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

Authors

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

* equal contribution

1. Department of Biology and Evolution of Marine Organisms (BEOM), 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; Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Paris, France

4. Department of Biology and Biotechnology "Charles Darwin", Sapienza University of Roma, Italy

5. BCM, Univ. Montpellier, CNRS, INSERM, Montpellier, France

6. Platform MGX - Montpellier GenomiX, IGF, Univ. Montpellier, France.

7. Institut des Biomolécules Max Mousseron, UMR 5247, Université Montpellier – CNRS, Montpellier, France

*__Corresponding author email__ mariavittoria.modica@szn.it.

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

Abstract

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 the 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.

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 6,000 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

66 Mediterranean Sea, gorgonians constitute up to 40% of the biomass in rocky sublittoral habitats [23].
67 The white gorgonian *Eunicella singularis* (Esper, 1791) is widespread and abundant in the north-western
68 Mediterranean [24,25]. Despite hosting *Symbiodinium* zooxanthellae, *E. singularis* displays low levels of
69 primary productivity and relies on heterotrophic nutrition to meet its metabolic needs, feeding mainly on
70 zooplankton integrated with particulated organic matter (POM) [26]. Predation on zooplankton,
71 intraspecific competition and defence from predators (mostly snails, crabs and fish) [27] are physiological
72 activities likely supported, in *E. singularis* and allies, by the production of a specialized venom. Indeed,
73 preliminary evidence have demonstrated that haemolytic peptides are present in different species of
74 octocorals [28], and the diversity and abundance of nematocysts could be higher than previously held [21].
75 However, toxin diversity has not been addressed in Octocorallia so far, hampering our complete
76 understanding of a crucial eco-evolutionary question concerning Cnidaria biology and evolution: to what
77 extent are main toxin families conserved across Anthozoa? To start answering this question, we provided a
78 first overview of the venom produced by an octocoral, applying to *E. singularis* an integrated proteo-
79 transcriptomic approach, which employed whole colony transcriptomic data to match proteomic data
80 obtained from chemical-induced nematocysts discharge or from the whole body of polyps. Our analyses
81 unveiled several previously unreported putative toxin families, providing the first comprehensive
82 description of the venom of an octocoral species. Leveraging these findings, we aimed to assess the degree
83 of conservation of specific toxin folds of Octocorallia in comparison to other Anthozoa, and to elucidate how
84 venom toxin families diversified throughout the evolutionary radiation of octocorals. Employing a
85 comparative approach on available Octocorallia genomes, we reconstructed the evolutionary patterns of
86 the main toxin families, highlighting distinctive evolutionary innovations specific to Octocorallia. Eventually,
87 by complementing our analyses with untemplated, deep-learning assisted structural modelling of the most
88 relevant toxins, we attempted to infer their function and adaptive value.
89

90 **Methods**

91 **(a) Specimens collection, venom sample preparation and RNA and proteins extraction**

92 Colonies of *Eunicella singularis* were collected by scuba diving at a depth of 25-30 m around Giannutri Island,

93 in the Tuscany Archipelago (Italy) and kept alive in seawater until further processing. Polyps were dissected
94 from the colony under a stereomicroscope and pooled in three aliquots of about 20 each (about 15 mg of
95 tissue) that were preserved in Trizol at -80 °C until further processing. Total RNA was extracted using Trizol
96 (Thermo Fisher Scientific), following the manufacturer's instructions, obtaining 5-10,000 ng with RQN values
97 between 6.7 and 8.8. Proteins from 10 additional polyps were extracted through homogenization in milliQ
98 water, followed by centrifugation and analysis of the supernatant to produce the whole body proteome.
99 Additionally, 10 live colony fragments about 25 mm long were immersed for 30 seconds into 1 mL of EtOH
100 in a 1.5 mL tube to induce nematocyst discharge [29]. Samples were freeze-dried and resuspended in 100
101 µL milliQ water for proteomic analysis.
102
103 **(b) RNASeq**
104 To account for variations in venom expression among samples, three cDNA libraries for RNASeq were
105 prepared using an Illumina TruSeq Stranded mRNA Sample Preparation kit according to manufacturer's
106 instructions and validated qualitatively and quantitatively on a Fragment Analyser, as well as by qPCR on a
107 ROCHE LightCycler 480. The three libraries were then processed for paired-end sequencing-by-synthesis in
108 an Illumina HiSeq 2500 at the MGX facility (Montpellier, France), and quality controlled using FastQC. After
109 trimming, we obtained a number of reads longer than 30 bp and quality above Q20 comprised between 21
110 and 30 million.
111
112 **(c) Bioinformatic analyses of transcriptomic data**
113 Paired-end raw trimmed sequencing data obtained from the three libraries were pooled and *de novo*
114 assembled using Trinity v.2.5.1 [30] with default settings, allowing a minimum contig length of 200
115 nucleotides. The assembled transcriptome was virtually translated using TransDecoder v.5.01
116 (https://github.com/TransDecoder/), setting the minimum allowed ORF length to 100 codons.
117 Expression levels of toxin transcripts were calculated based on unique gene counts as Transcripts Per Million
118 (TPM), as this normalised unit allows an efficient and reliable comparison of gene expression levels both
119 within and between samples [31]. Briefly, trimmed reads from each library were separately mapped against

120 the transcriptome assembly, previously filtered to remove sequence redundancy by considering only the
121 longest transcript for each gene model.
122
123 **(d) Protein digestion**
124 Protein extracts were denatured, reduced and alkylated. Briefly, the volume of each sample was adjusted
125 to 100 µL with 100 mM triethylammonium bicarbonate (TEABC). 1M DTT (1 µL) was added to each sample,
126 which was then incubated 30 min at 60 °C before adding 10 µL 0.5 M Iodoacetoamide (IAA). Alkylation lasted
127 30 min in the dark. Enzymatic digestion was performed by adding 2 µg trypsin (Gold, Promega, Madison
128 USA) in TEABC 100 mM and incubating overnight at 30 °C. Then, peptides were purified and concentrated
129 using OMIX Tips C18 reverse-phase resin (Agilent Technologies Inc.) according to the manufacturer's
130 specifications. Peptides samples were finally dehydrated in a vacuum centrifuge.
131
132 **(e) Shotgun proteomics**
133 Samples were resuspended in 20 µL of 0.1% formic acid (buffer A) and 1 µL was loaded onto an analytical
134 50 cm reversed-phase column (0.75 mm i.d., Acclaim Pepmap 100® C18, Thermo Fisher Scientific) and
135 separated with an Ultimate 3000 RSLC system (Thermo Fisher Scientific) coupled to a Q Exactive HF (Thermo
136 Fisher Scientific) *via* a nano-electrospray source, using a 120-min gradient of 2 to 40% of buffer B (80% ACN,
137 0.1% formic acid) and a flow rate of 300 nL/min. MS/MS analyses were performed in a data-dependent
138 mode. Full scans (375 – 1,500 m/z) were acquired in the Orbitrap mass analyser with a 60,000 resolution at
139 200 m/z. For the full scans, 3 x 10^6 ions were accumulated within a maximum injection time of 60 ms and
140 detected in the Orbitrap analyser. The twelve most intense ions with charge states ≥ 2 were sequentially
141 isolated to a target value of 1 x 10^5 with a maximum injection time of 45 ms and fragmented by HCD (Higher-
142 energy collisional dissociation) in the collision cell (normalised collision energy of 28%) and detected in the
143 Orbitrap analyser at a resolution of 30,000.
144
145 **(f) Bioinformatic analyses of proteomic data**
146 PEAKS Studio 8.5 (Bioinformatics solutions, Waterloo, ON, Canada), a de novo assisted database software

147 [32-34] was used to analyse MS/MS data from *E. singularis* venom, with parameters set as in [35]. The
148 MS/MS spectra obtained from shotgun proteomics were matched to a custom database resulting from the
149 whole-body transcriptome of the same colony, translated in the six open reading frames.
150 Protein isoforms were considered only when their existence was supported by MS/MS fragmentation data,
151 whereas translated transcripts identified by PEAKS Studio as belonging to the same protein group were
152 clustered with CD-HIT [36] using an identity threshold of 90%. Only the most representative sequence was
153 retained for the subsequent annotation.
154 Protein annotation was performed as detailed in the Suppl. Met. and Fig.1. Toxin nomenclature of mature
155 polypeptides followed Oliveira et al. [37]
156 Structure predictions of mature toxins were performed with the colabfold implementation of AlphaFold2,
157 using MMSeqs2 for multiple sequence alignment generation [38,39]. To increase model confidence for the
158 most challenging targets, custom multiple sequence alignments with orthologue octocorallian sequences
159 were used as inputs and each model was built using 12 recycles. All predictions were performed without the
160 use of templates and energy minimization. The structures that were modelled with reasonable confidence
161 (average pLDDT ≥ 70, which is generally recognised as the threshold for reliable models) [40] were submitted
162 to the DALI [41] and RUPEE [42] servers for structural homology search against the Protein Data Bank.
163 Homology modelling of the saposin-like toxins in the closed conformation was performed with Modeller
164 v9.24 [43]. Electrostatic surfaces were calculated with APBS [44]. UCSF ChimeraX was used for model
165 analysis, visualization and graphics [45].
166
167 **(g) Comparative genomic analyses**
168 Toxins identified in *Eunicella singularis* were used as queries for sequence homology searches against a
169 selection of available anthozoan genomes, representative of the diversity of Octocorallia [46] and major
170 Hexacorallia lineages summarized in Suppl. Tab. 1.
171 Sequence homology searches were carried out with tBLAStn [47], based on an initial E-value significance
172 threshold arbitrarily set to 10. Potential hits were manually curated, by annotating the inferred CDS,
173 detecting putative splicing donor and acceptor sites with the aid of Genie [48]. The presence of *bona fide*

174 orthologues between *E. singularis* and the other species was assessed through combined observations, i.e.
175 (i) the conservation of the main topological feature of the encoded precursor proteins, such as the
176 identification of a signal peptide and the cysteine residues involved in the formation of the disulphide bonds
177 (whenever present); (ii) the shared presence of conserved protein domains, detected with a HMM profile
178 approach, based on the Pfam-A database; (iii) the conservation of gene architecture (e.g. intron number,
179 phase and position), with respect to the orthologous gene from the congeneric species *E. verrucosa*.
180 Whenever needed (i.e., U-GRTX-Esi20 and 21), a fine-scale classification of orthologous sequences was aided
181 by Maximum Likelihood phylogenetic inference analyses, carried out with IQ-Tree [49] with 1000 ultrafast
182 bootstrap replicates based on the best-fitting model of molecular evolution detected by ModelFinder [50].
183

