toxicity of fireworms to a clear gelatinous substance filling the chaetal core, while Schulze et al. (2017) using light microscopy, reported that the chaetal core was hollow and that a small amount of fluid was occasionally released from the tip of the chaeta. Additionally, scanning electron microscopy studies showed that some chaetae are grooved, adding another potential channel for toxin delivery (Schulze et al. 2017). However, to date, no specific secretory glands have been found near the base of the chaetae (Eckert 1985; Schulze et al. 2017). More recently, in their ultrastructural study of *Eurythoe complanata*, Tilic et al. (2017) found no evidence of chaeta functioning as hypodermic needles and suggested that the chaetae are not hollow, but filled with a calcium carbonate matrix that might dissolve leaving a central cavity when exposed to acidic fixatives.

The conflicting evidence regarding the possible function of the calcareous chaetae as a toxin delivery system and the fact that no secretory tissue has been identified near the base of the chaetae has generated doubt about the venomous nature of fireworms (Tilic et al. 2017). Here, we attempt to clarify this question by investigating the transcriptomes of three species of fireworms, *Hermodice carunculata*, *Eurythoe complanata*, and *Paramphinome jeffreysii* (Amphinomidae) following a venomics approach. We identified 34 clusters of orthologous genes that represent 13 distinct toxin classes that have been convergently recruited into a wide array of animal venoms (fig. 1A–C). Our findings provide compelling evidence that fireworms produce a venomous secretion composed of a diverse cocktail of toxins used for defense against predators, and offer insights into the potential functions of toxin genes in fireworms.

Materials and Methods

Sample Collection, RNA Extraction, and Sequencing

An individual of *Hermodice carunculata* was collected by SCUBA from Norman's Pond Cay Cave, Norman's Pond
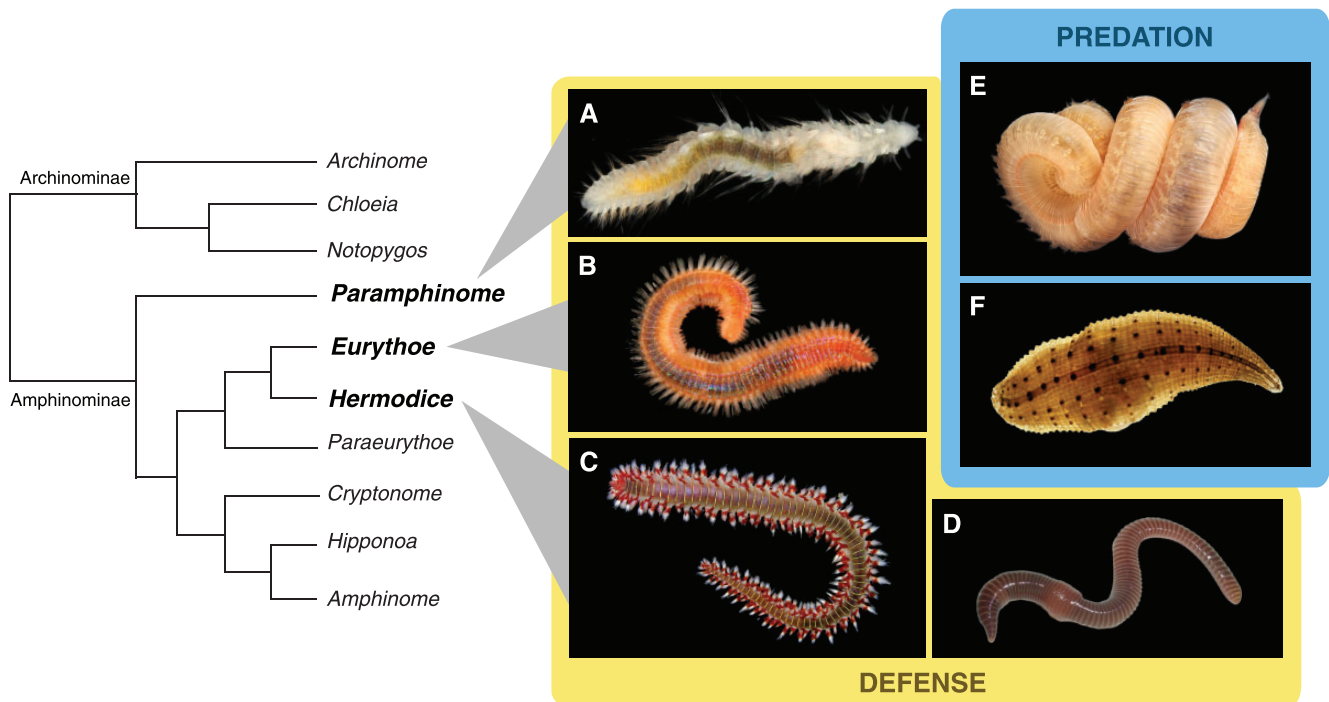
Fig. 1.—Phylogeny of the family Amphinomidae and venomous annelid representatives. Phylogenetic tree of the family Amphinomidae with the genera that include the three fireworm species investigated here, highlighted in bold. Cladogram based on phylogenetic reconstruction from Borda et al. (2015). Images of the fireworm species focus of this study are shown to the right: *Paramphinome jeffreysii* (A); *Eurythoe complanata* (B); and *Hermodice carunculata* (C). Fireworm species are grouped with earthworm (D) *Eisenia foetida* based on their defensive use of toxins and highlighted in yellow. Representatives of species that use venom for predation are also shown, highlighted in blue: bloodworm *Glycera capitata* (E) and leech *Helobdella europaea* (F). Images courtesy of Arne Nygren (A), Denis Riek (B), Arthur Anker (C), and Alexander Semenov (E). Remainder images (F and D) are publicly available under Creative Commons License.

Cay, Exumas, Bahamas (GPS N 23 47.181, W 076 08.428) and transported back to the field station where it was flash frozen in liquid nitrogen. A cross section of the posterior end of the animal was dissected for total RNA extraction. The tissue was homogenized in TriZol reagent (Life Technologies, NY) and total RNA was precipitated with isopropanol and dissolved in distilled water. Total RNA was used as template to perform polyA enriched first strand cDNA synthesis using the HiSeq RNA sample preparation kit for Illumina Sequencing (Illumina Inc., San Diego, CA) following manufacturer's instructions. Total RNA and the resulting cDNA library were assessed for quality and concentration with the Agilent 2100 BioAnalyzer and with agarose gel electrophoresis. The cDNA library was sequenced with Illumina HiSeq 1000 using a paired end flow cell and 80 x 2 cycle sequencing at the New York Genome Center. Transcriptomes of *P. jeffreysii* and *E. complanata* were generated as described in Weigert et al. (2014) from RNA extracted from the anterior end for *P. jeffreysii* and from a pool of individuals for *E. complanata*.

Raw Read Processing and Transcriptome Assembly

Raw reads for *H. carunculata* were visualized and quality checked with FastQC v0.11.5 (www.bioinformatics.

babraham.ac.uk; last accessed July 2017). Adapter sequences and low-quality reads were removed using Trimmomatic v0.36 (Bolger et al. 2014) and trimmed reads were re-evaluated with FastQC to ensure the high quality of the data after the trimming process. Due to the lack of a reference genome, the processed reads were de novo assembled using Trinity v2.4.0 (Grabherr et al. 2011; Haas et al. 2013). Transcriptomes for *E. complanata* and *P. jeffreysii* were generated as described in Weigert et al. (2014). See supplementary table S1, Supplementary Material online, for assembly statistics.

Identification and Annotation of Putative Toxins

The three de novo assembled fireworm transcriptomes were analyzed to identify putative toxins and venom components following a custom in silico venomics pipeline modified from Verdes et al. (2016) (fig. 2). Briefly, TransDecoder v3.0.1 (Haas et al. 2013) was first used to predict protein-coding regions within transcripts. As venom is mainly composed of secreted proteins and peptides, SignalP v4.1 (Petersen et al. 2011) was then used to predict signal peptide sequences in the putative coding regions.

Putative coding regions that included a signal peptide were extracted using a custom Perl script and analyzed following
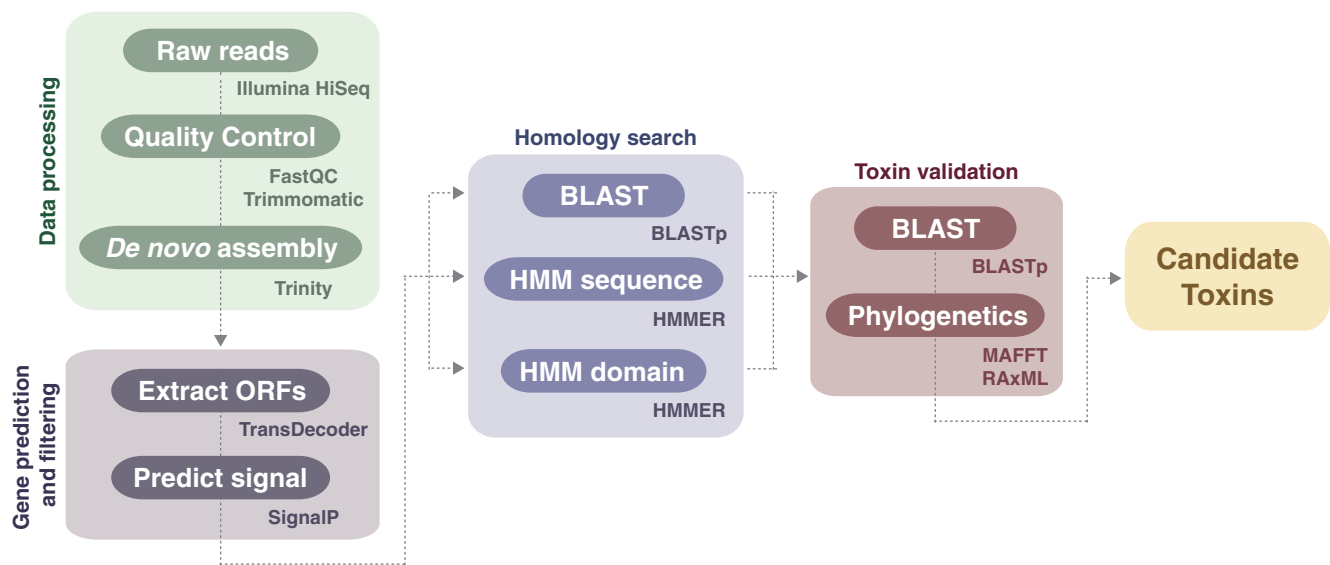
Fig. 2.—Venomics pipeline for the identification of toxin homologs. Diagram of the bioinformatics pipeline followed to identify putative venom components in fireworms. The different steps are highlighted with colored ovals and the corresponding software used indicated below each step. After RNA extraction and sequencing, the first step is data processing (green) and de novo assembly of the transcriptomes. Subsequently, a gene prediction and filtering step (gray) is performed, in which contigs are translated into amino acids, and open reading frames and signal sequences are predicted. The following step is a homology search (blue) using three different search strategies based on BLAST, HMMER-sequence and HMMER-domain. The results are merged and the putative toxins validated (maroon) through BLAST and phylogenetic reconstructions to generate a final list of candidate toxins (yellow).

two different homology search strategies, namely sequence similarity searches using the BLASTP tool of the NCBI BLAST v2.6.0 package (Altschul 1997; Johnson et al. 2008) and with HMMER v3.1b2 (Eddy 2009; Finn et al. 2015). Transcripts were searched using the BLASTP tool (Altschul 1997; Johnson et al. 2008) against an in-house database that includes all known toxins available in public databases such as ConoServer (Kaas et al. 2008, 2012) or Tox-Prot (Jungo and Bairoch 2005; Jungo et al. 2012) and additional putative toxins identified by the Holford group. In parallel, transcripts were also searched using HMMER (Finn et al. 2011) against hidden Markov models (HMM) profiles built from alignments of 24 known venom protein classes derived from a recent study of bloodworm (Glyceridae) venom gland transcriptomics (von Reumont et al. 2014b) and from public databases as mentioned earlier. The HMM profiles were built using complete sequence alignments (HMM sequence) and alignments of specific domain regions of each venom protein class (HMM domain) that had been previously identified with InterProScan v64.0 (Zdobnov and Apweiler 2001; Mulder and Apweiler 2008) in Geneious R10 (Kearse et al. 2012).

Hits with an e-value of 1e-5 or less derived from each independent search (BLASTP, HMMER sequence, HMMER domain) were considered putative toxins and, after removing redundant sequences, searched against the UniProtKB/Swiss-Prot nonredundant protein sequence database using BLASTP (Altschul 1997; Johnson et al. 2008) for cross validation. Transcripts initially identified as putative toxins were included in the final candidate toxin list only if 1) the best hit

from the UniProtKB/Swiss-Prot database had a higher e-value than the initial hit from the venom database or HMM profile, or 2) the best hit from the UniProtKB/Swiss-Prot database was labeled as a toxin. Transcripts initially identified as putative toxins that did not meet these criteria were considered false positives and excluded from downstream analyses. The final databases of candidate toxins for each species were manually curated and orthologous clusters of putative toxins shared by the three species were identified and annotated with OrthoVenn (Wang et al. 2015).

