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 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, 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 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 allied, 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 This study endeavours to address significant eco-evolutionary questions concerning Octocorallia biology and
76 evolution. Our first aims were to ascertain the capacity of octocorals for venom production, investigate its
77 composition, and compare it to other Cnidaria. To achieve these objectives, we utilized an integrated
78 proteo-transcriptomic approach, which employed whole colony transcriptomic data to match proteomic
79 data obtained from chemical-induced nematocysts discharge or from the whole body of polyps. Our
80 analyses unveiled several previously unreported toxin families, providing the first comprehensive
81 description of the venom of an octocoral species. Leveraging these findings, we aimed to assess the degree
82 of conservation of specific toxin folds of Octocorallia in comparison to other Anthozoa, and to elucidate how
83 venom toxin families diversified throughout the evolutionary radiation of octocorals. Employing a
84 comparative approach on available Octocorallia genomes, we reconstructed the evolutionary patterns of
85 the main toxin families, highlighting distinctive evolutionary innovations specific to Octocorallia. Ultimately,
86 by complementing our analyses with untemplated, deep-learning assisted structural modelling of the most
87 relevant toxins, we attempted to infer their function and adaptive value.
88

89 **Methods**

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

91 Colonies of *Eunicella singularis* were collected by scuba diving at a depth of 25-30 m around Giannutri Island,
92 in the Tuscany Archipelago (Italy) and kept alive in seawater until further processing. Polyps were dissected
93 from the colony under a stereomicroscope and pooled in three aliquots of about 20 each (about 15 mg of

94 tissue) that were preserved in Trizol at -80 °C until further processing. Total RNA was extracted using Trizol
95 (Thermo Fisher Scientific), following the manufacturer's instructions, obtaining 5-10,000 ng with RQN values
96 between 6.7 and 8.8.
97 Additionally, 10 live colony fragments about 2.5 cm long were immersed for 30'' into 1 ml of EtOH in a 1.5
98 ml tube to induce nematocyst discharge [29]. Samples were freeze-dried and resuspended in 100 µg milliQ
99 water for proteomic analysis.
100
101 **(b) RNASeq**
102 To account for variations in venom expression among samples, three cDNA libraries for RNASeq were
103 prepared using an Illumina TruSeq Stranded mRNA Sample Preparation kit according to manufacturer's
104 instructions and validated qualitatively and quantitatively on a Fragment Analyser, as well as by qPCR on a
105 ROCHE LightCycler 480. The three libraries were then processed for paired-end sequencing-by-synthesis in
106 an Illumina HiSeq 2500 at the MGX facility (Montpellier, France), and quality controlled using FastQC. After
107 trimming, we obtained a number of reads longer than 30 bp and quality above Q20 comprised between 21
108 and 30 million.
109
110 **(c) Bioinformatic analyses of transcriptomic data**
111 Paired-end raw trimmed sequencing data obtained from the three libraries were pooled and *de novo*
112 assembled using Trinity v.2.5.1 [30] with default settings, allowing a minimum contig length of 200
113 nucleotides. The assembled transcriptome was virtually translated using TransDecoder v.5.01
114 (https://github.com/TransDecoder/), setting the minimum allowed ORF length to 100 codons.
115 Expression levels of toxin transcripts were calculated based on unique gene counts as Transcripts Per Million
116 (TPM), as this normalised unit allows an efficient and reliable comparison of gene expression levels both
117 within and between samples [31]. Briefly, trimmed reads from each library were separately mapped against
118 the transcriptome assembly, previously filtered to remove sequence redundancy by considering only the
119 longest transcript for each gene model.
120

121 **(d) Protein digestion**

122 Protein extracts were denatured, reduced and alkylated. Briefly, the volume of each sample was adjusted
123 to 100 µL with 100 mM triethylammonium bicarbonate (TEABC). 1M DTT (1 µL) was added to each sample,
124 which was then incubated 30 min at 60 °C before adding 10 µL 0.5 M Iodoacetoamide (IAA). Alkylation lasted
125 30 min in the dark. Enzymatic digestion was performed by adding 2 µg trypsin (Gold, Promega, Madison
126 USA) in TEABC 100 mM and incubating overnight at 30 °C. Then, peptides were purified and concentrated
127 using OMIX Tips C18 reverse-phase resin (Agilent Technologies Inc.) according to the manufacturer's
128 specifications. Peptides samples were finally dehydrated in a vacuum centrifuge.

129

130 **(e) Shotgun proteomics**

131 Samples were resuspended in 20 µL of 0.1% formic acid (buffer A) and 1 µL was loaded onto an analytical
132 50 cm reversed-phase column (0.75 mm i.d., Acclaim Pepmap 100® C18, Thermo Fisher Scientific) and
133 separated with an Ultimate 3000 RSLC system (Thermo Fisher Scientific) coupled to a Q Exactive HF (Thermo
134 Fisher Scientific) *via* a nano-electrospray source, using a 120-min gradient of 2 to 40% of buffer B (80% ACN,
135 0.1% formic acid) and a flow rate of 300 nl/min. MS/MS analyses were performed in a data-dependent
136 mode. Full scans (375 – 1,500 m/z) were acquired in the Orbitrap mass analyser with a 60,000 resolution at
137 200 m/z. For the full scans, 3 x 10^6 ions were accumulated within a maximum injection time of 60 ms and
138 detected in the Orbitrap analyser. The twelve most intense ions with charge states ≥ 2 were sequentially
139 isolated to a target value of 1 x 10^5 with a maximum injection time of 45 ms and fragmented by HCD (Higher-
140 energy collisional dissociation) in the collision cell (normalised collision energy of 28%) and detected in the
141 Orbitrap analyser at a resolution of 30,000.

142

143 **(f) Bioinformatic analyses of proteomic data**

144 PEAKS Studio 8.5 (Bioinformatics solutions, Waterloo, ON, Canada), a de novo assisted database software
145 [32-34] was used to analyse MS/MS data from *E. singularis* venom, with parameters set as in [35]. The
146 MS/MS spectra obtained from shotgun proteomics were matched to a custom database resulting from the
147 whole-body transcriptome of the same colony, translated in the six open reading frames.