184 **RESULTS AND DISCUSSION**

185 ***(a) E. singularis* produces a complex venom**

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

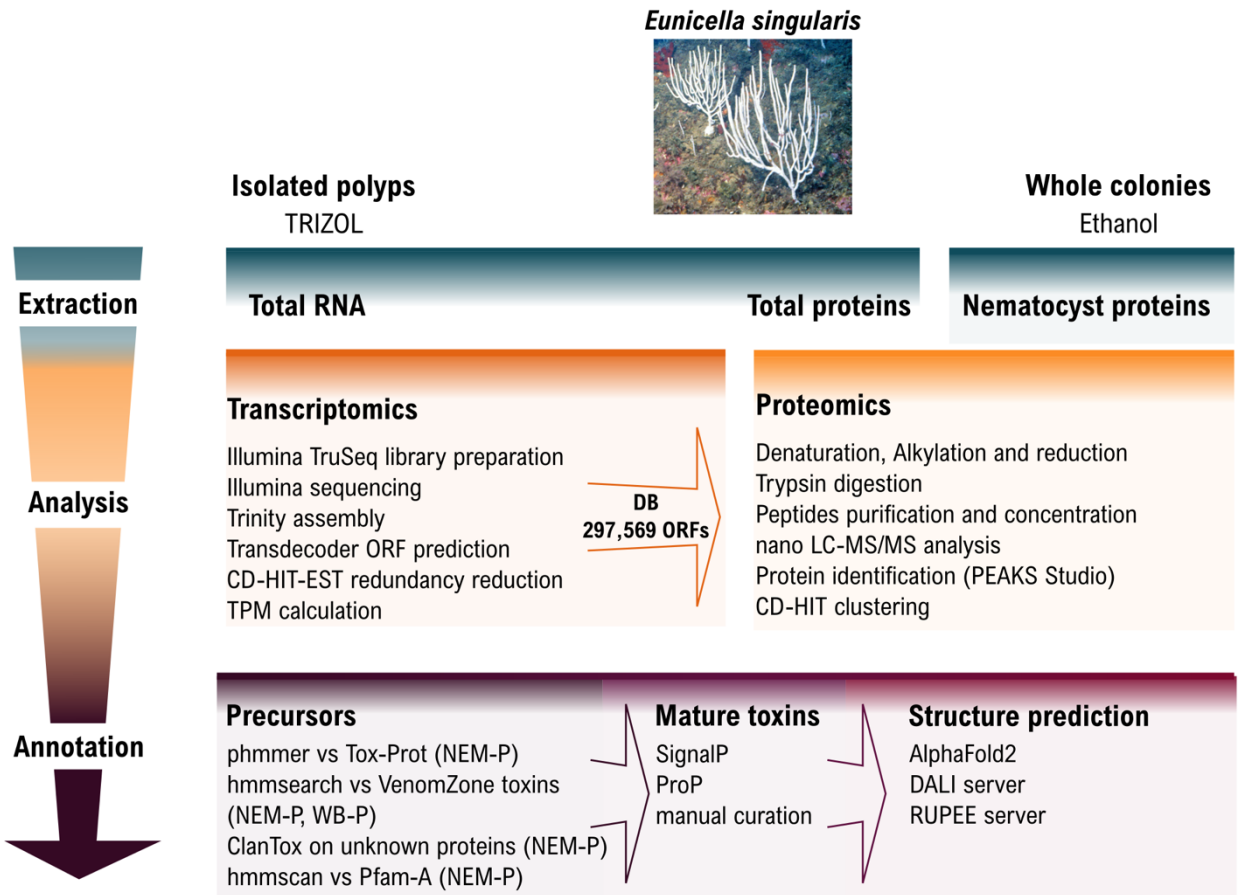
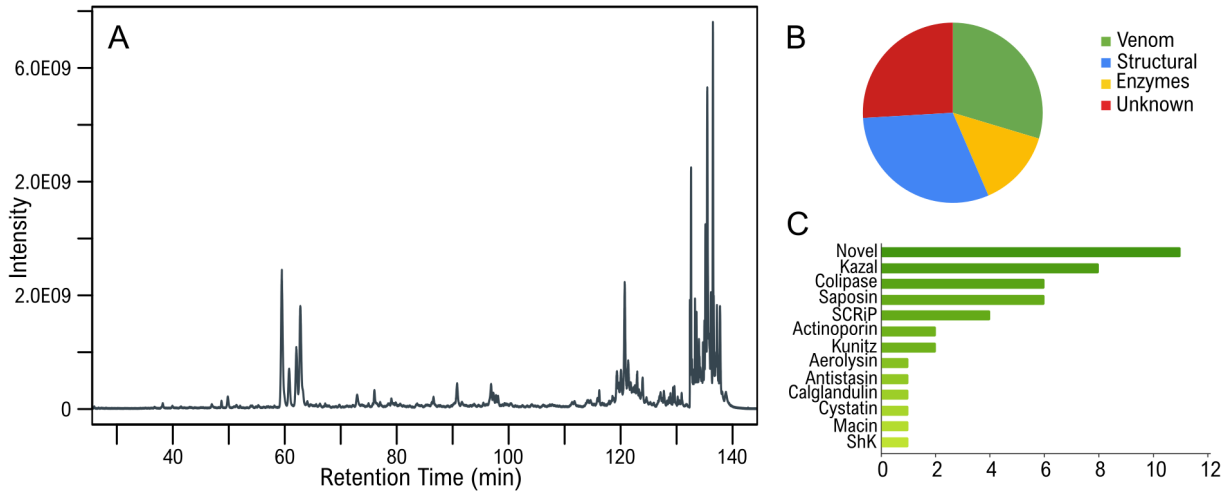


Figure 1. Schematic representation of proteotranscriptomics and bioinformatic annotation pipelines employed in this work.

Only 17 of them could be also detected also in the whole-body proteome (WB-P, Suppl. Tab. 3) 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 Proteins (SCRiPs) 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 the ectodermal glands, as previously demonstrated for sea anemones [51], that were not affected by ethanol exposure. In

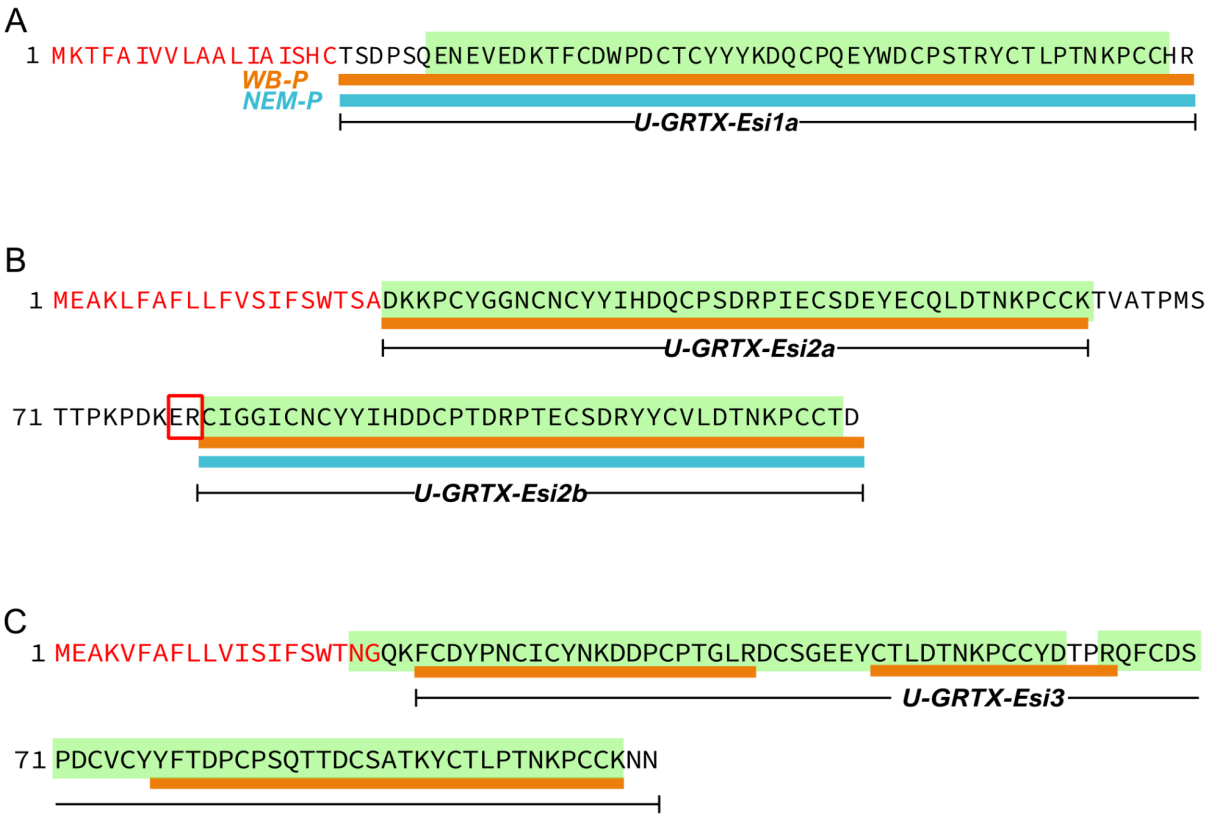
214 alternative, these proteins may lack a venom function despite having toxin-like folds (e.g. they might be
215 involved in neurotransmission or digestion). The NEM-P comprised some transcripts encoding precursors
216 of multiple mature toxins, which were further classified by predicted fold (Fig. 1C and Suppl. Tab. 4). In total,
217 we identified 44 mature toxins, including a calglandulin, a protein characterized by multiple EF-hand motifs
218 with an accessory role in venom production and processing, omitted from the detailed discussion below.



219

220 **Figure 2.** Structural and functional annotation of the venom from *E. singularis.* (A) Total Ion Chromatogram (TIC) of the
221 ethanol extract; (B) Functional annotation of the translated transcripts confirmed at the protein level in the NEM-P; (C)
222 Fold distribution of the mature toxins produced by *E. singularis* based on HMM profile search; the x axis represents the
223 number of proteomics-validated sequences.
224
225 Following [37], the newly characterized toxins were collectively termed "Gorgotoxins" (GRTX-Esi), according
226 to the most conservative placement of *Eunicella* in the family Gorgoniidae, although recent work has
227 proposed a separate family for this genus (Eunicellidae) [46].
228
229 ***(b)* Established toxin folds are present in the venom of *E. singularis***
230 **Small Cystein Rich Proteins-like (octo-SCRiPs)**
231 We retrieved four toxins with homology to hexacorallian SCRiPs, which are known to induce severe
232 neurotoxicity in zebrafish larvae, possess antimicrobial activity and potentiate the transient receptor
233 potential ankyrin 1 (TRPA1) in mouse [52-54]. These comprised U-GRTX-Esi1 (particularly abundant at the
234 transcript level, TPM 1585), 2a-b and 3 (Fig. 3). These proteins exhibited a peptide coverage of the predicted
235 mature sequence exceeding 95%, shared a significant homology and up to 68% pairwise sequence similarity
236 with hexacoral SCRiPs. Despite this similarity, each toxin showcased distinct features.

237 In particular, U-GRTX-Esi2a-b derived from a precursor containing two consecutive highly conserved SCRiP-
238 like domains (pairwise identity 78.9%), separated by a protease cleavage site. Interestingly, while both forms
239 could be detected in the WB-P, only U-GRTX-Esi2b was detectable in the nematocyst discharge (Fig. 3B and
240 Suppl. Fig. 1).
241 U-GRTX-Esi3, solely detectable in the WB-P, featured a more complex architecture, comprising two SCRiP-
242 like domains in the mature polypeptide, confirmed by the proteomic coverage (Fig. 3C). The absence of
243 predicted cleavage sites between these domains would qualify U-GRTX-Esi3 as a Secreted Cysteine-Rich
244 Repeat Protein (SCREP), an emerging typology of venom proteins found in various organisms, often with
245 high affinity and avidity towards their biological target and a multi-modal action [55].
246
247



248
249 **Figure 3.** Tissue-specific maturation patterns of the SCRiP-like toxins- from *E. singularis.* The sequences of the
250 precursors of U-GRTX-Esi1 (A), 2a-b (B) and 3 (C) are reported. Red residues correspond to predicted signal peptide,
251 green boxes define the SCRiP domains, the red rectangle defines a PQM protease recognition site [56], orange and blue
252 lines indicate the peptide coverage of the precursors in the WB-P and NEM-P, respectively.
253
254
255 Although SCRiPs-encoding genes were identified, often in large numbers and associated with tandemly

