BMC Genomics

RESEARCH

Open Access

Multi-omics characterization of toxin expression and producing organs in the predatory gastropods *Monoplex corrugatus* and *Stramonita haemastoma*

Allan Ringeval[1], Maria Vittoria Modica[1,2], Yuri Kantor[3], Manuel Jimenez Tenorio[1,4], Juan Carlos G. Galindo[5], Nicolas Puillandre[1] and Sarah Farhat[1*]

Abstract

Background The exploration of toxin diversity is crucial for understanding the evolutionary adaptation of venomous taxa. Despite being active venomous predators, neogastropods are largely understudied beyond Conidae. This study targets two predatory gastropods, *Monoplex corrugatus* and *Stramonita haemastoma*, aiming to characterize their toxin-producing tissues, evaluate the diversity and function of their toxins, and compare gene expression profiles across tissues.
Specimens of both species were dissected to isolate multiple replicates of secretory glands and other tissues. Transcriptomic data were complemented by shotgun proteomics for *S. haemastoma* and used to identify putative toxin genes using the DeTox pipeline. Differentially expressed genes were identified and putative toxins were manually annotated.

Results The study identified 2,565 and 1,777 putative toxins in *S. haemastoma* and *M. corrugatus*, respectively. Salivary glands were the major toxin-producing organ in both species, with additional toxin expression in mid-esophageal and accessory salivary glands. Manual annotation confidently identified 115 –*S. haemastoma*– and 143 –*M. corrugatus*– venom proteins, highlighting significant interspecies and inter-tissue differences. Functional categorization revealed the presence of enzymatic and peptide toxins, as well as venom-processing proteins, with *M. corrugatus* showing expression in non-secretory tissues. Despite their phylogenetic distance, shared orthologs were identified between the two species, namely for venom-processing proteins like calglandulin and disulfide isomerases, suggesting conserved functions. Toxins unique to each species analyzed, including echotoxins and plancitoxins in *M. corrugatus*, indicate lineage-specific venom adaptations. Proteomic validation supported transcriptomic predictions in *S. haemastoma*.

Conclusions These findings underscore the value of multi-omics approaches for toxin discovery and for investigating the complexity of gastropod venom evolution and expand our understanding of how venom systems evolve and

*Correspondence:
Sarah Farhat
sarah.farhat@mnhn.fr

Full list of author information is available at the end of the article

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

diversify in marine snails, highlighting both shared and unique toxin strategies that may reflect different ecological adaptations.

Keywords Neogastropoda, Venom, Differential gene expression, Proteotranscriptomics

Background

Animal toxins have garnered significant interest due to their potential therapeutic applications as analgesics, coagulation factors, or anticancer agents. In this context, neogastropods (Mollusca, Gastropoda), which encompass most of the predatory gastropods, represent an exceptional reservoir of toxins, making them highly relevant for both evolutionary and pharmaceutical research [1, 2]. However, the studies within this lineage covers only a small fraction of species, and little is known regarding the diversity and origin of their toxins. Current research has primarily focused on cone snails (family Conidae), which encompass over 1,000 species [3]. Cone snail toxins, termed "conotoxins" due to their historical association with this family, form a large group of structurally and functionally well-characterized neurotoxic peptides along with other toxic peptides as conoinsulins [4] and pore-forming conoporins [5]. These peptides primarily serve either to paralyze prey, or to deter predators, with estimates suggesting that each species produces more than 200 unique toxins [6, 7]. In contrast, knowledge of toxin diversity in other predatory gastropods families remains sparse, with only a few species studied [8, 9]. For example, in *Cumia reticulata* (Blainville, 1829) (Colubrariidae, now *Cumia intertexta* (Hebling, 1779)), the salivary glands produce anticoagulant compounds [10], while *Vexillum* species (Costellariidae) express "vexitoxins" structurally similar to conotoxins [11]. In contrast, *Monoplex parthenopeus* (Salis Marschlins, 1793) (Cymatiidae) has been investigated using more traditional biochemical approaches, which revealed the presence of pore-forming proteins in its salivary secretion [9], but without the benefit of modern transcriptomic tools. These findings emphasize the necessity for broader and more comprehensive studies to fully characterize toxin diversity, identify the organs responsible for toxin production, and elucidate how they affect predation strategy and feeding efficiency across a wider range of predatory gastropod lineages.

Venoms are complex mixtures of bioactive compounds that exhibit remarkable variability, largely driven by shifting selective pressures – positive selection drives early diversification linked to ecological expansion, while purifying selection later maintains functional stability [9]. This evolutionary dynamic helps to explain the wide range of venom compositions observed across species with different diets, predatory strategies, and developmental stages. For example, piscivorous *Conus* species that target fast-moving prey typically secrete venoms capable of inducing paralysis [12]. This contrasts with hematophagous species where anticoagulant toxins facilitate blood ingestion through the proboscis [10]. Variations also occur between families with similar prey preferences. For example, toxin sequences in Turridae (vermivorous) have been shown to be longer than those in Conidae (also mostly vermivorous), which may imply structural variability [13, 14]. Additionally, venom composition and predation strategy can vary between life stages as shown in *Conus magus* Linnaeus, 1758 [15], where juveniles and adults express different toxin profiles and feed on different preys. Collectively, these variations underscore the challenges of comparing toxins across different lineages and highlight the limitations of similarity-based research. Nevertheless, in the absence of in vitro assays, similarity approaches, recently complemented with methods based on structural analyses [16, 17], constitute the preferred strategy to detect toxins in poorly studied lineages. Indeed, toxins so far identified in predatory gastropods share several features, such as a conserved signal peptide [18] and (generally) a cysteine pattern - characterized by the number and position of disulfide bridges [19] - that enable their classification into families associated with potential functions [20]. Combining sequence similarity with structural identifications provides strong evidence to annotate new candidate toxins, although these predictions remain hypothetical until confirmed by proteomic analysis and functional validation.

Apart from Conoidea (Neogastropoda), where toxins are produced in the venom gland and that have been extensively studied, especially Conidae [21, 22], the organs responsible for toxin production in other predatory gastropods remain largely unknown. Most predatory gastropods are characterized by a modified digestive system with an extensible proboscis, that facilitates capture and consumption of prey [23]. The glandular structures associated with the anterior foregut include primary and accessory salivary glands. In predatory gastropods, studies suggest that these salivary glands, which discharge their secretion at the tip of the proboscis, in close contact with the prey, may play a role in toxin production [11, 24].

Similarly, the mid-esophageal gland plays a central role in digestion by secreting a range of digestive enzymes [25]. However, evidence also suggests that this gland may be evolutionarily homologous to the venom apparatus found in conoideans [26]. In fact, it has been hypothesized that ancestral digestive functions were co-opted

for toxin production [11] through the redistribution of ancestral digestive roles to other tissues, enabling exclusive specialization in toxin production via enhanced secretory and cellular stress pathways [26]. In most non-conoidean predatory gastropods, the mid-esophageal gland is thought to have evolved into the gland of Leiblein, whose secretions are discharged in the mid esophagus, reflecting a complex evolutionary history of foregut specialization [26]. These observations indicate that the toxin production mechanisms and their associated organs could be more diverse than currently known.

In this study, we performed toxin identification and characterization across various tissues in two phylogenetically distant [27] and poorly studied species of predatory gastropods, namely *Monoplex corrugatus* (Lamarck, 1816) and *Stramonita haemastoma* (Linnaeus, 1767), in order to uncover lineage-specific toxin evolution and gain broader insights into venom diversification. *Monoplex corrugatus* is a predatory species belonging to the Cymatiidae family within the superfamily Tonnoidea (Littorinimorpha), recently confirmed as a member of the sister lineage of neogastropods [28]. Although information on its diet is limited, one study identified bivalves and ascidians as prey for another species within the *Monoplex* genus [29]. A preliminary screening identified toxins in the genome of *M. corrugatus*, some of which resembled conotoxins [30]. *Stramonita haemastoma* belongs to the Muricidae family within the superfamily Muricoidea (Neogastropoda). Its diet appears to be quite diverse, including bivalves, barnacles, and limpets [31], as well as polychaetes, including *Phragmatopoma lapidosa*, identified as a preferred prey in environments where other food items are scarce [32]. *Stramonita haemastoma* possesses the ability to bore holes into bivalve shells, although this hunting strategy is less energetically efficient and more time-consuming than feeding on tubeworms [32].

In order to characterize toxin-producing organs, we analyzed several organs that play diverse functions: in *M. corrugatus*, beside heart and kidney, transcriptomes were obtained also for the salivary glands, the mid-esophageal gland and the ctenidium, while for *S. haemastoma*, the transcriptomes of the salivary glands, the gland of Leiblein, the accessory salivary glands and the osphradium were obtained. The osphradium is a sensory structure placed on the internal side of the mantle, able to detect chemical elements in the water and crucial for prey detection. In close proximity with the osphradium, the ctenidium is a gill, a lamellar structure responsible of respiratory exchanges [33].

Our primary aims were to identify toxin-producing tissues, characterize conserved toxin genes shared between the two species, evaluate interspecies differences in toxin expression profiles, and elucidate the putative biological roles of newly identified toxins. Based on recent literature, we hypothesized that toxins are predominantly produced in secretory glands associated with the foregut. Despite their phylogenetic distance, we also anticipated that *M. corrugatus* and *S. haemastoma* might share a core set of common toxins, while simultaneously possessing unique toxin repertoires influenced by their genetic divergence, morphological differences, prey diversity, and predation strategies.

To test these hypotheses, we implemented an integrated genomic, transcriptomic, and proteomic approach. We analyzed multiple tissues –including specialized secretory glands and other tissues– selected specifically to capture the broadest possible range of functional diversity and to clearly highlight tissue-specific toxin expression patterns. This combined strategy allowed us to annotate putative toxin genes within existing genomic data, quantify differential gene expression using replicated transcriptomic analyses, validate many of these sequences by shotgun proteomics, and infer their biological functions. Overall, this work represents an initial, yet comprehensive, exploration of toxin diversity in these understudied species, shedding light on the specific tissues responsible for toxin production and identifying proteins potentially involved in toxin-like biological functions.

Methods

Biological material

Genome assemblies and transcriptomes of *Monoplex corrugatus* and *Stramonita haemastoma* from our previous publication [30] were used for this study. In addition, the following tissue samples were dissected (Supplementary file 1, Table S1), flash-frozen in liquid nitrogen and then stored at – 80 °C and sent to the Institut du Cerveau et de la Moëlle épinière (ICM, Paris, France) for library preparation and RNA sequencing: for *M. corrugatus*, the anterior lobe of the salivary gland (samples 1020.6, 1020.21, and 0521.112), the posterior lobe of the salivary gland (samples 1020.7, 1020.22, 0521.113) and the mid-esophageal (OSP) (samples 1020.8, 1020.20 and 0521.114) of three specimens (MNHN-IM-2019-21838, MNHN-IM-2019-21839 and MNHN-IM-2019-21861, respectively) and the ctenidium, heart and kidney (samples 0521.107, 0521.110 and 0521.11) of the specimen MNHN-IM-2019-21861; for *S. haemastoma*, the gland of Leiblein (GL) (samples 0521.4, 0521.100 and 0521.133), the salivary gland (SG) (samples 0521.5, 0521.101 and 0521.134), the osphradium (OSP) (samples 0521.3, 0521.95 and 0521.128) and the accessory salivary gland (ASG) (samples 0521.9, 0521.102 and 0521.135) of three specimens (MNHN-IM-2019-21862, MNHN-IM-2019-21863 and MNHN-IM-2019-21864, respectively), the accessory boring organs (sample 0521.10)

from the specimen MNHN-IM-2019-21862 and the whole bodies (samples 0521.145, 0521.146 and 0521.154) of three additional specimens (MNHN-IM-2019-21865, MNHN-IM-2019-21866 and MNHN-IM-2019-21867, respectively). For a summary of all specimens collected and tissues processed, see Supplementary file 1, Table S1.

RNA sequencing and transcriptome assemblies

Total RNA library preparation was carried out following the manufacturer's recommendations (Illumina Stranded Total RNA Prep with Ribo-Zero Plus). The final pooled library prep was sequenced on an Illumina NovaSeq 6000 using a 300-cycle cartridge.

The raw reads were quality-filtered and trimmed using fastp [v0.23.4] with "-w 16 --detect_adapter_for_pe --cut_front --cut_front_window_size 1 --cut_tail --cut_tail_window_size 1 --cut_right --cut_right_window_size 4 --cut_mean_quality 15 --length_required 25" parameters [34] and then assembled using Trinity [v2.15.1] with "--seqType fq --full_cleanup –no_salmon" [35] on the IFB (Institut français de bioinformatique) cluster. BUSCO scores, N50, transcript counts, and GC content were calculated for each assembly using BUSCO [v5.5.0] with "-m transcriptome -l metazoa_odb10" parameters and Transrate [v1.0.3] with default parameters [36]. As the metrics of the posterior lobe of the salivary gland were too low, they were combined with the anterior lobe of the salivary gland in *M. corrugatus*. In addition, the sequence identifiers for each sample were modified to include the code of the sample, the unique number for the individual and an abbreviation of the tissue name.

Repeated elements and gene re-annotation

Repeat families found in the genome assemblies of *M. corrugatus* and *S. haemastoma* were identified *de novo* and classified according to existing pipeline from other publications [30, 37, 38]. Briefly, the first step consisted in annotating and masking the predicted LTR-retrotransposons detected within the genomes by LTRHarvest from GenomeTools [v1.6.2] [39] with the parameters "-minlenltr 80 -maxlenltr 1200 -mindistltr 2500 -maxdistltr 11000 -similar 80.0". These sequences were then clustered using uclust with "-cluster_fast -id 0.8 -sort length -strand both" [40] and inserts were removed within each cluster. This procedure was done twice to maximize the clustering of repeated elements. Then, only consensus from each cluster was given as a library for RepeatMasker [v4.1.2.p1] with default parameters [41] to retrieve putatively LTR-retrotransposon copies. The last step uses RepeatModeler [v2.0.1] with default parameters [42] generating the library for RepeatMasker to complete the annotation of all other transposable element types. "Concatenate_sequences.py" [43] was used in concatenating hits closer than 500 bp and removing hits smaller than 300 bp in length.

As we had new transcriptomic data for both genomes, we reannotated the genes following the same methodology as in [30] with the addition of the previous gene annotation provided to gmove [44] with the "–annot" option.

