OPEN BIOLOGY

royalsocietypublishing.org/journal/rsob

Research

Cite this article: Modica MV, Leone S, Gerdol M, Greco S, Aurelle D, Oliverio M, Fassio G, El Koulali K, Barrachina C, Dutertre S. 2025 The proteotranscriptomic characterization of venom in the white seafan *Eunicella singularis* elucidates the evolution of Octocorallia arsenal. *Open Biol.* **15**: 250015.
https://doi.org/10.1098/rsob.250015

Received: 15 January 2025
Accepted: 11 February 2025

Subject Areas:
genetics, genomics, molecular biology

Keywords:
venom, Cnidaria, Octocorallia, transcriptomics, proteomics, toxins

Author for correspondence:
Maria Vittoria Modica
e-mails: mariavittoria.modica@szn.it; mvittoria.modica@gmail.com

†These authors contributed equally to the study.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7702728.

THE ROYAL SOCIETY PUBLISHING

The proteotranscriptomic characterization of venom in the white seafan *Eunicella singularis* elucidates the evolution of Octocorallia arsenal

Maria Vittoria Modica[1,†], Serena Leone[1,†], Marco Gerdol[1,2], Samuele Greco[2], Didier Aurelle[3,4], Marco Oliverio[5], Giulia Fassio[5], Khadija El Koulali[6], Célia Barrachina[7] and Sebastien Dutertre[8]

[1]Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Italy
[2]Department of Life Sciences, University of Trieste, Trieste, Italy
[3]Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, Marseille, France
[4]Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France
[5]Department of Biology and Biotechnology 'Charles Darwin', Sapienza University of Rome, Rome, Italy
[6]BioCampus, University of Montpellier, CNRS, INSERM, Montpellier, France
[7]Platform MGX, IGF, University of Montpellier, Montpellier, France
[8]IBMM, University of Montpellier, CNRS, ENSCM, Montpellier, France

MVM, 0000-0003-2532-6763; SL, 0000-0003-4787-0882; MG, 0000-0001-6411-0813; GF, 0000-0003-1767-1452; SD, 0000-0002-2945-1484

All the members of the phylum Cnidaria are characterized by the production of venom in specialized structures, the nematocysts. Venom of jellyfish (Medusozoa) and sea anemones (Anthozoa) has been investigated since the 1970s, revealing a remarkable molecular diversity. Specifically, sea anemones harbour a rich repertoire of neurotoxic peptides, some of which have been developed in drug leads. However, venoms of the vast majority of Anthozoa species remain uncharacterized, particularly in the class Octocorallia. To fill this gap, we applied a proteo-transcriptomic approach to investigate venom composition in *Eunicella singularis*, a gorgonian species common in Mediterranean hard-bottom benthic communities. Our results highlighted the peculiarities of the venom of *E. singularis* with respect to sea anemones, which is reflected in the presence of several toxins with novel folds, worthy of functional characterization. A comparative genomic survey across the octocoral radiation allowed us to generalize these findings and provided insights into the evolutionary history, molecular diversification patterns and putative adaptive roles of venom toxins. A comparison of whole-body and nematocyst proteomes revealed the presence of different cytolytic toxins inside and outside the nematocysts. Two instances of differential maturation patterns of toxin precursors were also identified, highlighting the intricate regulatory pathways underlying toxin expression.

1. Introduction

Cnidaria is the most ancient venomous lineage among the extant Metazoa, having originated 600–800 million years ago [1,2]. Venom is produced in the nematocysts, specialized cells with high morphological diversity [3,4] and is involved in different functions including prey capture, defence from predators, and possibly also *intra-* and *inter*-specific competition [5]. In addition, in sea anemones the production of venom occurs also in ectodermal

gland cells, which may represent an ancestral secretory pathway for these animals [6].

Within Cnidaria, the Anthozoa is a large, well-supported clade comprising over 6000 species, and it includes both hexacorals (e.g. sea anemones and reef-building corals) and octocorals (e.g. soft corals, sea fans and sea pens) [7,8]. Anthozoa are the subject of active investigation on venoms, but traditionally most studies have been focused on sea anemones (Hexacorallia, Actiniaria) [9–15]. Despite a reduced taxon sampling, these studies highlighted Actiniaria produce complex venoms, including over 15 different protein families with neurotoxic or cytolytic action [14–16]. Comparative analyses have demonstrated that sea anemones produce a neurotoxin-rich venom, while the venom of jellyfish (Scyphozoa) and hydroids (Hydrozoa) is abundant in cytolysins and enzymes [17]. These neurotoxins have been used since the 1970s to investigate nervous system functioning, and some of them possess a remarkable biomedical potential as drug leads. For instance, a synthetic analogue of the ShK toxin from *Stichodactyla helianthus* (ShK 186 or dalazatide) is in phase II trial for the treatment of several autoimmune diseases [18].

Much less is known about the venoms of Hexacorallia other than sea anemones (e.g. reef-building corals and black corals) and of the other major lineages within Anthozoa: tube anemones (Ceriantharia) and Octocorallia [19]. Venom composition was investigated in four Ceriantharia species only and found predominantly cytolytic [20]. Octocorals, instead, have been completely neglected to date by venom studies, probably since they were supposed to possess few, small and poorly diversified nematocysts [21]. Only their secondary metabolites have been characterized: these molecules, mostly produced by bacteria of the coral microbiomes, play key roles in defence [22]. Octocorallia include some of the most conspicuous organisms of rocky subtidal communities, both in tropical and temperate seas. These ecosystem engineers play important ecological roles and contribute substantially to benthic communities' diversity: in the Mediterranean Sea, gorgonians constitute up to 40% of the biomass in rocky sublittoral habitats [23].

The white gorgonian *Eunicella singularis* (Esper, 1791) is widespread and abundant in the north-western Mediterranean [24,25]. Despite hosting *Symbiodinium* zooxanthellae, *E. singularis* displays low levels of primary productivity and relies on heterotrophic nutrition to meet its metabolic needs, feeding mainly on zooplankton integrated with particulated organic matter (POM) [26]. Predation on zooplankton, intraspecific competition and defence from predators (mostly snails, crabs and fish) [27] are physiological activities probably supported, in *E. singularis* and allies, by the production of a specialized venom. Indeed, preliminary evidence has demonstrated that haemolytic peptides are present in different species of octocorals [28], and the diversity and abundance of nematocysts could be higher than previously held [21].

However, toxin diversity has not been addressed in Octocorallia so far, hampering our complete understanding of a crucial eco-evolutionary question concerning Cnidaria biology and evolution: to what extent are main toxin families conserved across Anthozoa? To start answering this question, we provided a first overview of the venom produced by an octocoral, applying to *E. singularis* an integrated proteo-transcriptomic approach, which employed whole colony transcriptomic data to match proteomic data obtained from chemical-induced nematocysts discharge or from the whole body of polyps. Our analyses unveiled several previously unreported putative toxin families, providing the first comprehensive description of the venom of an octocoral species. Leveraging these findings, we aimed to assess the degree of conservation of specific toxin folds of Octocorallia in comparison to other Anthozoa, and to elucidate how venom toxin families diversified throughout the evolutionary radiation of octocorals. Employing a comparative approach on available Octocorallia genomes, we reconstructed the evolutionary patterns of the main toxin families, highlighting distinctive evolutionary innovations specific to Octocorallia. Eventually, by complementing our analyses with untemplated, deep learning-assisted structural modelling of the most relevant toxins, we attempted to infer their function and adaptive value.

2. Methods

2.1. Specimens collection, venom sample preparation and RNA and proteins extraction

Colonies of *E. singularis* were collected by scuba diving at a depth of 25–30 m around Giannutri Island, in the Tuscany Archipelago (Italy) and kept alive in seawater until further processing. Polyps were dissected from the colony under a stereomicroscope and pooled in three aliquots of about 20 each (about 15 mg of tissue) that were preserved in Trizol at −80°C until further processing. Total RNA was extracted using Trizol (Thermo Fisher Scientific), following the manufacturer's instructions, obtaining 5–10 000 ng with RQN values between 6.7 and 8.8. Proteins from 10 additional polyps were extracted through homogenization in milliQ water, followed by centrifugation and analysis of the supernatant to produce the whole-body proteome (WB-P).

Additionally, 10 live colony fragments about 25 mm long were immersed for 30 s into 1 ml of EtOH in a 1.5 ml tube to induce nematocyst discharge [29]. Samples were freeze-dried and resuspended in 100 µl milliQ water for proteomic analysis.

2.2. RNASeq

To account for variations in venom expression among samples, three cDNA libraries for RNASeq were prepared using an Illumina TruSeq Stranded mRNA Sample Preparation kit according to manufacturer's instructions and validated qualitatively and quantitatively on a Fragment Analyser, as well as by qPCR on a ROCHE LightCycler 480. The three libraries were then processed for paired-end sequencing-by-synthesis in an Illumina HiSeq 2500 at the MGX facility (Montpellier, France), and quality controlled using FastQC. After trimming, we obtained a number of reads longer than 30 bp and quality above Q20 comprised between 21 and 30 million.

2.3. Bioinformatic analyses of transcriptomic data

Paired-end raw trimmed sequencing data obtained from the three libraries were pooled and de novo assembled using Trinity v. 2.5.1 [30] with default settings, allowing a minimum contig length of 200 nucleotides. The assembled transcriptome was virtually translated using TransDecoder v. 5.01 (https://github.com/TransDecoder/), setting the minimum allowed ORF length to 100 codons.

Expression levels of toxin transcripts were calculated based on unique gene counts as transcripts per million (TPM), as this normalized unit allows an efficient and reliable comparison of gene expression levels both within and between samples [31]. Briefly, trimmed reads from each library were separately mapped against the transcriptome assembly, previously filtered to remove sequence redundancy by considering only the longest transcript for each gene model.

2.4. Protein digestion

Protein extracts were denatured, reduced and alkylated. Briefly, the volume of each sample was adjusted to 100 µl with 100 mM triethylammonium bicarbonate (TEABC). 1 M DTT (1 µl) was added to each sample, which was then incubated for 30 min at 60°C before adding 10 µl 0.5 M iodoacetoamide (IAA). Alkylation lasted 30 min in the dark. Enzymatic digestion was performed by adding 2 µg trypsin (Gold; Promega, Madison, WI, USA) in TEABC 100 mM and incubating overnight at 30°C. Then, peptides were purified and concentrated using OMIX Tips C18 reverse-phase resin (Agilent Technologies Inc.) according to the manufacturer's specifications. Peptides samples were finally dehydrated in a vacuum centrifuge.

2.5. Shotgun proteomics

Samples were resuspended in 20 µl of 0.1% formic acid (buffer A) and 1 µl was loaded onto an analytical 50 cm reversed-phase column (0.75 mm i.d., Acclaim Pepmap 100 C18; Thermo Fisher Scientific) and separated with an Ultimate 3000 RSLC system (Thermo Fisher Scientific) coupled to a Q Exactive HF (Thermo Fisher Scientific) via a nano-electrospray source, using a 120 min gradient of 2–40% of buffer B (80% ACN, 0.1% formic acid) and a flow rate of 300 nl min^{-1}. MS/MS analyses were performed in a data-dependent mode. Full scans (375–1500 m z^{-1}) were acquired in the Orbitrap mass analyser with a 60 000 resolution at 200 m z^{-1}. For the full scans, 3×10^6 ions were accumulated within a maximum injection time of 60 ms and detected in the Orbitrap analyser. The 12 most intense ions with charge states ≥2 were sequentially isolated to a target value of 1×10^5 with a maximum injection time of 45 ms and fragmented by higher-energy collisional dissociation in the collision cell (normalized collision energy of 28%) and detected in the Orbitrap analyser at a resolution of 30 000.

2.6. Bioinformatic analyses of proteomic data

PEAKS Studio 8.5 (Bioinformatics Solutions, Waterloo, ON, Canada), a de novo assisted database software [32–34] was used to analyse MS/MS data from *E. singularis* venom, with parameters set as in [35]. The MS/MS spectra obtained from shotgun proteomics were matched to a custom database resulting from the whole-body transcriptome of the same colony, translated in the six open reading frames.

Protein isoforms were considered only when their existence was supported by MS/MS fragmentation data, whereas translated transcripts identified by PEAKS Studio as belonging to the same protein group were clustered with CD-HIT [36] using an identity threshold of 90%. Only the most representative sequence was retained for the subsequent annotation.

Protein annotation was performed as detailed in the electronic supplementary material and figure 1. Toxin nomenclature of mature polypeptides followed Oliveira *et al*. [37].