148 Protein isoforms were considered only when their existence was supported by MS/MS fragmentation data,
149 whereas translated transcripts identified by PEAKS Studio as belonging to the same protein group were
150 clustered with CD-HIT [36] using an identity threshold of 90%. Only the most representative sequence was
151 retained for the subsequent annotation.
152 Protein annotation was performed with HMMER v3.1b2 (www.hmmer.org), as detailed in Met.S1. Toxin
153 nomenclature of mature polypeptides followed Oliveira et al. [37]
154 Structure predictions of mature toxins were performed with the colabfold implementation of AlphaFold2,
155 using MMSeqs2 for multiple sequence alignment generation [38,39]. To increase model confidence for the
156 most challenging targets, custom multiple sequence alignments with orthologue octocorallian sequences
157 were used as inputs and each model was built using 12 recycles. All predictions were performed without the
158 use of templates and energy minimization. The structures that were modelled with reasonable confidence
159 (average pLDDT ≥ 70, which is generally recognised as the threshold for reliable models)[40] were submitted
160 to the DALI [41] and RUPEE [42] servers for structural homology search against the Protein Data Bank.
161 Homology modelling of the saposin-like toxins in the closed conformation was performed with Modeller
162 v9.24 [43]. Electrostatic surfaces were calculated with APBS [44]. UCSF ChimeraX was used for model
163 analysis, visualization and graphics [45].
164
165 **(g) Comparative genomic analyses**
166 Toxins identified in *Eunicella singularis* were used as queries for sequence homology searches against a
167 selection of available anthozoan genomes, representative of the diversity of Octocorallia [46] and major
168 Hexacorallia lineages summarized in Tab. S1.
169 Sequence homology searches were carried out with tBLAStn [47], based on an initial E-value significance
170 threshold arbitrarily set to 10. Potential hits were manually curated, by annotating the inferred CDS,
171 detecting putative splicing donor and acceptor sites with the aid of Genie [48]. The presence of *bona fide*
172 orthologues between *E. singularis* and the other species was assessed through combined observations, i.e.
173 (i) the conservation of the main topological feature of the encoded precursor proteins, such as the
174 identification of a signal peptide and the cysteine residues involved in the formation of the disulphide bonds

175 (whenever present); (ii) the shared presence of conserved protein domains, detected with a HMM profile
176 approach, based on the Pfam-A database; (iii) the conservation of gene architecture (e.g. intron number,
177 phase and position), with respect to the orthologous gene from the congeneric species *E. verrucosa*.
178 Whenever needed (i.e., U-GRTX-Esi20 and 21), a fine-scale classification of orthologous sequences was aided
179 by Maximum Likelihood phylogenetic inference analyses, carried out with IQ-Tree [49] with 1000 ultrafast
180 bootstrap replicates based on the best-fitting model of molecular evolution detected by ModelFinder [50].
181

182 **RESULTS AND DISCUSSION**

183

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

185 Proteomic analyses revealed a complex profile in the venom of *Eunicella singularis:* 150 translated
186 transcripts were confirmed by MS/MS analysis, and subsequently reduced to 108 non-redundant precursor
187 protein sequences, which defined the nematocysts proteome (NEM-P, Tab. S2). Based on an extensive
188 bioinformatic annotation, 33 of these sequences were associated to structural or basal physiology functions,
189 15 were identified as enzymes, 28 did not display significant homology to known proteins in the databases
190 employed and were therefore classified as unknown and 32 were identified as venom components (Fig. 1A).
191 Only 17 among the sequences annotated as venom components in the NEM-P could be also detected in the
192 whole body proteome (WB-P) of *E. singularis,* which, by comparison, comprised over 1500 translated
193 transcripts. To identify venom components produced by the ectodermal glands, that could be absent in the
194 nematocysts, we also performed a targeted HMM search in the WB-P, looking for known cnidarian toxin
195 families, as defined by VenomZone (https://venomzone.expasy.org). This search produced only four
196 sequences previously undetected in the NEM-P: a neurotoxin of the Small Cysteine Rich Proteins (SCRiPs)
197 family and three cytolysins (see below). The NEM-P comprised some transcripts encoding precursors of
198 multiple mature toxins, which were further classified by predicted fold (Fig. 1B). In total, we identified 45
199 mature toxins, including a calglandulin, a protein characterized by multiple EF-hand motifs with an accessory
200 role in venom production and processing, omitted from the detailed discussion below.

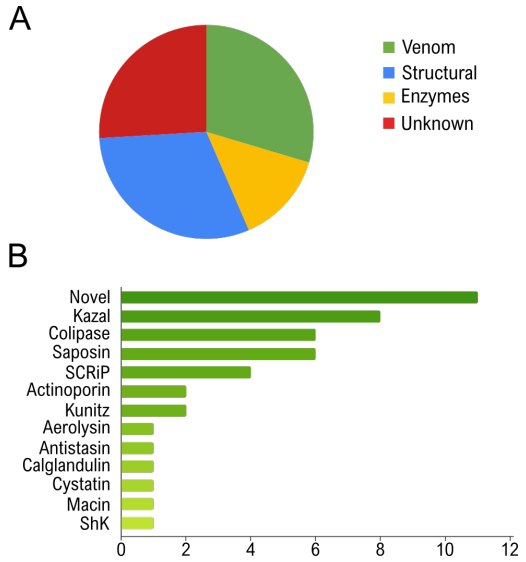


Figure 1. Structural and functional annotation of the venom from *E. singularis.* (A) Functional annotation of the translated transcripts confirmed at the protein level in the NEM-P; (B) Fold distribution of the mature toxins produced by *E. singularis* based on HMM profile search; the x axis represents the number of proteomics-validated transcripts.

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 an own family for this genus (Eunicellidae) [46].