Toxin annotation

In order to quantify the expression levels of toxin coding genes, we specifically annotated them using transcriptomic data, to complement the available annotation of both species' genomes [30]. The DeTox pipeline [v1.2.0] with "contamination_evalue: 1e-10" parameter [17] was first run on all 22 and 15 transcriptome assemblies from *S. haemastoma* and *M. corrugatus*, respectively. A selection of sequences was done based on the markers assigned by DeTox. The following flags were used for sequence selection: S = presence of a signal peptide, C = presence of a cysteine pattern, D = presence of a functional domain, B = match against the toxin reference database, and T = TPM value above 100. For all tissues except salivary glands, we retained all sequences whose tags contained a "B," along with sequences carrying the "SC" and "SCD" tags, in order to reduce computational time by limiting the number of sequences to be mapped against the genome. For salivary glands, the same sequences were selected, with the addition of those marked with "ST", as the salivary glands are considered the main tissue of toxin production. Since "ST" represents secreted sequences that are highly expressed, there should be more toxin candidates in the salivary glands than in other tissues. All sequences retained after this initial step, from all tissues, are subsequently mapped to the genome, ensuring consistency and comparability across the dataset. Then, DeTox was run on the two FASTA files containing the predicted nucleotide sequences from the re-annotation of the genomes, in order to qualify toxins already found in the genomes. Only sequences flagged by DeTox with "B", "SC", and "SCD" were retained for further analysis. DeTox detected 94,863 sequences from the transcriptomes and 265 sequences from the genome annotation for *S. haemastoma*, while it found 64,459 sequences from the transcriptomes and 250 sequences from the genome annotation for *M. corrugatus*. The resulting 95,128 and 64,709 sequences were clustered using CD-HIT [v4.8.1] [45] with the parameters "-c 0.95 -M 0 -T 16 -d 100 -g 1 -p 1". The clustering was used to group highly similar sequences and to identify whether potential toxin transcripts from the RNAseq data were already present in the genome annotation or represented additional candidates.

This resulted in three categories of clusters: Category 1, which corresponds to toxin sequences from the

transcriptome (57,031 clusters from 94,109 sequences for *S. haemastoma*, and 30,969 clusters from 63,649 sequences for *M. corrugatus*); Category 2, which corresponds to toxin sequences from the transcriptomes that were already annotated in the genome (99 clusters from 861 sequences for *S. haemastoma*, and 108 clusters from 928 sequences for *M. corrugatus*); and finally, Category 3, which contains only sequences from the genome annotation (150 clusters from 158 sequences for *S. haemastoma*, and 130 clusters from 132 sequences for *M. corrugatus*). As the number of sequences in the Category 1 is large, only the sequences considered to be the most relevant were retained by filtering them on the DeTox flags: sequences with an 'SC' flag and a TPM < 3 were deleted. These sequences were considered to be peptides from the secretome with very low expression or assembly chimeras. The resulting 21,593 (*S. haemastoma*) and 12,571 (*M. corrugatus*) sequences were supplemented with 579 sequences from Category 2 that do not share the same ORF as in the genome annotation file. To remove redundant sequences, we performed another clustering at 99% similarity using CD-HIT (-c 0.99 -M 0 -T 16 -d 100 -g 1 -p 1), retaining only the reference sequences from the 16,314 (*S. haemastoma*) and 7,867 (*M. corrugatus*) clusters obtained. Before annotating these sequences on the genome, repetitive elements in the genome were masked to avoid overlap. To do this, the repetitive element annotation file was converted to the bed format using gff2bed from BEDOPS (v2.4.41) with default parameters [46]. The bedtools [v2.30.0] with maskfasta option was then run to mask these elements in the genome based on the bed annotation file [47].

To annotate the new putative toxin genes in the genome, a protein sequence alignment against the genome was performed using BLAST [v2.14.0] with "-evalue 0.001 -outfmt 6" parameters [48]. Only matches with at least 50% identity were retained. Segments belonging to the same High-scoring Segment Pair identifier and located on the same chromosome within 100,000 base pairs were grouped together and the average BLAST score was assigned to each group. For each putative toxin, only the top three matches with the highest scores were retained. Regions were extended by 5,000 bp both before the start position and after the end position, ensuring that chromosome boundaries were not exceeded. Only segments with a BLAST score of 100 or greater were kept. Exonerate [v2.4.0] with "-m protein2genome -Q protein -T dna --percent 50 -c 1 –showtargetgff" parameters was then used to refine the determination of transcripts boundaries on the genomes, using the restricted region detected by BLAST to speed up its execution. The Exonerate results were combined to process a new GFF file corresponding to the potential toxin genes identified by DeTox. This file was then provided to Gmove [v1.2] with "--cds 30" parameter [44] to improve the gene prediction for the genomes. As a result, a GFF file was generated, identifying 2,027 and 1,670 potential toxin genes in *S. haemastoma* and *M. corrugatus* genomes, respectively, based on the transcripts identified in Categories 1 and 2 (as Category 3 sequences were already included in the global annotation).

A final manual filtering step was performed to verify whether the genes had already been annotated in the genome or overlapped with regions containing repetitive elements. This was based on the alignment of reads produced with STAR [v2.7.11a] with "--alignIntronMax 1000, --alignMatesGapMax 10000 and --outSAMtype BAM SortedByCoordinate" parameters [49] against the genome annotation file. Newly annotated genes were excluded if more than 50% of their exons, with at least 80% of their length, overlapped with a repetitive element or a previously annotated gene. Following the filtering process, 1,690 and 1,390 new putative toxin genes identified by DeTox were retained for *S. haemastoma* and *M. corrugatus*, respectively. These new annotations were merged with the initial genome annotation file using the AGAT_SP_MERGE_ANNOTATIONS.PL script from AGAT tools [v1.0.0] with default parameter [50]. This resulted in the generation of a complete GFF file, which included the potential toxin genes alongside the initial global annotation, totaling 41,780 and 40,837 genes for *S. haemastoma* and *M. corrugatus*, respectively. At this stage, to ensure that the analysis was performed on the correctly translated sequences, we reran DeTox on the final gene sets to get the summary of all the information given by DeTox, i.e., the presence of cysteine pattern, a hit against the toxin database, the detection of domains, etc [17].

Differential gene expression

The newly completed genome annotation file, generated by DeTox, was used to align each transcriptome to the genome using STAR [v2.7.11a] with "--alignIntronMax 1000, --alignMatesGapMax 10000 and --outSAMtype BAM SortedByCoordinate" parameters. These alignments were subsequently used to create gene count tables using featureCount [51] from subread [v2.0.6] with "-p -F GTF" parameters. The resulting tables were then processed through SARTOOLS [v1.8.1] [52] which uses DESeq2 [v.1.46.0] [98] (https://doi.org/10.1186/s13059-014-0550-8) with default parameters to identify differentially expressed genes in each tissue. Based on the statistics obtained from SARTOOLS, the transcriptomes from the whole bodies from *S. haemastoma* were not retained for the downstream analysis due to the excessive heterogeneity among the samples, which was observed through the normalized read count values, principal component analysis, and the resulting dendrogram (Supplementary

file 1, Figure S1). In addition, the BUSCO scores (using both metazoan and Mollusca databases v.10 and v12, respectively) for these datasets were markedly lower than expected for whole-body RNA of a metazoan species (typically > 80%), suggesting potential biases in the data. Such discrepancies may arise from technical issues during sequencing (e.g., flowcell capture), sample handling, or limitations in the software and lineage dataset used. Given these limitations, we chose not to include the whole-body transcriptomes in the final analyses. Additionally, to homogenize the data and ensure consistency in protocols and sequencing methods, only 3 SG samples (0521.5, 0521.101, and 0521.134) over a total of 6 and 3 OSP samples (0521.3, 0521.95, and 0521.128) from the newly sequenced datasets of *S. haemastoma* were retained, reducing the initial set of 6 transcriptomes per organ. The remaining samples were excluded due to differences in sequencing protocols done in a different place than all other samples (Supplementary file 1, Table S1). For *M. corrugatus*, MXMOG3 (0521.114) and MXSG3 (merging of 0521.112 and 0521.113) were not retained for analysis because the median raw counts and normalized counts were close to zero (Supplementary file 1, Figure S2).

Also, the complete set of proteins from *S. haemastoma*, *M. corrugatus*, and the toxin database used by DeTox were merged into a single file and then clustered using OrthoFinder v2.5.2 with default parameters [53] to identify orthologous genes between the two species, as well as groups of orthologous genes with potentially similar functions. Finally, violin plots were generated using Seaborn [v0.12.2] and Matplotlib [v3.8.0] on Python [v3.11.7].

Proteomic analysis

Extraction of the components of the salivary gland of S. haemastoma

The frozen salivary glands of two specimens of *S. haemastoma* (see samples identifier in supplementary file 1, Table S1) were freeze-dried, and then separately extracted with 1 mL of H_2O/MeCN 8:2 (0.1% formic acid). After homogenization, the extracts were centrifuged for 15 min at 16,100 g and the supernatant solution separated. Each pellet underwent two rounds of extraction and centrifugation following the same procedure, and the resulting extracts from each sample were pooled. The final extracts were freeze-dried and stored at -20 °C until further use.

RP-UHPLC/MS analysis

Reverse phase ultrahigh performance liquid chromatography (RP-UHPLC) analyses were conducted at the Central Service for Science and Technology (SCCyT) of the University of Cádiz using a Waters ACQUITY H-Class UHPLC system (Milford, MA, USA). The setup included a binary solvent system and an autosampler coupled with an ACQUITY UPLC BEH C18 column (2.1 × 50 mm, 1.7 µm), maintained at 50 °C. The mobile phases consisted of eluent A (0.1% formic acid in water, v/v) and eluent B (0.1% formic acid in acetonitrile, v/v). These phases were delivered at a flow rate of 0.6 ml/min by using a linear gradient program as follows: 0–14 min, 95 – 5%A; 14–16 min, 5–95% A; 16–18 min, 95 – 5% A. The injection volume was 2 µl. The UHPLC system was interfaced with a XEVO G2-S QTOF tandem quadrupole time-of-flight mass spectrometer (Waters) equipped with an electrospray ionization (ESI) source. The operating parameters in ESI were set as follows in positive mode: sample cone voltage of 20 V; source temperature of 120 °C; cone gas flow of 10 L/h; desolvation gas flow of 850 L/h. Acquisitions were carried out over the range 100 to 3000 Da. The m/z leucine-enkephalin (m/z 556.2771, positive ion mode) was continuously infused at 5 µL/min as an external LockSpray reference. Data acquisition and visualization was performed using MassLynx [v4.1] software (Waters, Manchester, UK) in positive-ion mode. Raw data files (.raw) were converted to mgf files using MSConvert from the ProteoWizard software package [54]. Chromatogram analyses and processing was carried out with MZmine 3 [55].

Shotgun proteomics (LC-MS/MS)

The analysis was carried out at the Servei Central de Support a la Investigació Experimental (SCSIE-UV, University of Valencia, Spain). The lyophilized venom extracts were resuspended in 100 µl of 50 mM ammonium bicarbonate (AB) and quantified using Qubit (Invitrogen) following manufacturer instructions. Upon quantification, a volume equivalent to 2 µg of protein was taken, and diluted to 20 µl of 50 mM ammonium bicarbonate buffer. Cysteine residues were reduced by 2 mM DTT in 50 mM ammonium bicarbonate buffer at 60 °C for 20 min. Sulfhydryl groups were alkylated with 5 mM iodoacetamide in 50 mM ammonium bicarbonate in the dark at room temperature for 30 min. Residual iodoacetamide was quenched using 10 mM DTT in 50 mM ammonium bicarbonate for 30 min at room temperature.

Proteins were then digested overnight at 37 °C with 100 ng of sequencing-grade modified trypsin (Promega, Madison, WI, USA) in 50 mM ammonium bicarbonate (protein/trypsin ratio 20/1). A 1 µl of this solution was diluted to 20 µl in 0.1% FA and loaded into an Evotip pure tip (EvoSep) according to the manufacturer protocols (https://www.evosep.com/wp-content/uploads/2020/03/Sample-loading-protocol.pdf). Liquid Chromatography (LC) coupled to tandem Mass Spectrometry analysis (LC-MS/MS) was performed on a TimsTOF flex (Bruker) system. Samples loaded onto Evotip Pure tips were eluted onto

an analytical column (EvoSep 15 cm × 150 µm, 1.5 µm; EvoSep) using the EvoSep One system and separated according to the manufacturer's 100 samples-per-day (30 SPD) chromatographic method. The eluted peptides were ionized in a captive Spray with 1700 V at 200 °C, and analyzed in a ddaPASEF mode with the following settings: TIMS settings: Mode: custom; 1/K0: 0.6–1.4 V.s/cm^2; ramp time: 100 ms; Duty Cycle: 100%; Ramp Rate: 9.42 Hz; Ms Averaging: 1; Auto Calibration: off. MS settings: Scan: 100–1700 m/z; Ion Polarity: Positive; Scan Mode: PASEF. MSMS: Number of PASEF ramps: 4; Total Cycle time: 0.53 s; Charge Minimum: 0 (unknown); Charge maximum: 5; Scheduling: target Intensity: 12,500, Intensity Threshold: 1000. Active exclusion: ON. The system sensitivity was controlled with 50 ng of HeLa digested proteins commercially available (Thermo). A total of 1,950 proteins were identified using the 60 SPD gradient method.

Bioinformatic integration of proteomic and transcriptomic data

PEAKS Online XPro software (Bioinformatics solutions, Waterloo, ON, Canada) was used to match MS/MS spectra obtained from proteomic analysis of the salivary gland extracts of *S. haemastoma*. MS spectra of peptides were elucidated based on three separate databases: (i) one containing 101,897 protein sequences from the pooled transcriptomes of the salivary glands of 6 specimens of *S. haemastoma*; (ii) another one containing 41,743 protein sequences from the genome of *S. haemastoma*; (iii) and a self-tailored one containing 35,791 sequences of toxins and venom proteins from cone snails. The latter includes sequences obtained from Conoserver (www.conoserver.org), sequences of all published cone snail venom gland transcriptomes up to date, sequences searched in GenBank and UniProt databases using the term "conotoxin", plus a series of highly-conserved protein sequences recently found in venomous amniotes [56].