Structure predictions of mature toxins were performed with the colabfold implementation of AlphaFold2, using MMSeqs2 for multiple sequence alignment generation [38,39]. To increase model confidence for the most challenging targets, custom multiple sequence alignments with orthologue octocorallian sequences were used as inputs and each model was built using 12 recycles. All predictions were performed without the use of templates and energy minimization. The structures that were modelled with reasonable confidence (average pLDDT ≥ 70, which is generally recognized as the threshold for reliable models) [40] were submitted to the DALI [41] and RUPEE [42] servers for structural homology search against the Protein Data Bank. Homology modelling of the saposin-like toxins in the closed conformation was performed with Modeller v. 9.24 [43]. Electrostatic surfaces were calculated with APBS [44]. UCSF ChimeraX was used for model analysis, visualization and graphics [45].

2.7. Comparative genomic analyses

Toxins identified in *E. singularis* were used as queries for sequence homology searches against a selection of available anthozoan genomes, representative of the diversity of Octocorallia [46] and major Hexacorallia lineages summarized in electronic supplementary material, table S1.

Sequence homology searches were carried out with tBLAStn [47], based on an initial E-value significance threshold arbitrarily set to 10. Potential hits were manually curated, by annotating the inferred CDS, detecting putative splicing donor and acceptor sites with the aid of Genie [48]. The presence of *bona fide* orthologues between *E. singularis* and the other species was assessed through combined observations: (i) the conservation of the main topological feature of the encoded precursor

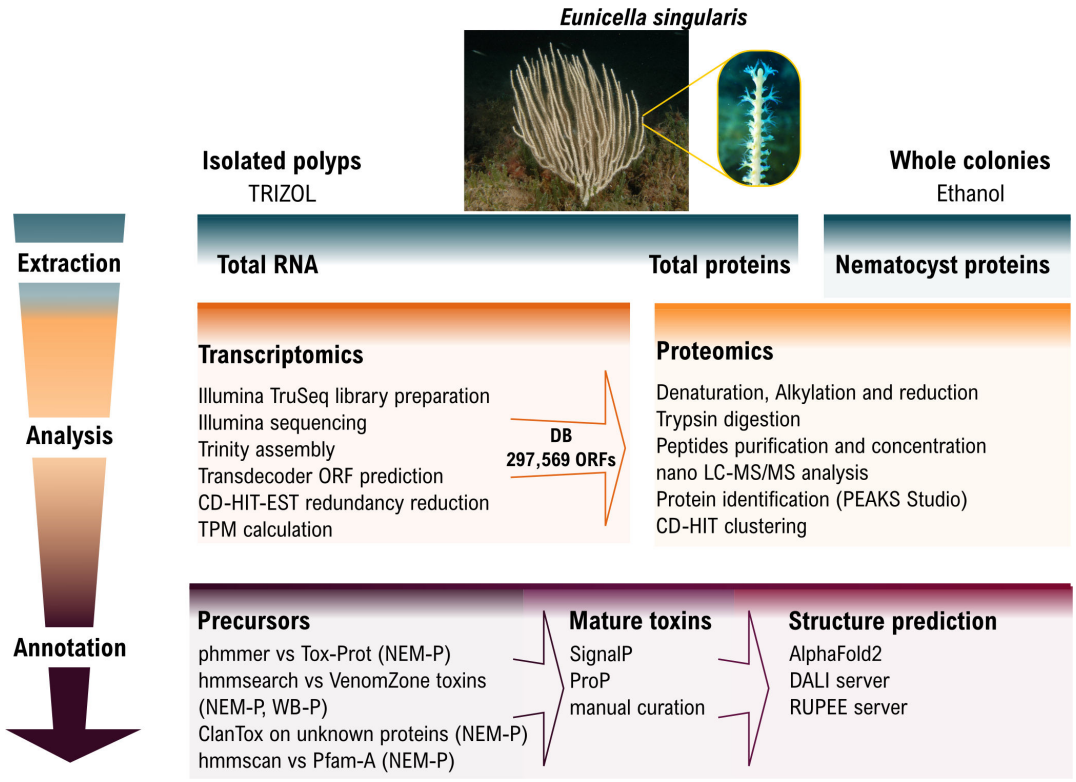

Figure 1. Schematic representation of proteotranscriptomics and bioinformatic annotation pipelines employed in this work.

proteins, such as the identification of a signal peptide and the cysteine residues involved in the formation of the disulfide bonds (whenever present); (ii) the shared presence of conserved protein domains, detected with a hidden Markov model (HMM) profile approach, based on the Pfam-A database; (iii) the conservation of gene architecture (e.g. intron number, phase and position), with respect to the orthologous gene from the congeneric species *Eunicella verrucosa*. Whenever needed (i.e. U-GRTX-Esi20 and 21), a fine-scale classification of orthologous sequences was aided by maximum likelihood phylogenetic inference analyses, carried out with IQ-Tree [49] with 1000 ultrafast bootstrap replicates based on the best-fitting model of molecular evolution detected by ModelFinder [50].

3. Results and discussion

3.1. *E. singularis* produces a complex venom

The extraction of venom from cnidarians, like *E. singularis*, is performed by stimulating the animals to fire their nematocysts and then purifying the toxins. Traditional methods use electrical or mechanical stimulation followed by toxin isolation from seawater. Alternatively, nematocysts can be isolated by density gradient centrifugation and mechanically disrupted to release their contents. Electrical stimulation of *E. singularis* colonies proved insufficient to recover enough crude venom for subsequent analysis. Hence, we attempted an ethanol-induced discharge method. While this has been validated for the box jellyfish *Chironex fleckeri* [29], its use for Anthozoan venoms has never been reported. Our results showed that the protein fraction recovered from ethanol-immersed *E. singularis* colonies had a composition comparable to the contents of nematocysts: proteomic analysis confirmed 162 translated transcripts (NEM-P; electronic supplementary material, table S2), reduced to 108 non-redundant precursor protein sequences. Based on an extensive bioinformatic annotation (see §2; electronic supplementary material; figure 1), 33 of these sequences were associated with structural components or basal physiology functions, 15 were enzymes, 28 were classified as unknown and 32 were identified as putative toxins (figure 2B).

Only 17 of them could be detected also in the WB-P (electronic supplementary material, table S3) of *E. singularis*, which, by comparison, comprised ca. 1400 translated transcripts. As the NEM-P constitutes a subset of the WB-P, its enrichment in putative venom components served as indirect confirmation of nematocysts firing due to the ethanol treatment. To identify additional putative toxins from other tissues (e.g. ectodermal glands), we also performed a targeted HMM search in the WB-P, looking for known cnidarian toxin families, as defined by VenomZone (https://venomzone.expasy.org). This search produced only four sequences previously undetected in the NEM-P: a protein with homology to the small cysteine-rich protein (SCRiP) family and three cytolysins (see below). The absence of these proteins in the NEM-P could be due either to incomplete ethanol-induced nematocyst firing due to differential reactivity of diverse types of nematocysts, or to the existence of additional venom delivery systems, such as ectodermal glands, as previously demonstrated for sea anemones [51], that were not affected by ethanol exposure. In alternative, these proteins may lack a venom function despite having toxin-like folds (e.g. they might be involved in neurotransmission or digestion). The NEM-P comprised some transcripts encoding precursors of multiple mature toxins, which were further classified by predicted fold (figure 1C; electronic supplementary material, table S4). In total, we

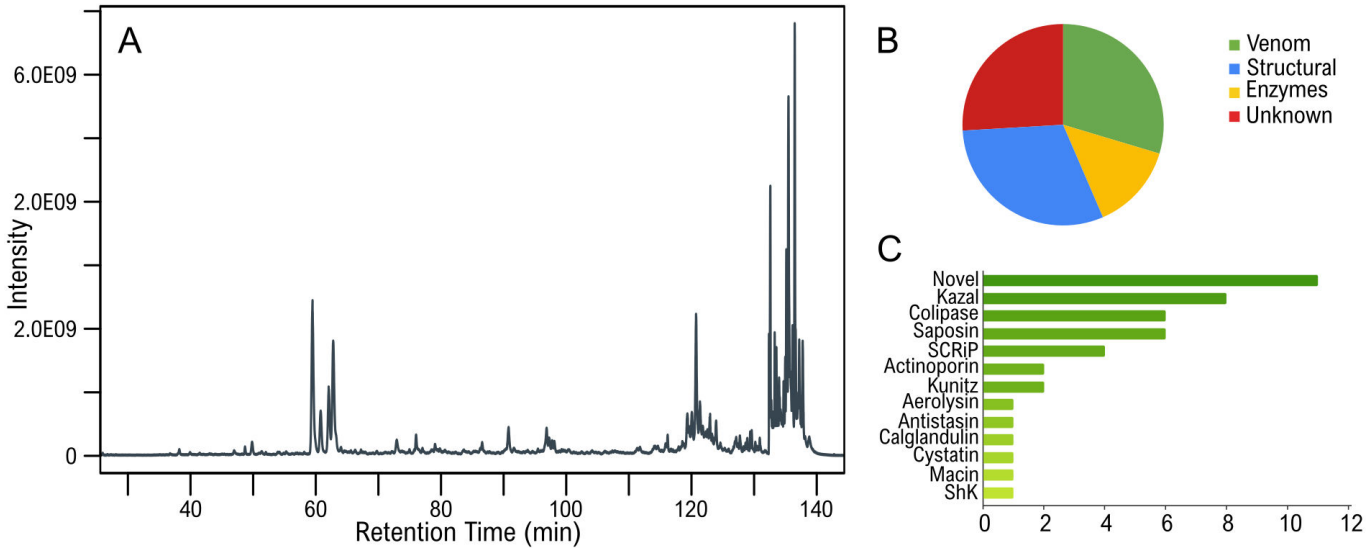

Figure 2. Structural and functional annotation of the venom from *E. singularis*. (A) Total ion chromatogram (TIC) of the ethanol extract. (B) Functional annotation of the translated transcripts confirmed at the protein level in the NEM-P. (C) Fold distribution of the mature toxins produced by *E. singularis* based on HMM profile search; the *x*-axis represents the number of proteomics-validated sequences.

identified 44 mature toxins, including a calglandulin, a protein characterized by multiple EF-hand motifs with an accessory role in venom production and processing, omitted from the detailed discussion below.

Following [37], the newly characterized toxins were collectively termed 'Gorgotoxins' (GRTX-Esi), according to the most conservative placement of *Eunicella* in the family Gorgoniidae, although recent work has proposed a separate family for this genus (Eunicellidae) [46].

3.2. Established toxin folds are present in the venom of *E. singularis*

3.2.1. Small cystein rich proteins-like (octo-SCRiPs)

We retrieved four toxins with homology to hexacorallian SCRiPs, which are known to induce severe neurotoxicity in zebrafish larvae, possess antimicrobial activity and potentiate the transient receptor potential ankyrin 1 (TRPA1) in mouse [52–54]. These comprised U-GRTX-Esi1 (particularly abundant at the transcript level, TPM 1585), 2a–b and 3 (figure 3). These proteins exhibited a peptide coverage of the predicted mature sequence exceeding 95%, shared a significant homology and up to 68% pairwise sequence similarity with hexacoral SCRiPs. Despite this similarity, each toxin showcased distinct features.

In particular, U-GRTX-Esi2a–b derived from a precursor containing two consecutive highly conserved SCRiP-like domains (pairwise identity 78.9%), separated by a protease cleavage site. Interestingly, while both forms could be detected in the WB-P, only U-GRTX-Esi2b was detectable in the nematocyst discharge (figure 3B; electronic supplementary material, figure S1).

U-GRTX-Esi3, solely detectable in the WB-P, featured a more complex architecture, comprising two SCRiP-like domains in the mature polypeptide, confirmed by the proteomic coverage (figure 3C). The absence of predicted cleavage sites between these domains would qualify U-GRTX-Esi3 as a secreted cysteine-rich repeat protein (SCREP), an emerging typology of venom proteins found in various organisms, often with high affinity and avidity towards their biological target and a multi-modal action [56].

Although SCRiPs-encoding genes were identified, often in large numbers and associated with tandemly duplicated gene clusters, in the genomes of all the octocoral species investigated in this study, they have a patchy occurrence in Hexacorallia, supporting a complex evolutionary history characterized by lineage-specific expansions and rapid diversification [56]. It is unclear whether SCREPs with two consecutive SCRiP domains were already present in the latest common ancestor of all Anthozoans, or independently acquired in different lineages. Indeed, clear orthologues of U-GRTX-Esi3 could be identified only in *Paramuricea clavata, Trachytela* sp. and *Renilla reniformis*: this architecture had been previously identified also in the hexacoral *Orbicella faveolata* [53]. However, while the only SCRiP so far reported in Octocorallia (in *Scleronephthya gracillima*) presented a conventional cysteine framework [56], the domains here identified display a shift in the cysteine pattern, which nonetheless did not affect the overall protein fold (electronic supplementary material, figure S2).