256 duplicated gene clusters, in the genomes of all the octocoral species investigated in this study, they have a
257 patchy occurrence in Hexacorallia, supporting a complex evolutionary history characterized by lineage-
258 specific expansions and rapid diversification [55]. It is unclear whether SCREPs with two consecutive SCRiP
259 domains were already present in the latest common ancestor of all Anthozoans, or independently acquired
260 in different lineages. Indeed, clear orthologues of U-GRTX-Esi3 could be identified only in *P. clavata,*
261 *Trachytela* sp. and *R. reniformis*: this architecture had been previously identified also in the hexacoral
262 *Orbicella faveolata* [53]. However, while the only SCRiP so far reported in Octocorallia (in *Scleronephthya*
263 *gracillima*) presented a conventional cysteine framework [55], the domains here identified display a shift in
264 the cysteine pattern, which nonetheless did not affect the overall protein fold (Suppl. Fig. 2).
265
266 **ShK-containing toxins**
267 The ShK toxins, first identified in the venom of the sea anemone *Stichodactyla helianthus*, are potent K_V1
268 channels blockers [57]. Our search for ShK domains in the proteomes of *E. singularis* led to the identification
269 of U-GRTX-Esi4 in the NEM-P (TPM 103), with three consecutive ShK domains separated by short linker
270 regions. Only the C-terminal domain has the typical ShKT cysteine pattern, whereas the first two domains
271 present an unusual shift of the C-terminal cysteine. None of the domains of U-GRTX-Esi4 presents the key
272 Lys residue necessary for binding $K_V1.2$ and $K_V1.3$, while the subsequent Tyr residue, also important for
273 binding K_V1, is extremely conserved (Suppl. Fig. S3). This architecture resembles the precursors of the NEP3
274 protein family from *N. vectensis* [58]. However, only the first ShK domain of NEP3 is retained in the mature
275 NEP3 toxin, while U-GRTX-Esi4 lacks putative cleavage sites and we found proteomic evidence for the first
276 and the last ShK domain, suggesting the occurrence of a multidomain protein in *E. singularis*.
277
278 **Protease inhibitors**
279 Toxins containing domains associated with protease inhibition are widespread across the tree of life. Among
280 them, Kunitz-like peptides are frequently found in venoms [59] where they can act as inhibitors of several
281 endogenous and exogenous proteases [60] but also show neurotoxic and neuromodulatory activity on a
282 variety of receptors and ion channels [61,62], sometimes with multimodal action [57,59]. Kunitz-like toxins

283 that bind ion channels can be used to paralyze prey and are being extensively studied for their analgesic and
284 neuroprotective properties [63-67]. Our HMM search revealed the occurrence of two Kunitz-like toxins. PI-
285 GRTX-Esi5, detectable in the NEM-P only, presented two Kunitz domains separated only by an Asp-Ile
286 dipeptide, an architecture previously detected in salivary proteins of blood-feeding animals [67,68]. Since
287 orthologous genes sharing the same 3 exons/2 introns organization of PI-GRTX-Esi5 were not detected in
288 Scleralcyonacea, this domain combination most likely represents a recent innovation of the
289 Malacalcyonacea. The second toxin, PI-GRTX-Esi6, with a single Kunitz domain, was detected both in the
290 NEM-P and in the WB-P, and derived from a precursor with a complex architecture, with the Kunitz-like
291 domain followed by three Kazal-type domains, all separated by propeptidase cleavage sites. Also in this case,
292 proteomic evidence indicated a differential maturation pattern of the precursor, with the Kazal-like peptides
293 from this precursor only retrieved in the WB-P (Suppl. Tab. 2). The Kunitz domain of PI-GRTX-Esi5 presented
294 the Arg residue (conventionally "P1") associated with the trypsin inhibition, whereas its replacement with a
295 Met in both domains of PI-GRTX-Esi6 suggests a chymotrypsin specificity [69]. In addition, PI-GRTX-Esi6
296 presented a Leu residue aligning with the crucial position for K_V1 binding of LmKTT, a scorpion Kunitz-like
297 toxin [70]. These similarities may hint to multimodal activity for PI-GRTX-Esi5, while the lack of key residue
298 conservation in PI-GRTX-Esi6 does not allow inferring any putative function (Suppl. Fig. 4)
299 Other peptides present in *E. singularis* nematocysts displaying protease inhibitory domains (Kazal-type,
300 cystatines, antistasins, and macins) were detected but did not present novelty elements; they are briefly
301 described in Suppl. Data.
302
303 **Colipases**
304 Finally, a striking feature within the NEM-P of *E. singularis* was the abundance of toxins with a predicted
305 colipase fold (U-GRTX-Esi14-18). Colipases are knottins containing five disulphide bridges, found in spider,
306 snake and tick venoms [71], but to date not in Cnidaria. The prototypical example, MIT-1 [72] interacts
307 through its N-terminal pentapeptide AVITG with prokineticin receptors, stimulating smooth muscle
308 contractions, and exhibiting hyperalgesic effects [73,74]. *E. singularis* colipases lacked the AVITG
309 pentapeptide, a feature shared with the atracotoxins of the spider *Hadronyche versuta,* whose

310 pharmacological target is unknown to date [75].

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

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

330 Notably, the transcripts for all these toxins had exceptionally high TPM values (1806, 569, 826 and 429,
331 respectively for the U-GRTX-Esi14 to 17/18), compared to other proteins detected in the NEM-P (cfr. Suppl.
332 Tab. 4).

333

334

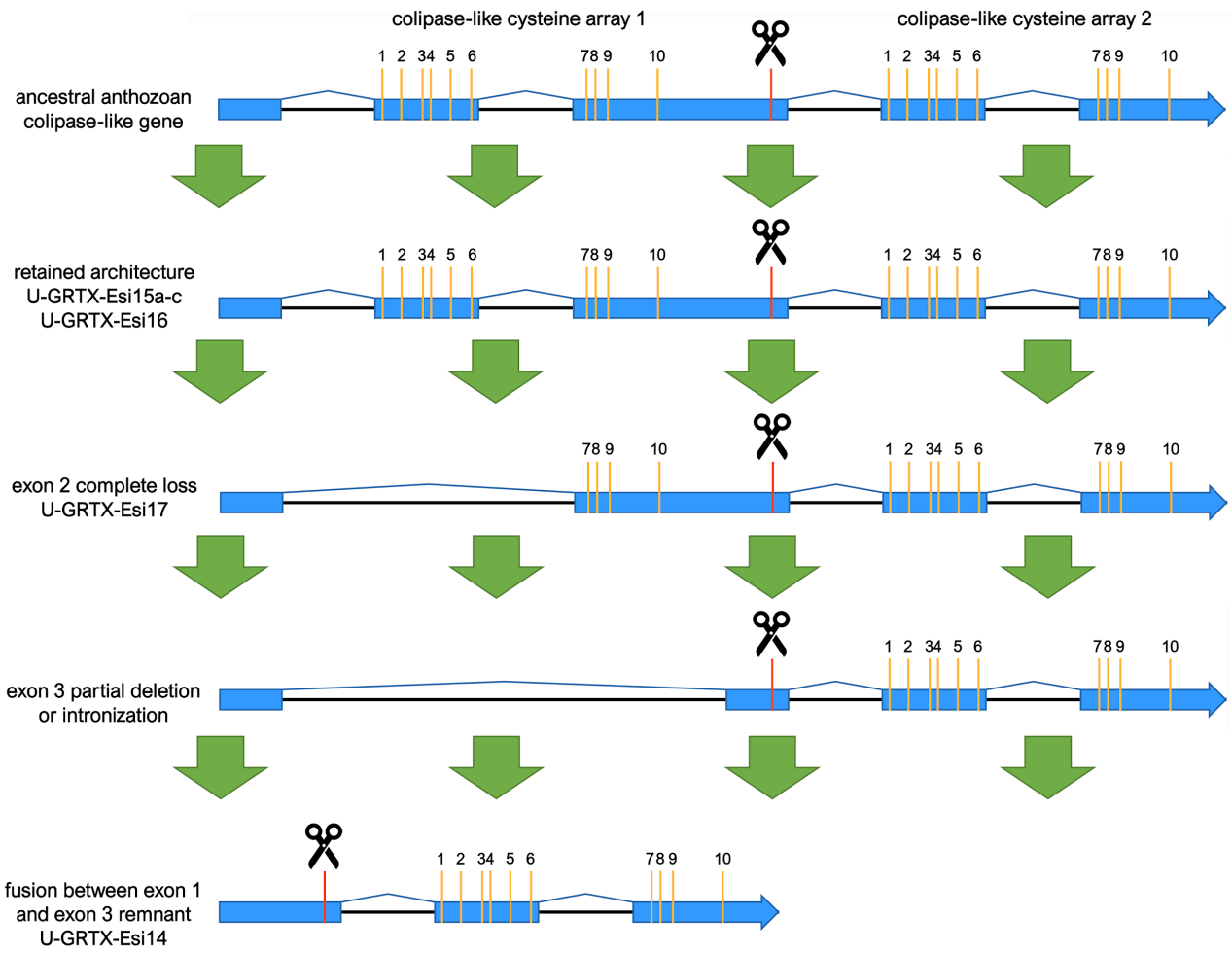


Figure 4. Schematic representation of the evolution of the colipase genes in Anthozoa, with references to specific *E. singularis* genes. Exons are in blue, cysteine residues are indicated by the yellow bars, scissors indicate cleavage sites.

(c) Novel putative toxins were retrieved in the NEM-P of *Eunicella 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 Suppl. Fig. 5. 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 (Fig.5).

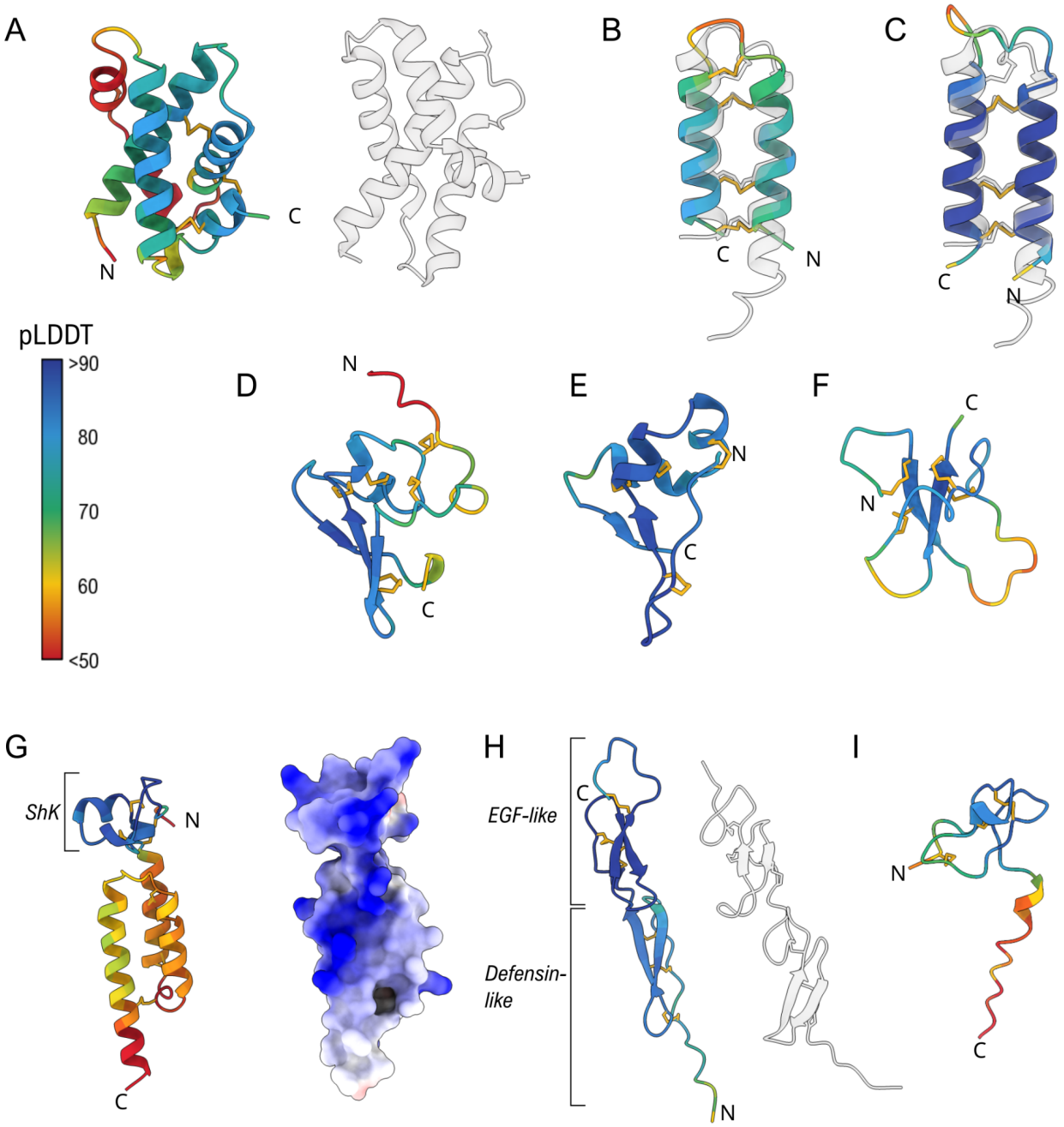


Figure 5. Overview of the structures predicted by AlphaFold2 for the novel toxins from *E. singularis.* Models are coloured according to pLDDT, with higher values (>70) indicating reliable prediction. (A) The putative cytolytic U-GRTX-Esi19 and, on the right, PulS from *K. oxytoca* (PDB 4A56); Overlay of the structures of the two disulphide 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 to the byssal protein Pvfp-5β from the mussel *Perna viridis* (PDB 7QAB, right); (I) The structure of U-GRTX-Esi28