Carbamidomethylation of cysteine was designated as a fixed modification, whereas methionine oxidation, C-terminal amidation, proline oxidation to pyroglutamic acid, and pyroglutamic from glutamic acid and glutamine were considered as variable modifications. Trypsin digestion allowed up to three missed cleavages. Parent mass tolerance was set at 25 ppm and fragment mass tolerance at 0.05 Da. False discovery rate (FDR) of 1% and unique peptide 1 were used for filtering out inaccurate proteins in all cases, which respectively corresponds to values of -10lgP > 157 for the salivary gland transcriptome database, -10lgP > 49 for the genome database, and – 10lgP > 57 for the conopeptide database. The Spider algorithm from PEAKS Online XPro software was used to find additional mutations or to correct the sequences. This algorithm corrects the sequences stored in transcriptomic database with *de novo* sequences based on MS/MS spectra, enabling the identification of additional post-translational modifications (PTMs) and mutations. Data for the identified proteins are included in Supplementary file 2.

Manual functional annotation of toxin DEGs

For each differentially expressed gene (DEG), we combined results from INTERPROSCAN (58), BLAST against the NR database, DeTox, proteomic data from *S. haemastoma*, and log fold-change values with orthogroup information. We, then, manually annotated this subset of genes to propose their hypothetical functions (Supplementary file 2), prioritizing specific characteristics, as described in detail in the flowchart in supplementary file 1, Figure S3. To assist in categorizing the confidence levels of toxin function assignments, we implemented a color-coding system. The confidence level annotation is based on multiple sources of information and is specific to each toxin. Each sequence was compared to presumed toxin sequences using several criteria: domain composition and organization, cysteine pattern, percentage of sequence identity, and the presence or absence of significant alignments in the NR database. If all criteria are consistent and the sequence has already been described as a toxin, it is annotated as "highly confident". When one or more criteria are missing, partially inconsistent, or when the sequence has been identified as a *putative* toxin or has not been previously reported in neogastropods, it is annotated in yellow. Sequences showing several inconsistencies, or those previously described both as putative toxins and as physiological proteins, are annotated in orange. Finally, when most features do not match or when the information suggests a stronger similarity to physiological proteins (particularly enzymes) than to known toxins, the sequence is annotated in red. This system was applied to each gene, providing a visual reference for subsequent functional analyses. The color-coded data can be found in Supplementary File 2, specifically in the "AC" column of the " *S. haemastoma* DEG" and "*M. corrugatus* DEG" sheets.

To facilitate the comparison of functional diversity, we categorized putative toxins into three distinct classes: venom processing, peptide toxins, and enzymatic toxins. The venom processing category corresponds to proteins whose function is associated with the propagation and maturation of venom peptides and proteins. Peptide toxins are short sequences targeting cell receptors, while enzymatic toxins are longer polypeptides with domains characterized by catalytic activities that disrupt the normal functioning in the prey or predator.

Results

In this study, we aimed to identify potential toxin-producing genes and determine which tissues are involved in toxin production in *S. haemastoma* and *M. corrugatus*. To achieve this, we have added transcriptomic material of different tissues to already sequenced transcriptomes from Farhat et al. 2023 [30] and reannotated the genomes. The final genome annotation yielded a total of 39,994 and 39,447 coding genes for *S. haemastoma* and *M. corrugatus*, respectively. Then, a differential expression analysis was conducted to identify genes with significant expression across tissues (Supplementary file 1, Table S1), providing new insights into the distribution of toxin production in these species.

Identification of toxins and differential gene expression between tissues

DeTox identified 2,565 putative toxins in *S. haemastoma* and 1,777 in *M. corrugatus*. The type of evidence used by DeTox differed between species: in *S. haemastoma*, 667 sequences were identified by homology alone, 919 by structure alone, and 979 by both criteria, while in *M. corrugatus* the majority were detected through structural information (1,105 sequences), compared with only 134 by homology and 538 by both. In *S. haemastoma*, 3,131 genes were differentially expressed, including 235 flagged by DeTox; of these, 154 were manually annotated as putative toxins, 115 classified as confident (orange), and 35 further validated by proteomics (Table 1). In *M. corrugatus*, 3,808 genes were differentially expressed, with 359 flagged by DeTox; 171 were manually annotated as putative toxins, 143 of them confident (Table 1).

Together, these results highlight a higher contribution of structural features in *M. corrugatus* toxin detection, whereas both homology and structure contributed substantially in *S. haemastoma*. The integration of DeTox predictions, differential expression, and proteomics thus provides a consistent framework for identifying toxin candidates in both species.

DEGs and toxins in *S. haemastoma*

The distribution of the DEGs in *S. haemastoma* showed only slight variation across tissue comparisons, with a relatively balanced distribution of up- and down-regulated genes (Fig. 1A). Interestingly, the SG vs. ASG comparison revealed a lower number of DEGs between tissues, indicating a substantial similarity in overall gene expression profiles between these two tissues, while the OSP tissue exhibits the lowest number of upregulated confidently identified toxins (Supplementary file 1, Table S2).

While confidently identified toxins were consistently found as DEGs across all tissue comparisons, the salivary gland had a higher number of upregulated confidently identified toxins, while the OSP exhibited the fewest (Supplementary file 1, Table S2 and Fig. 1B). Also, there is more upregulated genes in GL than in SG but a higher number of upregulated confidently identified toxins in SG than GL. This is even more obvious when looking at DEGs being always overexpressed in the same tissue compared to the others (Table 2). Particularly, SG have the highest number of upregulated confidently identified toxin genes (37). In contrast, the OSP demonstrated substantial downregulation of confidently identified toxin genes (22). The GL have the highest overall number of genes upregulated (356), with few upregulated confidently identified toxins (12). The ASG displayed the lowest number of differentially expressed genes compared to all other tissues.

Among the 115 genes annotated as confidently identified toxins, 55 (48%) were classified as "highly confident" (Table 2). Considering only these highly confident toxins, we can notice that when comparing OSP to all other tissues, most of the genes are downregulated (Fig. 2 and Supplementary file 1, Table S1) while in ASG, they are predominantly upregulated compared to OSP and GL, but not compared to SG.

DEGs and toxins in *M. corrugatus*

Gene expression variation in *M. corrugatus* was notably more pronounced compared to *S. haemastoma*, characterized by substantial heterogeneity in DEGs (Supplementary file 1, Table S3). This heterogeneity was especially pronounced in SG-related comparisons, evidenced by a higher density of downregulated confidently identified toxin genes (Supplementary file 1, Table S3 and Fig. 3A). For example, the MOG vs. H comparison yielded in 1,230 upregulated genes, significantly higher than the 437 genes identified in the MOG vs. K comparison. Further, the K vs. CT comparison identified 451 upregulated genes, including 19 confidently identified toxins, while the K vs. H comparison revealed 990

Table 1 Summary of detected sequences. Total genes annotated in the genomes of both *S. haemastoma* and *C. corrugatus*; number of differential expressed genes (DEGs) across all tissues and the number of which found by DeTox; Number of sequences found by DeTox by one or both strategies; toxin candidates manually annotated and the confident ones; and the number of sequences validated by proteomics, DEGs and found as putative toxins

	Total genes	DEGs (DeTox)	Homology; structure; both	Putative toxins(confident toxins)	Proteomics; DEGs; toxins
S. haemastoma	41,780	3,131 (235)	15; 99; 121	154 (115)	171; 95; 35
M. corrugastus	40,837	3,808 (359)	8; 216; 135	171 (143)	X

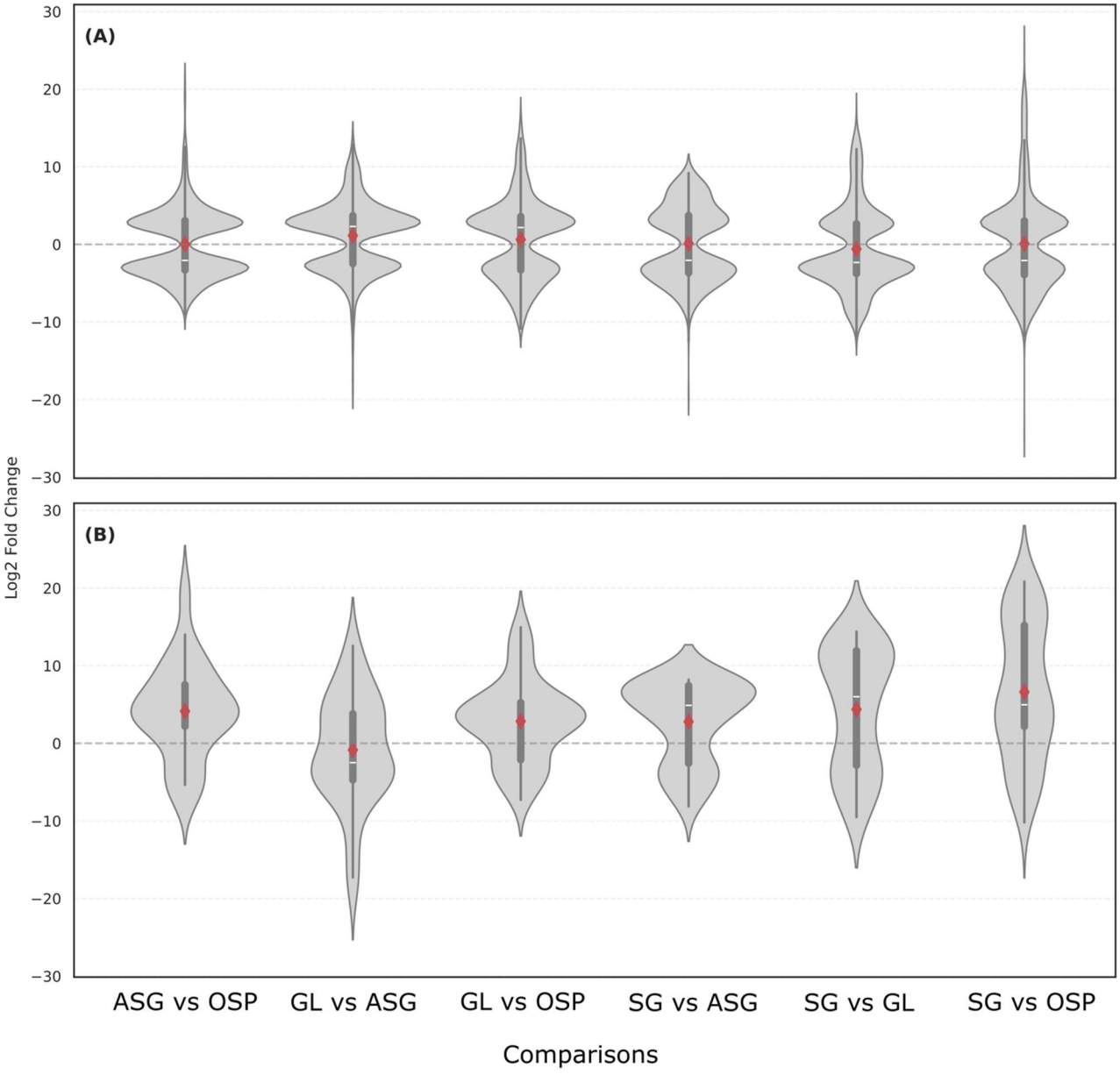

Fig. 1 Violin plot of the log2 fold change. **A** all differentially expressed genes (DEGs) and (**B**) toxin DEGs in *S. haemastoma*. Value equal to zero are not represented. SG= Salivary gland, GL=Gland of Leiblein, ASG=Accessory salivary gland, OSP=Osphradium

Table 2 Summary of the total number of genes overexpressed in exactly one tissue relative to all other tissues sampled for *S. haemastoma*. Numbers within parentheses represent predicted toxins after manual annotation, formatted as X/Y, where X indicates the total number of predicted confidently identified toxins, and Y the subset considered "highly confident". SG= Salivary gland, GL=Gland of Leiblein, ASG=Accessory salivary gland, OSP=Osphradium

	SG	GL	ASG	OSP
Upregulated	187 (37/20)	356 (12/5)	76 (8/6)	272 (6/0)
Downregulated	86 (1/0)	77 (1/1)	83 (1/0)	106 (22/11)

upregulated genes, with 45 confidently identified toxins. Consistent with observations in *S. haemastoma*, toxin-related genes were expressed across all analyzed tissues (Fig. 3B). Notably, the confidently identified toxin expression profile unexpectedly showed elevated expression in non-secretory organs, such as the ctenidium.

Gene expression results across tissue comparisons (H, CT, K, MOG, SG) were determined similarly as *S. haemastoma* (Table 3). Regarding confidently identified

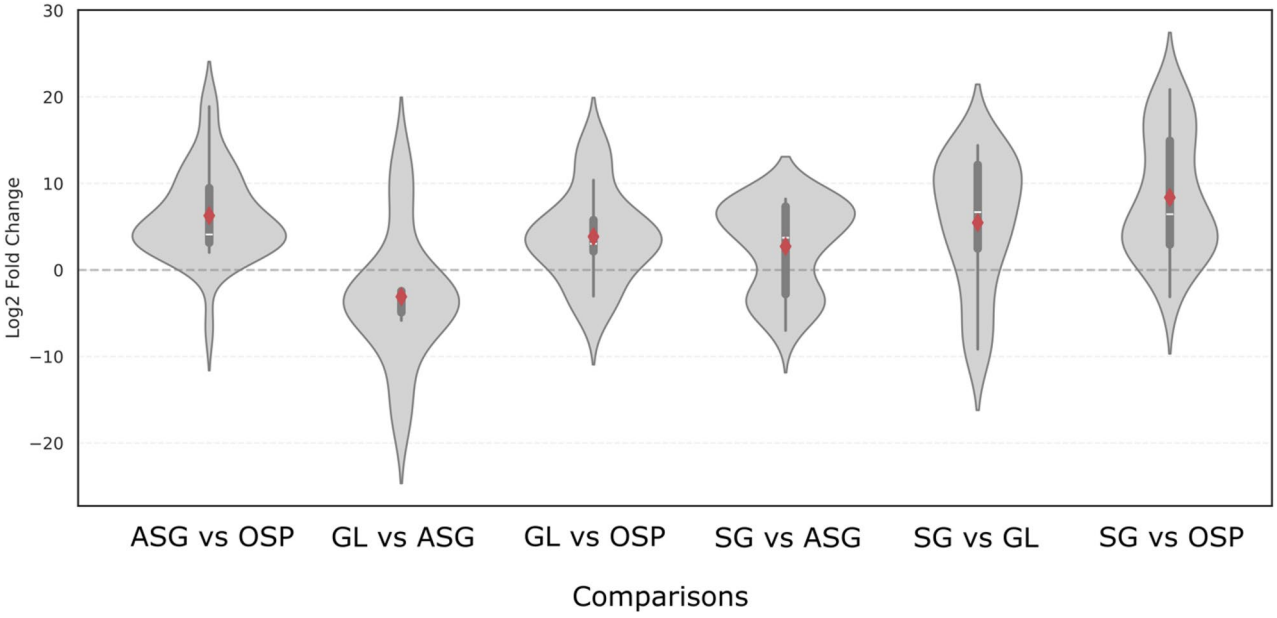

Fig. 2 Violin plot of the log2 fold change. Toxin differentially expressed genes (DEGs) in *S. haemastoma* which has been manually annotated as highly confident. Value equal to zero are not represented. SG= Salivary gland, GL=Gland of Leiblein, ASG=Accessory salivary gland, OSP=Osphradium

toxin genes, SG displays the highest level among tissues, followed by MOG and CT (Table 2).