Phylogenetic Reconstruction of Putative Toxins

Phylogenetic analyses of individual toxin families including amphinomid homologs were performed to validate the orthology predictions derived from the venomics pipeline and to investigate toxin evolutionary relationships and possible origins. Selected putative toxins (i.e., nonredundant contigs that were identified as a toxin by all search methods) and homologous sequences from venomous and nonvenomous taxa spanning the Metazoa were downloaded from NCBI GenBank and provided by von Reumont et al. (2014b) and aligned with MAFFT v7.310 (Katoh and Standley 2013). The best model of protein evolution for each family was selected with ProtTest 3 (Darriba et al. 2011) and used for phylogenetic tree reconstruction in RAxML-HPC-PTTHREADS v8.2.10 (Stamatakis 2006, 2014) with support values estimated through a rapid bootstrap algorithm and 1,000 pseudoreplicates. All trees where rooted with nonvenomous
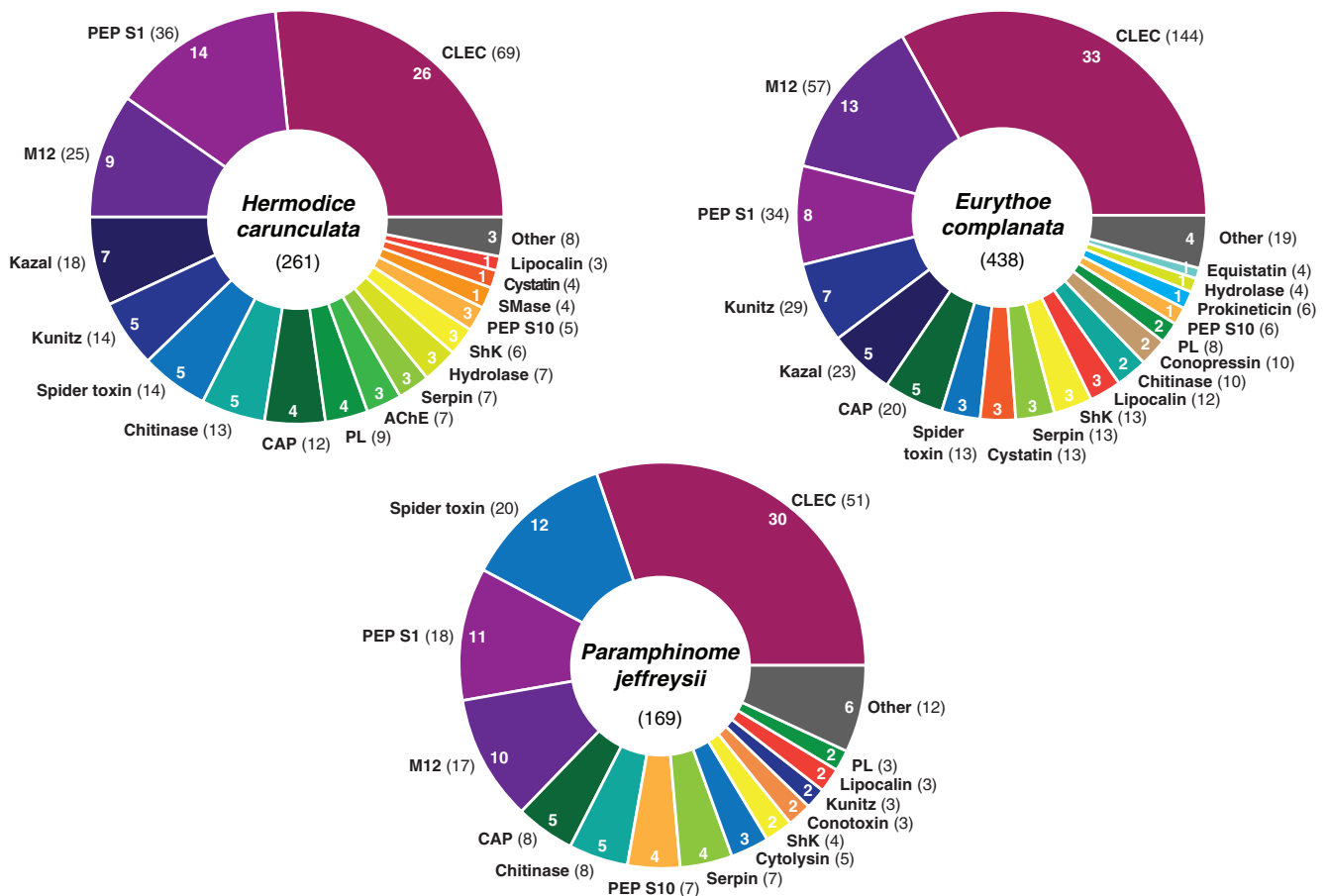
FIG. 3.—Diversity of toxin homologs identified in each fireworm species. Pie charts show the diversity of toxin homologs identified in each transcriptome. Toxin homologs are grouped by toxin class and numbers inside the pie chart indicate the relative abundance of each toxin class, as the percentage of the total number of toxins (i.e., 261 in *Hermodice carunculata*, 438 in *Eurythoe complanata*, and 169 in *Paramphinome jeffreysii*). The number of transcripts corresponding to each toxin class is indicated in parenthesis. Abbreviations are as follows: CLEC, C-type lectin; M12, metalloproteinase M12; PEP S1, peptidase S1; PEP S10, peptidase S10; PL, phospholipase; ShK, ShKT-domain containing peptides; AChE, acetylcholinesterase; SMase, sphingomyelinase. Other, groups a variety of less diverse toxin homologs (see supplementary files S2–S4, Supplementary Material online).

homologs. See supplementary table S2, Supplementary Material online, for accession numbers.

Results and Discussion

Diversity of Fireworm Toxin Genes

Transcriptomic analyses of fireworms, *H. carunculata, E. complanata, and P. jeffreysii*, revealed a wide diversity of putative toxin genes belonging to toxin classes that have been convergently recruited into other animal venoms (fig. 3 and supplementary files S2–S4, Supplementary Material online). The greatest diversity was found in *E. complanata* with 438 putative toxins representing 30 toxin classes, followed by *H. carunculata* with 261 putative toxins distributed in 24 toxin classes and *P. jeffreysii* with 169 putative toxins representing 22 toxin classes. Although differences in toxin composition are commonly found among different species (Colinet et al. 2013; Barghi et al. 2015; Phuong et al. 2016), the lower diversity of toxin homologs recovered

from some species might also be explained by differences in library construction methods and sequencing depth between the three transcriptomes (see supplementary table S1, Supplementary Material online).

The three fireworms share a similar toxin profile, with C-type lectin being the most diverse toxin class in all cases, and peptidase S1, metalloproteinase M12, spider toxin, and CAP found among the most highly expressed toxin homologs, altogether representing ∼60% of all putative toxins identified (fig. 3). Kazal-domain and Kunitz-type protease inhibitors are also among the most diverse toxin homologs expressed in the transcriptomes of *H. carunculata* and *E. complanata* (fig. 3A and B), whereas in *P. jeffreysii* spider neurotoxins including agatoxin, ctenitoxin, hainantoxin, and latrotoxin homologs are the second most diverse expressed putative toxins (fig. 3C). The evolutionary relationships among the species identified might explain the higher similarity in the toxin profiles of *H. carunculata* and *E. complanata* and the greater differences with *P. jeffreysii* (fig. 1).

A recent transcriptomic study found a similar profile of toxins expressed in the venom gland of the bloodworm *Glycera dibranchiata* (Annelida, Glyceridae) with peptidase S1 being the most abundant class in this species, and metalloproteinase M12, C-type lectin, and CAP found among the most abundantly expressed toxins, adding up to roughly 60% of the overall toxin precursors identified (von Reumont et al. 2014b). In addition, ~25% of the putative toxins identified in each of the three species investigated here are homologs of *Glycera* venom compounds (see supplementary files S2–S4, Supplementary Material online), suggesting that both glycerids and amphinomids have convergently recruited the same protein groups into their toxic secretions, as it has been demonstrated for multiple lineages of venomous animals (Fry et al. 2009).

Interspecific Variation in Toxin Composition

In addition to the more diverse toxin homologs found in *H. carunculata*, *E. complanata*, and *P. jeffreysii* fireworm transcriptomes, a variety of less common but particularly interesting venom toxin-like genes were identified in just one or two of the species, potentially reflecting interspecific variation in toxin composition (fig. 3). Inter- and intraspecific variation in venom composition have been documented as a result of diverse factors such as sex, ontogeny, diet, or geographical locality (Colinet et al. 2013; Sunagar et al. 2014; Barghi et al. 2015; Phuong et al. 2016; Cipriani et al. 2017). Therefore, it is reasonable to find differences in toxin composition in the fireworm species investigated here that may not be due to experimental procedures (fig. 1 and supplementary table S1, Supplementary Material online).

The putative species-specific toxins identified include most notably homologs of sea anemone and conoidean mollusk neurotoxins and several cytolysins (fig. 3). Among the neurotoxin homologs, we recovered transcripts with strong sequence similarity to conoidean neurotoxins, including two conotoxin-like peptides in *E. complanata* and four in *P. jeffreysii*, one teretoxin homolog in *H. carunculata* and two in *E. complanata*, and five turritoxin-like peptides in *E. complanata*. Conoidean venom peptides generally show remarkably high specificity to their molecular target and typically interfere with processes in the nervous system via inhibition of transmembrane ion channels and receptors (Teichert et al. 2015; Leffler et al. 2017). We also identified several cytolysins homologs, including one actinoporin-like transcript from each transcriptome, 4 gigantoxin homologs from *P. jeffreysii*, and one latarcin-like peptide from of *H. carunculata*. Actinoporin and gigantoxin are highly conserved pore-forming toxins initially identified in sea anemones that show cytolytic, hemolytic, and nerve stimulatory activity (Anderluh and Maček 2002; Hu et al. 2011). Gigantoxin also elicits acute pain by interfering with TRPV1 channels, which are ligand-gated nonselective cation channels expressed in the peripheral sensory neurons and involved in pain sensation (Cuypers et al. 2011). The other putative cytolytic toxin

identified is a latarcin-like peptide. Latarcins are antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaevi* that adopt an amphiphilic α-helical structure in contact with lipid membranes and show general cytotoxicity and hemolytic activity (Kozlov et al. 2006; Dubovskii et al. 2015).

Functional Categories and Biological Activity of Fireworm Toxins

Identifying the overlap of clusters of orthologous sequences across multiple species allows us to derive hypotheses about the function and evolution of proteins (Wang et al. 2015). The web platform OrthoVenn (Wang et al. 2015) was used to identify and annotate orthologous clusters among the fireworm toxin-like genes and inferred putative functions and biological activities that could explain the physiological symptoms observed after fireworm stings. The total 868 toxin-like genes recovered from the transcriptomes of the three species (438 from *E. complanata*, 261 from *H. carunculata* and 169 from *P. jeffreysii*) (fig. 3) were compared and 34 clusters of orthologous genes (supplementary table S3, Supplementary Material online) that represent 13 distinct toxin classes were identified (fig. 4). For convenience, we classified them into four functional categories: protease inhibitors (cystatin, lipocalin, serpin, and kazal-type and kunitz-type protease inhibitors), proteolytic enzymes (metalloproteinase M12, peptidase S1, and peptidase S10), neurotoxins (ShK-like, spider toxins, and turripeptide-like), and other proteins (CAP, C-type lectins, and phospholipases).