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

358 C-terminus with Amoebapore A, the major pore-forming protein of *Entamoeba histolytica*, suggestive of a
359 possible cytolytic function [76]. Most of the protein structure was modeled with high confidence (pLDDT >
360 80, Fig. 5A), highlighting a compact fold characterized by a knotted four-helix bundle stabilized by three
361 inter-helices disulphide bridges. A similar arrangement, although devoid of disulphide stabilization, is typical
362 of bacterial pilotins PulS and GspS, chaperone lipoproteins that play a role in the assembly of the T2SS
363 secretion system in pathogenic Gram-negative bacteria [77,78]. This putative toxin appears as an
364 evolutionary innovation of Malacalcyonacea, although orthologues are missing in *Phenganax subtilis* and in
365 *Xenia* sp. (Fig. 7, Suppl. Fig. 6). The cysteine framework is always conserved, while the length of some loops
366 varies across species. The abundance of this protein suggests a strong functional relevance, but activity
367 assays will be necessary to clarify its function.
368 The two putative toxins U-GRTX-Esi20 and 21 did not show any obvious sequence similarity, with a pairwise
369 identity of only 21%. U-GRTX-Esi20 is 48 amino acids long and derives from a longer propeptide
370 proteolytically cleaved at the C-terminus, whereas U-GRTX-Esi21 displays a shorter sequence with a very
371 high cysteine density (8 out of 36 amino acids). The models obtained (Fig. 5B and C) highlighted a structural
372 relationship, as both presented a disulphide-stabilized helical hairpin motif, recently characterized in
373 acrorhagin-1 (U-AITX-Aeq5a) [79], a toxin from the acrorhagi of *A. equina* used in conspecific aggressive
374 encounters [80]. Acrorhagin-1 can induce local tissue necrosis and displays toxicity on crabs, but its
375 mechanism of action has not been clarified, although it has been demonstrated that the toxin does not
376 interact with either potassium or sodium ion channels and it does not possess membranolytic activity
377 [79,80]. U-GRTX-Esi20 and 21 are stabilized by three and four disulphide bridges, respectively, which overlap
378 with those of acrorhagin-1, despite a sequence identity below 20% in both cases.

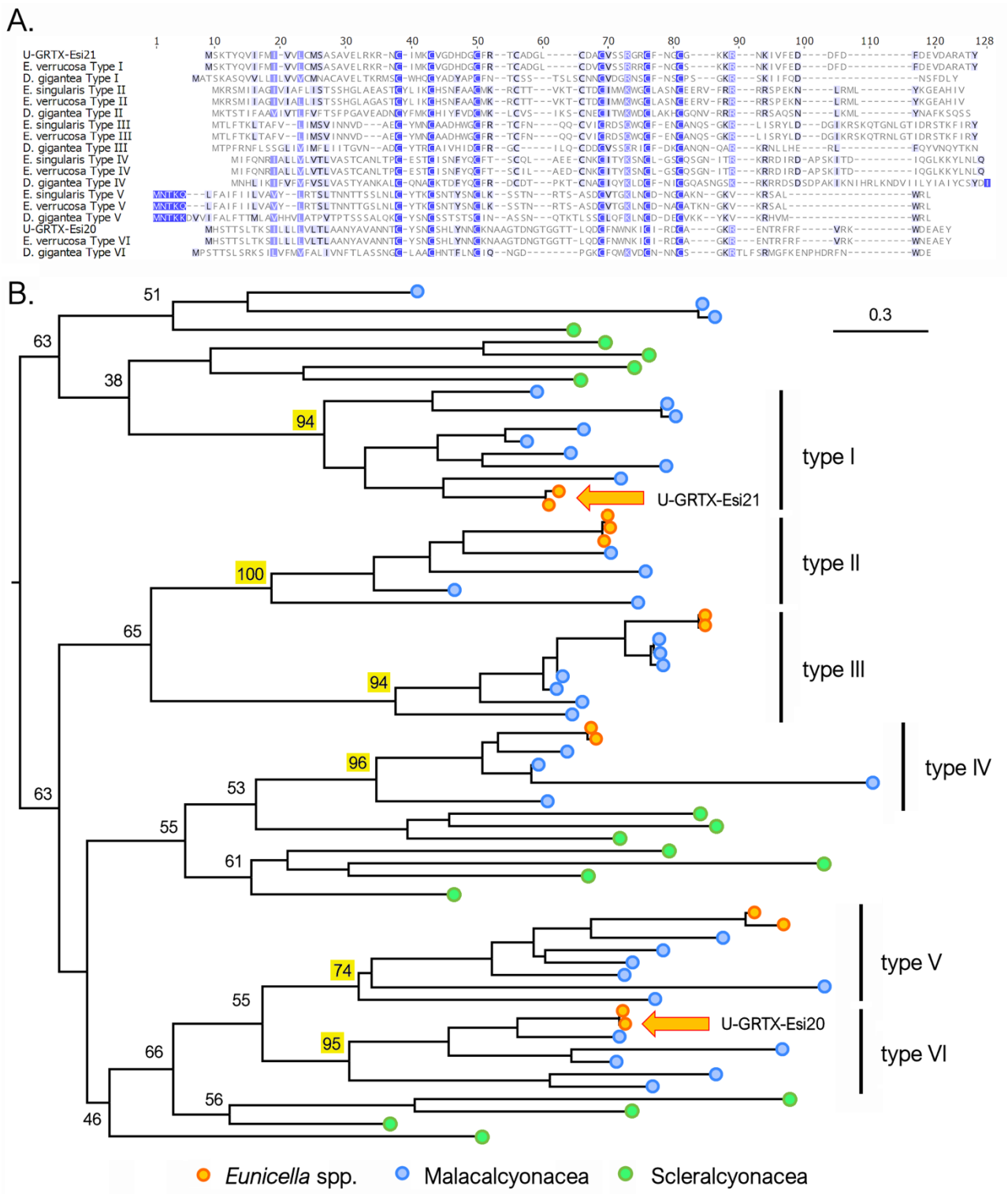
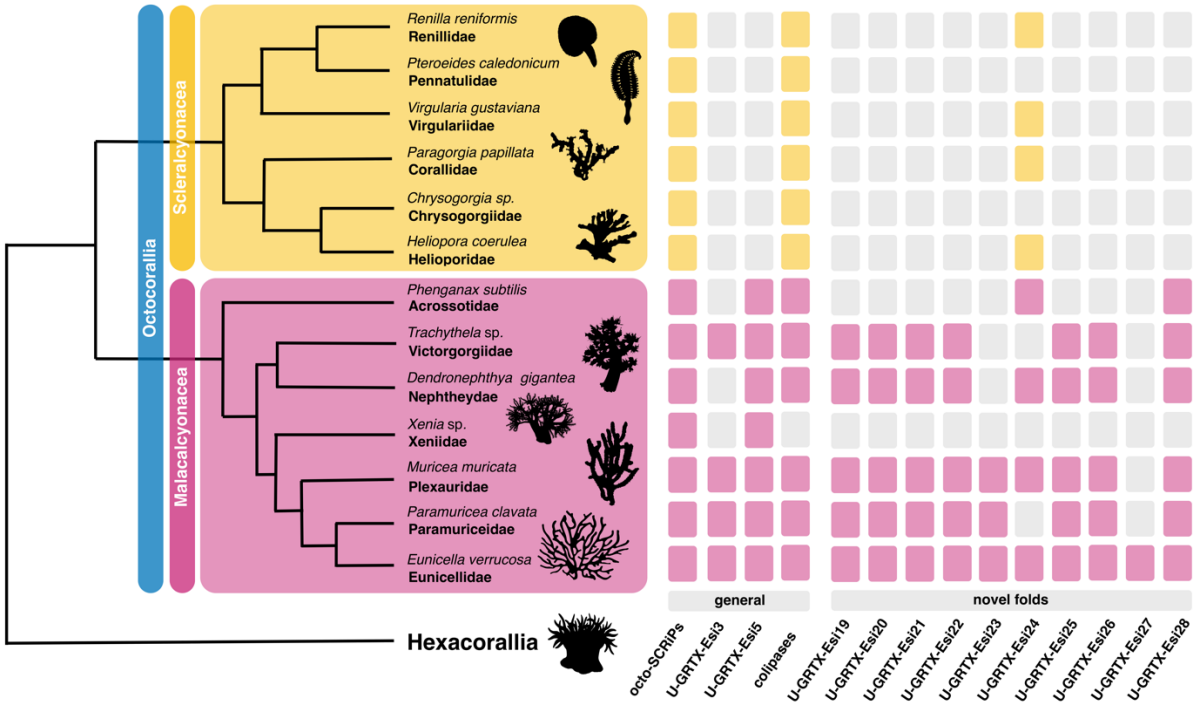


Figure 6. A) multiple sequence alignment of the complete precursor amino acid sequences of U-GRTX-Esi20, U-GRTX-Esi21, their paralogs, and their orthologs in *E. verrucosa* and *D. gigantea*. **B)** Maximum likelihood phylogeny of the sequences belonging to the same superfamily of U-GRTX-Esi20 and U-GRTX-Esi21 (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.

388 Our genomic survey indicates that these two toxins belong to two distinct monophyletic orthogroups within
389 a very large superfamily of cysteine-rich peptides, encoded by ancestrally duplicated paralogous genes with
390 intronless structures, that also include other members in *E. singularis* transcriptome, not detected in the
391 NEM-P (Fig. 6). High-confidence orthologues were exclusively found in Malacalcyonacea, except for *P.*
392 *subtilis* and *Xenia*, forming 6 well-supported clades (Type I to Type VI) (Fig.6). This suggests a rather recent
393 origin of these orthogroups, which derive from a large group of similar peptides also found in other
394 cnidarians, worthy of more extensive investigation. In fact, an ortholog of acrorhagin has been detected in
395 the sensory neurons of *N. vectensis* [81], suggesting an ancestral recruitment from the nervous to the venom
396 system for this class of peptides, followed by lineage-specific diversification in the Octocorallia (Fig. 7).
397



398
399 **Figure 7.** Distribution of orthologues of a subset of *E. singularis* toxins across the Octocorallia genomes here
400 investigated. Phylogenetic relationships of the Octocorallia after [46].
401

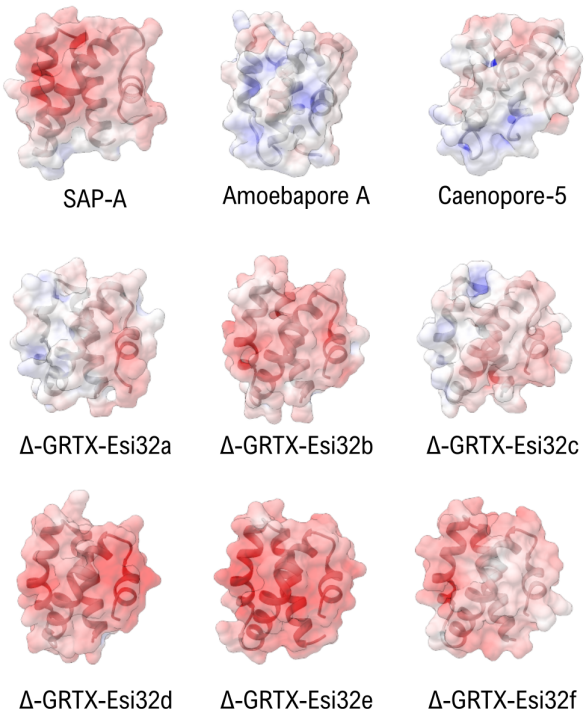
402 The next two putative toxins, U-GRTX-Esi22 (60 amino acids, 10 cysteines, Fig. 5D) and 23 (45 amino acids,
403 8 cysteines, Fig. 5E) did not present any sequence similarity with known proteins. In both cases, the models
404 highlighted an unusual disulphide-stabilized αβ fold, suggesting the existence of peculiar and to date
405 undiscovered cnidarian toxin folds. The gene encoding U-GRTX-Esi23 is present with two similar paralogous
406 genes in *E. singularis* and has clearly detectable orthologs in all Malacalcyonacea except *P. subtilis* and *Xenia*