3.2.2. ShK-containing toxins

The ShK toxins, first identified in the venom of the sea anemone *S. helianthus*, are potent K_V1 channels blockers [57]. Our search for ShK domains in the proteomes of *E. singularis* led to the identification of U-GRTX-Esi4 in the NEM-P (TPM 103), with three consecutive ShK domains separated by short linker regions. Only the C-terminal domain has the typical ShKT cysteine pattern, whereas the first two domains present an unusual shift of the C-terminal cysteine. None of the domains of U-GRTX-Esi4 presents the key Lys residue necessary for binding $K_V1.2$ and $K_V1.3$, while the subsequent Tyr residue, also important for binding K_V1, is extremely conserved (electronic supplementary material, figure S3). This architecture resembles the precursors

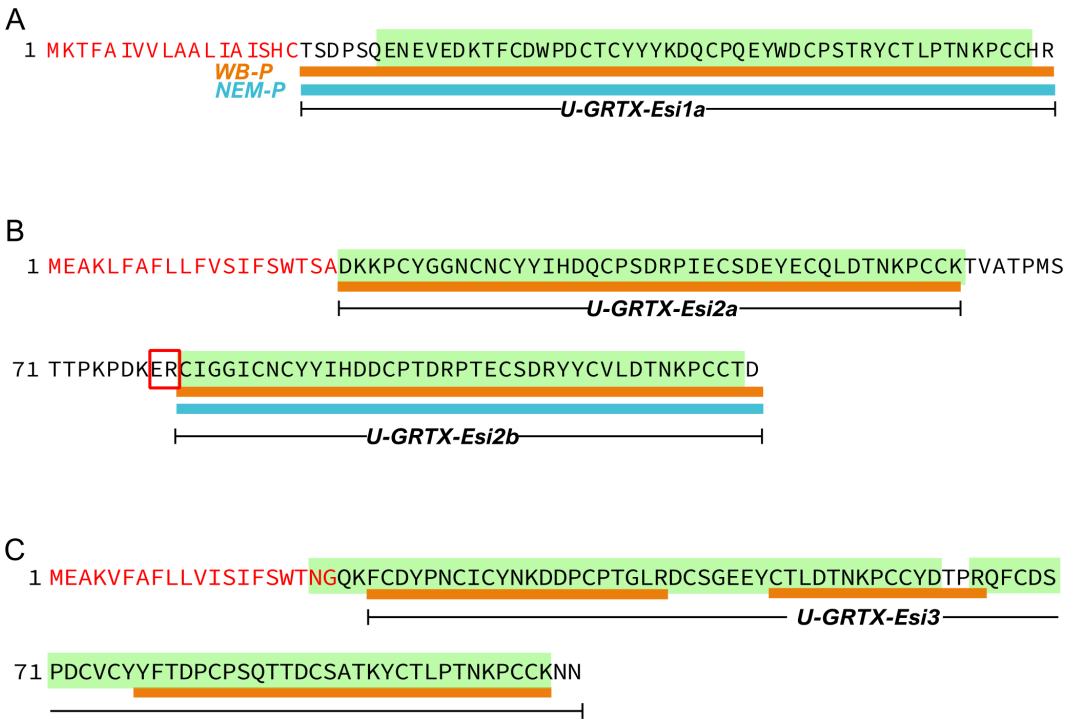

Figure 3. Tissue-specific maturation patterns of the SCRiP-like toxins from *E. singularis*. The sequences of the precursors of (A) U-GRTX-Esi1,(B) U-GRTX-Esi2a–b and (C) U-GRTX-Esi3 are reported. Red residues correspond to predicted signal peptide, green boxes define the SCRiP domains, the red rectangle defines a PQM protease recognition site [55], orange and blue lines indicate the peptide coverage of the precursors in the WB-P and NEM-P, respectively.

of the NEP3 protein family from *N. vectensis* [58]. However, only the first ShK domain of NEP3 is retained in the mature NEP3 toxin, while U-GRTX-Esi4 lacks putative cleavage sites and we found proteomic evidence for the first and the last ShK domain, suggesting the occurrence of a multidomain protein in *E. singularis*.

3.2.3. Protease inhibitors

Toxins containing domains associated with protease inhibition are widespread across the tree of life. Among them, Kunitz-like peptides are frequently found in venoms [59] where they can act as inhibitors of several endogenous and exogenous proteases [60] but also show neurotoxic and neuromodulatory activity on a variety of receptors and ion channels [61,62], sometimes with multimodal action [57,59]. Kunitz-like toxins that bind ion channels can be used to paralyse prey and are being extensively studied for their analgesic and neuroprotective properties [63–67]. Our HMM search revealed the occurrence of two Kunitz-like toxins. PI-GRTX-Esi5, detectable in the NEM-P only, presented two Kunitz domains separated only by an Asp-Ile dipeptide, an architecture previously detected in salivary proteins of blood-feeding animals [67,68]. Since orthologous genes sharing the same three exons/two introns organization of PI-GRTX-Esi5 were not detected in Scleralcyonacea, this domain combination probably represents a recent innovation of the Malacalcyonacea. The second toxin, PI-GRTX-Esi6, with a single Kunitz domain, was detected both in the NEM-P and in the WB-P, and derived from a precursor with a complex architecture, with the Kunitz-like domain followed by three Kazal-type domains, all separated by propeptidase cleavage sites. Also in this case, proteomic evidence indicated a differential maturation pattern of the precursor, with the Kazal-like peptides from this precursor only retrieved in the WB-P (electronic supplementary material, table S2). The Kunitz domain of PI-GRTX-Esi5 presented the Arg residue (conventionally 'P1') associated with the trypsin inhibition, whereas its replacement with a Met in both domains of PI-GRTX-Esi6 suggests a chymotrypsin specificity [69]. In addition, PI-GRTX-Esi6 presented a Leu residue aligning with the crucial position for K_V1 binding of LmKTT, a scorpion Kunitz-like toxin [70]. These similarities may hint to multimodal activity for PI-GRTX-Esi5, while the lack of key residue conservation in PI-GRTX-Esi6 does not allow inferring any putative function (electronic supplementary material, figure S4).

Other peptides present in *E. singularis* nematocysts displaying protease inhibitory domains (Kazal-type, cystatins, antistasins and macins) were detected but did not present novelty elements; they are briefly described in the electronic supplementary material.

3.2.4. Colipases

Finally, a striking feature within the NEM-P of *E. singularis* was the abundance of toxins with a predicted colipase fold (U-GRTX-Esi14–18). Colipases are knottins containing five disulfide bridges, found in spider, snake and tick venoms [71], but to date not in Cnidaria. The prototypical example, MIT-1 [72] interacts through its N-terminal pentapeptide AVITG with prokineticin receptors, stimulating smooth muscle contractions and exhibiting hyperalgesic effects [73,74]. *Eunicella singularis* colipases lacked the AVITG pentapeptide, a feature shared with the atracotoxins of the spider *Hadronyche versuta*, whose pharmacological target is unknown to date [75].

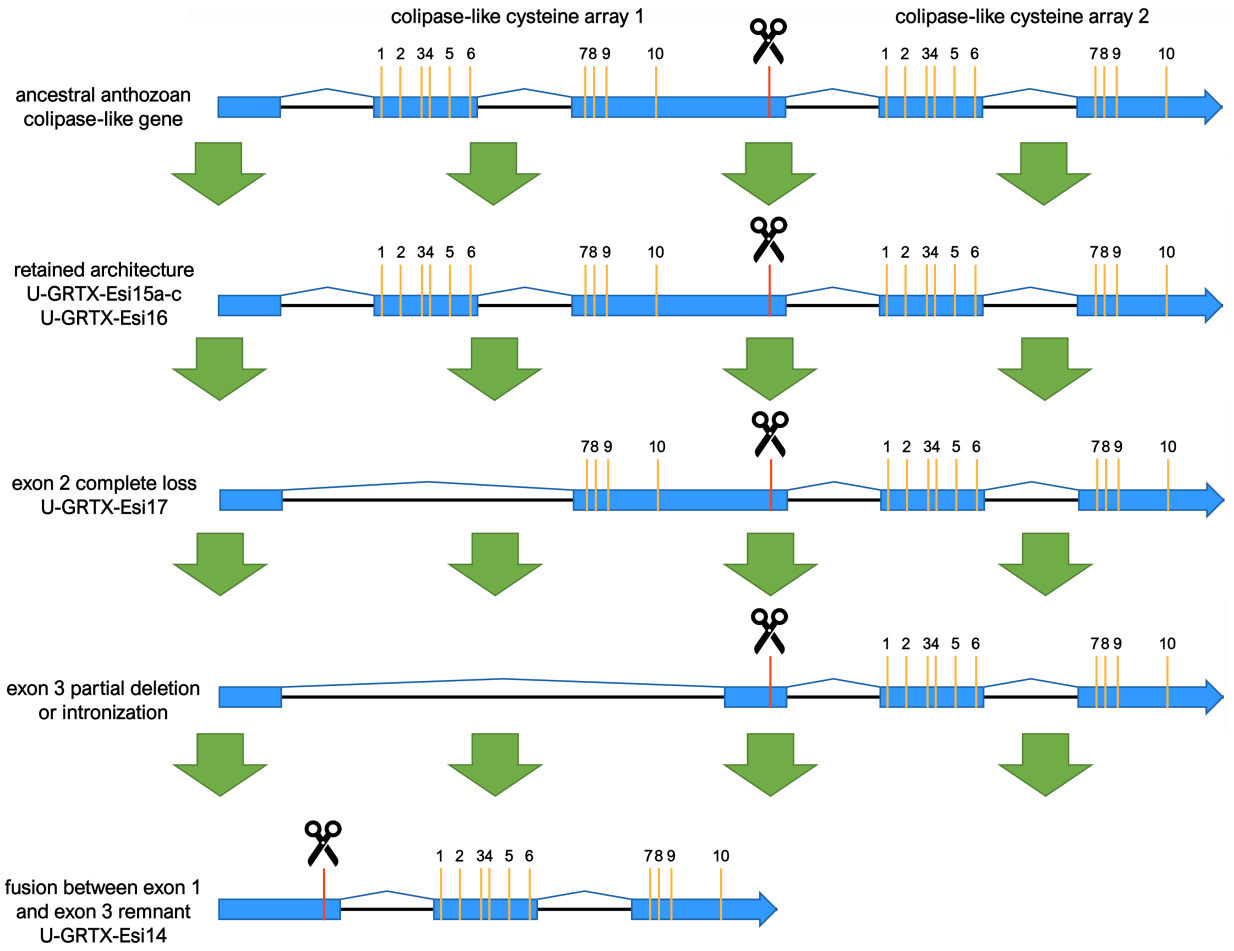

Figure 4. Schematic representation of the evolution of the colipase genes in Anthozoa, with references to specific *E. singularis* genes. Exons are in blue, cysteine residues are indicated by the yellow bars, scissors indicate cleavage sites.

Apart from the most abundantly expressed protein, U-GRTX-Esi14, all colipases are derived from modular precursors containing multiple domains separated by propeptidase cleavage sites. Both U-GRTX-Esi15a–c and U-GRTX-Esi16 were encoded by genes displaying a five exons/four introns architecture (figure 4), well-conserved in several orthologues from multiple Octocorallia and Hexacorallia species, pointing out a shared ancient origin with secondary losses in some lineages. U-GRTX-Esi14 and U-GRTX-Esi17 probably derive from paralogous gene copies, which underwent structural reorganization during evolution. In detail, U-GRTX-Esi17 entirely lost exon 2, which resulted in a shorter precursor protein with an incomplete disulfide pattern. U-GRTX-Esi14 further lost the colipase-fold encoded by exon 3, which was probably reduced and later fused with exon 1. This scenario is supported by the observations that exon 1 in U-GRTX-Esi14 is significantly longer than its paralogues and includes a dibasic propeptidase cleavage site usually present at the 3′ end of exon 3. Nevertheless, such remarkable modifications occurred in the latest common ancestor of all Malacalcyonacea, as evidenced by the identification of U-GRTX-Esi14 and U-GRTX-Esi17 orthologues in *Dendronephthya gigantea*, *Muricea muricata*, *P. clavata*, *Trachythela* sp. and *Phenganax subtilis* (the latter limited to orthologues of U-GRTX-Esi14).