Protease Inhibitors: Cystatin, Lipocalin, Serpin, Kazal Domain, and Kunitz-Type Protease Inhibitors

Cystatins are cysteine-proteinase inhibitors that play important regulatory roles in protein degradation processes (Kordis and Turk 2009). Cystatins have been found in the venoms of several organisms, such as reptiles, caterpillars, ticks, mosquitoes, nematodes, and segmented worms (Hartmann and Lucius 2003; Fry et al. 2009; Schwarz et al. 2012; Kvist et al. 2014, 2016; von Reumont et al. 2014b; Chmelař et al. 2017). In nematodes, cystatins modulate host immune response and can act as proinflammatory molecules whereas tick and leech cystatins are antihemostatic effectors involved in blood feeding (Hartmann and Lucius 2003; Kvist et al. 2016; Chmelař et al. 2017). Twelve cystatin homologs were recovered among the putative fireworm toxins that form an orthologous cluster. The *Eurythoe* and *Paramphinome* cystatins form a strongly supported clade and are closely related to the *Glycera* bloodworm cystatins and the cnidarian *Cyanea capillata* (fig. 5). The *Hermodice* cystatin-like genes are found elsewhere in the tree more closely related to other cystatins from nonvenomous taxa. Interestingly, the *Eurythoe* and *Paramphinome* putative cystatins do not cluster with other cystatin homologs from
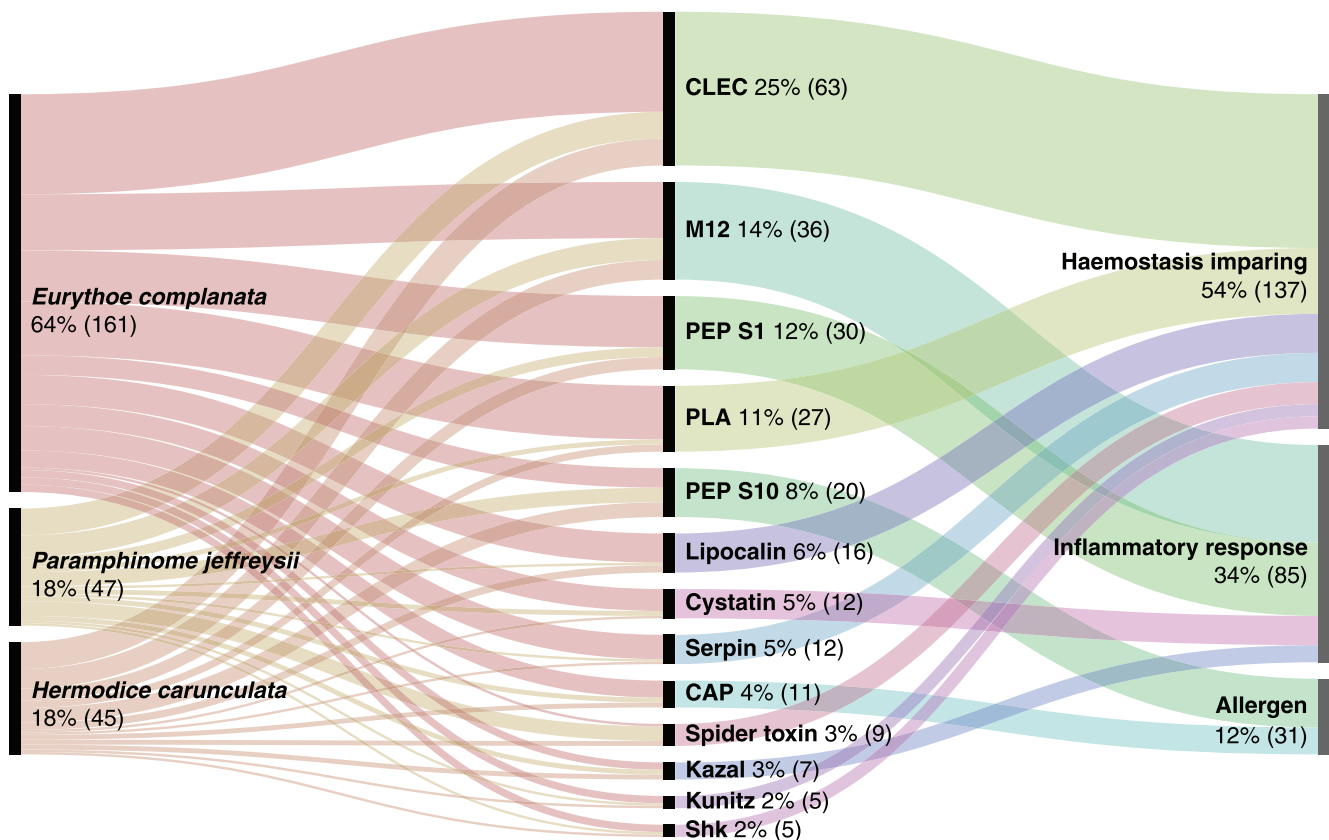
Fig. 4.—Orthologous toxins in fireworms and associated functions. Alluvial diagram showing the 253 orthologous toxins identified in the three fireworm species in the first node, the toxin class to which they belong in the second node, and the associated biological function in the third node. The relative abundance of toxin homologs corresponding to each category (species, toxin class, and function) is shown below the name as a percentage of the total. The number of transcripts that correspond to each category is shown in parenthesis. Abbreviations are as follows: CLEC, C-type lectin; M12, metalloproteinase M12; PEP S1, peptidase S1; PEP S10, peptidase S10; PL, phospholipase; ShK, ShKT-domain containing peptides.

nonvenomous annelids (*Capitella* and *Perinereis*) suggesting the involvement of this protein in fireworm toxicity.

Lipocalins are small extracellular proteins that bind small hydrophobic molecules and show great functional diversity including roles in retinol and pheromone transport, olfaction, prostaglandin synthesis, cell homeostasis, and modulation of the immune response (Flower 1996). Lipocalins have been recruited into a wide variety of animal venoms including vertebrates, such as bats and snakes (Wei and Chen 2012; Low et al. 2013; Junqueira-de-Azevedo et al. 2016), segmented worms (*Glycera*) (von Reumont et al. 2014b) and most notably, it has been independently recruited into the venom and salivary secretions of a variety of hematophagous arthropods such as ticks, kissing bugs, and blood feeding dipterans to serve a variety of antihemostatic functions (Mans et al. 2003; Andersen et al. 2005; Ribeiro et al. 2007; Santos et al. 2007; Fry et al. 2009). Lipocalins have also been identified as major allergens inducing severe anaphylactic reactions in humans after hematophagous insect bites (Paddock et al. 2001). Sixteen lipocalin-like genes belonging to two orthologous clusters were identified in the *H. carunculata*,

E. complanata, and *P. jeffreysii* transcriptomes. All putative fireworm lipocalins except one, cluster in a single clade (albeit not supported) that includes Lopap, a lipocalin from the *Lonomia obliqua* caterpillar venom that contributes to hemorrhagic syndrome through prothrombin activation (fig. 6) (Reis et al. 2006).

Serpin homologs were also recovered among the putative fireworm toxins. Serpins are a large and diverse family of protease inhibitors that use a conformational change to irreversibly inhibit serine proteases, although serpins that inhibit caspases and papain-like cysteine proteases have also been identified (Irving et al. 2000; Law et al. 2006). Serpins regulate important proteolytic cascades involved in a variety of biological processes like coagulation, inflammation, and immune response (Law et al. 2006). Serpins have been recruited into venoms and toxic secretions of numerous taxa, including vertebrates like male platypus, snakes, and frogs (Braud et al. 2000; Wu et al. 2011; Rokyta et al. 2012; Wong et al. 2012), cnidarians (Liu et al. 2015), annelids such as leeches and bloodworms (Kvist et al. 2014, 2016; von Reumont et al. 2014b) and arthropods including remipede crustaceans,
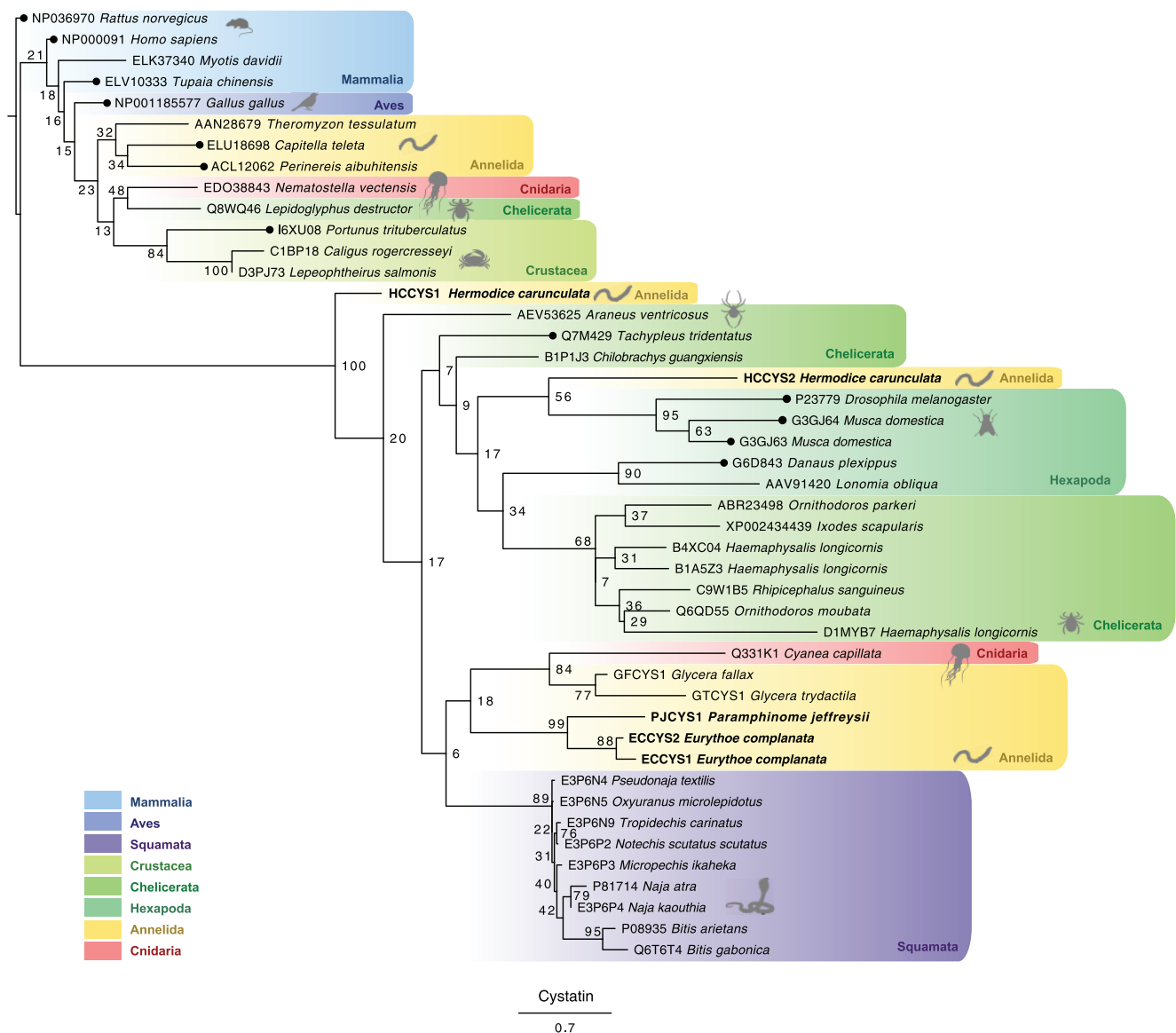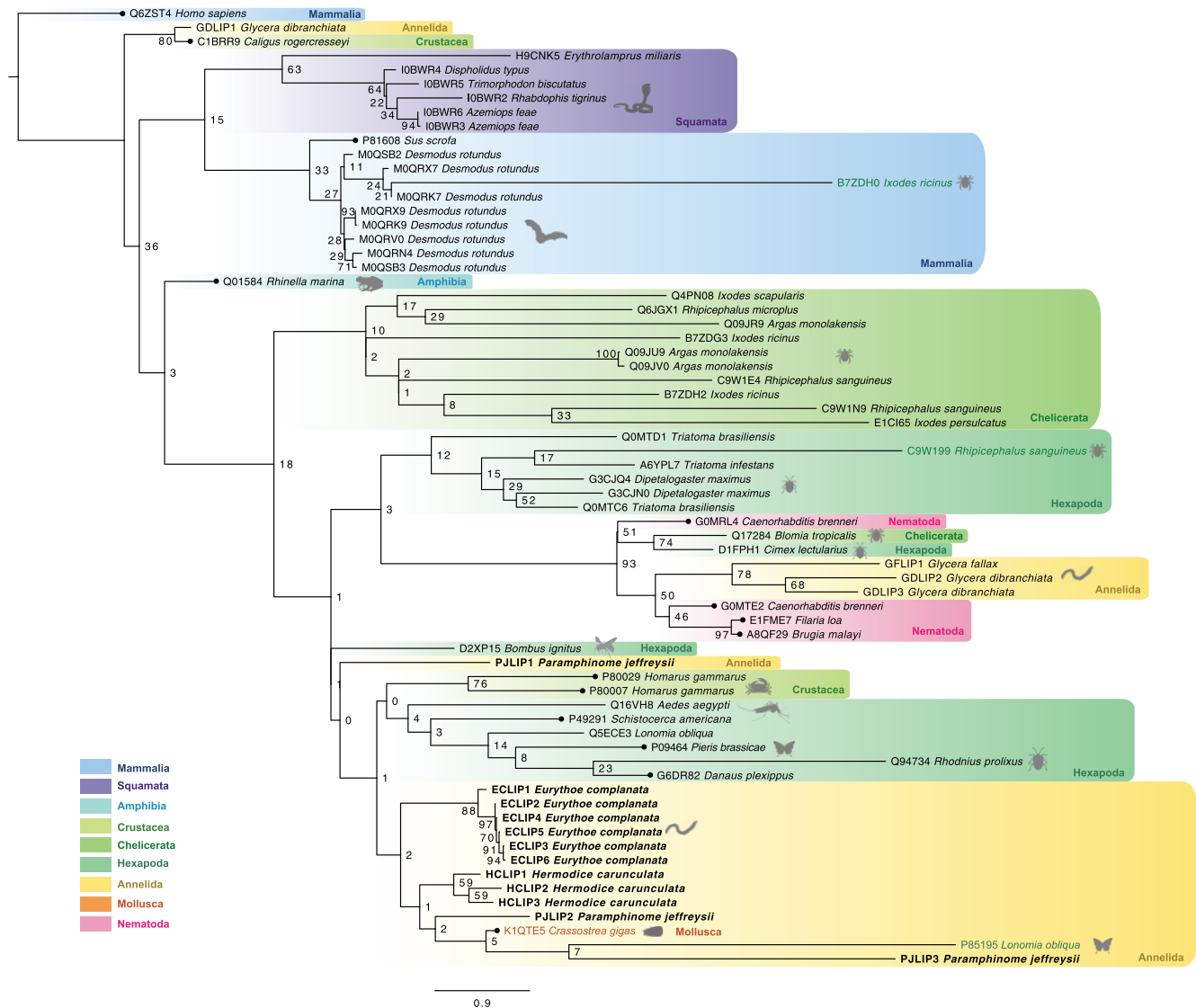
Fig. 5.—Phylogenetic tree of cystatin sequences. Cystatin phylogeny including sequences from venomous and nonvenomous taxa and fireworm homologs. Sequences from nonvenomous taxa are indicated with closed circles and fireworm homologs are highlighted in bold. Tree reconstructed with RAxML-HPC-PTTHREADS v8.2.10 and support values estimated through a rapid bootstrap algorithm and 1,000 pseudoreplicates. Bootstrap support values are given for all nodes and clade names are indicated by colored squares.