Among the 143 confidently identified toxin genes, 38 (26%) are categorized in "highly confident", a lower proportion than that observed in *S. haemastoma* (Supplementary file 1, Table S2). Regarding tissue-specific expression, the salivary gland (SG) is the tissue with the highest number of upregulated highly confident toxin genes (Fig. 4). The MOG also presents a substantial number, whereas tissue K shows considerably fewer. Furthermore, SG consistently demonstrates the highest mean log fold change across all comparisons (Supplementary file 2, sheet "*M. corrugatus* DEGs"). Notably, 6 of the 13 highly confident toxin genes are consistently and exclusively upregulated in SG, without differential expression in other tissues. In MOG, two highly confident toxin genes are consistently upregulated (with no differential expression in other tissues). These results also appear to support the involvement of digestive glands, particularly the SG, in the secretion of potential toxins in *M. corrugatus*.

Functional diversity among potential toxin candidates

When considering functional categories, in both species and overall, the largest proportion of annotated sequences corresponds to enzymatic toxins (Fig. 5). However, the distribution of the "venom processing" and "peptide toxin" categories shows moderate variation. Specifically, the proportion of "peptide toxin" sequences is higher in *M. corrugatus* than in *S. haemastoma*, whereas the "venom processing" category is more represented in *S. haemastoma* than in *M. corrugatus*. These results highlight a diverse venom cocktail in both species, combining enzymes, neurotoxic peptides, and binding proteins, suggesting complex envenomation strategies.

Enzymatic gene families may be related to toxins

Enzymatic genes constitute one of the largest and most diverse functional categories identified in *M. corrugatus* and *S. haemastoma* by DeTox (Fig. 5). Although it is known that proteins with normal physiological functions can be recruited into venom [57], since we are investigating different tissues, we cannot be certain of their specific involvement in venom production. Several enzymatic families are shared by both species, although their levels of representation vary ; for example, acetylcholinesterases are notably abundant in *M. corrugatus* (12), but also present in *S. haemastoma* (6). A similar pattern is observed for hydrolase-like genes, with 13 and 12 annotated ORFs, respectively. By contrast, serine-type peptidases are more numerous in *S. haemastoma* (17) compared to *M. corrugatus* (6), although they remain consistently identified in both species. Phospholipases (types B, A2, and C) and phosphodiesterases show an identical number of annotations across species, with three copies each. Six Phospholipase-A2 similar to those reported for the venomous lizard *Heloderma suspectum* (coverage from 21 to 46%) [58]were detected in the transcriptome of both specimens analyzed.

Certain enzymatic toxin families were identified in orthogroups (supplementary file 2 sheet "orthogroups"): Venom prothrombin activators, categorized broadly as serine proteases, clustered within the same orthogroup

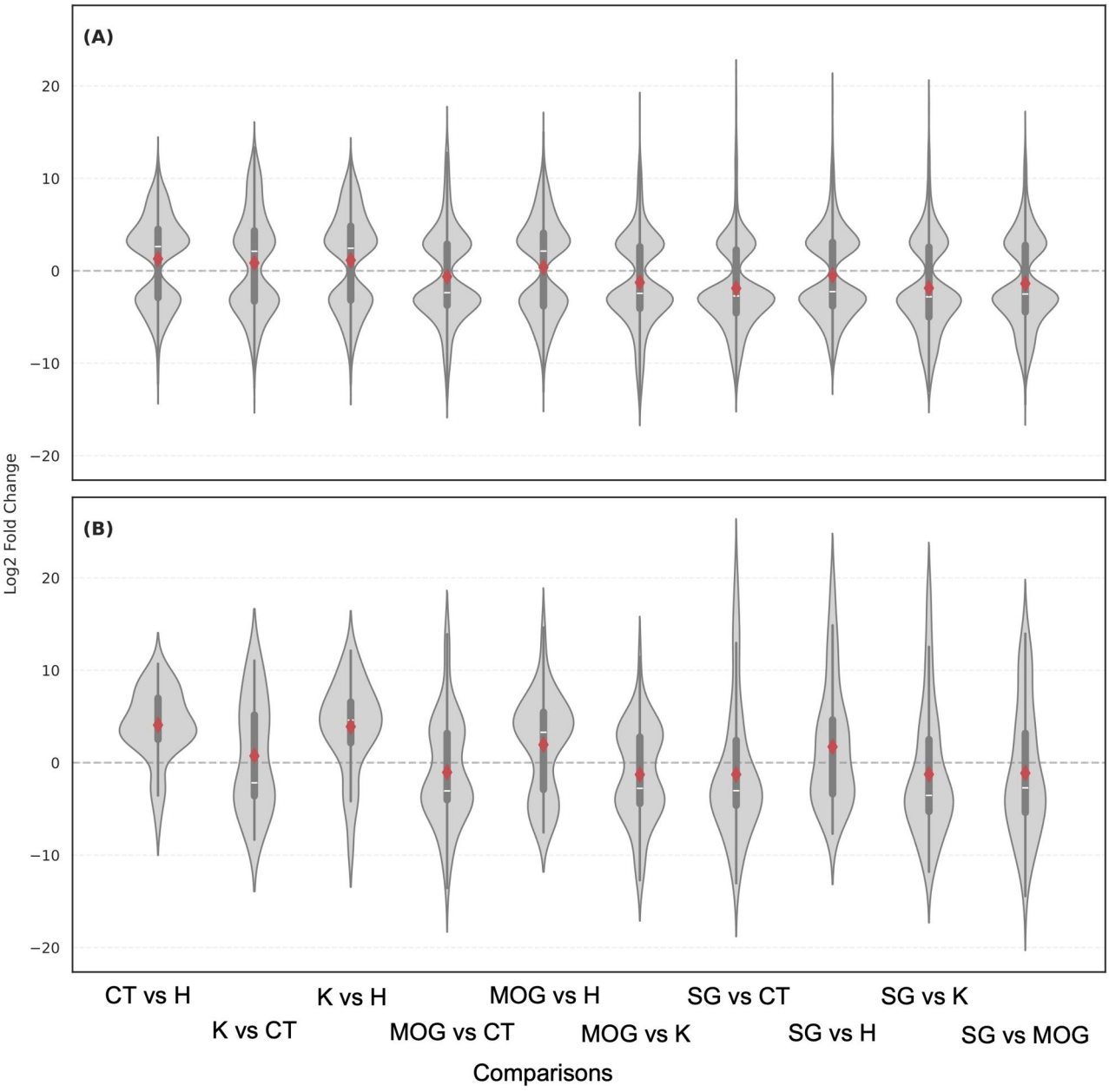

Fig. 3 Violin plot of the log2 fold change. **A** All differentially expressed genes (DEGs) and (**B**) toxin DEGs in *M. corrugatus*. Value equal to zero were not considered. H=Heart, CT=Ctenidium, K=Kidney, MOG=Mid-esophageal gland, SG=salivary gland

Table 3 Summary of the total number of genes overexpressed in exactly one tissue relative to all other tissues sampled for *M. corrugatus*. Numbers within parentheses represent predicted toxins after manual annotation, formatted as X/Y, where X indicates the total number of predicted confidently identified toxins, and Y the subset considered "highly confident". H=Heart, CT=Ctenidium, K=Kidney, MOG=Mid-esophageal gland, SG=salivary gland

	SG	MOG	K	CT	H
Upregulated	193 (19/13)	168 (13/2)	168 (6/0)	143 (10/1)	398 (5/1)
Downregulated	82 (6/1)	74 (1/0)	14 (0/0)	4 (0/0)	249 (7/1)

OG0000033, including two genes strongly overexpressed in *M. corrugatus* salivary glands (log_2 fold-change 11 and 20) showing limited percentage of identical positions (PID) (~34%) to known venom activators from *Cryptophis nigrescens* (snake) and *Bombus ignitus* (Hymenoptera). In contrast, eight genes from *S. haemastoma* belong to this orthogroup, and seven of their corresponded encoded proteins detected by proteomics

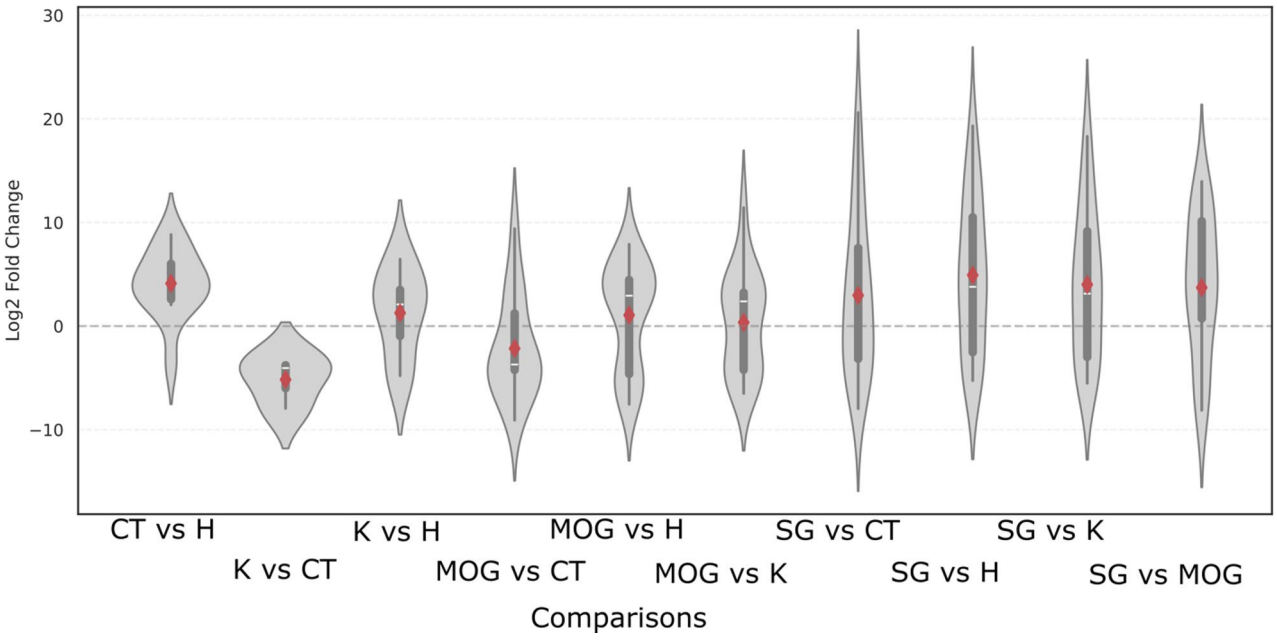

Fig. 4 Violin plot of the log2 fold change. Toxin differentially expressed genes (DEGs) in *M. corrugatus* which has been manually annotated as highly confident. Value equal to zero are not represented. H=Heart, CT=Ctenidium, K=Kidney, MOG=Mid-esophageal gland, SG=salivary gland

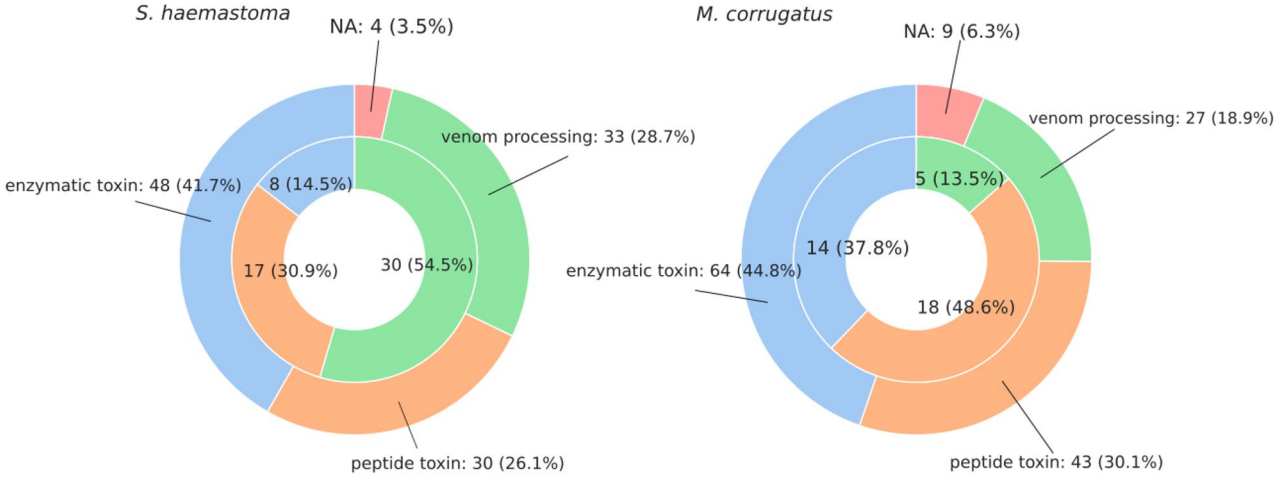

Fig. 5 Functional distribution of potential toxin candidates in *S. haemastoma* and *M. corrugatus*. Outer segments represent all annotated sequences, while inner segments show highly confident annotations

in both specimens (coverage up to 71.6%) and strongly expressed in the salivary gland and accessory salivary gland. Another shared orthogroup OG0000015 contains one gene per species encoding putative venom prothrombin activators, although their functional annotation differs notably between the species: in *S. haemastoma*, the gene possesses a conserved "trypsin-like serine protease" domain (48 PID), whereas in *M. corrugatus* it lacks clear toxin-related annotation and shows expression predominantly in non-secretory tissues (CT, K, H). It is worth to note that the gene was also detected in the proteome of the two specimens with a low, though similar relative abundance, thus suggesting their role in the envenomation process. Certain enzymatic toxin families were detected in *M. corrugatus* but not in *S. haemastoma*, including ADAMTS metalloproteinases, lipases, and hyaluronidases.