(b) Established toxins folds are present in the venom of *E. singularis*

We retrieved numerous familiar toxins folds within the venom proteome of *E. singularis*, notably Small Cysteine Rich Proteins (SCRiPs), ShK and Kunitz domain containing proteins, macins and colipases.

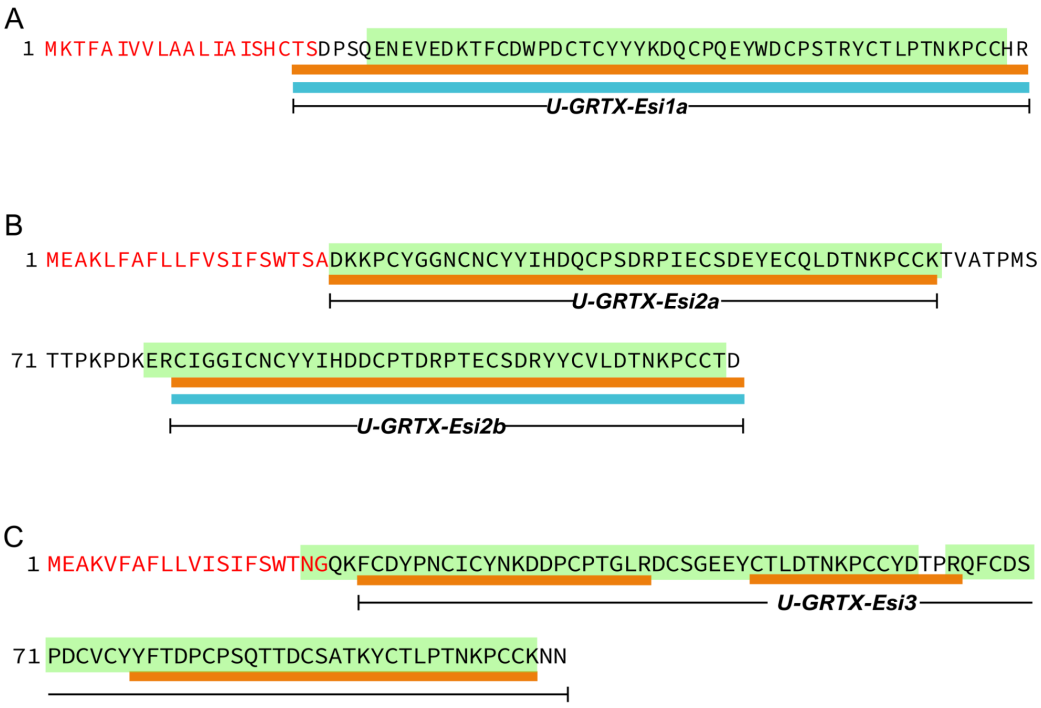
SCRiPs are known to induce severe neurotoxicity in zebrafish larvae, possess antimicrobial activity and potentiate the transient receptor potential ankyrin 1 (TRPA1) in mouse [51,52]. Four toxins with homology to SCRiPs were identified: U-GRTX-Esi1 (particularly abundant at the transcript level, TPM 1585), 2a-b and 3 (Fig. 2). 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 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, implying different maturation pathways of the proprotein in different cell types (Fig. 2B).

223 U-GRTX-Esi3, solely detectable in the WB-P, featured a more complex architecture, comprising two SCRiPs
224 domains in the mature polypeptide, confirmed by the proteomic coverage (Fig. 2C). The absence of
225 predicted cleavage sites between these domains would qualify U-GRTX-Esi3 as a Secreted Cysteine-Rich
226 Repeat Protein (SCREP), an emerging typology of venom proteins found in various organisms, often with
227 high affinity and avidity towards their biological target and a multi-modal action [53].

228

229



230

231 **Figure 2.** Tissue-specific maturation patterns of the SCRiPs from *E. singularis.* The sequences of the precursors of U-
232 GRTX-Esi1 (A), 2a-b (B) and 3 (C) are reported. Red residues correspond to predicted signal peptide, green boxes define
233 the SCRiP domains, orange and blue lines indicate the peptide coverage of the precursors in the WB-P and NEM-P,
234 respectively.

235

236

237 Although SCRiPs-encoding genes were identified, often in large numbers and associated with tandemly
238 duplicated gene clusters, in the genomes of all the octocoral species investigated in this study, they have a
239 patchy occurrence in Hexacorallia, supporting a complex evolutionary history characterized by lineage-
240 specific expansions and rapid diversification [53]. It is unclear whether SCREPs with two consecutive SCRiP
241 domains were already present in the latest common ancestor of all Anthozoans, or independently acquired
242 in different lineages. Indeed, clear orthologues of U-GRTX-Esi3 could be identified only in *P. clavata,*