caterpillars, parasitoid wasps, and blood feeding insects like ticks and mosquitoes (Veiga et al. 2005; Kato et al. 2010; Chagas et al. 2013; Poirié et al. 2014; von Reumont et al. 2014c; Chmelař et al. 2017; Yan et al. 2017). We have identified 12 serpins among the putative fireworm toxins that form a single orthologous cluster. About 10 out of the 12 toxins group in one clade in the serpin tree whereas the remaining two are clustered in another clade that includes human, crustacean, and parasitoid wasp serpins (supplementary fig. S1, Supplementary Material online).

Kazal-type inhibitors are serine protease inhibitors characterized by the presence of one or more Kazal domains,

40–60 amino acid long domains with a conserved structure of six cysteine residues forming three disulfide bridges (Laskowski and Kato 1980; Rimphanitchayakit and Tassanakajon 2010). Kazal-type inhibitors act on a variety of serine proteases including thrombin, trypsin, factor XIIa, subtilisin A, elastase, chymotrypsin, and plasmin (Sigle and Ramalho-Ortigão 2013) and have been reported from the venoms of snakes and bats (Francischetti et al. 2013; Low et al. 2013; Fernández et al. 2016), jellyfish (Liu et al. 2015), leeches, and bloodworms (Kvist et al. 2014; von Reumont et al. 2014b) and an array of insects including caterpillars, termites, bees, parasitoid wasps, and blood feeding
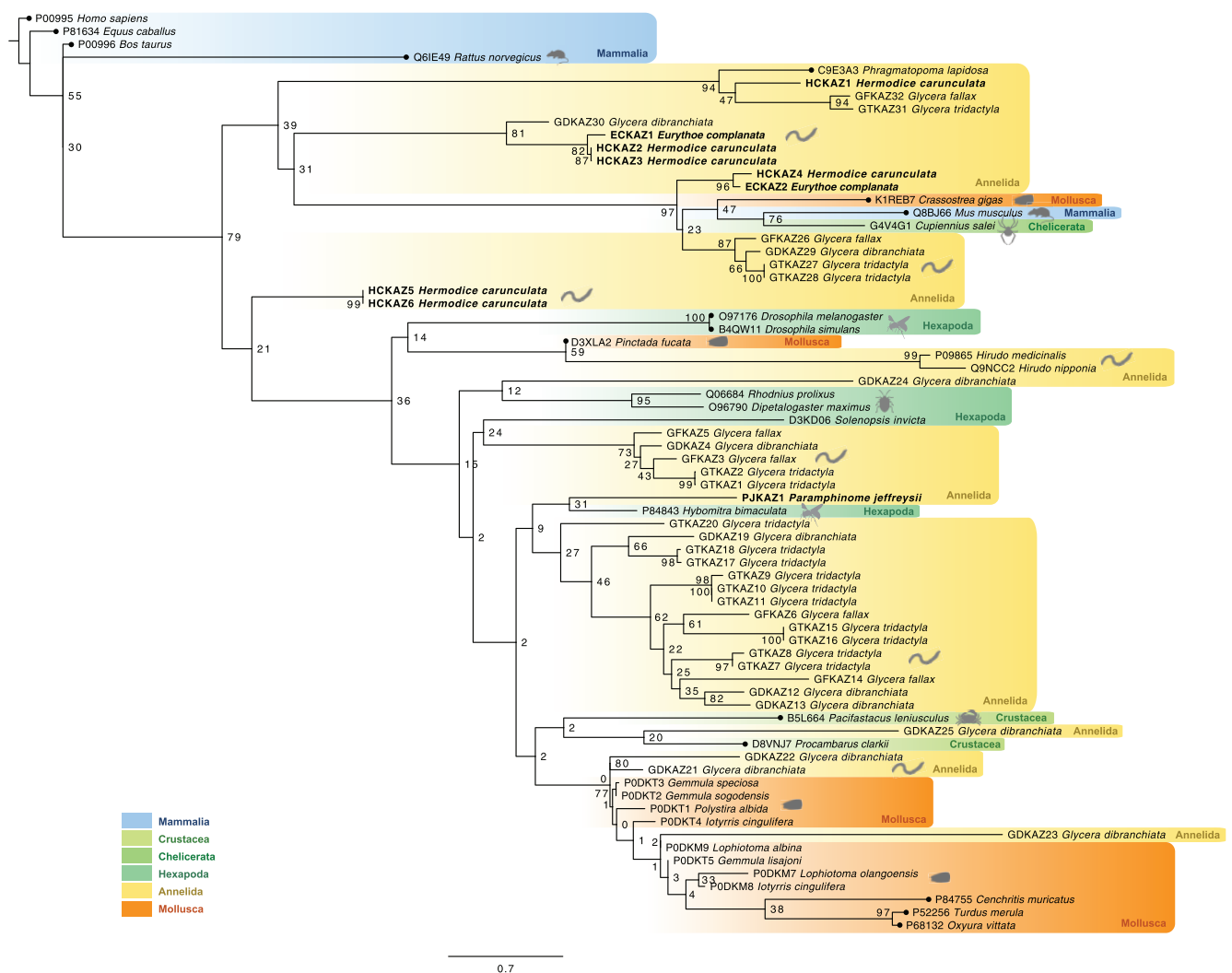
FIG. 6.—Phylogenetic tree of lipocalin sequences. Lipocalin phylogeny including sequences from venomous and nonvenomous taxa and fireworm homologs. Sequences from nonvenomous taxa are indicated with closed circles and fireworm homologs are highlighted in bold. Tree reconstructed with RAxML-HPC-PTTHREADS v8.2.10 and support values estimated through a rapid bootstrap algorithm and 1,000 pseudoreplicates. Bootstrap support values are given for all nodes and clade names are indicated by colored squares.

arthropods like ticks, flies, and mosquitoes (Veiga et al. 2005; Takác et al. 2006; Ribeiro et al. 2007; Kim et al. 2013; Sigle and Ramalho-Ortigão 2013; Negulescu et al. 2015; Qian et al. 2015). We identified two orthologous clusters of putative fireworm toxins with Kazal domains, one with 3 sequences showing similarity to kazal-type inhibitors and the other with 4 sequences resembling tur-ripeptides (the latter will be discussed in the neurotoxin section). The putative kazal-type fireworm toxins are found in several clades across the tree, mainly clustering with kazal-type toxins recovered from the venom gland of *Glycera* bloodworms (von Reumont et al. 2014b) (fig. 7). Interestingly, one *Paramphinome* transcript (PJKAZ1)

clusters with Vasotab, a kazal-type inhibitor identified from the salivary glands of the horse fly *Hybomitra bima-culata* that functions as a vasodilator (Takác et al. 2006) (fig. 7).

Kunitz-type protease inhibitors were also identified among the putative fireworm toxins. Kunitz-type inhibitors are ubiquitous serine protease inhibitors of approximately 60 amino acids with three conserved disulfide bridges that show inhibitory activity against trypsin, *chymotrypsin*, or both (Laskowski and Kato 1980; Wan et al. 2013). They have been identified from a variety of animal venoms including snakes and frogs (Calvete et al. 2007; Yuan et al. 2008; Wang et al. 2012), leeches and bloodworms (von Reumont et al. 2014b;

FIG. 7.—Phylogenetic tree of Kazal-type protease inhibitors. Kazal-type inhibitor phylogeny including sequences from venomous and nonvenomous taxa and fireworm homologs. Sequences from nonvenomous taxa are indicated with closed circles and fireworm homologs are highlighted in bold. Tree reconstructed with RAxML-HPC-PTTHREADS v8.2.10 and support values estimated through a rapid bootstrap algorithm and 1,000 pseudoreplicates. Bootstrap support values are given for all nodes and clade names are indicated by colored squares.

Kvist et al. 2016), sea anemones and cone snails (Schweitz et al. 1995; Bayrhuber et al. 2005; Isaeva et al. 2012), scorpions and spiders (Yuan et al. 2008; Ding et al. 2015), and a variety of insects like bees, wasps, ticks, flies (Yang et al. 2009; Choo et al. 2012; Soares et al. 2012; Tsujimoto et al. 2012). They contribute to the toxicity of the venom functioning as ion-channel blockers or interfering with physiological processes like blood coagulation, fibrinolysis, and inflammation (Wan et al. 2013). One orthologous cluster comprising 5 kunitz-domain sequences was recovered among the putative toxins identified in *H. carunculata, E. complanata*, and *P. jeffreysii* transcriptomes. The deep structure of the kunitz-type protease inhibitor phylogeny is poorly supported and does not reveal a clear evolutionary pattern (supplementary fig. S2, Supplementary Material online). The fireworm kunitz-domain transcripts group with different nonvenomous taxa,

except for two transcripts that are found in a clade including scorpion venom kunitz-type inhibitors (supplementary fig. S2, Supplementary Material online).

Proteolytic Enzymes: Metalloproteinase M12, Peptidase S1, and Peptidase S10

The M12 family of zinc dependent metalloproteinases has been widely recruited into the venoms of many taxa including snakes, platypus, scorpions, spiders, cnidarians, cephalopods, mollusks, leeches, and bloodworms (Da Silveira et al. 2007; Wong et al. 2012; Xia et al. 2013; von Reumont et al. 2014b; Modica et al. 2015; Kvist et al. 2016) where they interfere with cellular processes inducing various effects such as skin damage, edema, inflammation, hemorrhage, hypotension, and necrosis (von Reumont et al. 2014b).

Metalloproteinase-like transcripts are among the most diverse putative fireworm toxins identified, with 36 transcripts distributed in 7 orthologous clusters representing both the astacin (M12a) and reprolysin (M12b) subfamilies. The fireworm metalloproteinase-like sequences are distributed in two clades with one including also the astacin-like annelid bloodworm sequences and the other including the reprolysin-like bloodworm sequences (fig. 8).

The S1 family is the largest family of peptidases and includes serine proteases involved in a number of key biological processes, paramount among them are blood coagulation and immune responses (Page and Di Cera 2008). It is one of the most widely recruited venom proteins, recovered from the toxic secretions of a variety of taxa such as, reptiles, shrews, platypuses, and vampire bats (Kratzschmar et al. 1991; Kita et al. 2004; Whittington et al. 2010; Vaiyapuri et al. 2011; Menaldo et al. 2012), mollusks, annelids, cephalopods (Ruder et al. 2013; von Reumont et al. 2014b; Modica et al. 2015), and numerous arthropods including centipedes, remipede crustaceans, bees, parasitic wasps, ticks, mosquitoes, and caterpillars (Amarant et al. 1991; Asgari et al. 2003; Xu et al. 2005; Ribeiro et al. 2007; Choo et al. 2010; Liu et al. 2012; von Reumont et al. 2014c). In the venom, peptidase S1 serine proteases show varied activities such as anticoagulation, vasodilation, smooth muscle contraction, pain, immunosuppression, and inflammation (von Reumont et al. 2014b). Peptidase S1 are among the most diversified class of putative toxins in the fireworm transcriptomes, with 30 transcripts belonging to three orthologous clusters. Most peptidase S1-like genes cluster with bloodworm sequences (supplementary fig. S3, Supplementary Material online) and interestingly, a particularly large clade that includes transcripts from all three fireworm species and two Glycera species also includes serine proteases identified from the posterior venom glands of the cephalopod Octopus kaurna (Fry et al. 2009).