Venom processing

Proteins involved in venom processing encompass various families related to toxin maturation, post-translational modifications, and venom secretion, several of which are shared between *M. corrugatus* and *S. haemastoma*. Among these shared families are calglandulin-like proteins, venom-processing metallopeptidases and protein disulfide isomerases.

Calglandulin-like proteins represent an abundant shared family, with seven annotated genes in *M. corrugatus* and six in *S. haemastoma*. Notably, six and five calglandulin-like genes from these species, respectively, clustered within the same orthogroup OG0000001 (supplementary material 2 sheet "orthogroups"). Proteomic analyses of *S. haemastoma* detected three of these proteins in one or the two specimens analyzed by proteomics, albeit with low coverage (6–13%) and very low levels of expression (supplementary material 2 sheet "S. haemastoma proteomic genome"). While one calglandulin-like gene in *S. haemastoma* (~200 amino acids) aligns closely with canonical calglandulins, the other orthologs are significantly longer (>500 amino acids) and share only low identity (PID 33–44%). Expression patterns for these genes differ between the two species, with a predominant expression in the salivary gland for *S. haemastoma*, and in the mid-esophageal gland and ctenidium for *M. corrugatus*.

Similarly, protein disulfide isomerases constitute another shared venom-processing family, with orthologs from both species clustered within orthogroup OG0000044 (supplementary material 2 sheet "orthogroups"). In *S. haemastoma*, two orthologous genes were detected using proteomic analysis with low coverage (8–11%), showing variable homology (PID 81% and 35%) to protein disulfide isomerases from *Conus vexillum* and *Conus quercinus*. Even though the proteomic analysis suggested the presence of the disulfide isomerase from *C. vexillum* in the venom of both specimens, both the percentage of coverage and the expression level in the proteome are very low. In contrast, *M. corrugatus* expresses two orthologs that align closely with sequences from *C. vexillum* (with one ortholog exhibiting a high percentage identity of 85%), predominantly in the salivary gland and mid-esophageal gland.

Peptide toxins

Peptide toxins constitute another major class of venom components identified in both *M. corrugatus* and *S. haemastoma*. Several functional families are shared between species and found in multiple orthogroups. For example, scoloptoxins (orthogroup OG0000131), characterized by the presence of gamma-glutamyltranspeptidase domains, showed limited sequence identity (<50% PID) to known toxins. Within this orthogroup, two genes in *M. corrugatus* exhibit high expression in the kidney, ctenidium, and mid-esophageal gland, while four genes from *S. haemastoma* show broader tissue distribution, particularly elevated in the gland of Leiblein. Two additional orthogroups (OG0000107 and OG0000284) contain conotoxin-like peptides homologous to sequences from *Hemifusus tuba*. Orthogroup OG0000107 includes peptides annotated as serine-type endopeptidase inhibitors (similar to serpins), featuring domains such as Kunitz, BPTI, and thrombospondin type-1. They are expressed in the accessory salivary gland and gland of Leiblein in *S. haemastoma*, but widely expressed across tissues, excluding mid-esophageal and salivary gland, in *M. corrugatus*. In contrast, OG0000284 sequences harbor distinct domain structures (LDLa and calcium-binding EGF), indicative of specialized conotoxin variants but with a higher expression in the ctenidium and kidney. As these domains are common in non-toxin proteins, we hypothesize that they are not toxins given their expression in multiple tissues.

A particularly notable serpin orthogroup (OG0000342) includes peptides characterized as serine protease inhibitors with high identity to serpin from *Profundiconus_vaubani* (Pvau2_Contig9600, DeTox tag "BD"; PID: 82% in *S. haemastoma*, 60% in *M. corrugatus*). Expression patterns for these serpins are highly specific and complementary, overexpressed exclusively in the mid-esophageal gland in *M. corrugatus* and in the salivary gland in *S. haemastoma*.

In addition to shared toxin families, toxin peptides, unique to each species analyzed, were also clearly identified. Among these, three echotoxins exhibit strong up-regulation in the salivary gland of *M. corrugatus* compared to all other tissues (with $\log_2$ fold changes ranging from 4 to 14). Two of these echotoxins show no differential expression between the other tissues. The third echotoxin displays a different pattern, being also upregulated in the mid-esophageal gland when compared specifically to heart and kidney tissues. These echotoxins show a high sequence identity (PID 85–90%) to echotoxins from *M. parthenopeus*. Similarly, the conkunitzins genes display moderate to high homology (PID 49–87%) with conkunitzins identified in various *Conus* species and are predominantly upregulated in the salivary gland and mid-esophageal gland, with one transcript solely expressed in the salivary gland; however, one gene exhibits distinct expression in non-venom tissues (ctenidium, heart, and kidney). Furthermore, a toxin with low sequence homology (PID 29) to plancitoxin from the starfish *Acanthaster planci* was identified in *M. corrugatus*. This toxin was upregulated in the mid-esophageal gland compared to all other while it displayed strong up-regulation in ctenidium compared to heart and salivary. It is characterized by a clear cysteine pattern ("C-C-C-C"). Conversely, *S. haemastoma* exhibits two terepressin/tereryphin genes that are upregulated in the accessory salivary gland and the Leiblein gland relative to the salivary gland and osphradium, as well as two augerpeptide genes that are upregulated in the salivary gland, accessory salivary gland, and Leiblein gland but only compared to the osphradium. These genes were not detected in *M. corrugatus*.

In addition to these families, an orthogroup containing insulin-like growth factor binding proteins (IGFBPs) was also identified in both species (orthogroup OG0000231). The corresponding sequences closely match venom IGFBPs from *Profundiconus neocaledonicus*, *Conus ebraeus*, and *C. judaeus* (PID: 71% in *M. corrugatus*, 50% in *S. haemastoma*). Expression profiles again differ between species, with the *M. corrugatus* gene expressed in non-venom glands (CT, H, and K), while in *S. haemastoma*, expression is restricted to the ASG. Additionally, InterProScan detected Kazal-type serine protease inhibitor and immunoglobulin domains specifically in the *M. corrugatus* IGFBP, suggesting potential functional diversification.

Other functional families

A few shared families do not fall clearly within the three major venom categories. These include gelsolin-like and growth factor-like proteins, suggesting they may have accessory or poorly characterized roles. Some families remain unassigned but are found exclusively in one species; for instance, a unique redox-activity family appears only in *M. corrugatus*. Additionally, the cysteine-rich peptide CreCAP-ShK, similar to peptides identified in *C. reticulata* [59], is strongly upregulated in the SG of both species ($log_2FC > 14$ in *M. corrugatus*, > 7 in *S. haemastoma*) and to a lesser extent in other glands (mid-esophageal gland of *M. corrugatus*, accessory salivary gland and gland of Leiblein of *S. haemastoma*). In addition, it was detected in proteomic analysis with low coverage (up to 21%) in *S. haemastoma*. Although no proteomic results are available for *M. corrugatus*, this species displays an unusual cysteine pattern of "C-C-C-C-C" with an odd number of cysteines, which is uncommon and not referenced in the Conoserver database.

Proteomic analysis of the salivary glands of *S. haemastoma*

Examination by UHPLC-MS of the salivary gland extracts from two specimens of *S. haemastoma* yielded very similar, rather simple chromatograms (Fig. 6, supplementary file 2, sheet "*S. haemastoma* proteomics") with few peaks, corresponding to proteins with molecular masses in the range 3,400 to 7,100 Da. The most intense peak in both samples corresponds to a peptide with a molecular mass of 5,659.46 Da, which however was not identified.

The same venom extracts used for UHPLC-MS were also analyzed by tandem nanoUHPLC-MS/MS. A search against a customized database containing 101,897 protein sequences from the transcriptome of the salivary gland of 6 specimens of *S. haemastoma* yielded 788 hits corresponding to 240 protein groups. Protein groups were defined as those sequences from the database that gave rise upon trypsin digestion to common peptides. Hence, any of the sequences in the groups might be consistent with the observed sets of peptides detected by MS/MS analysis. Of the 788 sequences identified, 557 sequences were found in common in the two samples analyzed, whereas 109 and 122 sequences, respectively, were present in only one of each the two specimens. 35% of the identified sequences did not have an annotation match in the Pfam domain database. The remaining ones correspond to a total of 42 domain groups. The most abundant

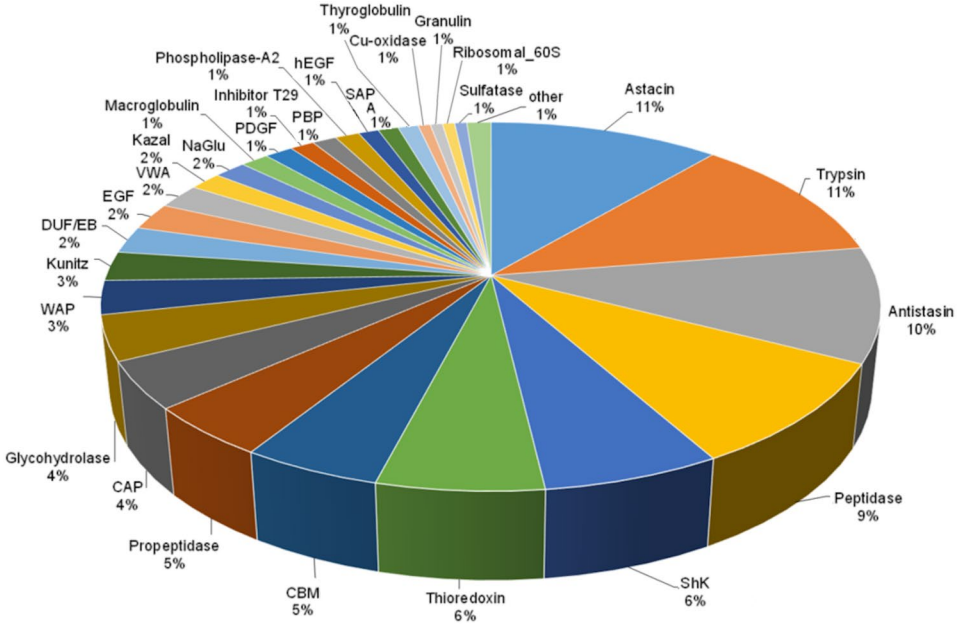

Fig. 6 Distribution of annotated salivary gland transcriptome protein types identified in the proteome of the salivary gland extracts of *Stramonita haemastoma*

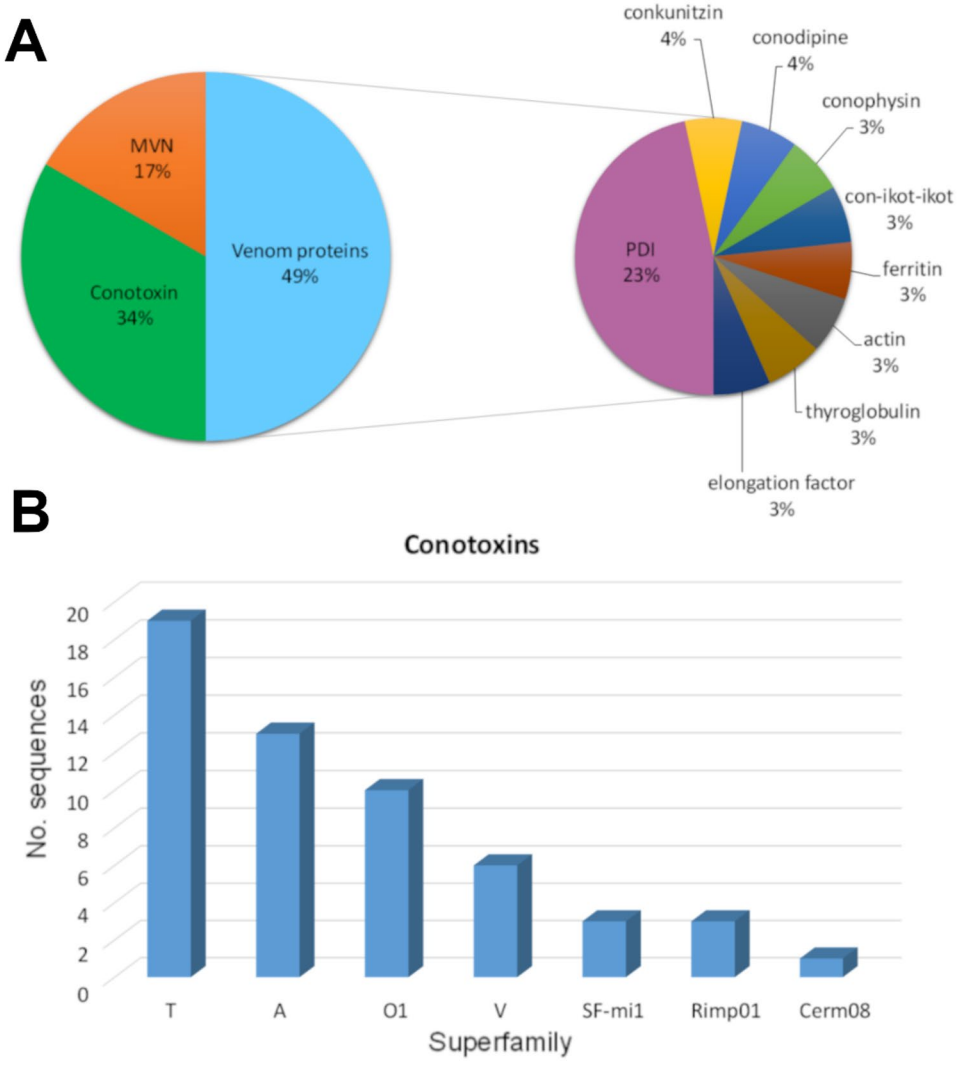

Fig. 7 **A** Conopeptides present in the proteome of the salivary gland extracts of *Stramonita haemastoma*; (**B**) distribution of superfamilies for the identified conotoxins

ones were astacin, trypsin, antistasin, peptidases, potassium channel inhibitor ShK, thioredoxins, carbohydrate binding module and propeptidases, as shown in Fig. 6.

In addition to this analysis, we carried out parallel searches against databases consisting of sequences extracted from the genome of *S. haemastoma* (with 41,743 records), and conotoxins and other venom proteins from cone snails (35,791 records). The analysis using the genome database identified 151 proteins in 107 groups, whereas the search against the conopeptide database yielded 185 hits in 30 groups of proteins. About one half of the total conopeptides were venom proteins, mainly protein disulfide isomerases (PDI), but also conkunitzins, conodipines and conophysins among others (Fig. 7, top).