243 *Trachytela* sp. and *R. reniformis*: this architecture had been previously identified also in the hexacoral
244 *Orbicella faveolata* [51]. However, while the only SCRiP so far reported in Octocorallia (in *Scleronephthya*
245 *gracillima*) presented a conventional cysteine framework [53], the domains here identified display a shift in
246 the cysteine pattern, which nonetheless did not affect the overall protein fold (Fig. S1).
247 The ShK toxins, first identified in the venom of the sea anemone *Stichodactyla helianthus*, are potent K_V1
248 channels blockers [54]. Our search for ShK domains in the proteomes of *E. singularis* led to the identification
249 of U-GRTX-Esi4 in the NEM-P (TPM 103), with three consecutive ShK domains with peculiar characteristics
250 (Suppl. Fig. 2).
251 Toxins containing domains associated with protease inhibition are widespread across the tree of life. Among
252 them, Kunitz-like peptides are frequently found in venoms [55] where they can act as inhibitors of several
253 endogenous and exogenous proteases [56] but also show neurotoxic and neuromodulatory activity on a
254 variety of receptors and ion channels [57,58], sometimes with multimodal action [56,58]. Kunitz-like toxins
255 that bind ion channels can be used to paralyze prey and are being extensively studied for their analgesic and
256 neuroprotective properties [59-62]. Our HMM search revealed the occurrence of two Kunitz-like toxins. PI-
257 GRTX-Esi5, detectable in the NEM-P only, presented two Kunitz domains separated only by an Asp-Ile
258 dipeptide, an architecture previously detected in salivary proteins of blood-feeding animals [63,64]. Since
259 orthologous genes sharing the same 3 exons/2 introns organization of PI-GRTX-Esi5 were not detected in
260 Scleralcyonacea, this domain combination most likely represents a recent innovation of the
261 Malacalcyonacea. The second toxin, PI-GRTX-Esi6, with a single Kunitz domain, was detected both in the
262 NEM-P and in the WB-P, and derived from a precursor with a complex architecture, with the Kunitz-like
263 domain followed by three Kazal-type domains, all separated by propeptidase cleavage sites. Also in this case,
264 proteomic evidence indicated a differential maturation pattern of the precursor, with the Kazal-like peptides
265 from this precursor only retrieved in the WB-P (Tab. S2). The Kunitz domain of PI-GRTX-Esi5 presented the
266 Arg residue (conventionally "P1") associated with the trypsin inhibition, whereas its replacement with a Met
267 in both domains of PI-GRTX-Esi6 suggests a chymotrypsin specificity [65]. In addition, PI-GRTX-Esi6
268 presented a Leu residue aligning with the crucial position for K_V1 binding of LmKTT, a scorpion Kunitz-like
269 toxin [66]. These similarities may hint to multimodal activity for PI-GRTX-Esi5, while the lack of key residue

270 conservation in PI-GRTX-Esi6 does not allow inferring any putative function (Suppl. Fig. 3)

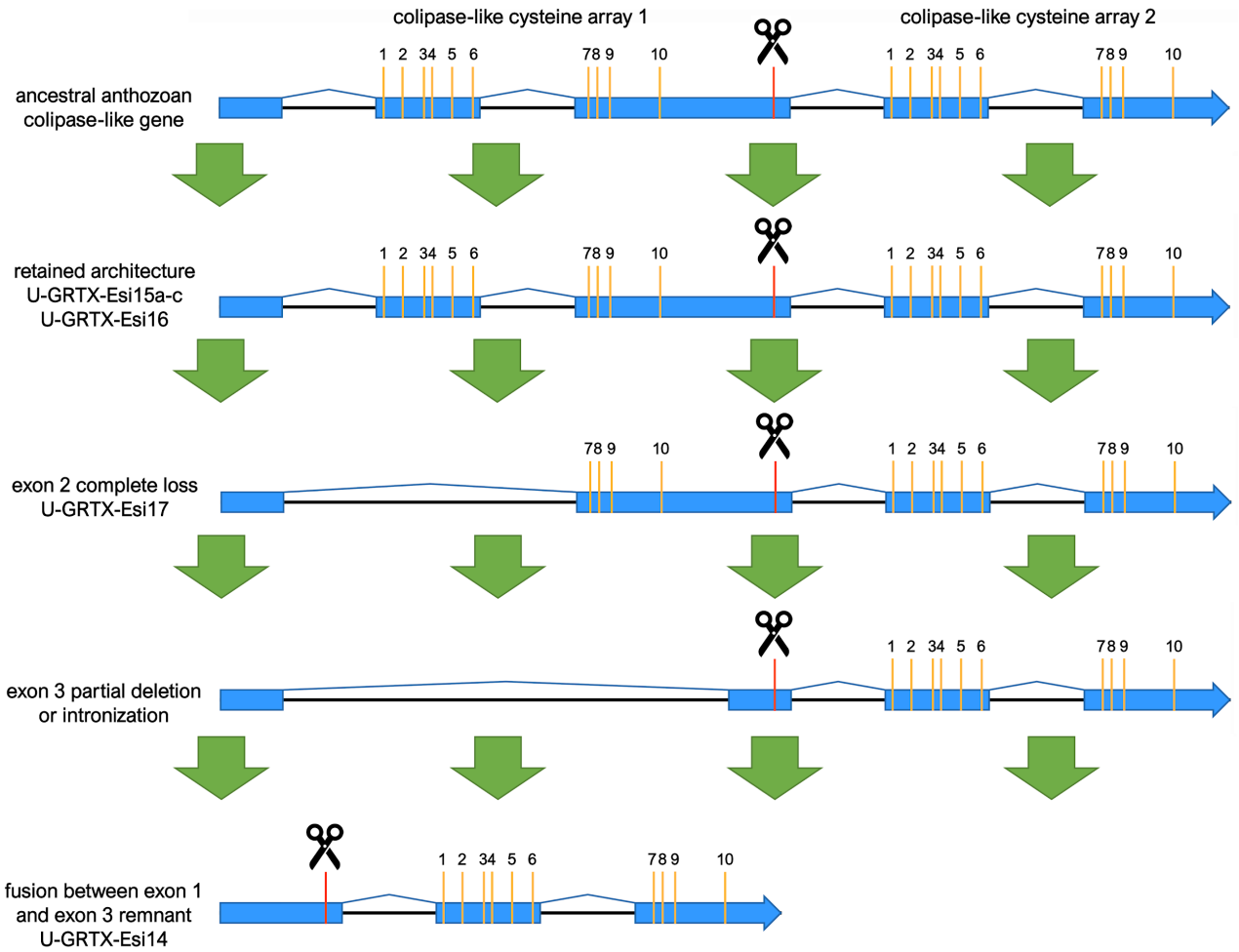
271 Other peptides present in *E. singularis* nematocysts and displaying protease inhibitory domains, but likely
272 lacking a toxin function (Kazal-type, cystatines, antistasins, and macins) are described in detail in Data.S1.

273 Finally, a striking feature of the venom from *E. singularis* was the abundance of toxins with a predicted
274 colipase fold (U-GRTX-Esi14-18). Colipases are knottins containing five disulphide bridges, found in spider,
275 snake and tick venoms [67], but to date not in Cnidaria. The prototypical example, MIT-1 [68] interacts
276 through its N-terminal pentapeptide AVITG with prokineticin receptors, stimulating smooth muscle
277 contractions, and exhibiting hyperalgesic effects [69,70]. *E. singularis* colipases lacked the AVITG
278 pentapeptide, a feature shared with the atracotoxins of the spider *Hadronyche versuta,* whose
279 pharmacological target is unknown to date [71].