Unlike peptidase S1, the S10 family of serine proteases has not been as widely recruited into animal venoms. Venom serine carboxypeptidases have been found in the venoms of bees (Li et al. 2013), parasitoid wasps (Perkin et al. 2015), assassin bugs (Walker et al. 2017), remipede crustaceans (von Reumont et al. 2017), the giant triton snail (Bose et al. 2017), and glycerid bloodworms (von Reumont et al. 2014b). Venom serine carboxypeptidases have various physiological functions including enhancing the release of histamine, neurotransmitter degradation, and neurotoxicity, mediation of immune response and phosphorylation of venom proteins (Li et al. 2013; Bose et al. 2017). S10 peptidases are also recognized allergens of bee venom, inducing strong systemic IgE-mediated allergic reactions (Li et al. 2013). We recovered 20 peptidase S10-like transcripts that represent 3 orthologous clusters among the putative fireworm toxins. All fireworm sequences cluster in one large clade that also includes the bloodworm sequences, which might suggest an annelid specific radiation of this protein class (supplementary fig. S4, Supplementary Material online).

Neurotoxins: ShKT, Spider Toxins, and Turripeptide-Like

ShKT-domain containing peptides where recovered from the three fireworm transcriptomes. ShKT-domains were initially discovered from sea anemone peptide toxins, which are potent inhibitors of potassium channels (Castañeda et al. 1995; Castañeda and Harvey 2009) and have been posteriorly reported from a variety of proteins including venom toxins from other cnidarians, snakes, ribbon worms, vampire snails, and bloodworms (von Reumont et al. 2014b; Whelan et al. 2014; Modica et al. 2015; Ponce et al. 2016). ShKT domain-containing transcripts are one of the most expressed toxins in the venom glands of the bloodworm Glycera dibranchiata (von Reumont et al. 2014b). Most cnidarian ShKT toxins cause paralysis, although some show hemolytic effects (von Reumont et al. 2014b). In the vampire snail Colubraria reticulata, ShKT toxins might act as anesthetics (Modica et al. 2015) whereas in Glycera bloodworms they are thought to be involved in paralysis and death of prey (von Reumont et al. 2014b). We have identified one orthologous cluster of ShKT-domain toxins including 5 transcripts among the fireworm toxin-like genes. The fireworm ShKT-domain toxins are grouped in a large clade that includes the bloodworm sequences and the cnidarian toxins (fig. 9). The fireworm ShKT-domain transcripts contain the six conserved cysteines characteristic of this domain and thus are able to form the three disulfide bonds that stabilize sea anemone toxins (Castañeda and Harvey 2009).

We have also identified a variety of spider toxin-like genes in the fireworm transcriptomes, including homologs of agatoxin, atracotoxin, ctenitoxin, hainantoxin, and latrotoxin. The majority of these spider toxins are potent neurotoxins that block different voltage-gated potassium, sodium, and calcium ion channels, causing paralysis and severe pain (Kubista et al. 2007; Vassilevski et al. 2009). We recovered one orthologous cluster that includes 9 ctenitoxin-like sequences. Ctenitoxins are neurotoxins that contain an inhibitor cysteine-knot motif, first identified from the venom of the hunting spider Cupiennius salei that act as L-type calcium channel inhibitors but also show cytolytic activity (Kubista et al. 2007; Kuhn-Nentwig et al. 2012).

Turripeptides are short peptides with conserved cysteine patterns, discovered from the venom of turrids, a diverse group of conoidean mollusks related to cone snails and terebrids (Watkins et al. 2006). Turripeptide-like toxins have also been reported from the venom glands of Glycera bloodworms, box jellyfish, and zoanthids (von Reumont et al. 2014b; Brinkman et al. 2015; Huang et al. 2016). Turripeptides function as neurotoxins modulating ion channels and causing disruption of the neuromuscular transmission, which leads to paralysis of prey (Aguilar et al. 2009;
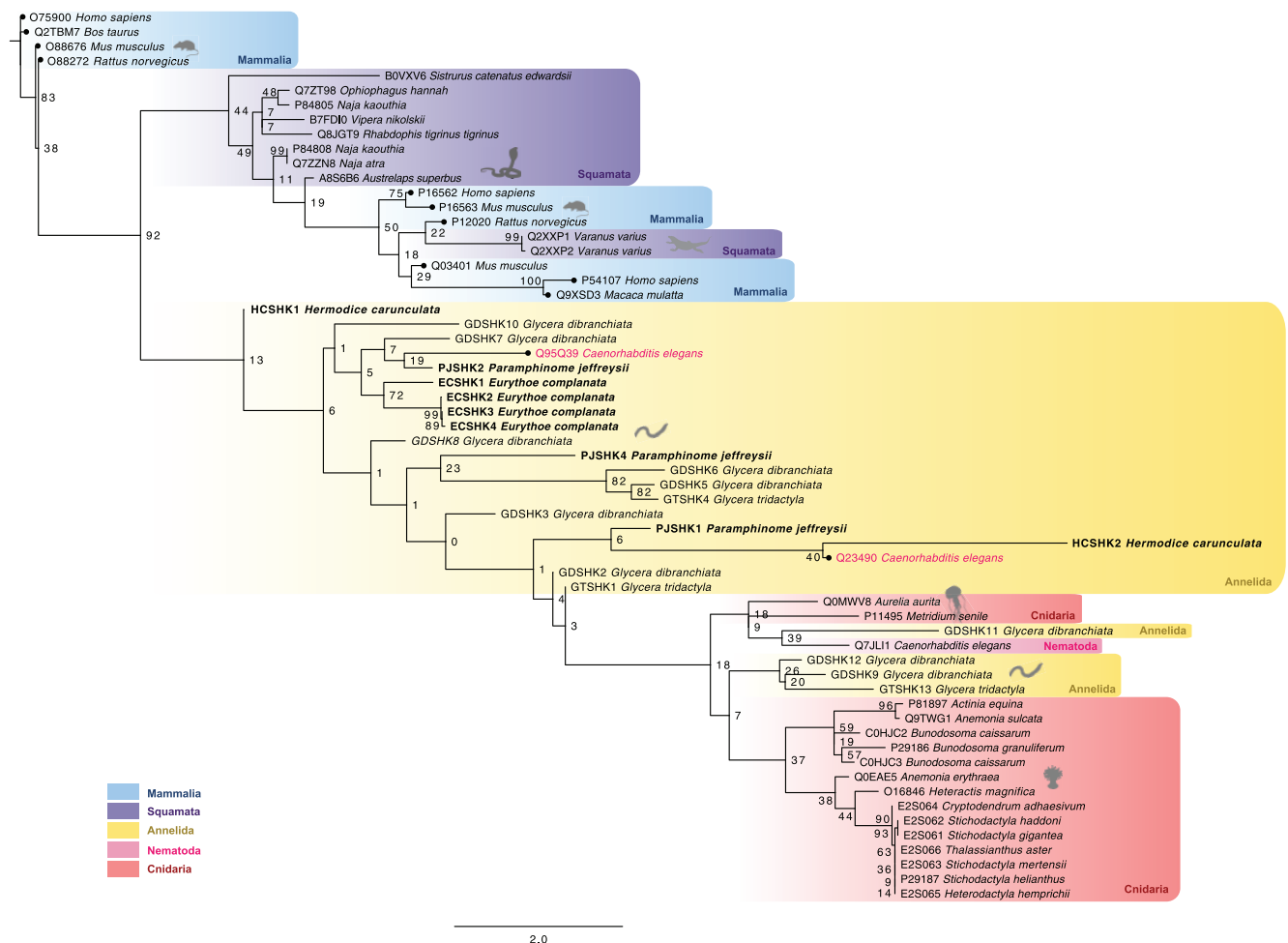
Fig. 8.—Phylogenetic tree of metalloproteinase M12 sequences. Metalloproteinase M12 phylogeny including sequences from venomous and non-venomous taxa and fireworm homologs. Sequences from nonvenomous taxa are indicated with closed circles and fireworm homologs are highlighted in bold. Tree reconstructed with RAxML-HPC-PTTHREADS v8.2.10 and support values estimated through a rapid bootstrap algorithm and 1,000 pseudo-replicates. Bootstrap support values are given for all nodes and clade names are indicated by colored squares.

Gonzales and Saloma 2014). One orthologous cluster including 4 turripeptide-like transcripts with the Kazal-domain characteristic of turripeptides was identified among the fireworm putative toxin genes. The turripeptide-like toxins have a predicted signal sequence with high sequence identity to that of turripeptides and the same conserved cysteine pattern (fig. 10).

Other Proteins: C-Type Lectins, CAP, and Phospholipases

The superfamily of C-type lectin-like domain (CTLD) proteins is a large group of extracellular proteins with varied functions (Zelensky and Gready 2005) that has been recruited into the venoms of numerous taxa including snakes, stonefish, cnidarians, crustaceans, blood feeding insects, caterpillars, leeches, and bloodworms (Morita 2005; Veiga et al. 2005;

Fig. 9.—Phylogenetic tree of ShKT-domain containing peptides. ShKT-domain containing peptide phylogeny including sequences from venomous and nonvenomous taxa and fireworm homologs. Sequences from nonvenomous taxa are indicated with closed circles and fireworm homologs are highlighted in bold. Tree reconstructed with RAxML-HPC-PTTHREADS v8.2.10 and support values estimated through a rapid bootstrap algorithm and 1,000 pseudoreplicates. Bootstrap support values are given for all nodes and clade names are indicated by colored squares.
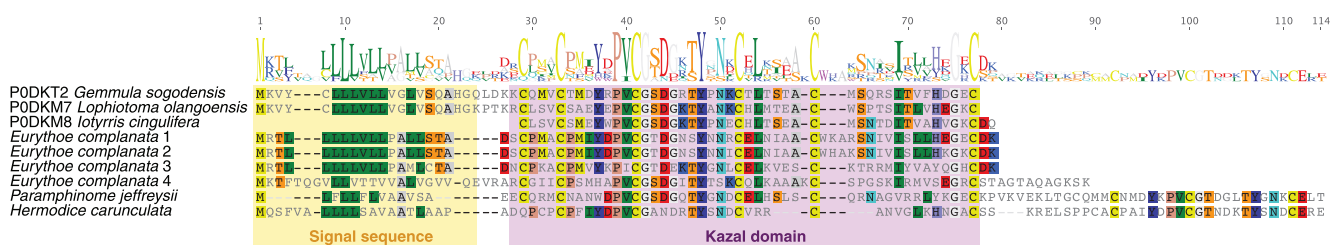
Fig. 10.—Multiple sequence alignment of turripeptide-like sequences. Multiple sequence alignment of turripeptides and fireworm homologs generated with MAFFT v7.310. Conserved residues are colored, including the conserved cysteines, which are highlighted in yellow. The predicted signal sequence is delimited by a yellow square and the Kazal domain is delineated by a purple square.

Magalhães et al. 2006; Ribeiro et al. 2007; Lopes-Ferreira et al. 2011; von Reumont et al. 2014b, 2017; Huang et al. 2016; Kvist et al. 2016). CTLD proteins are the most diverse toxin homolog recovered from the fireworm transcriptomes, with 63 toxins distributed in 9 orthologous clusters revealing a great diversity of paralogs that cluster with *Glycera*

bloodworm sequences in several clades across the phylogenetic tree (although many clades are not well supported) (supplementary fig. S6, Supplementary Material online). Venom C-type lectins are associated with hemagglutination, hemostasis, and inflammatory response (Huang et al. 2016). For example, snake venom CTLD proteins interfere with blood

coagulation pathways, whereas those in caterpillar venom have a potent anticoagulant effect and the homologs in blood feeding insects might have an antihemostatic function (Morita 2005; Veiga et al. 2005; Fry et al. 2009).

The CAP protein superfamily includes cysteine-rich secretory proteins (CRISPs), antigen 5 (Ag5), and pathogenesis-related 1 proteins (Pr-1) (Gibbs et al. 2008). CAP proteins have been recruited into the venoms of many taxa including reptiles, bats, monotremes, cnidarians, mollusks, crustaceans, spiders, scorpions, nemerteans, and bloodworms (Fry et al. 2009; Wong et al. 2012; Low et al. 2013; Moran et al. 2013; von Reumont et al. 2014b, 2017; Whelan et al. 2014). In snake venom, CAP proteins function as ion channel modulators inhibiting smooth muscle contraction (Yamazaki and Morita 2004), whereas in cone snails they have proteolytic activity (Milne et al. 2003) and in hymenopteran venoms they represent major allergens (Fang et al. 1988). We have identified 11 CAP homologs that belong to 2 orthologous clusters among the fireworm putative toxins. The fireworm CAP homologs cluster in a clade (although with no support) that includes the rest of the annelid sequences and a few arthropod sequences (supplementary fig. S5, Supplementary Material online).