The loss of exon 2 in the gene encoding for U-GRTX-Esi17 results in a shorter precursor protein, with an N-terminal incomplete colipase domain. The maturation of this precursor leads to the release of a polypeptide lacking 6 out of the 10 cysteines (U-GRTX-Esi18), whose presence was also confirmed in the NEM-P. The lack of most of the cysteine scaffold has a major impact on the protein structure: the model predicted with good confidence by AlphaFold2 shows in fact only a long β-hairpin flanked by disulfide stabilized tails.

Notably, the transcripts for all these toxins had exceptionally high TPM values (1806, 569, 826 and 429, respectively, for the U-GRTX-Esi14 to 17/18), compared with other proteins detected in the NEM-P (cf. electronic supplementary material, table S4).

3.3. Novel putative toxins were retrieved in the NEM-P of *E. singularis*

We identified in the NEM-P 10 additional proteins possessing the typical sequence features of toxins (small size, abundance of Cys, presence of a signal peptide), lacking sequence similarity with any known protein and often encoded by highly expressed transcripts. Their genomic organization is summarized in electronic supplementary material, figure S5. High-confidence structural models obtained with Alphafold2 were used to retrieve structural homologues in the Brookheaven Protein Data Bank, in the attempt to infer their biological functions (figure 5).

Among them, U-GRTX-Esi19 was encoded by an extremely abundant transcript (TPM 2898). The mature protein (100 amino acids, 6 Cys) was detected in both the NEM-P and WB-P and had a weak similarity at the C-terminus with Amoebapore A, the major pore-forming protein of *Entamoeba histolytica*, suggestive of a possible cytolytic function [76]. Most of the protein structure

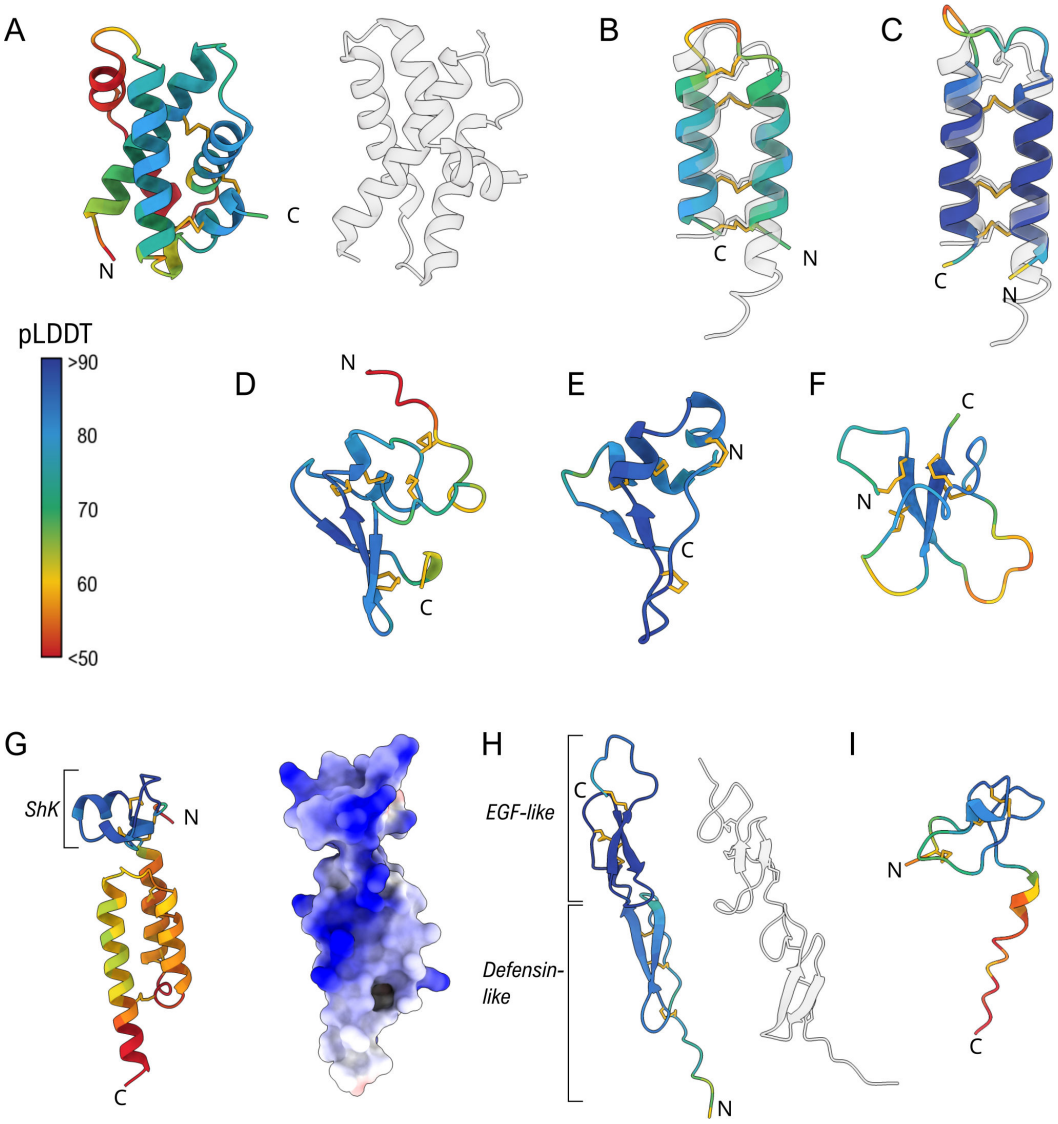

Figure 5. Overview of the structures predicted by AlphaFold2 for the novel toxins from *E. singularis*. Models are coloured according to pLDDT, with higher values (>70) indicating reliable prediction. (A) The putative cytolytic U-GRTX-Esi19 and, on the right, PulS from *Kebsiella oxytoca* (PDB 4A56). Overlay of the structures of the two disulfide stabilized helical hairpin toxins U-GRTX-Esi20 (B) and U-GRTX-Esi21 (C) with acrorhagin-1 (PDB 6UX5, right); The predicted structures of the two novel αβ toxins U-GRTX-Esi22 (D) and U-GRTX-Esi23 (E). (F) The SCRiP-like arrangement of U-GRTX-Esi24. (G) The two-domain organization of U-GRTX-Esi25 (left) and its surface charge distribution (right). The ShK domain is adjacent to an all-helical domain. The extended blue patch on the protein surface is common in cytolytic proteins. (H) U-GRTX-Esi26 (left) compared with the byssal protein Pvfp-5β from the mussel *Perna viridis* (PDB 7QAB, right). (I) The structure of U-GRTX-Esi28.

was modelled with high confidence (pLDDT>80; figure 5A), highlighting a compact fold characterized by a knotted four-helix bundle stabilized by three inter-helices disulfide bridges. A similar arrangement, although devoid of disulfide stabilization, is typical of bacterial pilotins PulS and GspS, chaperone lipoproteins that play a role in the assembly of the T2SS secretion system in pathogenic Gram-negative bacteria [77,78]. This putative toxin appears as an evolutionary innovation of Malacalcyonacea, although orthologues are missing in *P. subtilis* and in *Xenia* sp. (figure 6; electronic supplementary material, figure S6). The cysteine framework is always conserved, while the length of some loops varies across species. The abundance of this protein suggests a strong functional relevance, but activity assays will be necessary to clarify its function.

The two putative toxins U-GRTX-Esi20 and 21 did not show any obvious sequence similarity, with a pairwise identity of only 21%. U-GRTX-Esi20 is 48 amino acids long and derives from a longer propeptide proteolytically cleaved at the C-terminus, whereas U-GRTX-Esi21 displays a shorter sequence with a very high cysteine density (8 out of 36 amino acids). The models obtained (figure 5B,C) highlighted a structural relationship, as both presented a disulfide-stabilized helical hairpin motif, recently characterized in acrorhagin-1 (U-AITX-Aeq5a) [79], a toxin from the acrorhagi of *Actinia equina* used in conspecific aggressive encounters [80]. Acrorhagin-1 can induce local tissue necrosis and displays toxicity on crabs, but its mechanism of action has not been clarified, although it has been demonstrated that the toxin does not interact with either potassium or sodium ion channels and it does not possess membranolytic activity [79,80]. U-GRTX-Esi20 and 21 are stabilized by three and four disulfide bridges, respectively, which overlap with those of acrorhagin-1, despite a sequence identity below 20% in both cases.

Our genomic survey indicates that these two toxins belong to two distinct monophyletic orthogroups within a very large superfamily of cysteine-rich peptides, encoded by ancestrally duplicated paralogous genes with intronless structures, that also include other members in *E. singularis* transcriptome, not detected in the NEM-P (figure 7). High-confidence orthologues were

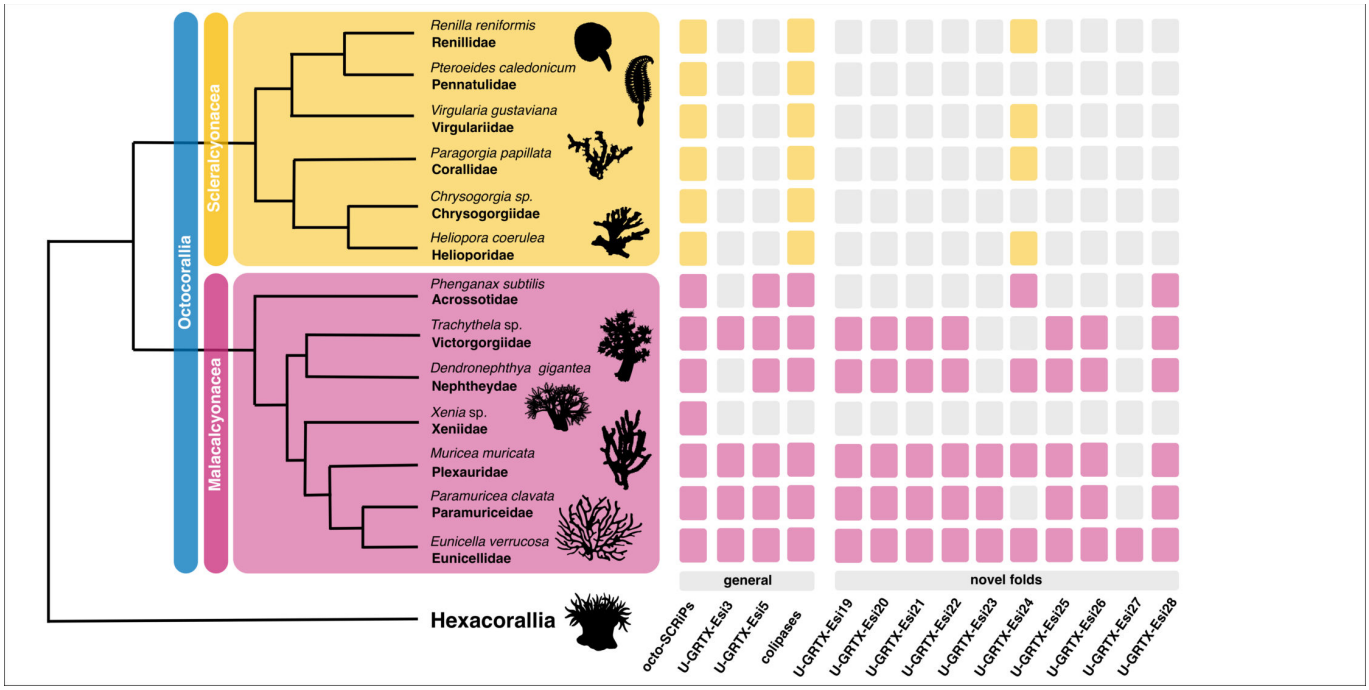

Figure 6. Distribution of orthologues of a subset of *E. singularis* toxins across the Octocorallia genomes here investigated. Phylogenetic relationships of the Octocorallia after [46].

exclusively found in Malacalcyonacea, except for *P. subtilis* and *Xenia*, forming six well-supported clades (Type I–Type VI) (figure 7). This suggests a rather recent origin of these orthogroups, which derive from a large group of similar peptides also found in other cnidarians, worthy of more extensive investigation. In fact, an orthologue of acrorhagin has been detected in the sensory neurons of *Nematostella vectensis* [81], suggesting an ancestral recruitment from the nervous to the venom system for this class of peptides, followed by lineage-specific diversification in the Octocorallia (figure 6).