407 sp. (Suppl. Fig. 6, Fig. 7) A divergent paralogue, with a slightly different cysteine pattern, is found in
408 Scleralcyonacea. This toxin most likely constitutes an evolutionary innovation of Malacalcyonacea but could
409 derive from an ancestral sequence with a similar cysteine connectivity present in the latest common
410 ancestor of all Octocorallia. Conversely, U-GRTX-Esi22 presents a single-copy gene, with a detectable
411 orthologue only in *P. clavata* and *M. muricata,* and is probably a recent innovation, only shared by a few
412 related orders within Malacalcyonacea (Suppl. Fig. 7, Fig. 7).
413 The structure of U-GRTX-Esi24, a 45 amino acids protein with 8 cysteines and a 50% sequence identity with
414 Conotoxin Cltx-2 from *Conus californicus,* could also be modelled with acceptable confidence (Fig. 5F).
415 Despite undetected by HMM profile analysis, the structure presented a SCRiP-like W-shaped fold. This
416 domain has a broad but scattered distribution in Octocorallia, suggesting an evolutionary origin slightly older
417 than most other toxins described in this work. Nevertheless, it is most often found repeated multiple times
418 within the same protein (also in *E. singularis*), with several independent occurrences of single-domain
419 peptides scattered across the tree of life (Fig. 7; Suppl. Fig. 8).
420 Unique case among the novel toxins from *E. singularis,* U-GRTX-Esi25 was predicted to possess a modular
421 architecture: a recognizable N-terminal ShK domain followed by a region lacking any significant homology.
422 The two domains were clearly distinguishable in the three-dimensional prediction (Fig. 5G), where they
423 appeared as directly connected, without spacer sequences or putative cleavage sites. The unknown C-
424 terminal domain was modelled as an asymmetrical three-helices arrangement, with three disulphide bridges
425 connecting and stabilizing the secondary structure elements. Interestingly, despite the complex architecture
426 predicted, several basic residues from both domains in the model aligned perfectly, forming an extended
427 positively charged patch (Fig. 5G), a feature that suggests a possible membranolytic and/or antimicrobial
428 action [82]. This peculiar combination of domains, paired with a shared 3 exons/2 introns gene architecture,
429 was detected in Malacalcyonacea (except *P. subtilis* and *Xenia* sp). A single orthologous gene was found in
430 *Paragorgia papillata* (Scleralcyonacea), suggesting that the appearance of this protein could predate the
431 split between Malacolcyonacea and Scleralcyonacea (Fig. 7; Suppl. Fig. 9).
432 The model predicted for U-GRTX-Esi26 (77 residues, 12 Cys) presented an elongated structure made of two
433 domains: an N-terminal region in which two antiparallel beta strands are connected by three disulphide

434 bridges to a proline-rich stretch, in a defensin-like fold found in other cnidarian toxins, and a C-terminal EGF-
435 like domain (Fig. 5H). A similar modular structure was recently described in the byssal adhesive protein Pvfp-
436 5β from the mussel *Perna viridis* [83]; despite the absence of sequence similarity, this structural similarity
437 might indicate for this protein a structural or adhesive function. However, orthologues of the gene were
438 found only in Malacalcyonacea (apart from *Dendronephthya* and *Xenia)*, all sharing a common 3 exons/2
439 introns gene structure. Three divergent paralogues were found In *P. clavata*, probably originated through a
440 species- or genus-specific gene duplication process (Suppl. Fig. 10).
441 The last two novel toxins identified in the NEM-P of *E. singularis,* U-GRTX-Esi27 and 28, were short peptides
442 associated with relatively low transcription levels but a high proteomic coverage. In particular, U-GRTX-Esi27
443 displayed a weak sequence similarity to both SCRiPs and BDS toxins from *A. sulcata*, which are potent ligands
444 of K_v3 channels [84]. The mature sequence of U-GRTX-Esi27, a 39 amino acids long peptide, contained 8
445 cysteines, 6 of which correspond to the BDS framework. An additional cysteine pair occupied the position
446 of two extremely conserved residues in sea anemone BDS toxins, *i.e.*, a tryptophan and a tyrosine residue,
447 otherwise present in all other members of the family (Suppl. Fig. 11). Only low confidence models could be
448 predicted by AlphaFold2, thus preventing any function inference. Despite the similarity with BDS,
449 orthologous genes encoding for similar proteins were not detected in the genomes of the other Octocorallia
450 considered in this study except from *E. verrucosa.* U-GRTX-Esi27 might therefore constitute an evolutionary
451 innovation of the genus *Eunicella* or the family Eunicellidae (Fig. 7). Finally, U-GRTX-Esi28 possessed a weak
452 sequence homology with the ω-hexatoxin-Hi1g. The modelled structure confirmed this similarity, revealing
453 a cystine knotted core (Fig. 5I), with the addition of an extended and disordered C-terminus, possibly
454 proteolitically cleaved. Orthologues coding for this protein are shared by nearly all Malacalcyonacea but are
455 lacking in *Xenia* (Suppl. Fig. S12). This finding, along with the conservation of the gene structure, in which
456 two exons are spaced by an intron with an equal placement in all species, the signal peptide sequence, and
457 the cleavage sites, strongly indicated that this protein represents an evolutionary innovation of
458 Malacalcyonacea (Fig. 7).
459
460 ***(d)* A diverse range of cytolytic proteins are produced both inside and outside the nematocysts**

461 Our analysis highlighted the absence, in the NEM-P of *E. singularis*, of the common cnidarian cytolytic
462 proteins. A targeted hmmscan search in the WB-P yielded instead positive hits for profiles associated with
463 both the aerolysin (PF01117) and the full anemone cytotoxin (PF06369) families. Coherently with the
464 approach adopted for the NEM-P annotation, we considered only the proteins encoded by complete
465 transcripts presenting a signal peptide; this approach resulted in the identification of three pore forming
466 toxins (PFTs), characterized by diverse structural features. The first one, Δ-GRTX-Esi29, was an aerolysin-like
467 protein, structurally related to the β-PFT Hydralysin [85]; the other two, Δ-GRTX-Esi30 and 31, which shared
468 ~50% sequence similarity, belonged to the actinoporins family, the most abundant family of cnidarian porins
469 [86]. Interestingly, we observed relevant non-conservative substitutions in *E. singularis* actinoporins, which
470 may indicate selectivity towards different lipids (Suppl. Fig. 13).
471 Finally, the highest scoring match in the proteomic analysis of the nematocysts content of *E. singularis*
472 corresponded to a 699 amino-acid translated transcript with homology to prosaposin (pSAP) [87]. In
473 humans, the maturation of the orthologous pSAP produces four globular saposins (SAP-A/D), which are co-
474 factors of sphingolipid degrading enzymes. The saposin fold is very ancient and characterizes all the saposin-
475 like proteins (SAPLIPs), which include the amoebapores [88-90]. SAPs and SAPLIPs have a distinctive cysteine
476 framework but display minimal conservation (<25% sequence identity). They are often released in several
477 copies from the same precursor gene [91], assuming a closed, monomeric conformation at lysosomal pH
478 [88,92]. All these proteins are involved in membrane destabilization/solubilization, acting through different
479 mechanisms: lysosomal SAPs undergo a conformational change upon pH shift or in the presence of lipids or
480 detergents, exposing the residues of the hydrophobic core and encapsulating a large number of lipids [93-
481 95]. Other SAPLIPs, like the amoebapores from *E. histolytica* and the caenopores from *C. elegan*s, form
482 oligomeric pores [96,97]. Finally, basic SAPLIPs, like NK-lysin and granulysin, act by "molecular
483 electroporation", inserting within the membranes [98,99]. In rare cases, the saposin fold has also been
484 detected in non-cytolytic neurotoxins, like the macro-conotoxins con-ikot-ikot and Mu8.1 [100,101], but
485 these proteins are highly dissimilar from SAPLIPs in terms of sequence and cysteine pattern and are encoded
486 by single domain (non pSAP-like) transcripts. The six sequences deriving from the maturation of the pSAP
487 precursor of *E. singularis* (Δ-GRTX-Esi32a-f) were all identified at the proteomic level in the NEM-P. These

488 proteins shared a 24%-65% similarity, higher than what generally observed for homologous SAPLIPs, and
489 were more similar to human saposins SAP-A and C (similarity 26-40%) than to antimicrobial porogenic
490 peptides like amoebapore A (16-25%) or caenopore-5 (13-27%).
491 We examined the surface potential distribution of the SAPLIPs from *E. singularis*, since the activity of SAPLIPs
492 correlates to their surface electrostatics [102,103]. Except Δ-GRTX-Esi32a and c, they all shared with human
493 SAP-A extended patches of negative charge (Fig. 8), which hints to a pH-dependent mechanism for activation
494 and membrane interaction [104]. In contrast, Δ-GRTX-Esi32a and c presented a more balanced distribution
495 of basic and acidic regions and extended neutral patches, like caenopore-5, suggesting a different, possibly
496 non-monomeric, native arrangement and a pH-independent mechanism of action, possibly through pore
497 formation.



498
499 **Figure 8**. Comparison of the surface electrostatic potentials of the three model saposin-like proteins SAP-A (4UEX),
500 Amoebapore A (1OF9) and Caenopore-5 (2JSA) with the six SAPLIPs identified in the venom of *E. singularis.* The
501 extended patches of negative charge (red) observed in Δ-GRTX-Esi32b, d, e, f are observed in human SAP-A and are
502 indicative of a pH-dependent conformational switch. The lower charge density of Δ-GRTX-Esi32a and c may indicate a
503 different mode of action.
504

505 **Conclusions**

506 We have provided the first preliminary depiction of the venomous arsenal of a widespread octocoral.
507 Although experimental limitations did not allow us to conduct our analysis on an isolated venom fraction,

508 by exposing whole polyps of the gorgonian *Eunicella singularis* to ethanol [29] we isolated a sample enriched
509 in proteins presenting typical toxin features. Among them, we detected several folds previously identified
510 in cnidarian venoms, some proteins with structural similarities to venom components of unrelated
511 organisms, and a few with new folds. Even in the absence of functional characterization, the NEM-P seems
512 to provide a good snapshot of the nematocysts content in *E. singularis,* although we could not rule out the
513 possibility that these proteins derive from other tissues. The scarce similarity observed between octocoral
514 and hexacoral putative toxins repertoire is not surprising, given the ancient divergence between these two
515 lineages that recent works dated about 500-700 MYA [2]. In comparison, Elapidae and Viperidae snakes,
516 which present remarkable differences in their venom composition, have diverged only 50 MYA [105].
517 However, our results clearly indicate, within the NEM-P of *E. singularis,* an elevated content of putative
518 neurotoxic peptides, a feature previously reported for sea anemones. Cytolytic toxins with different
519 structures (and, likely, specificity/function) seem distributed both inside and outside the nematocysts.
520 The transcriptionally more abundant components of *E. singularis* venom are the potentially cytolytic U-
521 GRTX-Esi19 and the acrorhagin-like U-GRTX-Esi20, the six colipases, the SCRiP-like U-GRTX-Esi1a and the
522 SCREP U-GRTX-Esi4 with three ShK domains. With the exception of the "octo-SCRiPs", which, however, are
523 characterized by sequence peculiarities, these toxins represent evolutionary innovations of Octocorallia,
524 either shared by the entire lineage (*e.g.,* some colipases), or restricted to the Malacalcyonacea (*e.g.*, U-
525 GRTX-Esi19), or even exclusive of a group of sublineages within the Malacalcyonacea (*e.g.*, U-GRTX-Esi20)
526 (Fig.7).
527 Putative neurotoxins in *E. singularis* included both novel and established cnidarian folds, the latter often
528 detected within multidomain SCREPs [106]. This is the case, *e.g.,* of the highly expressed U-GRTX-Esi4, with
529 three ShK domains: multidomain ShK proteins have been so far reported in the transcriptomes of parasitic
530 roundworms, in *Hydra* and *Nematostella* [58], and in the vampire snails [107]. The toxin U-GRTX-Esi3
531 represents another case of SCREP, presenting a double SCRiP domain in the mature polypeptide: this
532 architecture had not been documented in Cnidaria mature toxins to date, although it has been recognized
533 at the transcript level (e.g. in *Orbicella faveolata*). Two domains were also retrieved in the Kunitz BPTI-like
534 PI-GRTX-Esi5, an architecture so far investigated in detail in hematophagous invertebrates only, but