A variety of phospholipase homologs have also been detected among the putative fireworm toxins. Phospholipases catalyze the hydrolysis of phospholipids generating free fatty acids and lysophospholipids, and thus inducing indirect hemolysis (Dennis 1983; Kini 2003). Phospholipases have been convergently recruited into the venoms of reptiles, cnidarians, cephalopods, insects, scorpions, spiders, and annelids and they have multiple pharmacological effects including neurotoxicity, cytotoxicity, and myotoxicity (King et al. 1996; Kini 2003; Veiga et al. 2005; Nevalainen 2008; Fry et al. 2009; von Reumont et al. 2014b; Huang et al. 2016). Venom phospholipases can induce necrosis, inflammation, inhibition of blood coagulation, and blocking of the neuromuscular transmission (Kini 2003; Fry et al. 2009; Huang et al. 2016). We have recovered 27 phospholipase-like genes that belong to two orthologous clusters, one of them corresponding to phospholipase B homologs and the other including phospholipase A2 homologs. Several of the fireworm phospholipases cluster with *Glycera* bloodworm sequences whereas others form a clade more closely related with snake phospholipases (supplementary fig. S7, Supplementary Material online).

Evidence of the Venomous Nature of Fireworms

We have identified a great number and diversity of toxin homologs in the transcriptomes of *E. complanata*, *H. carunculata*, and *P. jeffreysii*. However, these transcriptomes are not tissue specific since no venom producing tissue has been yet identified in the fireworms, and therefore distinguishing

between putative toxins and related nontoxin homologues is a challenge. Despite this difficulty, we provide several lines of evidence that strongly suggest the sequences we identified here are likely coding for putative toxins. First, It has been shown that gene models predicted to encode toxin-like proteins by similarity searches matched known venom components more likely if they are overexpressed in the venom gland and if their best BLAST hit is a known venom protein (as opposed to another protein from the UniProtKB database) (Rodríguez De La Vega and Giraud 2016). We initially identified putative toxins based on sequence similarity to known toxins and venom proteins, but because this method can also yield nontoxin homologs, we further validated these candidates by comparing them to the entire UniProtKB database. After this second comparison, we filter out the toxin-like transcripts that are more similar to other genes with nonvenom-related functions. Thus, guaranteeing the best BLAST hit for all our toxin candidates is a known venom protein. Second and importantly, we applied our venomics pipeline to identify putative toxins in the transcriptomes of two related nonvenomous annelids (*Chaetopterus* sp. and *Sipunculus nudus*) and a mollusk (*Chiton olivaceus*), recovering only 18 putative toxins in *Chaetopterus* sp. and none in *S. nudus* or *C. olivaceus* (see supplementary table S1, Supplementary Material online). We believe the findings in the nonvenomous organisms provide validation for the specificity of our pipeline and indicate that most of the sequences identified in the fireworms are more likely coding for venom toxins rather than for nontoxin homologs, suggesting only a minor percentage might represent false positives.

It has been long hypothesized that fireworms are venomous and use their dorsal calcareous chaetae to inject toxins. However, conflicting evidence regarding the function of chaetae as a venom delivery apparatus and the inability to identify glandular tissue near the base of the chaetae has generated some doubts (Eckert 1985; Schulze et al. 2017; Tilic et al. 2017). Tilic et al. (2017) concluded that the intrachaetal cavity is artificial and found no evidence of associated venom glands, which led them to suggest that the physiological reactions triggered by fireworm stings result from direct injuries rather than venom injection. Conversely, Schulze et al. (2017) found evidence of fluid being released from chaetal tips and identified grooved chaetae under the scanning electron microscope, indicating a possible alternative for toxin delivery. It has also been suggested that fireworms might sequester toxins from their diet, including palytoxin (PTX), the most potent nonprotein toxin known, produced by zoanthids and a few other organisms (Gleibs et al. 1995; Stoner and Layman 2015; Schulze et al. 2017). This evidence coupled with the uncertainty of a toxin delivery system may suggest that fireworms could be poisonous. As a consequence, the toxicity of fireworms is unclear and up to date it is still debated whether they are venomous, poisonous, or if they are toxic whatsoever.

Although the presence of a venom delivery mechanism, which is a clear feature distinguishing venomous versus poisonous animals, has yet to be confirmed in fireworms, other criteria can provide evidence of their venomous nature. These criteria include the specific molecular composition of the toxins, and how the toxins are produced. Venoms tend to be complex proteinaceous mixtures of larger compounds that must be injected into the prey, whereas poisons are usually composed of small chemical compounds such as alkaloids that can be readily absorbed (Daly 1995; Casewell et al. 2013). Additionally, most poisonous organisms such as puffer fish, poison dart frogs, or birds like the hooded pitohui, sequester the toxins from their diet or from symbiotic microorganisms (Dumbacher et al. 2004; Savitzky et al. 2012). Venomous animals on the contrary, produce venom proteins endogenously by expressing specific gene products (Fry et al. 2008; Perkin et al. 2015; Gorson et al. 2016) that are thought to originate from the neofunctionalization of regular body proteins, such as hormones (Fry et al. 2009; Undheim et al. 2015).

Following these criteria, our findings provide strong evidence supporting that fireworms are venomous organisms that endogenously produce a proteinaceous mixture of peptide toxins and venom protein homologs. Our results show that the species *E. complanata*, *H. carunculata*, and *P. jeffreysii* express a wide diversity of transcripts coding for toxin homologs that have been repeatedly recruited into animal venoms, including protease inhibitors, proteolytic enzymes, and neurotoxins (figs. 3 and 4). For instance, we identified homologs of Kazal-type protease inhibitors, zinc dependent metalloproteinases, and turripeptides found in the venoms of a wide diversity of animals spanning the Metazoa, including worms, snails, reptiles, platypuses, and vampire bats (figs. 5–9). The scale and scope of the toxin homologs identified in the transcriptomes of *H. carunculata*, *E. complanata*, and *P. jeffreysii* strongly suggests that venom has convergently evolved in fireworms as a defensive mechanism. Additionally, similar to other venomous animals, the toxin homologs found in fireworms appear to have originated through neofunctionalization of body proteins that target major molecular pathways in the circulatory and nervous systems. The putative toxins identified in this study are predominantly involved in three major biological activities: hemostasis, inflammatory response, and allergic reactions, all processes that are commonly disrupted after fireworm envenomations (Smith 2002; Ottuso 2013; Haddad 2015) (fig. 4).

Based on the known activities of fireworm toxin homologs in other venomous species, we hypothesize that the diverse cocktail of putative toxins identified here is responsible for the symptoms and physiological reactions observed after fireworm stings. Fireworm stings produce localized reactions with a variety of symptoms that include strong skin inflammation or edema, acute and severe pain, intense itching, partial paresthesia, or numbness of the affected area and sometimes serious wound infections and local necrosis (Smith 2002; Nakamura et al. 2008; Ottuso 2013; Haddad 2015; Schulze et al. 2017). Although rare, there are also reports of systemic reactions including nausea, cardiac, and respiratory anaphylactic reactions (Ottuso 2013). The symptoms can last several days or even weeks and the recommended treatment includes applying vinegar and hot water to the affected area (Smith 2002; Schulze et al. 2017), a common remedy for marine envenomations as evidence suggests that heat inactivates key components of the venom (Wilcox and Yanagihara 2016). These symptoms are consistent with the expression of the putative peptide toxins and venom proteins identified in this study.

The numerous protease inhibitors (i.e., cystatin, lipocalin, serpin, kazal, and kunitz-type), proteolytic enzymes (metalloproteinase M12 and peptidase S1), and C-type lectins found in *H. carunculata*, *E. complanata*, and *P. jeffreysii* fireworms mainly interfere with hemostatic processes and the inflammatory response (fig. 4), and therefore their activities are consistent with the strong skin inflammation, edema, and intense pain that occur after fireworm stings. The expression of metalloproteinase M12 and phospholipases could also explain the reported cases of local necrosis. The intense itching and the occasional reports of anaphylactic reactions can be explained by the activities of S10 peptidases and CAP proteins, which are major allergens in hymenopteran venoms. Lastly, the partial paresthesia and numbness reported in the affected area is consistent with the expression of several neurotoxins (ShKT, spider toxins, and turripeptides) that can cause paralysis, block the neuromuscular transmission or function as anesthetics, and also contribute to the severe pain induced by fireworm stings.

Conclusion

This is the first molecular comparative analysis identifying toxin-like genes in fireworm transcriptomes. Our results show that the fireworm species *H. carunculata*, *E. complanata*, and *P. jeffreysii* express a wide diversity of transcripts coding for toxin homologs that represent 13 known toxin classes. These toxins have been convergently recruited into a wide array of animal venoms and their activities explain the symptoms and physiological reactions observed after fireworm stings. Although the debate of whether fireworms use their calcareous dorsal chaetae or an alternative delivery system to inject toxins still remains to be clarified (see Schulze et al. 2017; Tilic et al. 2017), we provide compelling evidence suggesting that fireworms are venomous animals that use a complex mixture of toxins as a defensive mechanism. The toxin homologs identified in this study represent a valuable resource to further investigate fireworm venom systems and greatly contribute to the venomics toolkit for deciphering venom diversity and evolution in the animal kingdom.

Supplementary Material

Acknowledgments

Literature Cited

Aguilar MB, de la Rosa RAC, Falcón A, Olivera BM, Heimer de la Cotera EP. 2009. Peptide pal9a from the venom of the turrid snail *Polystira albida* from the Gulf of Mexico: purification, characterization, and comparison with P-conotoxin-like (framework IX) conoidean peptides. Peptides 30(3):467–476.

Ahrens J, et al. 2013. The curious case of *Hermodice carunculata* (Annelida: Amphinomidae): evidence for genetic homogeneity throughout the Atlantic Ocean and adjacent basins. Mol Ecol. 22(8):2280–2291.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389–3402.

Amarant T, Burkhart W, LeVine H, Arocha-Pinango CL, Parikh I. 1991. Isolation and complete amino acid sequence of two fibrinolytic proteinases from the toxic saturnid caterpillar *Lonomia achelous*. Biochim Biophys Acta 1079(2):214–221.

Amorim AM, et al. 2015. Transcripts involved in hemostasis: exploring salivary complexes from *Haementeria vizottoi* leeches through transcriptomics, phylogenetic studies and structural features. Toxicon 106:20–29.

Anderluh G, Maček P. 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). Toxicon 40(2):111–124.

Andersen JF, Gudderra NP, Francischetti IMB, Ribeiro JMC. 2005. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. Arch Insect Biochem Physiol. 58(2):97–105.

Andrade SCS, et al. 2015. Articulating 'Archiannelids': phylogenomics and annelid relationships, with emphasis on meiofaunal taxa. Mol Biol Evol. 32(11):2860–2875.

Asgari S, Zhang G, Zareie R, Schmidt O. 2003. A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph. Insect Biochem Mol Biol. 33(10):1017–1024.

Barghi N, Concepcion GP, Olivera BM, Lluisma AO. 2015. Comparison of the venom peptides and their expression in closely related conus species: insights into adaptive post-speciation evolution of Conus exogenomes. Genome Biol Evol. 7(6):1797–1814.

Barroso R, Klautau M, Solé-Cava AM, Paiva PC. 2010. *Eurythoe complanata* (Polychaeta: Amphinomidae), the 'cosmopolitan' fireworm, consists of at least three cryptic species. Mar Biol. 157(1):69–80.

Bayrhuber M, et al. 2005. Conkunitzin-S1 is the first member of a new Kunitz-type neurotoxin family: structural and functional characterization. J Biol Chem. 280(25):23766–23770.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120.

Bon C, Saliou B, Thieffry M, Manaranche R. 1985. Partial purification of alpha-glycerotoxin, a presynaptic neurotoxin from the venom glands of the polychaete annelid *Glycera convoluta*. Neurochem Int. 7(1):63–75.

Borda E, et al. 2015. Revamping Amphinomidae (Annelida: Amphinomida), with the inclusion of Notopygos. Zool Scr. 44(3):324–333.

Borda E, Kudenov JD, Bienhold C, Rouse GW. 2012. Towards a revised Amphinomidae (Annelida, Amphinomida): description and affinities of a new genus and species from the Nile Deep-sea Fan, Mediterranean Sea. Zool Scr. 41(3):307–325.

Bose U, et al. 2017. Multiomics analysis of the giant triton snail salivary gland, a crown-of-thorns starfish predator. Sci Rep. 7(1):1–14.

Braud S, Bon C, Wisner A. 2000. Snake venom proteins acting on hemostasis. Biochimie 82(9-10):851–859.

Brinkman DL, et al. 2015. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. BMC Genomics 16:407.

Calvete JJ, Marcinkiewicz C, Sanz L. 2007. Snake venomics of *Bitis gabonica gabonica*. Protein family composition, subunit organization of venom toxins, and characterization of dimeric disintegrins bitisgabonin-1 and bitisgabonin-2. J Proteome Res. 6(1):326–336.

Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. 2013. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol Evol. 28(4):219–229.

Castañeda O, et al. 1995. Characterization of a potassium channel toxin from the Caribbean Sea anemone *Stichodactyla helianthus*. Toxicon 33(5):603–613.

Castañeda O, Harvey AL. 2009. Discovery and characterization of cnidarian peptide toxins that affect neuronal potassium ion channels. Toxicon 54(8):1119–1124.

Chagas AC, et al. 2013. A deep insight into the sialotranscriptome of the mosquito, *Psorophora albipes*. BMC Genomics 14:875.

Chmelař J, Kotál J, Langhansová H, Kotsyfakis M. 2017. Protease inhibitors in tick saliva: the role of serpins and cystatins in tick-host-pathogen interaction. Front Cell Infect Microbiol. 7:216.

Choo YM, et al. 2010. Dual function of a bee venom serine protease: prophenoloxidase-activating factor in arthropods and fibrin(ogen)-olytic enzyme in mammals. PLoS One 5:e10393.

Choo YM, et al. 2012. Antifibrinolytic role of a bee venom serine protease inhibitor that acts as a plasmin inhibitor. PLoS One 7:e32269.

Cipriani V, et al. 2017. Correlation between ontogenetic dietary shifts and venom variation in Australian brown snakes (*Pseudonaja*). Comp Biochem Physiol C Toxicol Pharmacol. 197:53–60.

Colinet D, et al. 2013. Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: the case of Leptopilina parasitoids of Drosophila. Insect Biochem Mol Biol. 43(7):601.

Cuypers E, Peigneur S, Debaveye S, Shiomi K, Tytgat J. 2011. TRPV1 channel as new target for marine toxins: example of gigantoxin i, a sea

anemone toxin acting via modulation of the PLA2 pathway. Acta Chim Slov. 58(4):735–741.

Da Silveira RB, et al. 2007. Identification, cloning, expression and functional characterization of an astacin-like metalloprotease toxin from *Loxosceles intermedia* (brown spider) venom. Biochem J. 406(2):355–363.

Daly JW. 1995. The chemistry of poisons in amphibian skin. Proc Natl Acad Sci U S A. 92(1):9–13.

Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27(8):1164–1165.

Dennis EA. 1983. Phospholipases. Enzym 16:307–353.

Ding L, et al. 2015. A new Kunitz-type plasmin inhibitor from scorpion venom. Toxicon 106:7–13.

Dubovskii PV, et al. 2015. Latarcins: versatile spider venom peptides. Cell Mol Life Sci. 72(23):4501–4522.

Dumbacher JP, et al. 2004. Melyrid beetles (Choresine): a putative source for the batrachotoxin alkaloids found in poison-dart frogs and toxic passerine birds. Proc Natl Acad Sci U S A. 101(45):15857–15860.

Eckert GJ. 1985. Absence of toxin-producing parapodial glands in amphinomid polychaetes (fireworms). Toxicon 23(2):350–353.

Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Informatics 23(1):205–211.

Fang KS, Vitale M, Fehlner P, King TP. 1988. cDNA cloning and primary structure of a white-face hornet venom allergen, antigen 5. Proc Natl Acad Sci U S A. 85(3):895–899.

Fernández J, Gutiérrez JM, Calvete JJ, Sanz L, Lomonte B. 2016. Characterization of a novel snake venom component: Kazal-type inhibitor-like protein from the arboreal pitviper *Bothriechis schlegelii*. Biochimie 125:83–90.

Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39(Suppl):W29.

Finn RD, et al. 2015. 2015. HMMER web server: 2015 Update. Nucleic Acids Res. 43(W1):W30–W38.

Flower DR. 1996. The lipocalin protein family: structure and function. Biochem J. 318(1):1–14.

Francischetti IMB, et al. 2013. The 'Vampirome': transcriptome and proteome analysis of the principal and accessory submaxillary glands of the vampire bat *Desmodus rotundus*, a vector of human rabies. J Proteomics 82:288–319.

Fry BG, et al. 2008. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). Mol Cell Proteomics 7(2):215–246.

Fry BG, et al. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annu Rev Genomics Hum Genet. 10:483–511.

Fry BG, Roelants K, Norman JA. 2009. Tentacles of venom: toxic protein convergence in the kingdom animalia. J Mol Evol. 68(4):311–321.

Gibbs GM, Roelants K, O'Bryan MK. 2008. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins —roles in reproduction, cancer, and immune defense. Endocr Rev. 29(7):865–897.

Glasby CJ, et al. 2000. Class Polychaeta. In: Beesley PL, Ross GJB, Glasby CJ, editors. Polychaetes & allies: the southern synthesis. Fauna of Australia. Vol 4A. Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula. Melbourne: CSIRO Publishing. p. 1–296.

Gleibs S, Mebs D, Werding B. 1995. Studies on the origin and distribution of palytoxin in a Caribbean coral reef. Toxicon 33(11):1531–1537.

Gonzales DT. h T, Saloma CP. 2014. A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. Toxicon 92:66–74.

Gorson J, et al. 2015. Molecular diversity and gene evolution of the venom arsenal of terebridae predatory marine snails. Genome Biol Evol. 7(6):1761–1778.

Gorson JM, Verdes A, Shin H, Khawaja S, Holford M. 2016. Characterizing the molecular diversity and function of Terebridae venom. Integr Comp Biol. Vol. 56: p E77–E77.

Grabherr MG, et al. 2011. Trinity: recontructing a full-length transcriptome assembly without a genome from RNA-Seq data. Nat Biotechnol. 29(7):644–652.

Haas BJ, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8(8):1494–1512.

Haddad V. 2015. Miscellaneous marine toxins of medical significance. In: Gopalakrishnakone P, Vidal Haddad Jr., William R. Kem, Aurelia Tubaro, Euikyung Kim, editors. Marine and freshwater toxins. Dordrecht: Springer Netherlands. p. 1–15.

Hartmann S, Lucius R. 2003. Modulation of host immune responses by nematode cystatins. Int J Parasitol. 33(11):1291–1302.

Hu B, et al. 2011. Purification and characterization of gigantoxin-4, a new actinoporin from the sea anemone *Stichodactyla gigantea*. Int J Biol Sci. 7(6):729–739.

Huang C, et al. 2016. The transcriptome of the zoanthid Protopalythoa variabilis (Cnidaria, Anthozoa) predicts a basal repertoire of toxin-like and venom-auxiliary polypeptides. Genome Biol Evol. 8(9):3045–3064.

Irving JA, Pike RN, Lesk AM, Whisstock JC. 2000. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. Genome Res. 10(12):1845–1864.

Isaeva MP, et al. 2012. A new multigene superfamily of Kunitz-type protease inhibitors from sea anemone *Heteractis crispa*. Peptides 34(1):88–97.

Johnson M, et al. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36(Web Server issue):W5–W9.

Jungo F, Bairoch A. 2005. Tox-Prot, the toxin protein annotation program of the Swiss-Prot protein knowledgebase. Toxicon 45(3):293–301.

Jungo F, Bougueleret L, Xenarios I, Poux S. 2012. The UniProtKB/Swiss-Prot Tox-Prot program: a central hub of integrated venom protein data. Toxicon 60(4):551–557.

Junqueira-de-Azevedo ILM, Campos PF, Ching ATC, Mackessy SP. 2016. Colubrid venom composition: an -omics perspective. Toxins (Basel) 8(8):230.

Kaas Q, Westermann JC, Halai R, Wang CKL, Craik DJ. 2008. ConoServer, a database for conopeptide sequences and structures. Bioinformatics 24(3):445–446.

Kaas Q, Yu R, Jin AH, Dutertre S, Craik DJ. 2012. ConoServer: updated content, knowledge, and discovery tools in the conopeptide database. Nucleic Acids Res. 40(D1):D325.

Kato H, et al. 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(2):184–191.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kearse M, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647–1649.

Kim BY, et al. 2013. Antimicrobial activity of a honeybee (*Apis cerana*) venom Kazal-type serine protease inhibitor. Toxicon 76:110–117.

King TP, Lu G, Gonzalez M, Qian N, Soldatova L. 1996. Yellow jacket venom allergens, hyaluronidase and phospholipase: sequence similarity and antigenic cross-reactivity with their hornet and wasp homologs and possible implications for clinical allergy. J Allergy Clin Immunol. 98(3):588–600.

Kini RM. 2003. Excitement ahead: structure, function and mechanism of snake venom phospholipase A2 enzymes. Toxicon 42(8):827–840.

Kita M, et al. 2004. Blarina toxin, a mammalian lethal venom from the short-tailed shrew *Blarina brevicauda*: isolation and characterization. Proc Natl Acad Sci U S A. 101(20):7542–7547.

Kobayashi H, Ohta N, Umeda M. 2004. Biology of Lysenin, a protein in the coelomic fluid of the earthworm *Eisenia foetida*. Int Rev Cytol. 236:45–99.

Kordis D, Turk V. 2009. Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. BMC Evol Biol. 9:266.

Kozlov SA, et al. 2006. Latarcins, antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaevi* (Zodariidae) that exemplify biomolecular diversity. J Biol Chem. 281(30):20983–20992.

Kratzschmar J, et al. 1991. The plasminogen activator family from the salivary gland of the vampire bat *Desmodus rotundus*: cloning and expression. Gene 105(2):229–237.

Kubista H, et al. 2007. CSTX-1, a toxin from the venom of the hunting spider *Cupiennius salei*, is a selective blocker of L-type calcium channels in mammalian neurons. Neuropharmacology 52(8):1650–1662.

Kuhn-Nentwig L, et al. 2012. A venom-derived neurotoxin, Cstx-1, from the spider *Cupiennius salei* exhibits cytolytic activities. J Biol Chem. 287(30):25640–25649.

Kvist S, Brugler MR, Goh TG, Giribet G, Siddall ME. 2014. Pyrosequencing the salivary transcriptome of *Haemadipsa interrupta* (Annelida: Clitellata: Haemadipsidae): anticoagulant diversity and insight into the evolution of anticoagulation capabilities in leeches. Invertebr Biol. 133(1):74–98.

Kvist S, et al. 2016. When predator becomes prey: investigating the salivary transcriptome of the shark-feeding leech *Pontobdella macrothela* (Hirudinea: Piscicolidae). Zool J Linn Soc.179(4):725–737.

Kvist S, Min GS, Siddall ME. 2013. Diversity and selective pressures of anticoagulants in three medicinal leeches (Hirudinida: Hirudinidae, Macrobdellidae). Ecol Evol. 3(4):918–933.

Laskowski M, Kato I. 1980. Protein inhibitors of proteinases. Annu Rev Biochem. 49:593–626.

Law RHP, et al. 2006. An overview of the serpin superfamily. Genome Biol. 7(5):216.

Leffler AE, et al. 2017. Discovery of peptide ligands through docking and virtual screening at nicotinic acetylcholine receptor homology models. Proc Natl Acad Sci U S A. 114(38):201703952.

Li R, et al. 2013. Proteome and phosphoproteome analysis of honeybee (*Apis mellifera*) venom collected from electrical stimulation and manual extraction of the venom gland. BMC Genomics 14:766.

Liu G, et al. 2015. Global transcriptome analysis of the tentacle of the jellyfish *Cyanea capillata* using deep sequencing and expressed sequence tags: insight into the toxin-and degenerative disease-related transcripts. PLoS One 10:e0142680.

Liu ZC, et al. 2012. Venomic and transcriptomic analysis of centipede *Scolopendra subspinipes dehaani*. J Proteome Res. 11(12):6197–6212.

Lopes-Ferreira M, et al. 2011. Structural and biological characterization of Nattectin, a new C-type lectin from the venomous fish *Thalassophryne nattereri*. Biochimie 93(6):971–980.

Low DHW, et al. 2013. Dracula's children: molecular evolution of vampire bat venom. J Proteomics 89:95–111.

Magalhães GS, et al. 2006. Transcriptome analysis of expressed sequence tags from the venom glands of the fish *Thalassophryne nattereri*. Biochimie 88(6):693–699.

Mans BJ, Louw AI, Neitz AWH. 2003. The major tick salivary gland proteins and toxins from the soft tick, *Ornithodoros savignyi*, are part of the tick lipocalin family: implications for the origins of tick toxicoses. Mol Biol Evol. 20(7):1158–1167.

Martinson EO, Mrinalini, Kelkar YD, Chang CH, Werren JH. 2017. The evolution of venom by co-option of single-copy genes. Curr Biol. 27:2007–2013.e8.

Menaldo DL, et al. 2012. Biochemical characterization and comparative analysis of two distinct serine proteases from Bothrops pirajai snake venom. Biochimie 94(12):2545–2558.

Michel C, Keil B. 1975. Biologically active proteins in the venomous glands of the polychaetous annelid, *Glycera convoluta* Keferstein. Comp Biochem Physiol B Comp Biochem. 50(1):29–33.