The next two putative toxins, U-GRTX-Esi22 (60 amino acids and 10 cysteines; figure 5D) and 23 (45 amino acids and 8 cysteines; figure 5E) did not present any sequence similarity with known proteins. In both cases, the models highlighted an unusual disulfide-stabilized $\alpha\beta$ fold, suggesting the existence of peculiar and to date undiscovered cnidarian toxin folds. The gene encoding U-GRTX-Esi23 is present with two similar paralogous genes in *E. singularis* and has clearly detectable orthologues in all Malacalcyonacea except *P. subtilis* and *Xenia* sp. (electronic supplementary material, figure S6; figure 6). A divergent paralogue, with a slightly different cysteine pattern, is found in Scleralcyonacea. This toxin probably constitutes an evolutionary innovation of Malacalcyonacea but could derive from an ancestral sequence with a similar cysteine connectivity present in the latest common ancestor of all Octocorallia. Conversely, U-GRTX-Esi22 presents a single-copy gene, with a detectable orthologue only in *P. clavata* and *M. muricata*, and is probably a recent innovation, only shared by a few related orders within Malacalcyonacea (electronic supplementary material, figure S7; figure 6).

The structure of U-GRTX-Esi24, a 45 amino acids protein with 8 cysteines and a 50% sequence identity with Conotoxin Cltx-2 from *Conus californicus*, could also be modelled with acceptable confidence (figure 5F). Despite undetected by HMM profile analysis, the structure presented a SCRiP-like W-shaped fold. This domain has a broad but scattered distribution in Octocorallia, suggesting an evolutionary origin slightly older than most other toxins described in this work. Nevertheless, it is most often found repeated multiple times within the same protein (also in *E. singularis*), with several independent occurrences of single-domain peptides scattered across the tree of life (figure 6; electronic supplementary material, figure S8).

Unique case among the novel toxins from *E. singularis*, U-GRTX-Esi25 was predicted to possess a modular architecture: a recognizable N-terminal ShK domain followed by a region lacking any significant homology. The two domains were clearly distinguishable in the three-dimensional prediction (figure 5G), where they appeared as directly connected, without spacer sequences or putative cleavage sites. The unknown C-terminal domain was modelled as an asymmetrical three-helices arrangement, with three disulfide bridges connecting and stabilizing the secondary structure elements. Interestingly, despite the complex architecture predicted, several basic residues from both domains in the model aligned perfectly, forming an extended positively charged patch (figure 5G), a feature that suggests a possible membranolytic and/or antimicrobial action [82]. This peculiar combination of domains, paired with a shared three exons/two introns gene architecture, was detected in Malacalcyonacea (except *P. subtilis* and *Xenia* sp). A single orthologous gene was found in *Paragorgia papillata* (Scleralcyonacea), suggesting that the appearance of this protein could predate the split between Malacolcyonacea and Scleralcyonacea (figure 6; electronic supplementary material, figure S9).

The model predicted for U-GRTX-Esi26 (77 residues, 12 Cys) presented an elongated structure made of two domains: an N-terminal region in which two antiparallel beta strands are connected by three disulfide bridges to a proline-rich stretch, in a defensin-like fold found in other cnidarian toxins, and a C-terminal EGF-like domain (figure 5H). A similar modular structure was recently described in the byssal adhesive protein Pvfp-5β from the mussel *P. viridis* [83]; despite the absence of sequence similarity, this structural similarity might indicate for this protein a structural or adhesive function. However, orthologues of the gene were found only in Malacalcyonacea (apart from *Dendronephthya* and *Xenia*), all sharing a common three exons/two introns

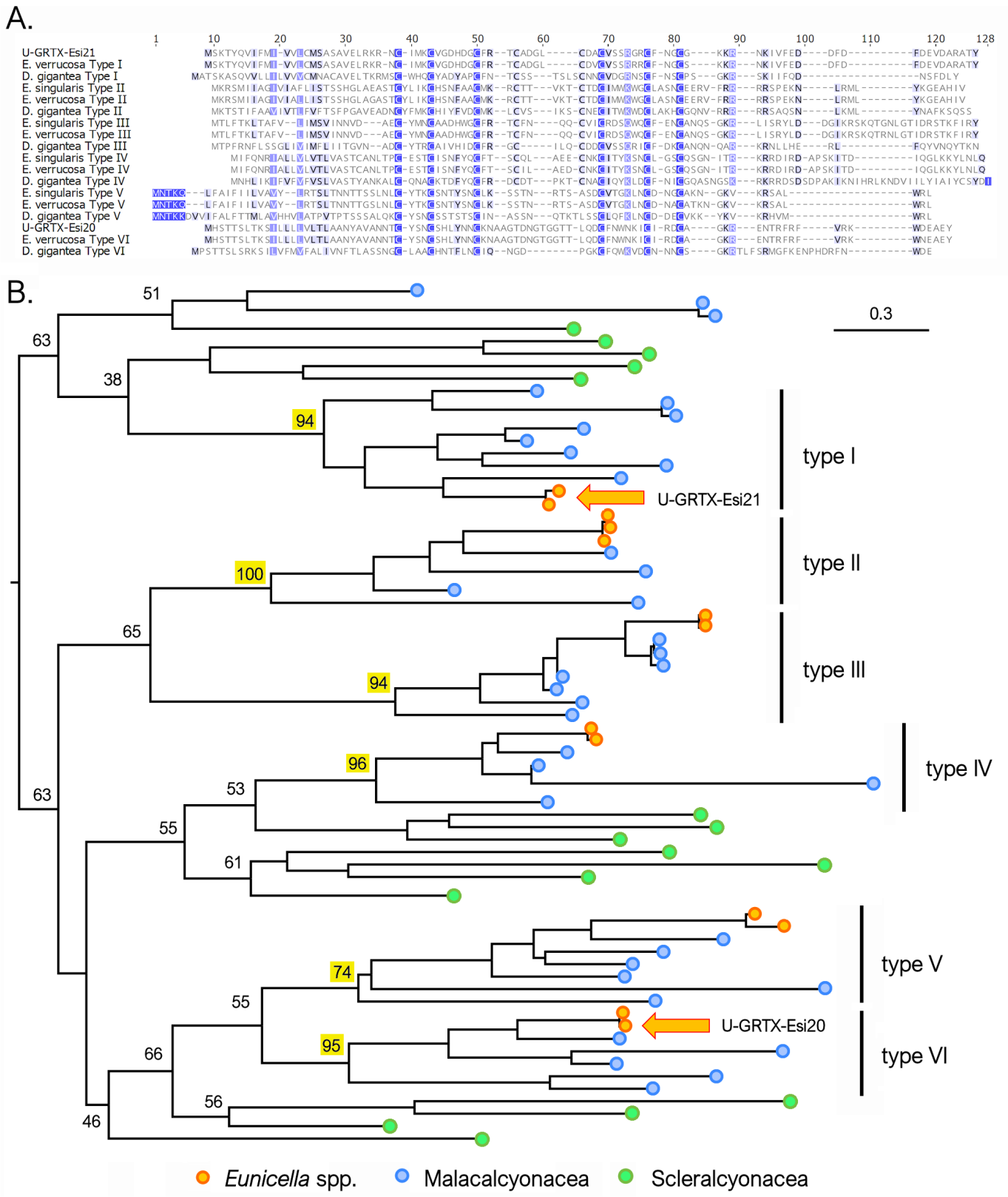

Figure 7. (A) Multiple sequence alignment of the complete precursor amino acid sequences of U-GRTX-Esi20, U-GRTX-Esi21, their paralogues and their orthologues in *E. verrucosa* and *D. gigantea*. (B) Maximum likelihood phylogeny of the sequences belonging to the same superfamily of U-GRTX-Esi20 and U-GRTX-Esi21 (indicated with orange arrows), identified in different Malacalcyoneacea and Scleralcyonacea species. All Malacalcyonacea sequences (excluding those from *Xenia*) can be classified within six highly supported monophyletic clades. Bootstrap support values are only shown for the major nodes of the tree, and those denoting the basal node of the six aforementioned clades are highlighted with a yellow background.

gene structure. Three divergent paralogues were found In *P. clavata*, probably originated through a species- or genus-specific gene duplication process (electronic supplementary material, figure S10).

The last two novel toxins identified in the NEM-P of *E. singularis*, U-GRTX-Esi27 and 28, were short peptides associated with relatively low transcription levels but a high proteomic coverage. In particular, U-GRTX-Esi27 displayed a weak sequence similarity to both SCRiPs and BDS toxins from *Anemonia sulcata*, which are potent ligands of K_v3 channels [84]. The mature sequence of U-GRTX-Esi27, a 39 amino acids long peptide, contained 8 cysteines, 6 of which correspond to the BDS framework.

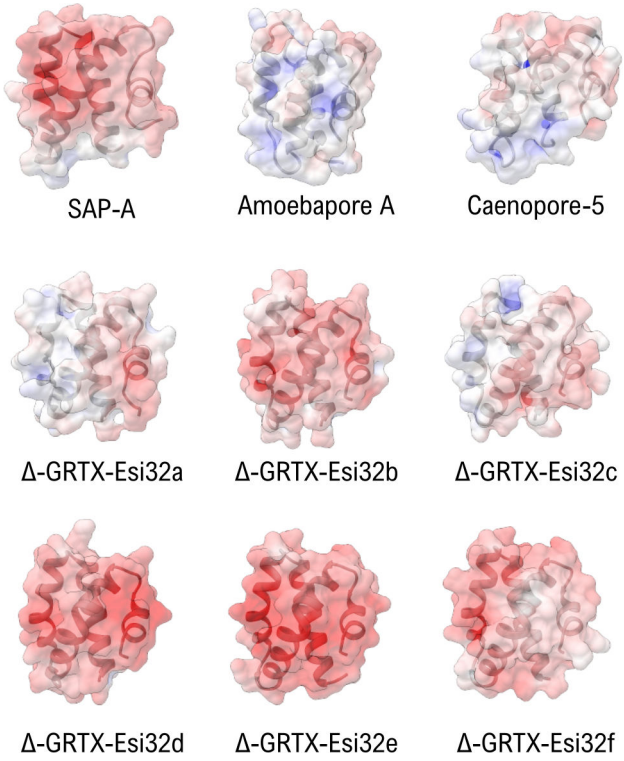

Figure 8. Comparison of the surface electrostatic potentials of the three model saposin-like proteins SAP-A (4UEX), Amoebapore A (1OF9) and Caenopore-5 (2JSA) with the six SAPLIPs identified in the venom of *E. singularis*. The extended patches of negative charge (red) observed in Δ-GRTX-Esi32b, d, e, f are observed in human SAP-A and are indicative of a pH-dependent conformational switch. The lower charge density of Δ-GRTX-Esi32a and c may indicate a different mode of action.

An additional cysteine pair occupied the position of two extremely conserved residues in sea anemone BDS toxins, i.e. a tryptophan and a tyrosine residue, otherwise present in all other members of the family (electronic supplementary material, figure S11). Only low confidence models could be predicted by AlphaFold2, thus preventing any function inference. Despite the similarity with BDS, orthologous genes encoding for similar proteins were not detected in the genomes of the other Octocorallia considered in this study except from *E. verrucosa*. U-GRTX-Esi27 might therefore constitute an evolutionary innovation of the genus *Eunicella* or the family Eunicellidae (figure 6). Finally, U-GRTX-Esi28 possessed a weak sequence homology with the ω-hexatoxin-Hi1g. The modelled structure confirmed this similarity, revealing a cystine knotted core (figure 5I), with the addition of an extended and disordered C-terminus, possibly proteolytically cleaved. Orthologues coding for this protein are shared by nearly all Malacalcyonacea but are lacking in *Xenia* (electronic supplementary material, figure S12). This finding, along with the conservation of the gene structure, in which two exons are spaced by an intron with an equal placement in all species, the signal peptide sequence, and the cleavage sites, strongly indicated that this protein represents an evolutionary innovation of Malacalcyonacea (figure 6).

3.4. A diverse range of cytolytic proteins are produced both inside and outside the nematocysts

Our analysis highlighted the absence, in the NEM-P of *E. singularis*, of the common cnidarian cytolytic proteins. A targeted hmmscan search in the WB-P yielded instead positive hits for profiles associated with both the aerolysin (PF01117) and the full anemone cytotoxin (PF06369) families. Coherently with the approach adopted for the NEM-P annotation, we considered only the proteins encoded by complete transcripts presenting a signal peptide; this approach resulted in the identification of three pore-forming toxins (PFTs), characterized by diverse structural features. The first one, Δ-GRTX-Esi29, was an aerolysin-like protein, structurally related to the β-PFT Hydralysin [85]; the other two, Δ-GRTX-Esi30 and 31, which shared ~50% sequence similarity, belonged to the actinoporins family, the most abundant family of cnidarian porins [86]. Interestingly, we observed relevant non-conservative substitutions in *E. singularis* actinoporins, which may indicate selectivity towards different lipids (electronic supplementary material, figure S13).