A number of highly conserved proteins in the so-called metavenom network (MVN, Barua and Mikheyev, 2021) were also found. Interestingly, standard cone snail conotoxins were also detected. These were mainly members of the T, A and O1 superfamilies, with presence also of members of other four superfamilies (Fig. 7, bottom). The conotoxins present in *S. haemastoma* match sequences in the database found in several species of Conidae, mostly in *Conus* (*Virgiconus*) *virgo*, but also in *Conus* (*Virroconus*) *ebraeus* and *Conus* (*Virroconus*) *judaeus*.

In summary, we were able to validate 1,124 sequences – 788 from the transcriptome and 151 from the genome, of *S. haemastoma*. Additionally, 185 conotoxin-like peptides were also detected. All relevant data for the identified proteins are included in Supplementary file 2.

Discussion

Salivary gland might be the primary organ of toxin expression

This study aimed to identify toxin-producing tissues, characterize conserved toxins shared among the studied species, evaluate their differential expression profiles, and clarify the biological roles of novel toxins identified in *M. corrugatus* and *S. haemastoma*. Our analyses demonstrated that confidently identified toxin genes were differentially expressed across all tissues in both *M. corrugatus* and *S. haemastoma*. However, genes annotated as toxins or toxin-processing enzymes appear to be predominantly localized in glands, especially in the primary salivary glands of both species. The expression of this genes in the proteome of *S. haemastoma* strongly supports their role in the envenomation process and that the salivary glands are the primary organ for toxin production. Expression profiles revealed that in *S. haemastoma*, 37 confidently identified toxin genes were consistently upregulated in the salivary gland compared to less than 12 in other tissues. Similarly, although the pattern is less contrasted, *M. corrugatus* exhibited 19 consistently upregulated confidently identified toxin genes in the salivary gland versus fewer than 13 in other tissues. This pattern becomes even more evident when focusing exclusively on highly confident toxins. In *S. haemastoma*, 20 highly confident toxin genes were consistently overexpressed in the salivary gland, while other tissues had less than six. Likewise, in *M. corrugatus*, 13 highly confident toxin genes were consistently overexpressed in the salivary gland, while other tissues had less than two.

Among these highly confident toxins, we identified echotoxins previously described in the genus *Monoplex* [60], as well as several metallopeptidases which have been reported in venoms of other venomous species [61], notably those containing "ShK" domains [59]. These findings align with previous studies reporting the salivary gland as a primary toxin-producing organ in other neogastropod lineages beyond conoideans [9, 62].

To expand the scope of our toxin identification and ensure comprehensive toxin detection, our approach incorporated an extensive database of toxin sequences derived from a wide range of distantly related groups including arachnids, snakes, neogastropods, and sea anemones, as well as from less-represented groups, such as ants. The majority of the putative toxins identified were detected based on their similarity with those from other taxa, underscoring the importance of including diverse venomous taxa in toxin reference databases.

Identification of orthologous toxins supports shared venom strategies

Although *M. corrugatus* and *S. haemastoma* belong to relatively distant lineages within caenogastropods, our transcriptomic analysis revealed a shared repertoire of confidently identified toxin-related genes, underscoring deep conservation of venom-associated pathways across predatory Caenogastropoda [27]. Among these, calglandulin, which have sequence similarity with calmodulin [63], has been identified as playing a role in venom secretion without being an integral component of the venom itself [64]. For example, in snake venom, calmodulin enhances the enzymatic capacity of the neurotoxin "ammodytoxins" (phospholipase A2) by 21-fold [65], and calglandulin has also been detected in the venom gland of the ant *Solenopsis invicta* [66]. Expression profiles of calglandulin genes in *S. haemastoma*, although relatively heterogeneous among individual genes, indicate a slightly higher expression level within the salivary gland. *Monoplex corrugatus* shows considerable variation among genes as well, but with higher expression levels predominantly observed in the mid-esophageal gland and the ctenidium. In neither species does a single tissue stand out clearly as the tissue producing this gene family. It can be either hypothesized, as observed in other species, that calglandulin proteins serve to potentiate or amplify the activity of other venom components, or that their expression in non-secretory organs like the ctenidium is linked to other functions.

Beyond calglandulin, our investigation revealed another component in the venom processing category: Protein Disulfide Isomerase (PDI), an enzyme well-known for catalyzing the formation of disulfide bonds between cysteine residues. Protein Disulfide Isomerase is essential for establishing a protein's three-dimensional structure, particularly during toxin folding [67]. Its involvement in venom maturation has been demonstrated in cone snails [68, 69] and snakes [70, 71], highlighting its importance in venom processing. Furthermore, in *S. haemastoma*, PDI genes showed increased expression in the accessory salivary gland compared to the osphradium and gland of Leiblein, as well as elevated expression in the salivary gland compared to both the gland of Leiblein and osphradium. In *M. corrugatus*, these genes were upregulated in the salivary gland relative to all other tissues and exhibited elevated expression in the mid-esophageal gland compared to the heart for one of the identified genes (Mocor.mRNA.JAPXFY010000016.1.962.1). These expression patterns further support the hypothesis that PDI genes are associated with other potential toxins, facilitating their proper folding. Consequently, these findings reinforce the role of the salivary gland in toxin secretion.

Expanding our toxin repertoire analysis, we identified a protein annotated as a "scoloptoxin" due to a significant match with the scoloptoxin SSD14 (UniProt ID P0DPU3). This scoloptoxin found in *Scolopendra dehaani* is described as having toxic activity, notably

inducing dose-dependent human platelet aggregation, and causing hemolysis of mouse and rabbit erythrocytes [72]. We observed that the scoloptoxin genes in *S. haemastoma* showed marked upregulation in the gland of Leiblein (GL), whose secretions are discharged in the mid esophagus. However, the protein matches do not provide sufficient functional information to infer toxicity or physiological role. Similarly, in *M. corrugatus*, one scoloptoxin-like gene exhibited notably elevated expression in the kidney, contrasting sharply with its lower expression in ctenidium, heart, salivary gland, and mid-esophageal gland. Although the scoloptoxin-like gene in *M. corrugatus* carries a gamma-glutamyltranspeptidase (GGT) domain matching the UniProt 'scoloptoxin' entry, the alignment score is low (39% pairwise identity) and its strongest expression is in the kidney, which is not a secretory tissue. The GGT domain, typically involved in amino acid metabolism and glutathione turnover, has been previously characterized in glycoproteins within rat kidneys [73, 74]. This observation suggests a potential physiological function rather than a strictly toxic one in both species, possibly reflecting a specialized role in digestive or renal function.

In addition to toxins, our comparative analysis identified serpins which, despite displaying only limited sequence homology with known toxins, play an essential regulatory role by inhibiting peptidases that control diverse biological processes, such as blood coagulation and inflammation [75]. Recently, a serpin was identified in the venom of a spider of the genus *Loxosceles*, where it can confer protection against the venom [76]. It is established that *S. haemastoma* employs its accessory boring organ and radula to drill holes into bivalve shells, where the proboscis can be inserted. However, limited information is available regarding the mechanism of prey consumption. Some authors have proposed that the proboscis may deliver a mixture of toxic compounds and proteolytic enzymes into the prey, subsequently aspirating the digested tissues through the proboscis [77]. Consequently, we hypothesize that serpins contribute to prey incapacitation by targeting proteases involved in defensive processes. This action could enhance venom efficacy by neutralizing key proteolytic enzymes, as proposed for scorpion venom [78].

Insulin-like growth factor binding proteins (IGFBPs) are key regulators of insulin-like growth factors (IGFs) and modulate IGF signaling. They have been repeatedly detected in scorpion venoms [79, 80], but their precise function in venom remains unclear. Of particular interest, the venom-derived IGFBPs from *M. corrugatus* carry a Kazal-type serine-protease-inhibitor domain, mirroring observations in IGFBPs from snake venoms [81, 82] and from several spider species [83]. Moreover, the genes matched IGFBP sequences associated with the venom of *Conus ebraeus* [84], further supporting a convergent recruitment of IGFBPs into venom across disparate lineages. The use of insulin itself as a toxin is well documented in cone snails, where "Con-Ins" peptides act as fast-acting hypoglycemic agents [85, 86]. In *M. corrugatus*, expression analyses revealed IGFBP over-expression in several non-glandular tissues, most notably the heart relative to all other tissues, the ctenidium relative to the mid-esophageal gland, and the kidney relative to the mid-esophageal gland. By contrast, in *S. haemastoma*, IGFBP transcripts were predominantly over-expressed in the accessory salivary gland compared with the salivary gland. Although cone-insulin was first discovered in piscivorous species, subsequent studies have demonstrated its presence in molluscivorous and vermivorous lineages as well [4]. In *S. haemastoma* and *M. corrugatus*, their physiological link to insulin, together with this differential tissue distribution, suggests that IGFBPs could simultaneously retain ancestral regulatory functions within the host and be co-opted as potential toxins, thereby highlighting the broader theme of repurposing endogenous molecules for envenomation.

In addition to the toxins assigned to orthogroups, our catalogue also contains a subset of well-characterized neogastropod toxins (conoporins, conotoxins, conkunitzins, ShK-type peptides and turritoxins). Focusing on the cysteine-rich venom proteins, specifically the "CreCAP-ShK" toxins originally described from the salivary glands of the haematophagous neogastropod *C. reticulata* [59], these peptides possess conserved ShK and CAP domains and are thought to modulate ion-channel activity [8]. Our analysis detected a confidently identified toxin expressed in both *M. corrugatus* and *S. haemastoma* that shares up to 59% pairise identity with this family. Notably, this toxin diverges structurally by lacking the ShK domain, retaining only the cysteine-rich CAP domain. A review suggests that CAP proteins derived from venomous animals exhibit diverse activities, including ion channel, inflammatory, proteolysis, and immune regulatory activities [87]. Moreover, in *S. haemastoma*, the cysteine-rich venom protein features a CAP domain with four cysteines (DeTox.CM060334.1.26.1), similar to the CAP domain of the cysteine-rich venom protein Mr30 (A1BQQ5) identified in *Conus marmoreus*. Interestingly, the *M. corrugatus* protein shows homology both to CreCAP-ShK5 and to a Mr30-like cysteine-rich venom protein from *Conus marmoreus*, whose functional characterization is likewise still pending [88]. As in *C. reticulata*, we observed very strong salivary-gland expression for this toxin ($\log_2$FC > 14, reaching 18 in *M. corrugatus*; $\log_2$FC > 7, reaching 18 in *S. haemastoma*), and notably, the corresponding protein was also detected in the *S. haemastoma* proteome. Although its precise biochemical role remains to be elucidated, this evidence supports

the hypothesis that the toxin may have conserved the same function across these phylogenetically distant neogastropods.

Toxins unique to each species analyzed reflect functional diversity

Although the salivary gland is the only overlapping tissue between the two species, our investigation revealed distinct repertoires of high-confidence toxins in each species. Focusing first on *M. corrugatus*, our screening approach detected echotoxin-2, echotoxin-A and echotoxin-B2 in *M. corrugatus*, which are toxins previously characterized in *M. echo* [89]. Consistent with those earlier findings, echotoxins share structural similarity with actinoporins (from sea anemones) [60] and exhibit potent hemolytic activity that is lethal in mice models [9]. Moreover, as reported for *M. echo*, we found these echotoxins to be strongly upregulated in the salivary glands of *M. corrugatus*. Together, these results point to a conserved toxin profile among closely related species and highlight their likely role in prey envenomation.

In addition to these salivary gland toxins, the Plancitoxin identified in *M. corrugatus* shares moderate sequence similarity with a toxin previously characterized in the crown-of-thorns starfish *Acanthaster planci*, which exhibits cytotoxic and hepatotoxic activities [90]. In *M. corrugatus*, the corresponding gene is markedly over-expressed in the mid-esophageal gland relative to all other tissues, and in the ctenidium compared with the heart and salivary gland. In *A. planci*, Plancitoxins constitute a key component of the defensive arsenal of the animal's venomous spines [91]. Taken together, these data suggest that this protein may participate either in the venom cocktail that promotes extracellular pre-digestion of prey tissues or in defense against potential predators.

Concerning *S. haemastoma*, two terepressin/terephysin-type toxins, homologous to those characterized in Terebridae, are markedly overexpressed in both the accessory salivary gland and the gland of Leiblein relative to the osphradium. For one of these sequences, no signal peptide was detected, preventing the identification of a mature region or an associated cysteine pattern. For the second sequence, although the overall sequence similarity is moderate (> 50%) with terepressin/terephysin-type peptides, the canonical cysteine framework and predicted disulfide bonds usually conserved in this family are absent, which raises the possibility that this protein may not function as a toxin but could instead represent a structurally related peptide with a different function. These terepressins/terephysins share homology with conopressins and conophysins [92], which act on vasopressin/oxytocin receptors [93]. Although the precise physiological role of terephysins remains unresolved, they have been proposed to act on hormonal neurotransmitter pathways to manipulate the functioning of nervous system [92]. By contrast, conopressins G and S function as vasopressin-receptor agonists, whereas conopressin T serves as a selective antagonist of V1a-type receptors involved in blood-pressure regulation and vasoconstriction [94]. Overexpression of these genes in *S. haemastoma's* salivary glands suggests they may fulfill a role analogous to that of conopressins in cone snails.

Based on our results, it appears that both our target species are capable of producing toxins within their secretory glands (accessory salivary gland, salivary gland, and mid-esophageal gland). While previous studies have already suggested a role for salivary gland in venom production [95, 96], our findings provide additional evidence supporting these hypotheses. Moreover, we demonstrated that despite the phylogenetic distance between *S. haemastoma* and *M. corrugatus*, they appear to share a set of proteins whose functions may be associated with toxin activity. These potential toxins, although highly diverse, encompass both previously characterized toxins in neogastropods, as well as novel candidates that have not been previously described. Importantly, our success in precisely identifying and characterizing these toxins was made possible through the integration of genomic, transcriptomic, and proteomic approaches, highlighting the critical value of multi-omics strategies in comprehensive toxin discovery. Building upon these findings, future research should focus on functional validation of these putative toxins through targeted experimental approaches to confirm their biological activities and mechanisms of action.

Conclusions

This study presents the first integrative transcriptomic analysis of venom components in *Monoplex corrugatus* and *Stramonita haemastoma*. We identified the salivary glands as the primary toxin-producing organs in both species, supported by transcriptomic and proteomic data. Despite their phylogenetic distance, the two species share several conserved venom-processing genes, suggesting a common molecular foundation, while also exhibiting lineage-specific toxins, such as echotoxins and plancitoxins. Functional annotation revealed a diverse venom composition, including peptide toxins, enzymatic components, and venom-processing proteins.