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

294 The loss of exon 2 in the gene encoding for U-GRTX-Esi17 results in a shorter precursor protein, with an N-
295 terminal incomplete colipase domain. The maturation of this precursor leads to the release of a polypeptide
296 lacking six out of the ten cysteines (U-GRTX-Esi18), whose presence was also confirmed in the NEM-P. The

297 lack of most of the cysteine scaffold has a major impact on the protein structure: the model predicted with
298 good confidence by AlphaFold2 shows in fact only a long β-hairpin flanked by disulphide stabilized tails, that
299 might have acquired peculiar biological activity.
300 Notably, the transcripts for all these toxins had exceptionally high TPM values (1806, 569, 826 and 429,
301 respectively for the U-GRTX-Esi14 to 17/18), which support their relevance for the predatory and/or
302 defensive strategies of *Eunicella singularis.*
303
304



305
306 **Figure 3.** Schematic representation of the evolution of the colipase genes in Anthozoa, with references to specific *E.*
307 *singularis* genes. Exons are in blue, cysteine residues are indicated by the yellow bars, scissors indicate cleavage sites..
308
309
310 **(c) Novel putative toxins were retrieved in the venom of *Eunicella singularis***

311 We identified in the NEM-P 10 additional proteins possessing the typical sequence features of toxins (small
312 size, abundance of Cys, presence of a signal peptide), lacking sequence similarity with any known protein
313 and often encoded by highly expressed transcripts. Their genomic organization is summarized in Suppl. Fig.
314 4. High-confidence structural models obtained with Alphafold2 were used to retrieve structural homologues
315 in the Brookheaven Protein Data Bank, in the attempt to infer their biological functions.
316 Among them, U-GRTX-Esi19 was encoded by an extremely abundant transcript (TPM 2898). The mature
317 toxin (100 amino acids, 6 Cys) was detected in both the NEM-P and WB-P and had a weak homology at the
318 C-terminus with Amoebapore A, the major pore-forming protein of *Entamoeba histolytica*, suggestive of a
319 possible cytolytic function [72]. Most of the protein structure was modeled with high confidence (pLDDT >
320 80, Fig. 4A), highlighting a compact fold characterized by a knotted four-helix bundle stabilized by three
321 inter-helices disulphide bridges. A similar arrangement, although devoid of disulphide stabilization, is typical
322 of bacterial pilotins PulS and GspS, chaperone lipoproteins that play a role in the assembly of the T2SS
323 secretion system in pathogenic Gram-negative bacteria [73,74]. This toxin appears as an evolutionary
324 innovation of Malacalcyonacea, although orthologues are missing in *Phenganax subtilis* and in *Xenia* sp. The
325 cysteine framework is always conserved, while the length of some loops varies across species. The
326 abundance of this protein suggests a strong functional relevance, but activity assays will be necessary to
327 clarify its function.
328 The two toxins U-GRTX-Esi20 and 21 did not show any obvious sequence similarity, with a pairwise identity
329 of only 21%. U-GRTX-Esi20 is 48 amino acids long and derives from a longer propeptide proteolytically
330 cleaved at the C-terminus, whereas U-GRTX-Esi21 displays a shorter sequence with a very high cysteine
331 density (8 out of 36 amino acids). The models obtained (Fig. 4B and C) highlighted a structural relationship,
332 as both presented a disulphide-stabilized helical hairpin motif, recently characterized in acrorhagin-1 from
333 *Actinia equina* (U-AITX-Aeq5a) [75], a toxin from the acrorhagi of *A. equina* used in conspecific aggressive
334 encounters [76]. Acrorhagin-1 is able to induce local tissue necrosis and displays toxicity on crabs, but its
335 mechanism of action has not been clarified, although it has been demonstrated that the toxin does not
336 interact with either potassium or sodium ion channels and it does not possess membranolytic activity
337 [75,76]. U-GRTX-Esi20 and 21 are stabilized by three and four disulphide bridges, respectively, which overlap

338 with those of acrorhagin-1, despite a sequence identity below 20% in both cases [75].

339 Our genomic survey indicates that these two toxins belong to two distinct monophyletic orthogroups within
340 a very large superfamily of cysteine-rich peptides, encoded by ancestrally duplicated paralogous genes with
341 intronless structures, that also include other members in *E. singularis*, not detected in the NEM-P. High-
342 confidence orthologues were exclusively found in Malacalcyonacea, even though, for U-GRTX-Esi21, these
343 were absent in *P. subtilis* and *Xenia*. This suggests a rather recent origin of these orthogroups, which derive
344 from a large group of similar peptides also found in other cnidarians, worthy of more extensive investigation.
345 The next two toxins, U-GRTX-Esi22 (60 amino acids, 10 cysteines, Fig. 4D) and 23 (45 amino acids, 8 cysteines,
346 Fig. 4E) did not present any sequence homology with known proteins. In both cases, the models highlighted
347 an unusual disulphide-stabilized αβ fold, suggesting the existence of peculiar and to date undiscovered
348 cnidarian toxin folds. The gene encoding U-GRTX-Esi23 is present with two similar paralogous genes in *E.*
349 *singularis* and has clearly detectable orthologs in all Malacalcyonacea except *P. subtilis* and *Xenia* sp. A
350 divergent paralogue, with a slightly different cysteine pattern, is found in Scleralcyonacea. This toxin most
351 likely constitutes an evolutionary innovation of Malacalcyonacea, but derives from an ancestral sequence
352 with a similar cysteine connectivity that was already present in the latest common ancestor of all
353 Octocorallia. Conversely, U-GRTX-Esi22 presents a single-copy gene, with a detectable orthologue only in *P.*
354 *clavata* and *M. muricata,* and is probably a recent innovation, only shared by a few related orders within
355 Malacalcyonacea.