Finally, the highest scoring match in the proteomic analysis of the nematocysts content of *E. singularis* corresponded to a 699 amino acid translated transcript with homology to prosaposin (pSAP) [87]. In humans, the maturation of the orthologous pSAP produces four globular saposins (SAP-A/D), which are co-factors of sphingolipid degrading enzymes. The saposin fold is very ancient and characterizes all the saposin-like proteins (SAPLIPs), which include the amoebapores [88–90]. SAPs and SAPLIPs have a distinctive cysteine framework but display minimal conservation (<25% sequence identity). They are often released in several copies from the same precursor gene [91], assuming a closed, monomeric conformation at lysosomal pH [88,92]. All these proteins are involved in membrane destabilization/solubilization, acting through different mechanisms: lysosomal SAPs undergo a conformational change upon pH shift or in the presence of lipids or detergents, exposing the residues of the hydrophobic core and encapsulating a large number of lipids [93–95]. Other SAPLIPs, like the amoebapores from *E. histolytica* and the caenopores from *Caenorhabditis elegans*, form oligomeric pores [96,97]. Finally, basic SAPLIPs, like NK-lysin and granulysin, act by 'molecular electroporation', inserting within the membranes [98,99]. In rare cases, the saposin fold has also

been detected in non-cytolytic neurotoxins, like the macro-conotoxins con-ikot-ikot and Mu8.1 [100,101], but these proteins are highly dissimilar from SAPLIPs in terms of sequence and cysteine pattern and are encoded by single domain (non-pSAP-like) transcripts. The six sequences deriving from the maturation of the pSAP precursor of *E. singularis* (Δ-GRTX-Esi32a–f) were all identified at the proteomic level in the NEM-P. These proteins shared 24–65% similarity, higher than what generally observed for homologous SAPLIPs, and were more similar to human saposins SAP-A and C (similarity 26–40%) than to antimicrobial porogenic peptides like amoebapore A (16–25%) or caenopore-5 (13–27%).

We examined the surface potential distribution of the SAPLIPs from *E. singularis*, since the activity of SAPLIPs correlates to their surface electrostatics [102,103]. Except Δ-GRTX-Esi32a and c, they all shared with human SAP-A extended patches of negative charge (figure 8), which hints to a pH-dependent mechanism for activation and membrane interaction [104]. In contrast, Δ-GRTX-Esi32a and c presented a more balanced distribution of basic and acidic regions and extended neutral patches, like caenopore-5, suggesting a different, possibly non-monomeric, native arrangement and a pH-independent mechanism of action, possibly through pore formation.

4. Conclusions

We have provided the first preliminary depiction of the venomous arsenal of a widespread octocoral. Although experimental limitations did not allow us to conduct our analysis on an isolated venom fraction, by exposing whole polyps of the gorgonian *E. singularis* to ethanol [29] we isolated a sample enriched in proteins presenting typical toxin features. Among them, we detected several folds previously identified in cnidarian venoms, some proteins with structural similarities to venom components of unrelated organisms, and a few with new folds. Even in the absence of functional characterization, the NEM-P seems to provide a good snapshot of the nematocysts content in *E. singularis*, although we could not rule out the possibility that these proteins derive from other tissues. The scarce similarity observed between octocoral and hexacoral putative toxins repertoire is not surprising, given the ancient divergence between these two lineages that recent works date to about 500–700 Ma [2]. In comparison, Elapidae and Viperidae snakes, which present remarkable differences in their venom composition, have diverged only around 50 Ma [105]. However, our results clearly indicate, within the NEM-P of *E. singularis*, an elevated content of putative neurotoxic peptides, a feature previously reported for sea anemones. Cytolytic toxins with different structures (and probably specificity/function) seem distributed both inside and outside the nematocysts.

The transcriptionally more abundant components of *E. singularis* venom are the potentially cytolytic U-GRTX-Esi19 and the acrorhagin-like U-GRTX-Esi20, the six colipases, the SCRiP-like U-GRTX-Esi1a and the SCREP U-GRTX-Esi4 with three ShK domains. With the exception of the 'octo-SCRiPs', which, however, are characterized by sequence peculiarities, these toxins represent evolutionary innovations of Octocorallia, either shared by the entire lineage (e.g. some colipases), or restricted to the Malacalcyonacea (e.g. U-GRTX-Esi19), or even exclusive of a group of sublineages within the Malacalcyonacea (e.g. U-GRTX-Esi20) (figure 6).

Putative neurotoxins in *E. singularis* included both novel and established cnidarian folds, the latter often detected within multidomain SCREPs [106]. This is the case, for example, of the highly expressed U-GRTX-Esi4, with three ShK domains: multidomain ShK proteins have been so far reported in the transcriptomes of parasitic roundworms, in *Hydra* and *Nematostella* [58] and in the vampire snails [107]. The toxin U-GRTX-Esi3 represents another case of SCREP, presenting a double SCRiP domain in the mature polypeptide: this architecture had not been documented in Cnidaria mature toxins to date, although it has been recognized at the transcript level (e.g. in *O. faveolata*). Two domains were also retrieved in the Kunitz BPTI-like PI-GRTX-Esi5, an architecture so far investigated in detail in hematophagous invertebrates only, but extremely abundant among the putative ion-channel impairing SCREP toxins [67,106]. While the abundance of the SCREPs is probably underestimated across venomous Metazoa, also due to the paucity of proteomic studies confirming the architecture of mature toxins, their abundance in the venom of *Eunicella* might indeed possess some functional significance.

Interestingly, conventional PFTs of Cnidaria, including the aerolysin-like Δ-GRTX-Esi29 and the two actinoporins Δ-GRTX-Esi30 and 31 were not retrieved in the NEM-P. As anticipated, this could depend on the intrinsic limitations of the venom collection procedure employed or reflect an extra-nematocystic production of these proteins. Indeed, finding PFTs outside the nematocysts is not unusual (e.g. hydralysin is an exclusively non-nematocystic PFT involved in prey digestion [84]). Actinoporins, on the other hand, have been to date detected both in the nematocysts and in other tissues. The extra-nematocystic localization retrieved in *E. singularis* might suggest that its actinoporins share the digestive function of the hydralysin-like Δ-GRTX-Esi29. Further histological and functional analysis are needed to confirm this hypothesis, and to evaluate the impact of the observed mutations at critical sites on chemical selectivity and pore formation mechanism of *Eunicella* actinoporins.

Among the proteins of the NEM-P, probably representing the nematocyst content, the cytolytic activity seems to be ascribable mostly to the six saposins Δ-GRTX-Esi32a–f, abundant at the proteomic level. The analysis of their surface charge hints to a considerable diversification of targets and modes of action for this novel group of cytolytic polypeptides: most of them might be inactive at the low pH typical of mature nematocysts but could be activated by the abrupt pH change which follows discharge [108]. On the contrary, the neutral surface electrostatic of Δ-GRTX-Esi32a and c might suggest a pH-independent activation. The novel toxin U-GRTX-Esi19, transcribed at high levels and with a fold compatible with a cytolytic activity, might complement saposins.

While some of these putative toxins are shared by all the investigated octocorals, others had a taxonomic range restricted to Malacalcoynacea and were not retrieved in the analysed Scleralcyonacea, suggesting they might constitute evolutionary innovations arisen after the split between the two major Octocorallia lineages. This is the case of the novel toxins U-GRTX-Esi19 and U-GRTX-Esi28 (figure 6). In addition, several toxins were present in a subset of Malacalcyonacea species and are frequently

lacking in the representatives of two specific families, Xeniidae (or pumping corals) and Acrossotiidae (i.e. tube corals) (figure 6). This peculiar taxonomic distribution might suggest a key role in predation for these toxins, since they are absent in species with a mixotrophic habit, in which autotrophy sustained by symbiotic dinoflagellates is the most relevant source of energy. In contrast, the other Malacalcyonacea included in our analysis (e.g. *Paramuricea*, *Trachytela*) are suspension feeders that largely rely on zooplankton intake [109]. Novel toxins, such as U-GRTX-Esi26, U-GRTX-Esi22, the acrorhagin-like U-GRTX-Esi21 and U-GRTX-Esi27 might thus constitute a predation-related group of putative toxins. In particular, U-GRTX-Esi27 appears to be restricted to *Eunicella* as a genus-specific adaptation (figure 6).

The inference on the potential function of octocoral putative toxins is, however, heavily affected by the sparse information available on the biology of these organisms. For example, some Octocorallia species possess sweeper and thread-like tentacles, nematocysts-rich structures that have been hypothesized to mediate interspecific competition, similarly to sea anemones acrorhagi [110]. The occurrence of these structures across the Octocorallia radiation has been poorly investigated to date: this lack of comprehensive information hampers the evaluation of a potential role of specific toxins in competition. Nevertheless, although our comparative genomic approach has revealed a powerful tool to restrict the range of potential functions, further functional characterization studies are needed to unequivocally establish the physiological role of these toxins.

Interestingly, our proteomic analysis pointed towards two cases of differential toxin maturation pathways in the nematocyst versus whole body. The first concerns the two SCRiP isotoxins U-GRTX-Esi2a–b, derived from a double-domain precursor, with both isoforms present in the WB-P while only U-GRTX-Esi2b represented in the NEM-P (electronic supplementary material, figure S1). The second corresponds to the toxin PI-GRTX-Esi6 and the three isotoxins PI-GRTX-Esi7a–c, encoded by a single transcript in which a Kunitz-like domain (retrieved both in the whole body and in the nematocyst proteomes) is followed by three Kazal-type PI domains only identified in the WB-P. The contribution of alternative cleavage sites to venom diversity has been demonstrated in conopeptides [111], but cell- or tissue-specific differential maturation patterns have not been so far documented for venom toxins, suggesting that the mechanisms regulating venom toxin expression could be more complex than previously assessed.

Through our multidisciplinary approach, we have demonstrated that Octocorallia produce a complex neurotoxin-rich venom, showcasing an overall scarce resemblance to the hexacoral venom repertoire. This distinctiveness is underscored by the presence of novel folds with a complex evolutionary history. These findings call for extensive work, focusing both on the biology of this understudied group of organisms and on the functional characterization of their toxins, to unravel the adaptive significance of Octocorallia arsenal.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

Data accessibility. The Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GKVI00000000. The version described in this paper is the first version, GKVI01000000. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [112] partner repository with the dataset identifier PXD061313 and 10.6019/PXD061313. Other data underlying this article are available in the electronic supplementary material [113].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. M.V.M.: conceptualization, data curation, funding acquisition, investigation, methodology, project administration, supervision, validation, visualization, writing—original draft, writing—review and editing; S.L.: conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing—original draft, writing—review and editing; M.G.: conceptualization, data curation, investigation, methodology, resources, validation, visualization, writing—original draft, writing—review and editing; S.G.: data curation, investigation; D.A.: writing—review and editing; M.O.: conceptualization, resources, writing—review and editing; G.F.: resources, writing—review and editing; K.E.K.: methodology, resources; C.B.: methodology, resources; S.D.: conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 748902. The MGX acknowledges financial support from France Génomique National infrastructure, funded as part of 'Investissement d'Avenir' program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09).

Acknowledgements. Mass spectrometry experiments were carried out using the facilities of the Montpellier Proteomics Platform (PPM, BioCampus Montpellier). Dr Dany Domínguez-Pérez is gratefully acknowledged for his help in data handling and curation. We thank Dr Simone Pietro Canese and Mr Cesare Fattori for kindly providing pictures of *Eunicella singularis*.