535 extremely abundant among the putative ion-channel impairing SCREP toxins [67,106]. While the abundance
536 of the SCREPs is most likely underestimated across venomous Metazoa, also due to the paucity of proteomic
537 studies confirming the architecture of mature toxins, their abundance in the venom of *Eunicella* might
538 indeed possess some functional significance.
539 Interestingly, conventional pore forming toxins (PFTs) of Cnidaria, including the aerolysin-like Δ-GRTX-Esi29
540 and the two actinoporins Δ-GRTX-Esi30 and 31 were not retrieved in the NEM-P. As anticipated, this could
541 depend on the intrinsic limitations of the venom collection procedure employed or reflect an extra-
542 nematocystic production of these proteins. Indeed, finding PFTs outside the nematocysts is not unusual:
543 e.g., hydralysin is an exclusively non-nematocystic PFT involved in prey digestion [84]. Actinoporins, on the
544 other hand, have been to date detected both in the nematocysts and in other tissues. The extra-
545 nematocystic localization retrieved in *E. singularis* might suggest that its actinoporins share the digestive
546 function of the hydralysin-like Δ-GRTX-Esi29. Further histological and functional analysis are needed to
547 confirm this hypothesis, and to evaluate the impact of the observed mutations at critical sites on chemical
548 selectivity and pore formation mechanism of *Eunicella* actinoporins.
549 Among the proteins of the NEM-P, likely representing the nematocyst content, the cytolytic activity seems
550 to be ascribable mostly to the six saposins Δ-GRTX-Esi32a-f, abundant at the proteomic level. The analysis
551 of their surface charge hints to a considerable diversification of targets and modes of action for this novel
552 group of cytolytic polypeptides: most of them might be inactive at the low pH typical of mature nematocysts
553 but could be activated by the abrupt pH change which follows discharge [108]. On the contrary, the neutral
554 surface electrostatic of Δ-GRTX-Esi32a and c might suggest a pH-independent activation. The novel toxin U-
555 GRTX-Esi19, transcribed at high levels and with a fold compatible with a cytolytic activity, might complement
556 saposins.
557 While some of these putative toxins are shared by all the investigated octocorals, others had a taxonomic
558 range restricted to Malacalcoynacea and were not retrieved in the analysed Scleralcyonacea, suggesting
559 they might constitute evolutionary innovations arisen after the split between the two major Octocorallia
560 lineages. This is the case of the novel toxins U-GRTX-Esi19 and U-GRTX-Esi28 (Fig. 7). In addition, several
561 toxins were present in a subset of Malacalcyonacea species and are frequently lacking in the representatives

562 of two specific families, Xeniidae (or pumping corals) and Acrossotiidae (i.e., tube corals) (Fig. 7). This
563 peculiar taxonomic distribution might suggest a key role in predation for these toxins, since they are absent
564 in species with a mixotrophic habit, in which autotrophy sustained by symbiotic dinoflagellates is the most
565 relevant source of energy. In contrast, the other Malacalcyonacea included in our analysis (e.g.,
566 *Paramuricea*, *Trachytela*) are suspension feeders that largely rely on zooplankton intake [109]. Novel toxins,
567 such as U-GRTX-Esi26, U-GRTX-Esi22, the acrorhagin-like U-GRTX-Esi21 and U-GRTX-Esi27 might thus
568 constitute a predation-related group of putative toxins. In particular, U-GRTX-Esi27 appears to be restricted
569 to *Eunicella* as a genus-specific adaptation (Fig. 7).
570 The inference on the potential function of octocoral putative toxins is, however, heavily affected by the
571 sparse information available on the biology of these organisms. For example, some Octocorallia species
572 possess sweeper and thread-like tentacles, nematocysts-rich structures that have been hypothesized to
573 mediate interspecific competition, similarly to sea anemones acrorhagi [110]. The occurrence of these
574 structures across the Octocorallia radiation has been poorly investigated to date: this lack of comprehensive
575 information hampers the evaluation of a potential role of specific toxins in competition. Nevertheless,
576 although our comparative genomic approach has revealed a powerful tool to restrict the range of potential
577 functions, further functional characterization studies are needed to unequivocally establish the physiological
578 role of these toxins.
579 Interestingly, our proteomic analysis pointed towards two cases of differential toxin maturation pathways
580 in the nematocyst vs whole body. The first concerns the two SCRiP isotoxins U-GRTX-Esi2a-b, derived from
581 a double-domain precursor, with both isoforms present in the WB-P while only U-GRTX-Esi2b represented
582 in the NEM-P (Suppl. Fig. 1). The second corresponds to the toxin PI-GRTX-Esi6 and the three isotoxins PI-
583 GRTX-Esi7a-c, encoded by a single transcript in which a Kunitz-like domain (retrieved both in the whole body
584 and in the nematocyst proteomes) is followed by three Kazal-type PI domains only identified in the whole-
585 body proteome. The contribution of alternative cleavage sites to venom diversity has been demonstrated in
586 conopeptides [111], but cell- or tissue-specific differential maturation patterns have not been so far
587 documented for venom toxins, suggesting that the mechanisms regulating venom toxin expression could be
588 more complex than previously assessed.

589 Through our multidisciplinary approach, we have demonstrated that Octocorallia produce a complex
590 neurotoxin-rich venom, showcasing an overall scarce resemblance to the hexacoral venom repertoire. This
591 distinctiveness is underscored by the presence of novel folds with a complex evolutionary history. These
592 findings call for extensive work, focusing both on the biology of this understudied group of organisms and
593 on the functional characterization of their toxins, to unravel the adaptive significance of Octocorallia arsenal.
594

595 **Data Availability**

596 The Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the
597 accession GKVI00000000. The version described in this paper is the first version, GKVI01000000. Proteomic
598 data have been submitted to PRIDE with reference 1-20240529-113849-2368759. Other data underlying this
599 article are available in the online supplementary material.
600

601 **Acknowledgements**

602 This work has received funding from the European Union's Horizon 2020 research and innovation
603 programme under the Marie Skłodowska-Curie grant agreement No 748902. The MGX acknowledges
604 financial support from France Génomique National infrastructure, funded as part of "Investissement
605 d'Avenir" program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09). Mass
606 spectrometry experiments were carried out using the facilities of the Montpellier Proteomics Platform
607 (PPM, BioCampus Montpellier). Dr. Dany Dominguez-Perez is gratefully acknowledged for his help in data
608 handling and curation.
609

610 **References**

611 1. Van Iten H, Marques AC, Leme J, Pacheco MLA, Simões MG. 2014 Origin and early diversification of
612 the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's proterozoic-cambrian
613 history. Palaeontology 57, 677–690. (doi:10.1111/pala.12116)
614 2. McFadden CS, Quattrini AM, Brugler MR, Cowman PF, Dueñas LF, Kitahara M V., Paz-García DA,
615 Reimer JD, Rodríguez E. 2021 Phylogenomics, Origin, and Diversification of Anthozoans (Phylum

616 Cnidaria). *Syst Biol* **70**, 635–647. (doi:10.1093/sysbio/syaa103)

617 3. David CN, Özbek S, Adamczyk P, Meier S, Pauly B, Chapman J, Hwang JS, Gojobori T, Holstein TW.
618 2008 Evolution of complex structures: minicollagens shape the cnidarian nematocyst. *Trends in*
619 *Genetics* **24**, 431–438. (doi:10.1016/j.tig.2008.07.001)

620 4. Fautin DG. 2009 Structural diversity, systematics, and evolution of cnidae. *Toxicon* **54**, 1054–1064.
621 (doi:10.1016/j.toxicon.2009.02.024)

622 5. Kass-Simon G, Scappaticci, Jr. AA. 2002 The behavioral and developmental physiology of
623 nematocysts. *Can J Zool* **80**, 1772–1794. (doi:10.1139/z02-135)

624 6. Moran Y, Genikhovich G, Gordon D, Wienkoop S, Zenkert C, Özbek S, Technau U, Gurevitz M. 2012
625 Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom
626 delivery in sea anemones. *Proc. Royal Soc B* **279**, 1351–1358. (doi:10.1098/rspb.2011.1731)

627 7. Marques AC, Collins AG. 2004 Cladistic analysis of Medusozoa and cnidarian evolution. *Inv Biol* **123**,
628 23–42. (doi:10.1111/j.1744-7410.2004.tb00139.x)

629 8. Kayal E, Bentlage B, Sabrina Pankey M, Ohdera AH, Medina M, Plachetzki DC, Collins AG, Ryan JF.
630 2018 Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the
631 origins of key organismal traits. *BMC Evol Biol* **18**, 68. (doi:10.1186/s12862-018-1142-0)

632 9. Norton RS. 1991 Structure and structure-function relationships of sea anemone proteins that
633 interact with the sodium channel. *Toxicon* **29**, 1051–1084. (doi:10.1016/0041-0101(91)90205-6)

634 10. Norton RS. 2009 Structures of sea anemone toxins. *Toxicon* **54**, 1075–1088.
635 (doi:10.1016/j.toxicon.2009.02.035)

636 11. Honma T, Shiomi K. 2006 Peptide toxins in sea anemones: Structural and functional aspects. *Mar*
637 *Biotech* **8**, 1–10. (doi:10.1007/s10126-005-5093-2)

638 12. Shiomi K. 2009 Novel peptide toxins recently isolated from sea anemones. *Toxicon* **54**, 1112–1118.
639 (doi:10.1016/j.toxicon.2009.02.031)

640 13. Jouiaei M, Yanagihara AA, Madio B, Nevalainen TJ, Alewood PF, Fry BG. 2015 Ancient venom
641 systems: A review on Cnidaria toxins. *Toxins* **7**, 2251–2271. (doi:10.3390/toxins7062251)

642 14. Madio B, King GF, Undheim EAB. 2019 Sea anemone toxins: A structural overview. *Mar Drugs* **17**,

643 325. (doi:10.3390/md17060325)

644 15. Menezes C, Thakur NL. 2022 Sea anemone venom: Ecological interactions and bioactive potential.
645 *Toxicon* **208**, 31–46. (doi:10.1016/j.toxicon.2022.01.004)

646 16. Prentis PJ, Pavasovic A, Norton RS. 2018 Sea anemones: Quiet achievers in the field of peptide
647 toxins. *Toxins* **10**, 36. (doi:10.3390/toxins10010036)

648 17. Rachamim T, Morgenstern D, Aharonovich D, Brekhman V, Lotan T, Sher D. 2015 The dynamically
649 evolving nematocyst content of an Anthozoan, a scyphozoan, and a hydrozoan. *Mol Biol Evol* **32**,
650 740–753. (doi:10.1093/molbev/msu335)

651 18. Pennington MW, Czerwinski A, Norton RS. 2018 Peptide therapeutics from venom: Current status
652 and potential. *Bioorg Med Chem* **26**, 2738–2758. (doi:10.1016/j.bmc.2017.09.029)

653 19. Schmidt CA, Daly NL, Wilson DT. 2019 Coral venom toxins. *Front Ecol Evol* **7**.
654 (doi:10.3389/fevo.2019.00320)

655 20. Klompen AML, Macrander J, Reitzel AM, Stampar SN. 2020 Transcriptomic analysis of four cerianthid
656 (Cnidaria, Ceriantharia) venoms. *Mar Drugs* **18**, 413. (doi:10.3390/MD18080413)

657 21. Yoffe C, Lotan T, Benayhau Y. 2012 A modified view on octocorals: *Heteroxenia fuscescens*
658 nematocysts are diverse, featuring both an ancestral and a novel type. *PLoS One* **7**, 1–5.
659 (doi:10.1371/journal.pone.0031902)

660 22. Raimundo I, Silva SG, Costa R, Keller-Costa T. 2018 Bioactive secondary metabolites from octocoral-
661 associated microbes. New chances for blue growth. *Mar Drugs* **16**, 485. (doi:10.3390/md16120485)

662 23. Ballestreros E. 2006 Mediterranean coralligenous assemblages: A synthesis of present knowledge.
663 In *Oceanography and marine biology: An annual review*, pp. 123–195.