356 The structure of U-GRTX-Esi24, a 45 amino acids protein with 8 cysteines and a 50% sequence identity with
357 Conotoxin Cltx-2 from *Conus californicus,* could also be modelled with acceptable confidence (Fig. 4F).
358 Despite undetected by HMM profile analysis, the structure presented a SCRiP-like W-shaped fold, indicative
359 of a potential neurotoxic function. This domain has a broad but scattered distribution in Octocorallia,
360 suggesting an evolutionary origin slightly older than most other toxins described in this work. Nevertheless,
361 it is most often found repeated multiple times within the same protein (also in *E. singularis*), with several
362 independent occurrences of single-domain peptides scattered across the tree of life.

363

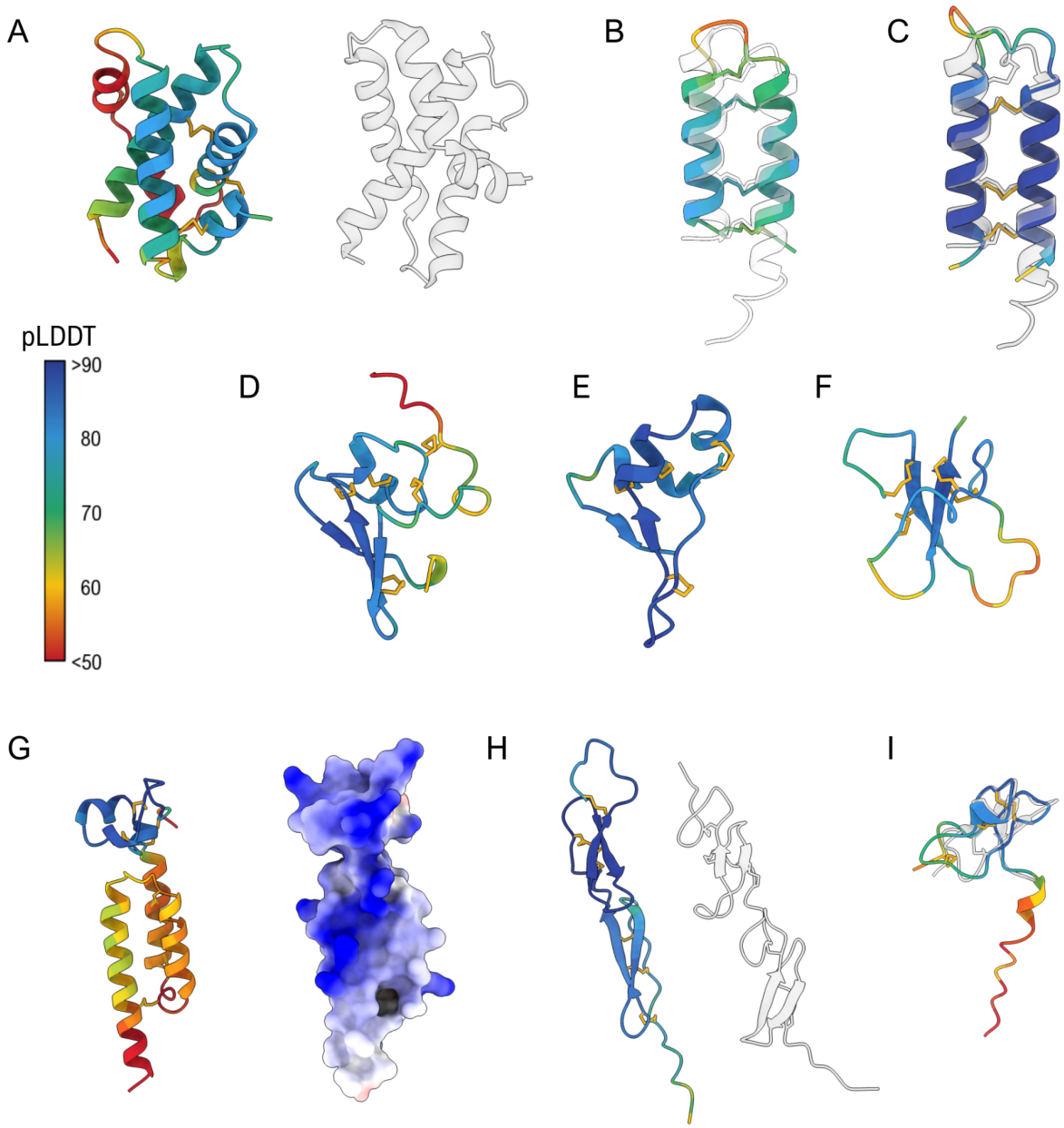


Figure 4. Overview of the structures predicted by AlphaFold2 for the novel toxins from the venom of *E. singularis.* Models are coloured according to pLDDT, with higher values (>70) indicating reliable prediction. (A) The putative cytolytic U-GRTX-Esi19 and, on the right, PulS from *K. oxytoca* (PDB 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) U-GRTX-Esi28 overlying the carboxypeptidase inhibitors from *Solanum tuberosum* (light gray, PDB 7EQZ).

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.