These findings highlight the power of multi-omics approaches to uncover toxin diversity in neglected Neogastropoda lineages and provide new insights into the evolution of venom systems across gastropods. By revealing both conserved and divergent venom strategies, this work advances our understanding of gastropod predation and opens new avenues for the discovery of bioactive compounds with potential biomedical applications.

Abbreviations

ICM	Institut du Cerveau et de la Moëlle épinière
GL	The gland of Leiblein
SG	The salivary gland
OSP	The osphradium
ASG	The accessory salivary gland
S	Presence of a signal peptide
C	Presence of a cysteine pattern
D	Presence of a functional domain
B	Match against the toxin reference database
T	TPM
TPM	Transcript per million
AB	Ammonium bicarbonate
FDR	False discovery rate
PTMs	Post-translational modifications
NR	Non-redundant
GO	Gene Ontology
CT	Ctenidium
H	Heart
K	Kidney
MOG	Mid-esophageal gland
log_2FC	log_2 Fold Change

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12592-3.

Supplementary Material 1. Supplementary figures and tables.

Supplementary Material 2. An Excel file with all the resulting data of the analysis. Sheet of the proteomic analysis of the salivary gland extracts of Stramonita haemastoma. Sheet of deTox and differentially expressed genes analysis results for Stramonita haemastoma and Monoplex corrugatus. And a sheet of all orthogroups having differentially expressed genes (DEGs) from S. haemastoma and/or M. corrugatus

Acknowledgements

Samples of *M. corrugatus* were collected during the expedition CORSICABENTHOS 2, conducted by Muséum National d'Histoire Naturelle in partnership with Parc Naturel Marin du Cap Corse et de l'Agriate and Université de Corse Pasquale Paoli. It is the marine component of the "Our Planet Reviewed" expedition in Corsica, made possible by fundings from Agence Française pour la Biodiversité and Collectivité Territoriale de Corse. Samples of *S. haemastoma* were collected at Capo Linaro (RM), Italy, and kindly provided by Prof. Marco Oliverio, Dept. of Biology and Biotechnologies Charles Darwin, Sapienza University of Roma, Italy. Sampling operated under the regulations then in force in the countries in question and satisfy the conditions set by the Nagoya Protocol for access to genetic resources. This work was supported by the iGenSeq facility (Genotyping and Sequencing) at ICM, which provided equipment and services. All biocomputing were done on IFB-core cluster, Institut Français de Bioinformatique (IFB), CNRS, INSERM, INRAE, CEA, 94800 Villejuif, France.

Authors' contributions

AR performed all analysis for differential gene expression and gene/toxin annotation; MVM and YK collected samples; MVM performed toxin annotation and analyzed it structures; MJT, JCGG performed proteomic analysis; NP and SF conceived the study; AR, MJT, JCGG, NP and SF wrote the manuscript. All authors edited and approved the final version of the manuscript.

Funding

European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 865101) to N.P. Ministerio de Ciencia e Innovación of Spain, MCIN/AEI/https://doi.org/10.13039/501100011033 grant number PID2022-138477NB-C21.

Data availability

The datasets generated and/or analyzed during the current study including *Monoplex corrugatus* and *Stramonita haemastoma* new transcriptomes raw sequences are available on SRA under the BioProject PRJNA912994 and PRJNA913004, respectively. See supplementary file 1, Table S1 for the SRR Accessions numbers for each sample. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [97] with the dataset identifier PXD069035: [https://ftp.pride.ebi.ac.uk/pride/data/archive/2025/12/PXD069035]

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France
[2]Department of Biology and Evolution of Marine Organisms (BEOM), Stazione Zoologica Anton Dohrn, Roma, Italy
[3]A.N.Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russian Federation
[4]Department CMIM y Química Inorgánica-INBIO, Facultad de Ciencias, Universidad de Cádiz, Puerto Real (Cádiz), Spain
[5]Department Química Orgánica-INBIO, Facultad de Ciencias, Universidad de Cádiz, Puerto Real (Cádiz), Spain

Received: 24 July 2025 / Accepted: 23 January 2026
Published online: 04 February 2026

References

1. Dave K, Lahiry A. Conotoxins: Review and Docking Studies to determine potentials of Conotoxin as an Anticancer Drug Molecule, Curr Top Med Chem, vol. 12, no. 8, pp. 845–851, Apr. 2012, https://doi.org/10.2174/156802612800166765
2. Li X, Tae HS, Chu Y, Jiang T, Adams DJ, Yu R. Medicinal chemistry, pharmacology, and therapeutic potential of α-conotoxins antagonizing the α9α10 nicotinic acetylcholine receptor. Pharmacol Ther. Jun. 2021;222:107792. https://doi.org/10.1016/J.PHARMTHERA.2020.107792.
3. WoRMS EB. World Register of Marine Species., VLIZ. [Online]. Available: https://www.marinespecies.org/
4. Safavi-Hemami H, et al. Venom insulins of cone snails diversify rapidly and track prey taxa. Mol Biol Evol. Nov. 2016;33(11):2924–34. https://doi.org/10.1093/MOLBEV/MSW174.
5. Ruparčič M et al. Jun., The Biological Role of Conoporins, Actinoporin-like Pore-Forming Toxins from Cone Snails, Toxins 2025, Vol. 17, Page 291, vol. 17, no. 6, p. 291, 2025, https://doi.org/10.3390/TOXINS17060291.
6. Olivera BM, Teichert RW. Diversity of the neurotoxic Conus peptides: A model for concerted pharmacological discovery, Mol Interv, vol. 7, no. 5, pp. 251–260, Oct. 2007, https://doi.org/10.1124/MI.7.5.7.
7. Dutertre S et al. Mar., Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails, Nature Communications 2014 5:1, vol. 5, no. 1, pp. 1–9, 2014, https://doi.org/10.1038/ncomms4521.
8. Modica MV, Lombardo F, Franchini P, Oliverio M. The venomous cocktail of the vampire snail Colubraria reticulata (Mollusca, Gastropoda). BMC Genomics. Dec. 2015;16(1):1–21. https://doi.org/10.1186/S12864-015-1648-4/FIGURES/10.
9. Shiomi K, Kawashima Y, Mizukami M, Nagashima Y. Properties of proteinaceous toxins in the salivary gland of the marine gastropod (Monoplex echo). Toxicon. May 2002;40(5):563–71. https://doi.org/10.1016/S0041-0101(01)00256-2.
10. Modica MV, Reinoso Sánchez J, Pasquadibisceglie A, Oliverio M, Mariottini P, Cervelli M. Anti-haemostatic compounds from the vampire snail Cumia reticulata: molecular cloning and in-silico structure-function analysis. Comput

Biol Chem. Aug. 2018;75:168–77. https://doi.org/10.1016/J.COMPBIOLCHEM.2018.05.014.
11. Kuznetsova KG et al. Vexitoxins: conotoxin-like venom peptides from predatory gastropods of the genus Vexillum, Proceedings of the Royal Society B, vol. 289, no. 1980, Aug. 2022. https://doi.org/10.1098/RSPB.2022.1152.
12. Koch TL et al. Aug., Prey Shifts Drive Venom Evolution in Cone Snails, Mol Biol Evol, vol. 41, no. 8, 2024, https://doi.org/10.1093/MOLBEV/MSAE120.
13. homas DT, Gonzales T, Saloma CP. A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. Toxicon. Dec. 2014;92:66–74. https://doi.org/10.1016/J.TOXICON.2014.10.003.
14. Watkins M, Hillyard DR, Olivera BM. Genes expressed in a turrid venom duct: Divergence and similarity to conotoxins, J Mol Evol, vol. 62, no. 3, pp. 247–256, Mar. 2006, https://doi.org/10.1007/S00239-005-0010-X/METRICS.
15. Rogalski A, Himaya SWA, Lewis RJ. Coordinated adaptations define the ontogenetic shift from worm- to fish-hunting in a venomous cone snail, Nature Communications 2023 14:1, vol. 14, no. 1, pp. 1–11, Jun. 2023, https://doi.org/10.1038/s41467-023-38924-5.
16. Macrander J, Panda J, Janies D, Daly M, Reitzel AM. Venomix: A simple bioinformatic pipeline for identifying and characterizing toxin gene candidates from transcriptomic data, PeerJ, vol. 2018, no. 7, p. e5361, Jul. 2018. https://doi.org/10.7717/PEERJ.5361/SUPP-6.
17. Ringeval A, et al. DeTox: a pipeline for the detection of toxins in venomous organisms. Brief Bioinform. Jan. 2024;25(2). https://doi.org/10.1093/BIB/BBAE094.
18. Robinson SD, Norton RS. Conotoxin Gene Superfamilies, Marine Drugs 2014, Vol. 12, Pages 6058–6101, vol. 12, no. 12, pp. 6058–6101, Dec. 2014. https://doi.org/10.3390/MD12126058.
19. Bulaj G, Olivera BM. Folding of conotoxins: formation of the native disulfide bridges during chemical synthesis and biosynthesis of Conus peptides. Antioxid Redox Signal. 2008;10(1):141–55. https://doi.org/10.1089/ars.2007.1856.
20. Kaas Q, Westermann JC, Craik DJ. Conopeptide characterization and classifications: An analysis using ConoServer, *Toxicon*, vol. 55, no. 8, pp. 1491–1509, Jul. 2010, https://doi.org/10.1016/J.TOXICON.2010.03.002.
21. Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol Evol. Apr. 2013;28(4):219–29. https://doi.org/10.1016/J.TREE.2012.10.020/ASSET/B8C49531-1CC9-40E3-AB5F-29CBAC2B4054/MAIN.ASSETS/GR1B2.SML.
22. Schendel V, Rash LD, Jenner RA, Undheim EAB. The Diversity of Venom: The Importance of Behavior and Venom System Morphology in Understanding Its Ecology and Evolution, Toxins 2019, Vol. 11, Page 666, vol. 11, no. 11, p. 666, Nov. 2019, https://doi.org/10.3390/TOXINS11110666.
23. Kantor, I. Yuri. Morphological prerequisites for understanding neogastropod phylogeny. Bollettino Malacologico. 2002;(38): p. 161-74.
24. Fedosov A, Tucci CF, Kantor Y, Farhat S, Puillandre N. Collaborative Expression: Transcriptomics of Conus virgo Suggests Contribution of Multiple Secretory Glands to Venom Production, J Mol Evol, vol. 91, no. 6, pp. 837–853, Dec. 2023, https://doi.org/10.1007/S00239-023-10139-8.
25. Andrews EB, Thorogood KE. An ultrastructural study of the gland of Leiblein of Muricid and Nassariid neogastropods in relation to function, with a discussion on its homologies in other caenogastropods, Journal of Molluscan Studies, vol. 71, no. 3, pp. 269–300, Aug. 2005, https://doi.org/10.1093/MOLLUS/EYI036.
26. Zancolli G et al. Apr., Redistribution of Ancestral Functions Underlies the Evolution of Venom Production in Marine Predatory Snails, Mol Biol Evol. 2025; 42(5): pp. 1–12. https://doi.org/10.1093/MOLBEV/MSAF095.
27. Lemarcis T, Fedosov AE, Kantor YI, Abdelkrim J, Zaharias P, Puillandre N. Neogastropod (Mollusca, Gastropoda) phylogeny: A step forward with mitogenomes, Zool Scr. 2022;51:5:550–561. https://doi.org/10.1111/ZSC.12552;PAGE:STRING:ARTICLE/CHAPTER.
28. Fedosov AE, et al. Phylogenomics of neogastropoda: the backbone hidden in the Bush. Syst Biol. Sep. 2024;73(3):521–31. https://doi.org/10.1093/SYSBIO/SYAE010.
29. Laxton JH. Feeding in some Australasian cymatiidae (Gastropoda: Prosobranchia). Zool J Linn Soc. Feb. 1971;50(1):1–9. https://doi.org/10.1111/J.1096-3642.1971.TB00748.X.
30. Farhat S, Modica MV, Puillandre N. Whole genome duplication and gene evolution in the hyperdiverse venomous gastropods. Mol Biol Evol. Aug. 2023;40(8). https://doi.org/10.1093/MOLBEV/MSAD171.
31. Giacoletti A, Rinaldi A, Mercurio M, Mirto S, Sarà G. Local consumers are the first line to control biological invasions: a case of study with the whelk Stramonita haemastoma (Gastropoda: Muricidae), Hydrobiologia, vol. 772, no. 1, pp. 117–129, Jun. 2016, https://doi.org/10.1007/S10750-016-2645-6/METRICS.
32. Watanabe JT, Young CM. Feeding habits and phenotypic changes in proboscis length in the southern oyster drill, Stramonita haemastoma (Gastropoda: Muricidae), on Florida sabellariid worm reefs, Mar Biol, vol. 148, no. 5, pp. 1021–1029, Mar. 2006, https://doi.org/10.1007/S00227-005-0152-9/METRICS.
33. Brown KM, Lydeard C. Mollusca: Gastropoda, Ecology and Classification of North American Freshwater Invertebrates, pp. 277–306, Jan. 2010, https://doi.org/10.1016/B978-0-12-374855-3.00010-8.
34. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. Sep. 2018;34(17):i884–90. https://doi.org/10.1093/BIOINFORMATICS/BTY560.
35. Haas, B., Papanicolaou, A., Yassour, M. et al. *De novo* transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. *Nat Protoc* . 2013;1494–1512. https://doi.org/10.1038/nprot.2013.084.
36. Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol. 2019;1962:227–45. https://doi.org/10.1007/978-1-4939-9173-0_14.
37. Filée J et al. Dec., Comparative genomic and transcriptomic analyses of transposable elements in polychaetous annelids highlight LTR retrotransposon diversity and evolution, Mob DNA, vol. 12, no. 1, pp. 1–24, 2021, https://doi.org/10.1186/S13100-021-00252-0/FIGURES/11.
38. Farhat S, et al. Identification of variants associated with hard clam, *Mercenaria mercenaria*, resistance to quahog parasite unknown disease. Genomics. 2020;112(6):4887–96. https://doi.org/10.1016/j.ygeno.2020.08.036.
39. Gremme G, Steinbiss S, Kurtz S. Genome tools: A comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Trans Comput Biol Bioinform. 2013;10(3):645–56. https://doi.org/10.1109/TCBB.2013.68.
40. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1. https://doi.org/10.1093/bioinformatics/btq461.
41. Smit A, Hubley R, Green P. RepeatMasker Open-4.0., 2015. http://www.repeatmasker.org.
42. Flynn JM, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020;117(17):9451–7. https://doi.org/10.1073/pnas.1921046117.
43. Thomas-Bulle C, Concatenate_sequences.py. 2017. https://github.com/CamilleTB/Concat_script/blob/master/Concatenate_sequences.py.
44. Dubarry M et al. Gmove a tool for Eukaryotic gene predictions using various evidence, F1000Res. 2017. https://doi.org/10.7490/f1000research.1111735.1.
45. Li W, Godzik A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics, vol. 22, no. 13, pp. 1658–1659, Jul. 2006, https://doi.org/10.1093/bioinformatics/btl158
46. Neph S, et al. BEDOPS: high-performance genomic feature operations. Bioinformatics. Jul. 2012;28(14):1919–20. https://doi.org/10.1093/BIOINFORMATICS/BTS277.
47. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics, vol. 26, no. 6, pp. 841–842, Mar. 2010, https://doi.org/10.1093/BIOINFORMATICS/BTQ033
48. Camacho C, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1–9. https://doi.org/10.1186/1471-2105-10-421.
49. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. Jan. 2013;29(1):15–21. https://doi.org/10.1093/BIOINFORMATICS/BTS635.
50. Dainat J et al. NBISweden/AGAT: AGAT-v1.4.1, 2024, https://doi.org/10.5281/ZENODO.13799920.
51. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics, vol. 30, no. 7, pp. 923–930, Apr. 2014, https://doi.org/10.1093/BIOINFORMATICS/BTT656.
52. Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA, SARTools:. A DESeq2- and EdgeR-Based R pipeline for comprehensive differential analysis of RNA-Seq data. PLoS ONE. Jun. 2016;11(6):e0157022. https://doi.org/10.1371/JOURNAL.PONE.0157022.
53. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20(1):1–14. https://doi.org/10.1186/s13059-019-1832-y.
54. Adusumilli R, Mallick P. Data conversion with proteowizard MsConvert. Methods Mol Biol. 2017;1550:339–68. https://doi.org/10.1007/978-1-4939-6747-6_23.
55. SchmidR, et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nat Biotechnol. 2023;41(4):447–9. https://doi.org/10.1038/s41587-023-01690-2.