Milne TJ, Abbenante G, Tyndall JDA, Halliday J, Lewis RJ. 2003. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. J Biol Chem. 278(33):31105–31110.

Min G-S, Sarkar IN, Siddall ME. 2010. Salivary transcriptome of the North American medicinal leech, *Macrobdella decora*. J Parasitol. 96(6):1211–1221.

Modica MV, Lombardo F, Franchini P, Oliverio M. 2015. The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). BMC Genomics 16:441.

Moran Y, et al. 2013. Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. Mar Biotechnol. 15(3):329–339.

Morita T. 2005. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. Toxicon 45(8):1099–1114.

Mulder NJ, Apweiler R. 2008. The InterPro database and tools for protein domain analysis. Curr Protoc Bioinformatics 21:2.7:2.7.1–2.7.18.

Nakamura K, et al. 2008. Complanine, an inflammation-inducing substance isolated from the marine fireworm *Eurythoe complanata*. Org Biomol Chem. 6(12):2058–2060.

Nakamura K, et al. 2010. Neocomplanines A and B, a complanine family isolated from the marine fireworm. J Nat Prod. 73(3):303–305.

Negulescu H, et al. 2015. A Kazal-type serine protease inhibitor from the defense gland secretion of the subterranean termite *Coptotermes formosanus* Shiraki. PLoS One 10:e0125376.

Nevalainen TJ. 2008. Phospholipases A2 in the genome of the sea anemone *Nematostella vectensis*. Comp Biochem Physiol. 3(3):226–233.

Ottuso P. 2013. Aquatic dermatology: encounters with the denizens of the deep (and not so deep) a review. Part I: the invertebrates. Int J Dermatol. 52(2):136–152.

Paddock CD, et al. 2001. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. J Immunol. 167(5):2694–2699.

Page MJ, Di Cera E. 2008. Serine peptidases: classification, structure and function. Cell Mol Life Sci. 65(7–8):1220–1236.

Perkin LC, Friesen KS, Flinn PW, Oppert B. 2015. Venom gland components of the ectoparasitoid wasp, *Anisopteromalus calandrae*. J Venom Res. 6:19–37.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8(10):785–786.

Phuong MA, Mahardika GN, Alfaro ME. 2016. Dietary breadth is positively correlated with venom complexity in cone snails. BMC Genomics 17:401.

Poirié M, Colinet D, Gatti JL. 2014. Insights into function and evolution of parasitoid wasp venoms. Curr Opin Insect Sci. 6:52.

Ponce D, Brinkman DL, Potriquet J, Mulvenna J. 2016. Tentacle transcriptome and venom proteome of the pacific sea nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa). Toxins (Basel) 8(4):102.

Qian C, Fang Q, Wang L, Ye GY. 2015. Molecular cloning and functional studies of two Kazal-type serine protease inhibitors specifically expressed by *Nasonia vitripennis* venom apparatus. Toxins (Basel) 7(8):2888–2905.

Reis CV, et al. 2006. Lopap, a prothrombin activator from *Lonomia obliqua* belonging to the lipocalin family: recombinant production, biochemical characterization and structure-function insights. Biochem J. 398(2):295–302.

Ribeiro JMC, et al. 2007. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. BMC Genomics 8:6.

Richter S, et al. 2017. Comparative analyses of glycerotoxin expression unveil a novel structural organization of the bloodworm venom system. BMC Evol Biol. 17(1):64.

Rimphanitchayakit V, Tassanakajon A. 2010. Structure and function of invertebrate Kazal-type serine proteinase inhibitors. Dev Comp Immunol. 34(4):377–386.

Rodríguez De La Vega RC, Giraud T. 2016. Intragenome diversity of gene families encoding toxin-like proteins in venomous animals. Integr Comp Biol. 56(5):938–949.

Rokyta DR, Lemmon AR, Margres MJ, Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). BMC Genomics 13:312.

Rouse G, Pleijel F. 2001. Polychaetes. Oxford: Oxford University Press.

Rouse GW, Fauchald K. 1997. Cladistics and polychaetes. Zool Scr. 26(2):139–204.

Ruder T, et al. 2013. Molecular phylogeny and evolution of the proteins encoded by Coleoid (Cuttlefish, Octopus, and Squid) posterior venom glands. J Mol Evol. 76(4):192–204.

Santos A, et al. 2007. The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). Insect Biochem Mol Biol. 37(7):702–712.

Savitzky AH, et al. 2012. Sequestered defensive toxins in tetrapod vertebrates: principles, patterns, and prospects for future studies. Chemoecology 22(3):141–158.

Schulze A, Grimes CJ, Rudek TE. 2017. Tough, armed and omnivorous: *Hermodice carunculata* (Annelida: Amphinomidae) is prepared for ecological challenges. J Mar Biol Assoc UK. 97(5):1–6.

Schwarz A, Valdés JJ, Kotsyfakis M. 2012. The role of cystatins in tick physiology and blood feeding. Ticks Tick Borne Dis. 3(3):117–127.

Schweitz H, et al. 1995. Kalicludines and Kaliseptine: two different classes of sea anemone toxins for voltage sensitive K+ cannels. J Biol Chem. 270(42):25121–25126.

Siddall ME, Brugler MR, Kvist S. 2016. Comparative transcriptomic analyses of three species of Placobdella (Rhynchobdellida: Glossiphoniidae) confirms a single origin of blood feeding in leeches. J Parasitol. 102(1):143–150.

Sigle LT, Ramalho-Ortigão M. 2013. Kazal-type serine proteinase inhibitors in the midgut of *Phlebotomus papatasi*. Mem Inst Oswaldo Cruz. 108(6):671–678.

Smith ML. 2002. Cutaneous problems related to coastal and marine worms. Dermatol Ther. 15(1):34–37.

Soares TS, et al. 2012. Expression and functional characterization of boophilin, a thrombin inhibitor from Rhipicephalus (Boophilus) microplus midgut. Vet Parasitol. 187(3-4):521–528.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21):2688–2690.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.

Stoner EW, Layman CA. 2015. Bristle worms attack: Benthic jellyfish are not trophic dead ends. Front Ecol Environ. 13(4):226–227.

Sunagar K, et al. 2014. Intraspecific venom variation in the medically significant Southern Pacific Rattlesnake (*Crotalus oreganus helleri*): biodiscovery, clinical and evolutionary implications. J Proteomics 99:68–83.

Takác P, et al. 2006. Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) salivary glands. J Exp Biol. 209(Pt 2):343–352.

Teichert RW, Olivera BM, McIntosh JM, Bulaj G, Horvath MP. 2015. The molecular diversity of conoidean venom peptides and their targets: from basic research to therapeutic applications. In: Glenn F. King, editor. Venoms to drugs: venom as a source for the development of human therapeutics. Cambridge: Royal Society of Chemistry. p. 163–203.

Thieffry M, Bon C, Manaranche R, Saliou B, Israël M. 1982. Partial purification of the *Glycera convoluta* venom components responsible for its presynaptic effects. J Physiol (Paris) 78:343–347.

Tilic E, Pauli B, Bartolomaeus T. 2017. Getting to the root of fireworms' stinging chaetae—chaetal arrangement and ultrastructure of *Eurythoe complanata* (Pallas, 1766) (Amphinomida). J Morphol. 278(6):865–876.

Tsujimoto H, et al. 2012. Simukunin from the salivary glands of the black fly *Simulium vittatum* inhibits enzymes that regulate clotting and inflammatory responses. PLoS One 7:e29964.

Undheim EAB, et al. 2015. Weaponization of a hormone: convergent recruitment of hyperglycemic hormone into the venom of arthropod predators. Structure 23(7):1283–1292.

Vaiyapuri S, Wagstaff SC, Harrison RA, Gibbins JM, Hutchinson EG. 2011. Evolutionary analysis of novel serine proteases in the Venom Gland transcriptome of *Bitis gabonica rhinoceros*. PLoS One 6:e21532.

Vassilevski AA, Kozlov SA, Grishin EV. 2009. Molecular diversity of spider venom. Biochemistry 74(13):1505–1534.

Veiga ABG, Ribeiro JMC, Guimarães JA, Francischetti IMB. 2005. A catalog for the transcripts from the venomous structures of the caterpillar *Lonomia obliqua*: identification of the proteins potentially involved in the coagulation disorder and hemorrhagic syndrome. Gene 355:11–27.

Verdes A, et al. 2016. From mollusks to medicine: a venomics approach for the discovery and characterization of therapeutics from terebridae peptide Toxins. Toxins (Basel) 8(4):117.

von Reumont BM, Campbell LI, Jenner RA. 2014a. Quo vadis venomics? A roadmap to neglected venomous invertebrates. Toxins (Basel) 6(12):3488–3551.

von Reumont BM, Campbell LI, Richter S, et al. 2014b. A polychaete's powerful punch: venom gland transcriptomics of Glycera reveals a complex cocktail of toxin homologs. Genome Biol Evol. 6(9):2406–2423.

von Reumont BM, Blanke A, et al. 2014c. 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(1):48–58.

von Reumont BM, Undheim EAB, Jauss RT, Jenner RA. 2017. Venomics of remipede crustaceans reveals novel peptide diversity and illuminates the venom's biological role. Toxins (Basel) 9(12):234.

Vonk FJ, et al. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proc Natl Acad Sci U S A. 110(51):20651–20656.

Walker AA, et al. 2017. Melt with this kiss: paralyzing and liquefying venom of the assassin bug *Pristhesancus plagipennis* (Hemiptera: Reduviidae). Mol Cell Proteomics 16(4):552–566.

Wan H, et al. 2013. A spider-derived kunitz-type serine protease inhibitor that acts as a plasmin inhibitor and an elastase inhibitor. PLoS One 8:e53343.

Wang H, et al. 2012. Functional peptidomics of amphibian skin secretion: a novel Kunitz-type chymotrypsin inhibitor from the African hyperoliid frog, *Kassina senegalensis*. Biochimie 94(3): 891–899.

Wang Y, Coleman-Derr D, Chen G, Gu YQ. 2015. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res. 43(W1): W78.

Watkins M, Hillyard DR, Olivera BM. 2006. Genes expressed in a turrid venom duct: divergence and similarity to conotoxins. J Mol Evol. 62(3):247–256.

Wei C, Chen J. 2012. A novel lipocalin homologue from the venom gland of *Deinagkistrodon acutus* similar to mammalian lipocalins. J Venom Anim Toxins Incl Trop Dis. 18:16–23.

Weigert A, Bleidorn C. 2016. Current status of annelid phylogeny. Org Divers Evol. 16(2):345–362.

Weigert A, et al. 2014. Illuminating the base of the annelid tree using transcriptomics. Mol Biol Evol. 31(6):1391–1401.

Whelan NV, Kocot KM, Santos SR, Halanych KM. 2014. Nemertean toxin genes revealed through transcriptome sequencing. Genome Biol Evol. 6(12):3314–3325.

Whittington CM, et al. 2010. Novel venom gene discovery in the platypus. Genome Biol. 11(9):R95.

Wiklund H, Nygren A, Pleijel F, Sundberg P. 2008. The phylogenetic relationships between Amphinomidae, Archinomidae and Euphrosinidae (Amphinomida: Aciculata: Polychaeta), inferred from molecular data. J Mar Biol Assoc UK. 88(03):509–513.

Wilcox CL, Yanagihara AA. 2016. Heated debates: hot-water immersion or ice packs as first aid for cnidarian envenomations? Toxins (Basel) 8(4):97.

Wong ESW, et al. 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(11):1354–1364.

Wu J, et al. 2011. Proteomic analysis of skin defensive factors of tree frog *Hyla simplex*. J Proteome Res. 10(9):4230–4240.

Xia X, et al. 2013. Cloning and molecular characterization of BumaMPs1, a novel metalloproteinase from the venom of scorpion *Buthus martensi* Karsch. Toxicon 76:234–238.

Xu Y, Bruno JF, Luft BJ. 2005. Identification of novel tick salivary gland proteins for vaccine development. Biochem Biophys Res Commun. 326(4):901–904.

Yamazaki Y, Morita T. 2004. Structure and function of snake venom cysteine-rich secretory proteins. Toxicon 44(3):227–231.

Yan Z, et al. 2017. A venom serpin splicing isoform of the endoparasitoid wasp *Pteromalus puparum* suppresses host prophenoloxidase cascade by forming complexes with host hemolymph proteinases. J Biol Chem. 292(3):1038–1051.

Yang X, et al. 2009. A novel serine protease inhibitor from the venom of *Vespa bicolor* Fabricius. Comp Biochem Physiol B Biochem Mol Biol. 153(1):116–120.

Yuan CH, et al. 2008. Discovery of a distinct superfamily of kunitz-type toxin (KTT) from Tarantulas. PLoS One 3(10):e3414.

Zdobnov EM, Apweiler R. 2001. InterProScan–an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17(9):847–848.

Zelensky AN, Gready JE. 2005. The C-type lectin-like domain superfamily. FEBS J. 272(24):6179–6217.

Associate editor: Maria Costantini