379 The two domains were clearly distinguishable in the three-dimensional prediction (Fig. 4G), where they
380 appeared as directly connected, without spacer sequences or putative cleavage sites. The unknown C-
381 terminal domain was modelled as an asymmetrical three-helices arrangement, with three disulphide bridges
382 connecting and stabilizing the secondary structure elements. Interestingly, despite the complex architecture
383 predicted, several basic residues from both domains in the model aligned perfectly, forming an extended
384 positively charged patch (Fig. 4G), a feature that suggests a possible membranolytic and/or antimicrobial
385 action [77]. This peculiar combination of domains, paired with a shared 3 exons/2 introns gene architecture,
386 was detected in Malacalcyonacea (except *P. subtilis* and *Xenia* sp). A single orthologous gene was found in
387 *Paragorgia papillata* (Scleralcyonacea), suggesting that the appearance of this protein could predate the
388 split between Malacolcyonacea and Scleralcyonacea.
389 The model predicted for U-GRTX-Esi26 (77, 12 Cys) presented an elongated structure made of two domains:
390 an N-terminal region in which two antiparallel beta strands are connected by three disulphide bridges to a
391 proline-rich stretch, in a defensin-like fold found in other cnidarian toxins, and a C-terminal EGF-like domain
392 (Fig. 4H). A similar modular structure was recently described in the byssal adhesive protein Pvfp-5β from the
393 mussel *Perna viridis* [78]; despite the absence of sequence homology, this structural similarity might indicate
394 for this protein a structural or adhesive function. However, orthologues of the gene were found only in
395 Malacalcyonacea (with the exception of *Dendronephthya* and *Xenia)*, all sharing a common 3 exons/2 introns
396 gene structure. Three divergent paralogues were found In *P. clavata*, probably originated through a species-
397 or genus-specific gene duplication process.
398 The last two novel toxins identified in the NEM-P of *E. singularis,* U-GRTX-Esi27 and 28, were short peptides
399 associated with relatively low transcription levels but a high proteomic coverage. In particular, U-GRTX-Esi27
400 displayed a weak sequence homology to both SCRiPs and BDS toxins from *A. sulcata*, which are potent
401 ligands of K_v3 channels [79]. The mature sequence of U-GRTX-Esi27, a 39 amino acids long peptide,
402 contained 8 cysteines, 6 of which correspond to the BDS framework. An additional cysteine pair occupied
403 the position of two extremely conserved residues in sea anemone BDS toxins, *i.e.*, a tryptophan and a
404 tyrosine residue, otherwise present in all other members of the family (Suppl. Fig. 5). Only low confidence
405 models could be predicted by AlphaFold2, thus preventing function inference. Despite the similarity with

406 BDS, orthologous genes encoding for similar proteins were not detected in the genomes of the other
407 Octocorallia considered in this study except from *E. verrucosa.* U-GRTX-Esi27 might therefore constitute an
408 evolutionary innovation of the genus *Eunicella* or the family Eunicellidae (Fig. 6). Finally, U-GRTX-Esi28
409 possessed a weak sequence homology with the ω-hexatoxin-Hi1g. The modelled structure confirmed this
410 homology, revealing a cystine knotted core (Fig. 4I), with the addition of an extended and disordered C-
411 terminus, possibly proteolitically cleaved. Orthologues coding for this protein are shared by nearly all
412 Malacalcyonacea, but are lacking in *Xenia*. This finding, along with the conservation of the gene structure,
413 in which two exons are spaced by an intron with an equal placement in all species, the signal peptide
414 sequence, and the cleavage sites, strongly indicated that this protein represents an evolutionary innovation
415 of Malacalcyonacea (Fig. 6).
416
417 **(d) A diverse range of cytolytic proteins are produced both inside and outside the nematocysts**
418 Our analysis highlighted the absence, in the NEM-P of *E. singularis*, of the common cnidarian cytolytic
419 proteins. A targeted hmmscan search in the WB-P yielded instead positive hits for profiles associated with
420 both the aerolysin (PF01117) and the full anemone cytotoxin (PF06369) families. Coherently with the
421 approach adopted for the NEM-P annotation, we considered only the proteins encoded by complete
422 transcripts presenting a signal peptide; this approach resulted in the identification of three pore forming
423 toxins (PFTs), characterized by diverse structural features. The first one, Δ-GRTX-Esi29, was an aerolysin-like
424 protein, structurally related to the β-PFT Hydralysin [80]; the other two, Δ-GRTX-Esi30 and 31, which shared
425 ~50% sequence similarity, belonged to the actinoporins family, the most abundant family of cnidarian porins
426 [81]. Interestingly, we observed relevant non-conservative substitutions in *E. singularis* actinoporins,
427 suggestive of a peculiar chemical selectivity towards different lipids (Fig. S6).
428 Finally, the highest scoring match in the proteomic analysis of the nematocysts content of *E. singularis*
429 corresponded to a 699 amino-acid translated transcript with homology to prosaposin (pSAP) [82]. In
430 humans, the maturation of the orthologous pSAP produces four globular saposins (SAP-A/D), which are co-
431 factors of sphingolipid degrading enzymes. The saposin fold is very ancient and characterizes all the saposin-
432 like proteins (SAPLIPs), which include the amoebapores [83-85]. SAPs and SAPLIPs have a distinctive cysteine

433 framework but display minimal conservation (<25% sequence identity). They are often released in several
434 copies from the same precursor gene [86], assuming a closed, monomeric conformation at lysosomal pH
435 [83, 87]. All these proteins are involved in membrane destabilization/solubilization, acting through different
436 mechanisms: lysosomal SAPs undergo a conformational change upon pH shift or in the presence of lipids or
437 detergents, exposing the residues of the hydrophobic core and encapsulating a large number of lipids [88-
438 90]. Other SAPLIPs, like the amoebapores from *E. histolytica* and the caenopores from *C. elegan*s, form
439 oligomeric pores [91-92]. Finally, basic SAPLIPs, like NK-lysin and granulysin, act by "molecular
440 electroporation", inserting within the membranes [93,94]. In rare cases, the saposin fold has also been
441 detected in non-cytolytic neurotoxins, like the macro-conotoxins con-ikot-ikot and Mu8.1 [95,96], but these
442 proteins are highly dissimilar from SAPLIPs in terms of sequence and cysteine pattern and are encoded by
443 single domain (non pSAP-like) transcripts. The six sequences deriving from the maturation of the pSAP
444 precursor of *E. singularis* (Δ-GRTX-Esi32a-f) were all identified at the proteomic level in the NEM-P. These
445 proteins shared a 24%-65% similarity, higher than what generally observed for homologous SAPLIPs, and
446 were more similar to human saposins SAP-A and C (similarity 26-40%) than to antimicrobial porogenic
447 peptides like amoebapore A (16-25%) or caenopore-5 (13-27%).
448 We examined the surface potential distribution of the SAPLIPs from *E. singularis*, since the activity of SAPLIPs
449 correlates to their surface electrostatics [97,98]. Except Δ-GRTX-Esi32a and c, they all shared with human
450 SAP-A extended patches of negative charge (Fig. 5), which hints to a pH-dependent mechanism for activation
451 and membrane interaction [99]. In contrast, Δ-GRTX-Esi32a and c presented a more balanced distribution
452 of basic and acidic regions and extended neutral patches, like caenopore-5, suggesting a different, possibly
453 non-monomeric, native arrangement and a pH-independent mechanism of action, possibly through pore
454 formation.