References

1. Van Iten H, Marques AC, Leme J de M, Pacheco M, Simões MG. 2014 Origin and early diversification of the phylum Cnidaria Verrill: major developments in the analysis of the taxon's Proterozoic–Cambrian history. *Palaeontology* **57**, 677–690. (doi:10.1111/pala.12116)
2. McFadden CS, Quattrini AM, Brugler MR, Cowman PF, Dueñas LF, Kitahara MV, Paz-García DA, Reimer JD, Rodríguez E. 2021 Phylogenomics, origin, and diversification of anthozoans (Phylum Cnidaria). *Syst. Biol.* **70**, 635–647. (doi:10.1093/sysbio/syaa103)
3. David CN, Ozbek S, Adamczyk P, Meier S, Pauly B, Chapman J, Hwang JS, Gojobori T, Holstein TW. 2008 Evolution of complex structures: minicollagens shape the cnidarian nematocyst. *Trends Genet.* **24**, 431–438. (doi:10.1016/j.tig.2008.07.001)
4. Fautin DG. 2009 Structural diversity, systematics, and evolution of cnidae. *Toxicon* **54**, 1054–1064. (doi:10.1016/j.toxicon.2009.02.024)
5. Kass-Simon G, Scappaticci AA. 2002 The behavioral and developmental physiology of nematocysts. *Can. J. Zool.* **80**, 1772–1794. (doi:10.1139/z02-135)
6. Moran Y, Genikhovich G, Gordon D, Wienkoop S, Zenkert C, Özbek S, Technau U, Gurevitz M. 2012 Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proc. R. Soc. B* **279**, 1351–1358. (doi:10.1098/rspb.2011.1731)
7. Marques AC, Collins AG. 2004 Cladistic analysis of Medusozoa and cnidarian evolution. *Invertebr. Biol.* **123**, 23–42. (doi:10.1111/j.1744-7410.2004.tb00139.x)

8. Kayal E, Bentlage B, Sabrina Pankey M, Ohdera AH, Medina M, Plachetzki DC, Collins AG, Ryan JF. 2018 Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol. Biol.* **18**, 68. (doi:10.1186/s12862-018-1142-0)
9. Norton RS. 1991 Structure and structure-function relationships of sea anemone proteins that interact with the sodium channel. *Toxicon* **29**, 1051–1084. (doi:10.1016/0041-0101(91)90205-6)
10. Norton RS. 2009 Structures of sea anemone toxins. *Toxicon* **54**, 1075–1088. (doi:10.1016/j.toxicon.2009.02.035)
11. Honma T, Shiomi K. 2006 Peptide toxins in sea anemones: structural and functional aspects. *Mar. Biotechnol.* **8**, 1–10. (doi:10.1007/s10126-005-5093-2)
12. Shiomi K. 2009 Novel peptide toxins recently isolated from sea anemones. *Toxicon* **54**, 1112–1118. (doi:10.1016/j.toxicon.2009.02.031)
13. Jouiaei M, Yanagihara AA, Madio B, Nevalainen TJ, Alewood PF, Fry BG. 2015 Ancient venom systems: a review on cnidaria toxins. *Toxins* **7**, 2251–2271. (doi:10.3390/toxins7062251)
14. Madio B, King GF, Undheim EAB. 2019 Sea anemone toxins: a structural overview. *Mar. Drugs* **17**, 325. (doi:10.3390/md17060325)
15. Menezes C, Thakur NL. 2022 Sea anemone venom: ecological interactions and bioactive potential. *Toxicon* **208**, 31–46. (doi:10.1016/j.toxicon.2022.01.004)
16. J. Prentis P, Pavasovic A, S. Norton R. 2018 Sea anemones: quiet achievers in the field of peptide toxins. *Toxins* **10**, 36. (doi:10.3390/toxins10010036)
17. Rachamim T, Morgenstern D, Aharonovich D, Brekhman V, Lotan T, Sher D. 2015 The dynamically evolving nematocyst content of an Anthozoan, a Scyphozoan, and a Hydrozoan. *Mol. Biol. Evol*. **32**, 740–753. (doi:10.1093/molbev/msu335)
18. Pennington MW, Czerwinski A, Norton RS. 2018 Peptide therapeutics from venom: current status and potential. *Bioorg. Med. Chem.* **26**, 2738–2758. (doi:10.1016/j.bmc.2017.09.029)
19. Schmidt CA, Daly NL, Wilson DT. 2019 Coral venom toxins. *Front. Ecol. Evol*. **7**. (doi:10.3389/fevo.2019.00320)
20. Klompen AML, Macrander J, Reitzel AM, Stampar SN. 2020 Transcriptomic analysis of four cerianthid (Cnidaria, Ceriantharia) venoms. *Mar. Drugs* **18**, 413. (doi:10.3390/md18080413)
21. Yoffe C, Lotan T, Benayhau Y. 2012 A modified view on octocorals: *Heteroxenia fuscescens* nematocysts are diverse, featuring both an ancestral and a novel type. *PLoS One* **7**, e31902. (doi:10.1371/journal.pone.0031902)
22. Raimundo I, Silva S, Costa R, Keller-Costa T. 2018 Bioactive secondary metabolites from octocoral-associated microbes—new chances for blue growth. *Mar. Drugs* **16**, 485. (doi:10.3390/md16120485)
23. Ballestreros E. 2006 Mediterranean coralligenous assemblages: a synthesis of present knowledge. In *Oceanography and marine biology: an annual review* (eds RN Gibson,R. J. A. Atkinson, JDM Gordon), pp. 123–195, vol. 44. Boca Raton, Florida, US: CRC Taylor & Francis.
24. Linares C, Coma R, Garrabou J, Díaz D, Zabala M. 2008 Size distribution, density and disturbance in two *Mediterranean gorgonians*: *Paramuricea clavata* and *Eunicella singularis*. *J. Appl. Ecol.* **45**, 688–699. (doi:10.1111/j.1365-2664.2007.01419.x)
25. Gori A, Rossi S, Berganzo E, Pretus JL, Dale MRT, Gili JM. 2011 Spatial distribution patterns of the gorgonians *Eunicella singularis*, *Paramuricea clavata*, and *Leptogorgia sarmentosa* (Cape of Creus, Northwestern Mediterranean Sea). *Mar. Biol.* **158**, 143–158. (doi:10.1007/s00227-010-1548-8)
26. Coma R, Llorente-Llurba E, Serrano E, Gili JM, Ribes M. 2015 Natural heterotrophic feeding by a temperate octocoral with symbiotic zooxanthellae: a contribution to understanding the mechanisms of die-off events. *Coral Reefs* **34**, 549–560. (doi:10.1007/s00338-015-1281-3)
27. Renzi JJ, Shaver EC, Burkepile DE, Silliman BR. 2022 The role of predators in coral disease dynamics. *Coral Reefs* **41**, 405–422. (doi:10.1007/s00338-022-02219-w)
28. Radwan FFY, Aboul-Dahab HM, Burnett JW. 2002 Some toxicological characteristics of three venomous soft corals from the Red Sea. *Comp. Biochem. Physiol. C* **132**, 25–35. (doi:10.1016/S1532-0456(02)00045-5)
29. Jouiaei M *et al*. 2015 Firing the sting: chemically induced discharge of cnidae reveals novel proteins and peptides from box jellyfish (*Chironex fleckeri*) venom. *Toxins* **7**, 936–950. (doi:10.3390/toxins7030936)
30. Grabherr MG *et al*. 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652. (doi:10.1038/nbt.1883)
31. Wagner GP, Kin K, Lynch VJ. 2012 Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285. (doi:10.1007/s12064-012-0162-3)
32. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G. 2003 PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342. (doi:10.1002/rcm.1196)
33. Zhang J *et al*. 2012 PEAKS DB: *de novo* sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell Proteomics* **11**, M111. (doi:10.1074/mcp.M111.010587)
34. Han Y, Ma B, Zhang K. 2005 SPIDER: software for protein identification from sequence tags with *de novo* sequencing error. *J. Bioinform. Comput. Biol*. **03**, 697–716. (doi:10.1142/s0219720005001247)
35. Degueldre M *et al*. 2017 Diversity in sequences, post-translational modifications and expected pharmacological activities of toxins from four *Conus* species revealed by the combination of cutting-edge proteomics, transcriptomics and bioinformatics. *Toxicon* **130**, 116–125. (doi:10.1016/j.toxicon.2017.02.014)
36. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010 CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682. (doi:10.1093/bioinformatics/btq003)
37. Oliveira JS, Fuentes-Silva D, King GF. 2012 Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. *Toxicon* **60**, 539–550. (doi:10.1016/j.toxicon.2012.05.020)
38. Jumper J *et al*. 2021 Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. (doi:10.1038/s41586-021-03819-2)
39. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022 ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682. (doi:10.1038/s41592-022-01488-1)
40. Tunyasuvunakool K *et al*. 2021 Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596. (doi:10.1038/s41586-021-03828-1)
41. Holm L. 2022 Dali server: structural unification of protein families. *Nucleic Acids Res*. **50**, W210–W215. (doi:10.1093/nar/gkac387)
42. Ayoub R, Lee Y. 2019 Rupee: a fast and accurate purely geometric protein structure search. *PLoS One* **14**, e0213712. (doi:10.1371/journal.pone.0213712)
43. Webb B, Sali A. 2016 Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform* **54**, 5. (doi:10.1002/cpbi.3)
44. Jurrus E *et al*. 2018 Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128. (doi:10.1002/pro.3280)
45. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023 UCSF ChimeraX: tools for structure building and analysis. *Protein Sci*. **32**, e4792. (doi:10.1002/pro.4792)
46. McFadden CS, Van Ofwegen LP, Quattrini AM. 2022 Revisionary systematics of Octocorallia (Cnidaria: Anthozoa) guided by phylogenomics. *Bull. Soc. Syst. Biol* **1**, 80–86. (doi:10.18061/bssb.v1i3.8735)