664 24. Linares C, Coma R, Garrabou J, Díaz D, Zabala M. 2008 Size distribution, density and disturbance in
665 two Mediterranean gorgonians: *Paramuricea clavata* and *Eunicella singularis*. *J Appl Ecol* **45**, 688–
666 699. (doi:10.1111/j.1365-2664.2007.01419.x)

667 25. Gori A, Rossi S, Berganzo E, Pretus JL, Dale MRT, Gili JM. 2011 Spatial distribution patterns of the
668 gorgonians *Eunicella singularis*, *Paramuricea clavata*, and *Leptogorgia sarmentosa* (Cape of Creus,
669 Northwestern Mediterranean Sea). *Mar Biol* **158**, 143–158. (doi:10.1007/s00227-010-1548-8)

670 26. Coma R, Llorente-Llurba E, Serrano E, Gili JM, Ribes M. 2015 Natural heterotrophic feeding by a
671 temperate octocoral with symbiotic zooxanthellae: a contribution to understanding the
672 mechanisms of die-off events. *Coral Reefs* **34**, 549–560. (doi:10.1007/s00338-015-1281-3)
673 27. Renzi JJ, Shaver EC, Burkepile DE, Silliman BR. 2022 The role of predators in coral disease dynamics.
674 *Coral Reefs* **41**, 405–422. (doi:10.1007/s00338-022-02219-w)
675 28. Radwan FFY, Aboul-Dahab HM, Burnett JW. 2002 Some toxicological characteristics of three
676 venomous soft corals from the Red Sea. *Comp Biochem Physiol C* **132**, 25–35. (doi:10.1016/S1532-
677 0456(02)00045-5)
678 29. Jouiaei M *et al.* 2015 Firing the sting: Chemically induced discharge of cnidae reveals novel proteins
679 and peptides from box jellyfish (*Chironex fleckeri*) Venom. *Toxins* **7**, 936–950.
680 (doi:10.3390/toxins7030936)
681 30. Grabherr MG *et al.* 2011 Full-length transcriptome assembly from RNA-Seq data without a reference
682 genome. *Nat Biotechnol* **29**, 644–652. (doi:10.1038/nbt.1883)
683 31. Wagner GP, Kin K, Lynch VJ. 2012 Measurement of mRNA abundance using RNA-seq data: RPKM
684 measure is inconsistent among samples. *Theor Biosci* **131**, 281–285. (doi:10.1007/s12064-012-0162-
685 3)
686 32. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G. 2003 PEAKS: Powerful software
687 for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectr* **17**,
688 2337–2342. (doi:10.1002/rcm.1196)
689 33. Zhang J *et al.* 2012 PEAKS DB: de novo sequencing assisted database search for sensitive and
690 accurate peptide identification. *Mol Cell Proteomics* **11**, M111.010587.
691 (doi:10.1074/mcp.M111.010587)
692 34. Han Y, Ma B, Zhang K. 2005 SPIDER: software for protein identification from sequence tags with de
693 novo sequencing error. *J Bioinform Comput Biol* **3**, 697–716. (doi:10.1142/s0219720005001247)
694 35. Degueldre M et al. 2017 Diversity in sequences, post-translational modifications and expected
695 pharmacological activities of toxins from four *Conus* species revealed by the combination of cutting-
696 edge proteomics, transcriptomics and bioinformatics. *Toxicon* **130**, 116–125.

697 (doi.org/10.1016/j.toxicon.2017.02.014)

698 36. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010 CD-HIT Suite: A web server for clustering and comparing
699 biological sequences. *Bioinformatics* **26**, 680–682. (doi:10.1093/bioinformatics/btq003)

700 37. Oliveira JS, Fuentes-Silva D, King GF. 2012 Development of a rational nomenclature for naming
701 peptide and protein toxins from sea anemones. *Toxicon* **60**, 539–550.
702 (doi:10.1016/j.toxicon.2012.05.020)

703 38. Jumper J *et al.* 2021 Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–
704 589. (doi:10.1038/s41586-021-03819-2)

705 39. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022 ColabFold: making
706 protein folding accessible to all. *Nat Methods* **19**, 679–682. (doi:10.1038/s41592-022-01488-1)

707 40. Tunyasuvunakool K *et al.* 2021 Highly accurate protein structure prediction for the human
708 proteome. *Nature* **596**, 590–596. (doi:10.1038/s41586-021-03828-1)

709 41. Holm L. 2022 Dali server: structural unification of protein families. *Nucleic Acids Res* **50**, W210–
710 W215. (doi:10.1093/nar/gkac387)

711 42. Ayoub R, Lee Y. 2019 Rupee: A fast and accurate purely geometric protein structure search. *PLoS*
712 *One* **14**, e0213712. (doi:10.1371/journal.pone.0213712)

713 43. Webb B, Sali A. 2016 Comparative protein structure modeling using MODELLER. *Curr Protoc*
714 *Bioinformatics* **2016**, 5.6.1-5.6.37. (doi:10.1002/cpbi.3)

715 44. Jurrus E *et al.* 2018 Improvements to the APBS biomolecular solvation software suite. *Protein*
716 *Science* **27**, 112–128. (doi:10.1002/pro.3280)

717 45. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023 UCSF
718 ChimeraX: Tools for structure building and analysis. *Protein Science* **32**, e4792.
719 (doi:10.1002/pro.4792)

720 46. McFadden CS, Van Ofwegen LP, Quattrini AM. 2022 Revisionary systematics of Octocorallia
721 (Cnidaria: Anthozoa) guided by phylogenomics. *Bull Soc Syst Biol* **1**. (doi:10.18061/bssb.v1i3.8735)

722 47. Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF. 2006 Composition-based statistics and
723 translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biol* **4**, 41.

724 (doi:10.1186/1741-7007-4-41)

725 48. Reese MG. 1997 Improved splice site detection in Genie. *J Comput Biol* **4**, 311–323.
726 (doi:10.1089/cmb.1997.4.311)

727 49. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-
728 TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. 2020.
729 *Mol Biol Evol.* **37**, 1530-1534 (doi: 10.1093/molbev/msaa015).

730 50. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017 ModelFinder: Fast model
731 selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587–589.
732 (doi:10.1038/nmeth.4285)

733 51. Sachkova MY, Singer SA, Macrander J, Reitzel AM, Peigneur S, Tytgat J, et al. 2019. The Birth and
734 Death of Toxins with Distinct Functions: A Case Study in the Sea Anemone Nematostella. *Mol Biol*
735 *Evol* **36**, 2001. (doi: 10.1093/molbev/msz132)

736 52. Jouiaei M, Sunagar K, Federman Gross A, Scheib H, Alewood PF, Moran Y, Fry BG. 2015. Evolution of
737 an ancient venom: recognition of a novel family of cnidarian toxins and the common evolutionary
738 origin of sodium and potassium neurotoxins in sea anemone. *Mol Biol Evol* **32**, 1598-610. (doi:
739 10.1093/molbev/msv050)

740 53. Sunagawa S, DeSalvo MK, Voolstra CR, Reyes-Bermudez A, Medina M. 2009 Identification and gene
741 expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals.
742 *PLoS One* **4**, e4865. (doi:10.1371/journal.pone.0004865)

743 54. Logashina YA *et al.* 2017 New disulfide-stabilized fold provides sea anemone peptide to exhibit both
744 antimicrobial and TRPA1 potentiating properties. *Toxins* **9**, 154. (doi:10.3390/toxins9050154)

745 55. Barroso RA, Ramos L, Moreno H, Antunes A. 2024 Evolutionary Analysis of Cnidaria Small Cysteine-
746 Rich Proteins (SCRiPs), an Enigmatic Neurotoxin Family from Stony Corals and Sea Anemones
747 (Anthozoa: Hexacorallia). *Toxins* **16**, 75. (doi:10.3390/toxins16020075)

748 56. Kozlov SA, Grishin EV. 2007. The universal algorithm of maturation for secretory and excretory
749 protein precursors. *Toxicon* **49**, 721–6.

750 57. Pennington MW, Czerwinski A, Norton RS. 2018. Peptide therapeutics from venom: Current status

751 and potential. *Bioorg. Med. Chem.* 26:2738–2758.

752 58. Columbus-Shenkar YY, Sachkova MY, Macrander J, Fridrich A, Modepalli V, Reitzel AM, Sunagar K,
753 Moran Y. 2018. Dynamics of venom composition across a complex life cycle**.** *Elife* **9,** e35014. (doi:
754 10.7554/eLife.35014)

755 59. Sintsova O *et al.* 2021 Sea anemone kunitz-type peptides demonstrate neuroprotective activity in
756 the 6-hydroxydopamine induced neurotoxicity model. *Biomedicines* **9**, 283.
757 (doi:10.3390/biomedicines9030283)

758 60. Peigneur S, Billen B, Derua R, Waelkens E, Debaveye S, Béress L, Tytgat J. 2011 A bifunctional sea
759 anemone peptide with Kunitz type protease and potassium channel inhibiting properties. *Biochem*
760 *Pharmacol* **82**, 81–90. (doi:10.1016/j.bcp.2011.03.023)

761 61. Nikolaev M V., Dorofeeva NA, Komarova MS, Korolkova Y V., Andreev YA, Mosharova I V., Grishin E
762 V., Tikhonov DB, Kozlov SA. 2017 TRPV1 activation power can switch an action mode for its
763 polypeptide ligands. *PLoS One* **12**, e0177077. (doi:10.1371/journal.pone.0177077)

764 62. García-Fernández R, Peigneur S, Pons T, Alvarez C, González L, Chávez MA, Tytgat J. 2016 The kunitz-
765 type protein ShPI-1 inhibits serine proteases and voltage-gated potassium channels. *Toxins* **8**, 110.
766 (doi:10.3390/toxins8040110)

767 63. Andreev YA *et al.* 2013 Polypeptide modulators of TRPV1 produce analgesia without hyperthermia.
768 *Mar Drugs* **11**, 5100–5115. (doi:10.3390/md11125100)

769 64. Wang X *et al.* 2021 Toxic Peptide From *Palythoa caribaeorum* Acting on the TRPV1 Channel Prevents
770 Pentylenetetrazol-Induced Epilepsy in Zebrafish Larvae. *Front Pharmacol* **12**, 763089.
771 (doi:10.3389/fphar.2021.763089)

772 65. Hwang SM, Jo YY, Cohen CF, Kim YH, Berta T, Park CK. 2022 Venom Peptide Toxins Targeting the
773 Outer Pore Region of Transient Receptor Potential Vanilloid 1 in Pain: Implications for Analgesic
774 Drug Development. *Int J Mol Sci* **23**, 5772. (doi:10.3390/ijms23105772)

775 66. Kvetkina A *et al.* 2022 Kunitz-Type Peptides from Sea Anemones Protect Neuronal Cells against
776 Parkinson's Disease Inductors via Inhibition of ROS Production and ATP-Induced P2X7 Receptor
777 Activation. *Int J Mol Sci* **23**, 5115. (doi:10.3390/ijms23095115)