56. Barua A, Mikheyev AS. An ancient, conserved gene regulatory network led to the rise of oral venom systems. Proc Natl Acad Sci U S A. Apr. 2021;118(14):e2021311118. https://doi.org/10.1073/PNAS.2021311118/SUPPL_FILE/PNAS.2021311118.SD08.TXT.
57. Almeida DD et al. May., Tracking the recruitment and evolution of snake toxins using the evolutionary context provided by the Bothrops jararaca genome, Proceedings of the National Academy of Sciences, vol. 118, no. 20, p. e2015159118, 2021, https://doi.org/10.1073/PNAS.2015159118
58. Vandermeers A, Vandermeers-piret M-C, Vigneron L, J. Rathe Mstievenart, Christophe J. Differences in primary structure among five phospholipases A2 from heloderma suspectum. Eur J Biochem. Mar. 1991;196(3):537–44. https://doi.org/10.1111/J.1432-1033.1991.TB15847.X.
59. Gerdol M, Cervelli M, Mariottini P, Oliverio M, Dutertre S, Modica MV. A Recurrent Motif: Diversity and Evolution of ShKT Domain Containing Proteins in the Vampire Snail Cumia reticulata, Toxins 2019, Vol. 11, Page 106, vol. 11, no. 2, p. 106, Feb. 2019, https://doi.org/10.3390/TOXINS11020106.
60. Kawashima Y, Nagai H, Ishida M, Nagashima Y, Shiomi K. Primary structure of echotoxin 2, an actinoporin-like hemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*, *Toxicon*, vol. 42, no. 5, pp. 491–497, Oct. 2003, https://doi.org/10.1016/S0041-0101(03)00226-5
61. Markland FS, Swenson S. Snake venom metalloproteinases. Toxicon. Feb. 2013;62:3–18. https://doi.org/10.1016/J.TOXICON.2012.09.004.
62. Chiappa G, Fassio G, Modica MV, Oliverio M. Potential ancestral conoidean toxins in the venom cocktail of the carnivorous snail Raphitoma purpurea (Montagu, 1803) (Neogastropoda: Raphitomidae). Toxins (Basel). Aug. 2024;16(8):348. https://doi.org/10.3390/TOXINS16080348/S1.
63. Junqueira-De-Azevedo IDLM, Pertinhez T, Spisni A, Carreño FR, Farah CS, Ho PL. Cloning and expression of calglandulin, a new EF-hand protein from the venom glands of Bothrops insularis snake in E. coli, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1648, no. 1–2, pp. 90–98, May 2003, https://doi.org/10.1016/S1570-9639(03)00111-0
64. St Pierre L, Woods R, Earl S, Masci PP, Lavin MF. Identification and analysis of venom gland-specific genes from the coastal taipan (Oxyuranus scutellatus) and related species, Cellular and Molecular Life Sciences, vol. 62, no. 22, pp. 2679–2693, Nov. 2005, https://doi.org/10.1007/S00018-005-5384-9/METRICS.
65. Kovačičc L, Novinec M, Petan T, Baici A, Krizaj I. Calmodulin is a nonessential activator of secretory phospholipase A 2, Biochemistry, vol. 48, no. 47, pp. 11319–11328, Dec. 2009, https://doi.org/10.1021/BI901244F/ASSET/IMAGES/MEDIUM/BI-2009-01244F_0013.GIF.
66. Cai L, et al. A combined protein toxin screening based on the transcriptome and proteome of solenopsis invicta. Proteome Sci. Dec. 2022;20(1):1–12. https://doi.org/10.1186/S12953-022-00197-Z/FIGURES/6.
67. Wilkinson B, Gilbert HF. Protein disulfide isomerase, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, vol. 1699, no. 1–2, pp. 35–44, Jun. 2004, https://doi.org/10.1016/J.BBAPAP.2004.02.017.
68. Wang L, et al. Oxidative folding of conopeptides modified by Conus protein disulfide isomerase. Protein J. Oct. 2017;36(5):407–16. https://doi.org/10.1007/S10930-017-9738-6/METRICS.
69. Figueroa-Montiel A, et al. In Silico identification of protein disulfide isomerase gene families in the de Novo assembled transcriptomes of four different species of the genus Conus. PLoS ONE. Feb. 2016;11(2):e0148390. https://doi.org/10.1371/JOURNAL.PONE.0148390.
70. Isomoto A, et al. Active expression of genes for protein modification enzymes in Habu venom glands. Toxins (Basel). May 2022;14(5):300. https://doi.org/10.3390/TOXINS14050300/S1.
71. Rokyta DR, Lemmon AR, Margres MJ, Aronow K. The venom-gland transcriptome of the Eastern Diamondback rattlesnake (Crotalus adamanteus). BMC Genomics. Jul. 2012;13(1):1–23. https://doi.org/10.1186/1471-2164-13-312/FIGURES/1.
72. Liu ZC, et al. Venomic and transcriptomic analysis of centipede *Scolopendra subspinipes Dehaani*. J Proteome Res. Dec. 2012;11(12):6197–212. https://doi.org/10.1021/PR300881D.
73. Horiuchi S, Inoue M, Morino Y. γ-Glutamyl transpeptidase: sidedness of its active site on renal Brush-Border membrane. Eur J Biochem. Jul. 1978;87(3):429–37. https://doi.org/10.1111/J.1432-1033. 1978.TB12392.X;CTYPE:STRING:JOURNAL.
74. Tate SS, Meister A. [50] γ-Glutamyl transpeptidase from kidney. Methods Enzymol. Jan. 1985;113:400–19. https://doi.org/10.1016/S0076-6879(85)13053-3.
75. Huntington JA. Serpin structure, function and dysfunction. J Thromb Haemost. Jul. 2011;9:1. https://doi.org/10.1111/J.1538-7836.2011.04360.X/ASSET/B5597F07-1297-44E2-9A45-7176B6806AC4/MAIN.ASSETS/JTH_4360_F6.JPG.
76. Schemczssen-Graeff Z et al. Jul., Description of a serpin toxin in Loxosceles (Brown spider) venoms: Cloning, expression in baculovirus-infected insect cells and functional characterization, Int J Biol Macromol, vol. 183, pp. 1607–1620, 2021, https://doi.org/10.1016/J.IJBIOMAC.2021.05.129.
77. Brown KM, Alexander JE. Group foraging in a marine gastropod predator: benefits and costs to individuals, Mar Ecol Prog Ser. 1994;112:1/2:97–105.
78. MA H, Xiao-Peng T, YANG SL, LU QM, LAI R. Protease inhibitor in Scorpion (Mesobuthus eupeus) venom prolongs the biological activities of the crude venom. Chin J Nat Med. Aug. 2016;14(8):607–14. https://doi.org/10.1016/S1875-5364(16)30071-1.
79. Cid-Uribe JI, Meneses EP, Batista CVF, Ortiz E, Possani LD. Dissecting toxicity: the venom gland transcriptome and the venom proteome of the highly venomous Scorpion centruroides limpidus (Karsch, 1879). Toxins 2019. Apr. 2019;11(5):247. https://doi.org/10.3390/TOXINS11050247.
80. García-Villalvazo PE et al. Aug., Unveiling the Protein Components of the Secretory-Venom Gland and Venom of the Scorpion Centruroides possanii (Buthidae) through Omic Technologies, Toxins (Basel), vol. 15, no. 8, p. 498, 2023, https://doi.org/10.3390/TOXINS15080498/S1.
81. Qian C, Fang Q, Wang L, Ye GY. Molecular Cloning and Functional Studies of Two Kazal-Type Serine Protease Inhibitors Specifically Expressed by Nasonia vitripennis Venom Apparatus, Toxins 2015, Vol. 7, Pages 2888–2905, vol. 7, no. 8, pp. 2888–2905, Aug. 2015, https://doi.org/10.3390/TOXINS7082888.
82. Fernández J, Gutiérrez JM, Calvete JJ, Sanz L, Lomonte B. Characterization of a novel snake venom component: Kazal-type inhibitor-like protein from the arboreal pitviper Bothriechis schlegelii. Biochimie. Jun. 2016;125:83–90. https://doi.org/10.1016/J.BIOCHI.2016.03.004.
83. Kuhn-Nentwig L et al. Nov., Purification, cDNA structure and biological significance of a single insulin-like growth factor-binding domain protein (SIBD-1) identified in the hemocytes of the spider Cupiennius salei, Insect Biochem Mol Biol, vol. 41, no. 11, pp. 891–901, 2011, https://doi.org/10.1016/J.IBMB.2011.08.003.
84. Pardos-Blas JR, Tenorio MJ, Galindo JCG, Zardoya R. Comparative venomics of the cryptic cone snail species virroconus ebraeus and virroconus Judaeus. Mar Drugs. Feb. 2022;20(2):149. https://doi.org/10.3390/MD20020149/S1.
85. Ahorukomeye P, et al. Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. Elife. Feb. 2019;8. https://doi.org/10.7554/ELIFE.41574.
86. Guo Q, et al. Diversity and evolutionary analysis of venom insulin derived from cone snails. Toxins 2024. Jan. 2024;16(1):34. https://doi.org/10.3390/TOXINS16010034.
87. Zhang Q, Xu J, Zhou X, Liu Z. CAP superfamily proteins from venomous animals: who we are and what to do? Int J Biol Macromol. Nov. 2022;221:691–702. https://doi.org/10.1016/J.IJBIOMAC.2022.09.079.
88. Qian J, Guo ZY, Chi CW. Cloning and isolation of a Conus cysteine-rich protein homologous to Tex31 but without proteolytic activity. Acta Biochim Biophys Sin (Shanghai). 2008;40(2):174–81. https://doi.org/10.1111/J.1745-7270.2008.00385.X.
89. Gunji K, Ishizaki S, Shiomi K. Cloning of complementary and genomic DNAs encoding echotoxins, proteinaceous toxins from the salivary gland of marine gastropod monoplex echo, Protein Journal, vol. 29, no. 7, pp. 487–492, Sep. 2010, https://doi.org/10.1007/S10930-010-9277-X/METRICS.
90. Shiomi K, Midorikawa S, Ishida M, Nagashima Y, Nagai H. Plancitoxins, lethal factors from the crown-of-thorns starfish Acanthaster planci, are deoxyribonucleases II, Toxicon, vol. 44, no. 5, pp. 499–506, Oct. 2004, https://doi.org/10.1016/j.toxicon.2004.06.012.
91. Hillberg AK, et al. Crown-of-thorns starfish spines secrete defence proteins. PeerJ. 2023;11. https://doi.org/10.7717/PEERJ.15689.
92. Gorson J et al. Jun., Molecular Diversity and Gene Evolution of the Venom Arsenal of Terebridae Predatory Marine Snails, Genome Biol Evol, vol. 7, no. 6, pp. 1761–1778, 2015, https://doi.org/10.1093/GBE/EVV104.
93. Giribaldi J et al. Synthesis, Pharmacological and Structural Characterization of Novel Conopressins from Conus miliaris, Marine Drugs 2020, Vol. 18, Page 150, vol. 18, no. 3, p. 150, Mar. 2020. https://doi.org/10.3390/MD18030150.
94. Lewis RJ, Dutertre S, Vetter I, Christie MJ. Conus venom peptide Pharmacology. Pharmacol Rev. Apr. 2012;64(2):259–98. https://doi.org/10.1124/PR.111.005322.

95. Ponte G, Modica MV. Salivary glands in predatory mollusks: Evolutionary considerations, Front Physiol, vol. 8, no. AUG, p. 266745, Aug. 2017, https://doi.org/10.3389/FPHYS.2017.00580/BIBTEX.
96. Turner AH, Craik DJ, Kaas Q, Schroeder CI. Bioactive compounds isolated from neglected predatory marine gastropods. Mar Drugs 2018. Apr. 2018;16(4):118. https://doi.org/10.3390/MD16040118. Page 118.
97. Perez-Riverol Y et al. Jan., The PRIDE database at 20 years: 2025 update, Nucleic Acids Res, vol. 53, no. D1, pp. D543–D553, 2025, https://doi.org/10.1093/NAR/GKAE1011.
98. Love MI, Huber W, Anders S. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology. 2014;550. https://doi.org/10.1186/s13059-014-0550-8.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.