47. Gertz EM, Yu YK, Agarwala R, Schäffer AA, Altschul SF. 2006 Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol.* **4**, 41. (doi:10.1186/1741-7007-4-41)
48. Reese MG, Eeckman FH, Kulp D, Haussler D. 1997 Improved splice site detection in Genie. *J. Comput. Biol*. **4**, 311–323. (doi:10.1089/cmb.1997.4.311)
49. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020 IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534. (doi:10.1093/molbev/msaa015)
50. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. (doi:10.1038/nmeth.4285)
51. Sachkova MY, Singer SA, Macrander J, Reitzel AM, Peigneur S, Tytgat J, Moran Y. 2019 The birth and death of toxins with distinct functions: a case study in the sea *Anemone Nematostella*. *Mol. Biol. Evol.* **36**, 2001–2012. (doi:10.1093/molbev/msz132)
52. Jouiaei M, Sunagar K, Federman Gross A, Scheib H, Alewood PF, Moran Y, Fry BG. 2015 Evolution of an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. *Mol. Biol. Evol*. **32**, 1598–1610. (doi:10.1093/molbev/msv050)
53. Sunagawa S, DeSalvo MK, Voolstra CR, Reyes-Bermudez A, Medina M. 2009 Identification and gene expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals. *PLoS One* **4**, e4865. (doi:10.1371/journal.pone.0004865)
54. Logashina YA *et al*. 2017 New disulfide-stabilized fold provides sea anemone peptide to exhibit both antimicrobial and TRPA1 potentiating properties. *Toxins* **9**, 154. (doi:10.3390/toxins9050154)
55. Kozlov SA, Grishin EV. 2007 The universal algorithm of maturation for secretory and excretory protein precursors. *Toxicon* **49**, 721–726. (doi:10.1016/j.toxicon.2006.11.007)
56. Barroso RA, Ramos L, Moreno H, Antunes A. 2024 Evolutionary analysis of cnidaria small cysteine-rich proteins (SCRiPs), an enigmatic neurotoxin family from stony corals and sea *Anemones* (Anthozoa: Hexacorallia). *Toxins* **16**, 75. (doi:10.3390/toxins16020075)
57. Pennington M, Czerwinski A, Norton R. 2018 Peptide therapeutics from venom: current status and potential. *Bioorganic Med. Chem*. **26**, 2738–2758.
58. Columbus-Shenkar YY, Sachkova MY, Macrander J, Fridrich A, Modepalli V, Reitzel AM, Sunagar K, Moran Y. 2018 Dynamics of venom composition across a complex life cycle. *eLife* **7**, e35014. (doi:10.7554/eLife.35014)
59. Sintsova O *et al*. 2021 Sea anemone kunitz-type peptides demonstrate neuroprotective activity in the 6-hydroxydopamine induced neurotoxicity model. *Biomedicines* **9**, 283. (doi:10.3390/biomedicines9030283)
60. Peigneur S, Billen B, Derua R, Waelkens E, Debaveye S, Béress L, Tytgat J. 2011 A bifunctional sea anemone peptide with Kunitz type protease and potassium channel inhibiting properties. *Biochem. Pharmacol.* **82**, 81–90. (doi:10.1016/j.bcp.2011.03.023)
61. Nikolaev MV, Dorofeeva NA, Komarova MS, Korolkova YV, Andreev YA, Mosharova IV, Grishin EV, Tikhonov DB, Kozlov SA. 2017 TRPV1 activation power can switch an action mode for its polypeptide ligands. *PLoS One* **12**, e0177077. (doi:10.1371/journal.pone.0177077)
62. García-Fernández R, Peigneur S, Pons T, Alvarez C, González L, Chávez MA, Tytgat J. 2016 The kunitz-type protein ShPI-1 inhibits serine proteases and voltage-gated potassium channels. *Toxins* **8**, 110. (doi:10.3390/toxins8040110)
63. Andreev YA *et al*. 2013 Polypeptide modulators of TRPV1 produce analgesia without hyperthermia. *Mar. Drugs* **11**, 5100–5115. (doi:10.3390/md11125100)
64. Wang X *et al*. 2021 Toxic peptide from *Palythoa caribaeorum* acting on the TRPV1 channel prevents pentylenetetrazol-induced epilepsy in zebrafish larvae. *Front. Pharmacol.* **12**, 763089. (doi:10.3389/fphar.2021.763089)
65. Hwang SM, Jo YY, Cohen CF, Kim YH, Berta T, Park CK. 2022 Venom peptide toxins targeting the outer pore region of transient receptor potential Vanilloid 1 in pain: implications for analgesic drug development. *Int. J. Mol. Sci.* **23**, 5772. (doi:10.3390/ijms23105772)
66. Kvetkina A *et al*. 2022 Kunitz-Type peptides from sea anemones protect neuronal cells against Parkinson's disease Inductors via Inhibition of ROS production and ATP-induced P2X7 receptor activation. *Int. J. Mol. Sci.* **23**, 5115. (doi:10.3390/ijms23095115)
67. Dai SX, Zhang AD, Huang JF. 2012 Evolution, expansion and expression of the Kunitz/BPTI gene family associated with long-term blood feeding in Ixodes Scapularis. *BMC Evol. Biol.* **12**, 4. (doi:10.1186/1471-2148-12-4)
68. Modica MV, Lombardo F, Franchini P, Oliverio M. 2015 The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). *BMC Genom*. **16**, 441. (doi:10.1186/s12864-015-1648-4)
69. Mishra M. 2020 Evolutionary aspects of the structural convergence and functional diversification of kunitz-domain inhibitors. *J. Mol. Evol.* **88**, 537–548. (doi:10.1007/s00239-020-09959-9)
70. Chen Z *et al*. 2013 Genomic and structural characterization of kunitz-type peptide LmKTT-1a highlights diversity and evolution of scorpion potassium channel toxins. *PLoS One* **8**, e60201. (doi:10.1371/journal.pone.0060201)
71. Fry BG *et al*. 2009 The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **10**, 483–511. (doi:10.1146/annurev.genom.9.081307.164356)
72. Boisbouvier J, Albrand JP, Blackledge M, Jaquinod M, Schweitz H, Lazdunski M, Marion D. 1998 A structural homologue of colipase in black mamba venom revealed by NMR floating disulphide bridge analysis. *J. Mol. Biol.* **283**, 205–219. (doi:10.1006/jmbi.1998.2057)
73. Kaser A, Winklmayr M, Lepperdinger G, Kreil G. 2003 The AVIT protein family. *EMBO Rep*. **4**, 469–473. (doi:10.1038/sj.embor.embor830)
74. Negri L, Lattanzi R, Giannini E, Melchiorri P. 2007 Bv8/Prokineticin proteins and their receptors. *Life Sci* **81**, 1103–1116. (doi:10.1016/j.lfs.2007.08.011)
75. Wen S *et al*. 2005 Discovery of an MIT-like atracotoxin family: spider venom peptides that share sequence homology but not pharmacological properties with AVIT family proteins. *Peptides* **26**, 2412–2426. (doi:10.1016/j.peptides.2005.05.012)
76. Hecht O, Van Nuland NA, Schleinkofer K, Dingley AJ, Bruhn H, Leippe M, Grötzinger J. 2004 Solution structure of the pore-forming protein of *Entamoeba histolytica*. *J. Biol. Chem.* **279**, 17834–17841. (doi:10.1074/jbc.M312978200)
77. Tosi T, Nickerson NN, Mollica L, Jensen MR, Blackledge M, Baron B, England P, Pugsley AP, Dessen A. 2011 Pilotin-secretin recognition in the type II secretion system of *Klebsiella oxytoca*. *Mol. Microbiol.* **82**, 1422–1432. (doi:10.1111/j.1365-2958.2011.07896.x)
78. Korotkov KV, Hol WGJ. 2013 Crystal structure of the pilotin from the enterohemorrhagic *Escherichia coli* type II secretion system. *J. Struct. Biol*. **182**, 186–191. (doi:10.1016/j.jsb.2013.02.013)
79. Krishnarjuna B, Sunanda P, Villegas-Moreno J, Csoti A, A V Morales R, Wai DCC, Panyi G, Prentis P, Norton RS. 2021 A disulfide-stabilised helical hairpin fold in acrorhagin I: an emerging structural motif in peptide toxins. *J. Struct. Biol.* **213**, 107692. (doi:10.1016/j.jsb.2020.107692)
80. Honma T, Minagawa S, Nagai H, Ishida M, Nagashima Y, Shiomi K. 2005 Novel peptide toxins from acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon* **46**, 768–774. (doi:10.1016/j.toxicon.2005.08.003)

81. Sachkova MY, Landau M, Surm JM, Macrander J, Singer SA, Reitzel AM, Moran Y. 2020 Toxin-like neuropeptides in the sea anemone *Nematostella* unravel recruitment from the nervous system to venom. *Proc. Natl Acad. Sci.* **117**, 27481–27492. (doi:10.1073/pnas.2011120117)
82. Anderluh G, Lakey JH. 2008 Disparate proteins use similar architectures to damage membranes. *Trends Biochem. Sci*. **33**, 482–490. (doi:10.1016/j.tibs.2008.07.004)
83. Morando MA, Venturella F, Sollazzo M, Monaca E, Sabbatella R, Vetri V, Passantino R, Pastore A, Alfano C. 2022 Solution structure of recombinant Pvfp-5β reveals insights into mussel adhesion. *Commun. Biol.* **5**, 739. (doi:10.1038/s42003-022-03699-w)
84. Yeung SYM, Thompson D, Wang Z, Fedida D, Robertson B. 2005 Modulation of Kv3 subfamily potassium currents by the sea anemone toxin BDS: significance for CNS and biophysical studies. *J. Neurosci.* **25**, 8735–8745. (doi:10.1523/JNEUROSCI.2119-05.2005)
85. Sher D, Fishman Y, Melamed-Book N, Zhang M, Zlotkin E. 2008 Osmotically driven prey disintegration in the gastrovascular cavity of the green hydra by a pore-forming protein. *FASEB J.* **22**, 207–214. (doi:10.1096/fj.07-9133com)
86. Macrander J, Daly M. 2016 Evolution of the cytolytic pore-forming proteins (actinoporins) in sea anemones. *Toxins* **8**, 1–16. (doi:10.3390/toxins8120368)
87. Darmoise A, Maschmeyer P, Winau F. 2010 The immunological functions of saposins. *Adv. Immunol.* **105**, 25–62. (doi:10.1016/S0065-2776(10)05002-9)
88. Ahn VE, Leyko P, Alattia JR, Chen L, Privé GG. 2006 Crystal structures of saposins A and C. *Protein Sci.* **15**, 1849–1857. (doi:10.1110/ps.062256606)
89. Munford RS, Sheppard PO, O'Hara PJ. 1995 Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *J. Lipid Res.* **36**, 1653–1663. (doi:10.1016/s0022-2275(20)41485-3)
90. Bruhn H. 2005 A short guided tour through functional and structural features of saposin-like proteins. *Biochem. J*. **389**, 249–257. (doi:10.1042/bj20050051)
91. Herbst R, Marciano-Cabral F, Leippe M. 2004 Antimicrobial and pore-forming peptides of free-living and potentially highly Pathogenic Naegleria fowleri Are released from the same precursor molecule. *J. Biol. Chem*. **279**, 25955–25958. (doi:10.1074/jbc.m401965200)
92. Hill CH, Read RJ, Deane JE. 2015 Structure of human saposin A at lysosomal pH. *Acta Crystallogr. Sect F* **71**, 895–900. (doi:10.1107/S2053230X15008584)
93. Hawkins CA, de Alba E, Tjandra N. 2005 Solution structure of human saposin C in a detergent environment. *J. Mol. Biol.* **346**, 1381–1392. (doi:10.1016/j.jmb.2004.12.045)
94. Alattia JR, Shaw JE, Yip CM, Privé GG. 2006 Direct visualization of saposin remodelling of lipid bilayers. *J. Mol. Biol*. **362**, 943–953. (doi:10.1016/j.jmb.2006.08.009)
95. Popovic K, Holyoake J, Pomès R, Privé GG. 2012 Structure of saposin a lipoprotein discs. *Proc. Natl Acad. Sci.* **109**, 2908–2912. (doi:10.1073/pnas.1115743109)
96. Leippe M, Bruhn H, Hecht O, Grötzinger J. 2005 Ancient weapons: the three-dimensional structure of amoebapore A. *Trends Parasitol.* **21**, 5–7. (doi:10.1016/j.pt.2004.10.009)
97. Mysliwy J, Dingley AJ, Stanisak M, Jung S, Lorenzen I, Roeder T, Leippe M, Grötzinger J. 2010 Caenopore-5: the three-dimensional structure of an antimicrobial protein from *Caenorhabditis elegans*. *Dev. Comp. Immunol.* **34**, 323–330. (doi:10.1016/j.dci.2009.11.003)
98. Liepinsh E, Andersson M, Ruysschaert JM, Otting G. 1997 Saposin fold revealed by the NMR structure of NK-lysin. *Nat. Struct. Biol.* **4**, 793–795. (doi:10.1038/nsb1097-793)
99. Anderson DH, Sawaya MR, Cascio D, Ernst W, Modlin R, Krensky A, Eisenberg D. 2003 Granulysin crystal structure and a structure-derived lytic mechanism. *J. Mol. Biol.* **325**, 355–365. (doi:10.1016/s0022-2836(02)01234-2)
100. Chen L, Dürr KL, Gouaux E. 2014 X-ray structures of AMPA receptor-cone snail toxin complexes illuminate activation mechanism. *Science* **345**, 1021–1026. (doi:10.1126/science.1258409)
101. Hackney CM *et al*. 2023 A previously unrecognized superfamily of macro-conotoxins includes an inhibitor of the sensory neuron calcium channel Cav2.3. *PLoS Biol*. **21**, e3002217. (doi:10.1371/journal.pbio.3002217)
102. Olmeda B, García-Álvarez B, Pérez-Gil J. 2013 Structure-function correlations of pulmonary surfactant protein SP-B and the saposin-like family of proteins. *EBJ* **42**, 209–222. (doi:10.1007/s00249-012-0858-9)
103. Olmeda B, García-Álvarez B, Gómez MJ, Martínez-Calle M, Cruz A, Pérez-Gil J. 2015 A model for the structure and mechanism of action of pulmonary surfactant protein B. *FASEB J.* **29**, 4236–4247. (doi:10.1096/fj.15-273458)
104. Garrido-Arandia M, Cuevas-Zuviría B, Díaz-Perales A, Pacios LF. 2018 A comparative study of human saposins. *Molecules* **23**, 422. (doi:10.3390/molecules23020422)
105. Yin W *et al*. 2016 Evolutionary trajectories of snake genes and genomes revealed by comparative analyses of five-pacer viper. *Nat. Commun*. **7**, 13107. (doi:10.1038/ncomms13107)
106. Maxwell M, Undheim EAB, Mobli M. 2018 Secreted cysteine-rich repeat proteins 'SCREPs': a novel multi-domain architecture. *Front. Pharmacol*. **9**, 1333. (doi:10.3389/fphar.2018.01333)
107. Gerdol M, Cervelli M, Mariottini P, Oliverio M, Dutertre S, Modica MV. 2019 A recurrent motif: diversity and evolution of ShKT domain containing proteins in the vampire snail *Cumia reticulata*. *Toxins* **11**, 106. (doi:10.3390/toxins11020106)
108. Berking S, Herrmann K. 2006 Formation and discharge of nematocysts is controlled by a proton gradient across the cyst membrane. *Helgol. Mar. Res.* **60**, 180–188. (doi:10.1007/s10152-005-0019-y)
109. Fabricius K, Dommisse M. 2000 Depletion of suspended particulated matter over coastal reef communities dominated by zooxanthellate soft corals. *MEPS* **196**, 157–167. (doi:10.3354/meps196157)
110. Sebens KP, Miles JS. 1988 Sweeper tentacles in a gorgonian octocoral: morphological modifications for interference competition. *Biol. Bull.* **175**, 378–387. (doi:10.2307/1541729)
111. Dutertre S, Jin A hua, Kaas Q, Jones A, Alewood PF, Lewis RJ. 2013 Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom. *Mol. Cell. Proteom* **12**, 312–329. (doi:10.1074/mcp.M112.021469)
112. Perez-Riverol Y *et al*. 2025 The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res.* **53**, D543–D553. (doi:10.1093/nar/gkae1011)
113. Modica MV, Leone S, Gerdol M, Greco S, Aurelle D, Oliverio M *et al*. 2025 Supplementary material from: The proteotranscriptomic characterization of venom in the white seafan *Eunicella singularis* elucidates the evolution of Octocorallia arsenal. Figshare (doi:10.6084/m9.figshare.c.7702728)