778 67. Dai SX, Zhang A Di, Huang JF. 2012 Evolution, expansion and expression of the Kunitz/BPTI gene
779 family associated with long-term blood feeding in *Ixodes scapularis*. *BMC Evol Biol* **12**, 4.
780 (doi:10.1186/1471-2148-12-4)
781 68. Modica MV, Lombardo F, Franchini P, Oliverio M. 2015 The venomous cocktail of the vampire snail
782 Colubraria reticulata (Mollusca, Gastropoda). *BMC Genomics* **16**, 441. (doi:10.1186/s12864-015-
783 1648-4)
784 69. Mishra M. 2020 Evolutionary Aspects of the Structural Convergence and Functional Diversification
785 of Kunitz-Domain Inhibitors. *J Mol Evol* **88**, 537–548. (doi:10.1007/s00239-020-09959-9)
786 70. Chen Z *et al.* 2013 Genomic and Structural Characterization of Kunitz-Type Peptide LmKTT-1a
787 Highlights Diversity and Evolution of Scorpion Potassium Channel Toxins. *PLoS One* **8**, e60201.
788 (doi:10.1371/journal.pone.0060201)
789 71. Fry BG *et al.* 2009 The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal
790 Venoms. *Annu Rev Genomics Hum Genet* **10**, 483–511.
791 (doi:10.1146/annurev.genom.9.081307.164356)
792 72. Boisbouvier J, Albrand J-P, Blackledge M, Jaquinod M, Schweitz H, Lazdunski M, Marion D. 1998 A
793 structural homologue of colipase in black mamba venom revealed by NMR floating disulphide bridge
794 analysis. *J Mol Biol* **283**, 205–219. (doi:10.1006/jmbi.1998.2057)
795 73. Kaser A, Winklmayr M, Lepperdinger G, Kreil G. 2003 The AVIT protein family. *EMBO Rep* **4**, 469–
796 473. (doi:10.1038/sj.embor.embor830)
797 74. Negri L, Lattanzi R, Giannini E, Melchiorri P. 2007 Bv8/Prokineticin proteins and their receptors. *Life*
798 *Sci* **81**, 1103–1116. (doi:10.1016/j.lfs.2007.08.011)
799 75. Wen S *et al.* 2005 Discovery of an MIT-like atracotoxin family: Spider venom peptides that share
800 sequence homology but not pharmacological properties with AVIT family proteins. *Peptides* **26**,
801 2412–2426. (doi:10.1016/j.peptides.2005.05.012)
802 76. Hecht O, van Nuland NA, Schleinkofer K, Dingley AJ, Bruhn H, Leippe M, Grötzinger J. 2004 Solution
803 Structure of the Pore-forming Protein of *Entamoeba histolytica*. *J Biol Chem* **279**, 17834–17841.
804 (doi:10.1074/jbc.M312978200)

805 77. Tosi T, Nickerson NN, Mollica L, Jensen MR, Blackledge M, Baron B, England P, Pugsley AP, Dessen
806 A. 2011 Pilotin-secretin recognition in the type II secretion system of *Klebsiella oxytoca*. *Mol*
807 *Microbiol* **82**, 1422–1432. (doi:10.1111/j.1365-2958.2011.07896.x)
808 78. Korotkov K V., Hol WGJ. 2013 Crystal structure of the pilotin from the enterohemorrhagic
809 *Escherichia coli* type II secretion system. *J Struct Biol* **182**, 186–191. (doi:10.1016/j.jsb.2013.02.013)
810 79. Krishnarjuna B, Sunanda P, Villegas-Moreno J, Csoti A, A.V. Morales R, Wai DCC, Panyi G, Prentis P,
811 Norton RS. 2021 A disulfide-stabilised helical hairpin fold in acrorhagin I: An emerging structural
812 motif in peptide toxins. *J Struct Biol* **213**, 107692. (doi:10.1016/j.jsb.2020.107692)
813 80. Honma T, Minagawa S, Nagai H, Ishida M, Nagashima Y, Shiomi K. 2005 Novel peptide toxins from
814 acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon* **46**, 768–774.
815 (doi:10.1016/j.toxicon.2005.08.003)
816 81. Sachkova MY, Landau M, Surm JM, Macrander J, Singer SA, Reitzel AM, et al. 2020. Toxin-like
817 neuropeptides in the sea anemone *Nematostella* unravel recruitment from the nervous system to
818 venom. *PNAS* **117**, 27481–92. (doi: 10.1073/pnas.2011120117)
819 82. Anderluh G, Lakey JH. 2008 Disparate proteins use similar architectures to damage membranes.
820 *Trends Biochem Sci* **33**, 482–490. (doi:10.1016/j.tibs.2008.07.004)
821 83. Morando MA, Venturella F, Sollazzo M, Monaca E, Sabbatella R, Vetri V, Passantino R, Pastore A,
822 Alfano C. 2022 Solution structure of recombinant Pvfp-5β reveals insights into mussel adhesion.
823 *Commun Biol* **5**, 1–10. (doi:10.1038/s42003-022-03699-w)
824 84. Yeung SYM, Thompson D, Wang Z, Fedida D, Robertson B. 2005 Modulation of Kv3 subfamily
825 potassium currents by the sea anemone toxin BDS: Significance for CNS and biophysical studies. *J*
826 *Neurosci* **25**, 8735–8745. (doi:10.1523/JNEUROSCI.2119-05.2005)
827 85. Sher D, Fishman Y, Melamed-Book N, Zhang M, Zlotkin E. 2008 Osmotically driven prey
828 disintegration in the gastrovascular cavity of the green hydra by a pore-forming protein. *The FASEB*
829 *Journal* **22**, 207–214. (doi:10.1096/fj.07-9133com)
830 86. Macrander J, Daly M. 2016 Evolution of the cytolytic pore-forming proteins (Actinoporins) in sea
831 anemones. *Toxins* **8**, 1–16. (doi:10.3390/toxins8120368)

832 87. Darmoise A, Maschmeyer P, Winau F. 2010 The Immunological Functions of Saposins. *Adv Immunol*
833 **105**, 25–62. (doi:10.1016/S0065-2776(10)05002-9)
834 88. Ahn VE, Leyko P, Alattia J-R, Chen L, Privé GG. 2006 Crystal structures of saposins A and C. *Protein*
835 *Sci* **15**, 1849–1857. (doi:10.1110/ps.062256606)
836 89. Munford RS, Sheppard PO, O'Hara PJ. 1995 Saposin-like proteins (SAPLIP) carry out diverse functions
837 on a common backbone structure. *J Lipid Res* **36**, 1653–1663.
838 90. Bruhn H. 2005 A short guided tour through functional and structural features of saposin-like
839 proteins. *Biochem J* **389**, 249–257. (doi:10.1042/BJ20050051)
840 91. Herbst R, Marciano-Cabral F, Leippe M. 2004 Antimicrobial and pore-forming peptides of free-living
841 and potentially highly pathogenic *Naegleria fowleri* are released from the same precursor molecule.
842 *Journal of Biological Chemistry* **279**, 25955–25958. (doi:10.1074/jbc.M401965200)
843 92. Hill CH, Read RJ, Deane JE. 2015 Structure of human saposin A at lysosomal pH. *Acta Crystallogr.*
844 *Sect F* **71**, 895–900. (doi:10.1107/S2053230X15008584)
845 93. Hawkins CA, De Alba E, Tjandra N. 2005 Solution structure of human saposin C in a detergent
846 environment. *J Mol Biol* **346**, 1381–1392. (doi:10.1016/j.jmb.2004.12.045)
847 94. Alattia JR, Shaw JE, Yip CM, Privé GG. 2006 Direct Visualization of Saposin Remodelling of Lipid
848 Bilayers. *J Mol Biol* **362**, 943–953. (doi:10.1016/j.jmb.2006.08.009)
849 95. Popovic K, Holyoake J, Pomès R, Privé GG. 2012 Structure of saposin A lipoprotein discs. *PNAS* **109**,
850 2908–2912. (doi:10.1073/pnas.1115743109)
851 96. Leippe M, Bruhn H, Hecht O, Grotzinger J. 2005 Ancient weapons: the three-dimensional structure
852 of amoebapore A. *Trends Parasitol* **21**, 5–7. (doi:10.1016/j.pt.2004.10.009)
853 97. Mysliwy J, Dingley AJ, Stanisak M, Jung S, Lorenzen I, Roeder T, Leippe M. 2010 Caenopore-5: The
854 three-dimensional structure of an antimicrobial protein from. *Dev Comp Immunol* **34**, 323-30 (doi:
855 10.1016/j.dci.2009.11.00
856 98. Liepinsh E, Andersson M, Ruysschaert JM, Otting G. 1997 Saposin fold revealed by the NMR
857 structure of NK-lysin. *Nat Struct Biol* **4**, 793–795. (doi:10.1038/nsb1097-793)
858 99. Anderson DH, Sawaya MR, Cascio D, Ernst W, Modlin R, Krensky A, Eisenberg D. 2003 Granulysin

859 Crystal Structure and a Structure-derived Lytic Mechanism. *J Mol Biol* **325**, 355–365.
860 (doi:10.1016/S0022-2836(02)01234-2)
861 100. Chen L, Dürr KL, Gouaux E. 2014 X-ray structures of AMPA receptor-cone snail toxin
862 complexes illuminate activation mechanism. *Science* **345**, 1021–1026.
863 (doi:10.1126/science.1258409)
864 101. Hackney CM *et al.* 2023 A previously unrecognized superfamily of macro-conotoxins
865 includes an inhibitor of the sensory neuron calcium channel Cav2.3. *PLoS Biol* **21**, e3002217.
866 (doi:10.1371/journal.pbio.3002217)
867 102. Olmeda B, García-Álvarez B, Pérez-Gil J. 2013 Structure-function correlations of pulmonary
868 surfactant protein SP-B and the saposin-like family of proteins. EBJ 42, 209–222.
869 (doi:10.1007/s00249-012-0858-9)
870 103. Olmeda B, García-Álvarez B, Gómez MJ, Martínez-Calle M, Cruz A, Pérez-Gil J. 2015 A model for
871 the structure and mechanism of action of pulmonary surfactant protein B. *FASEB Journal* **29**, 4236–
872 4247. (doi:10.1096/fj.15-273458)
873 104. Garrido-Arandia M, Cuevas-Zuviría B, Díaz-Perales A, Pacios LF. 2018 A comparative study of
874 human saposins. *Molecules* **23**, 422. (doi:10.3390/molecules23020422)
875 105. Yin W *et al.* 2016 Evolutionary trajectories of snake genes and genomes revealed by comparative
876 analyses of five-pacer viper. *Nat Commun* **7**, 13107. (doi:10.1038/ncomms13107)
877 106. Maxwell M, Undheim EAB, Mobli M. 2018 Secreted cysteine-rich repeat proteins 'SCREPs': A novel
878 multi-domain architecture. *Front Pharmacol* **9**, 1333. (doi:10.3389/fphar.2018.01333)
879 107. Gerdol M, Cervelli M, Mariottini P, Oliverio M, Dutertre S, Modica MV. 2019 A recurrent motif:
880 Diversity and evolution of ShKT domain containing proteins in the vampire snail *Cumia reticulata*.
881 *Toxins* **11**. (doi:10.3390/toxins11020106)
882 108. Berking S, Herrmann K. 2006 Formation and discharge of nematocysts is controlled by a proton
883 gradient across the cyst membrane. *Helgol Mar Res* **60**, 180–188. (doi:10.1007/s10152-005-0019-
884 y)
885 109. Fabricius KE, Dommisse M. 2000. Depletion of suspended particulated matter over coastal reef

886 communities dominated by zooxanthellate soft corals. *MEPS* **196**, 157-167
887 (doi:10.3354/meps196157)
888 110. Sebens KP, Miles JS. 1988 Sweeper Tentacles in a Gorgonian Octocoral: Morphological
889 Modifications for Interference Competition. *Biol Bull* **175**, 378–387. (doi:10.2307/1541729)
890 111. Dutertre S, Jin AH, Kaas Q, Jones A, Alewood PF, Lewis RJ. 2013 Deep venomics reveals the
891 mechanism for expanded peptide diversity in cone snail venom. *Mol Cell Proteomics* **12**, 312–329.
892 (doi:10.1074/mcp.M112.021469)
893