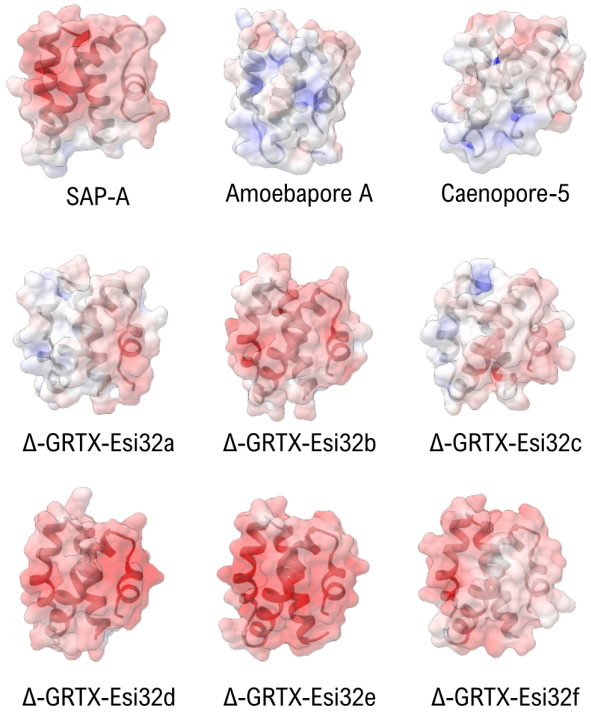


Figure 5. Comparison of the surface electrostatic potentials of the three model saposin-like proteins SAP-A (4UEX), Amoebapore A (1OF9) and Caenopore-5 (2JSA) with the six SAPLIPs identified in the venom of *E. singularis.* The extended patches of negative charge (red) observed in Δ-GRTX-Esi32b, d, e, 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.

Conclusions

Our investigation unequivocally demonstrated that Octocorallia do produce venom. The nematocyst venom of the gorgonian *Eunicella singularis* contains some toxins with folds previously identified in cnidarian venoms, others with structural similarities to venom components of unrelated organisms, and a few with completely new folds. The scarce similarity observed between octocoral and hexacoral venom composition is not surprising, given the ancient divergence between these two lineages that recent works dated about 500-700 MYA [2]. As comparison, Elapidae and Viperidae snakes, which present remarkable differences in their venom composition, have diverged only 50 MYA [100]. However, our results clearly indicate that the nematocyst venom of *E. singularis* has an elevated content of neurotoxic peptides, a feature previously reported for sea anemones, while cytolytic toxins with different structures (and, likely, specificity/function) are distributed both inside and outside the nematocysts.

The transcriptionally more abundant components of *E. singularis* venom are two novel toxins, the potentially cytolytic U-GRTX-Esi19 and the acrorhagin-like U-GRTX-Esi20, the six colipases, the SCRiPs U-

475 GRTX-Esi1a and the SCREP U-GRTX-Esi4 with three ShK domains. With the exception of the SCRiPs, which,
476 however, are characterized by sequence peculiarities, these toxins represent evolutionary innovations of
477 Octocorallia, either shared by the entire lineage (e.g., some colipases), or restricted to the Malacalcyonacea
478 (e.g., U-GRTX-Esi19), or even exclusive of a group of sublineages within the Malacalcyonacea (e.g., U-GRTX-
479 Esi20) (Fig. 6).
480 Peptide neurotoxins in *E. singularis* included both novel and established cnidarian folds, the latter often
481 detected within multidomain SCREPs [101]. This is the case, *e.g.,* of the highly expressed U-GRTX-Esi4, with
482 three ShK domains: multidomain ShK proteins have been so far reported in the transcriptomes of parasitic
483 roundworms, in *Hydra* and in the vampire snails [102]. The toxin U-GRTX-Esi3 represents another case of
484 SCREP, presenting a double SCRiP domain in the mature polypeptide: this architecture had not been
485 documented in Cnidaria mature toxins to date, although it has been recognized at the transcript level (e.g.
486 in *Orbicella faveolata*). Two domains were also retrieved in the Kunitz BPTI-like PI-GRTX-Esi5, an architecture
487 so far investigated in detail in hematophagous invertebrates only, but extremely abundant among the
488 putative ion-channel impairing SCREP toxins [63,101] While the abundance of the SCREPs is most likely
489 underestimated across venomous Metazoa, also due to the paucity of proteomic studies confirming the
490 architecture of mature toxins, their abundance in the venom of *Eunicella* might indeed possess some
491 functional significance.
492 Interestingly, conventional pore forming toxins (PFTs) of Cnidaria, including the aerolysin-like Δ-GRTX-Esi29
493 and the two actinoporins Δ-GRTX-Esi30 and 31 were not retrieved in the nematocysts' proteome. The finding
494 of PFTs outside the nematocysts is not unusual: e.g., hydralysin is an exclusively non-nematocystic PFT
495 involved in prey digestion [80]. Actinoporins, on the other hand, have been detected both in the
496 nematocysts and in other tissues: the results of our proteomics analysis indicate that both Δ-GRTX-Esi30 and
497 31 localize exclusively outside the nematocysts in *E. singularis*, thus likely sharing the digestive function of
498 the hydralysin-like Δ-GRTX-Esi29. In addition, the observed mutations at critical sites suggest, for *Eunicella*
499 actinoporins, peculiar chemical selectivity and pore formation mechanism, calling for experimental
500 characterization.
501

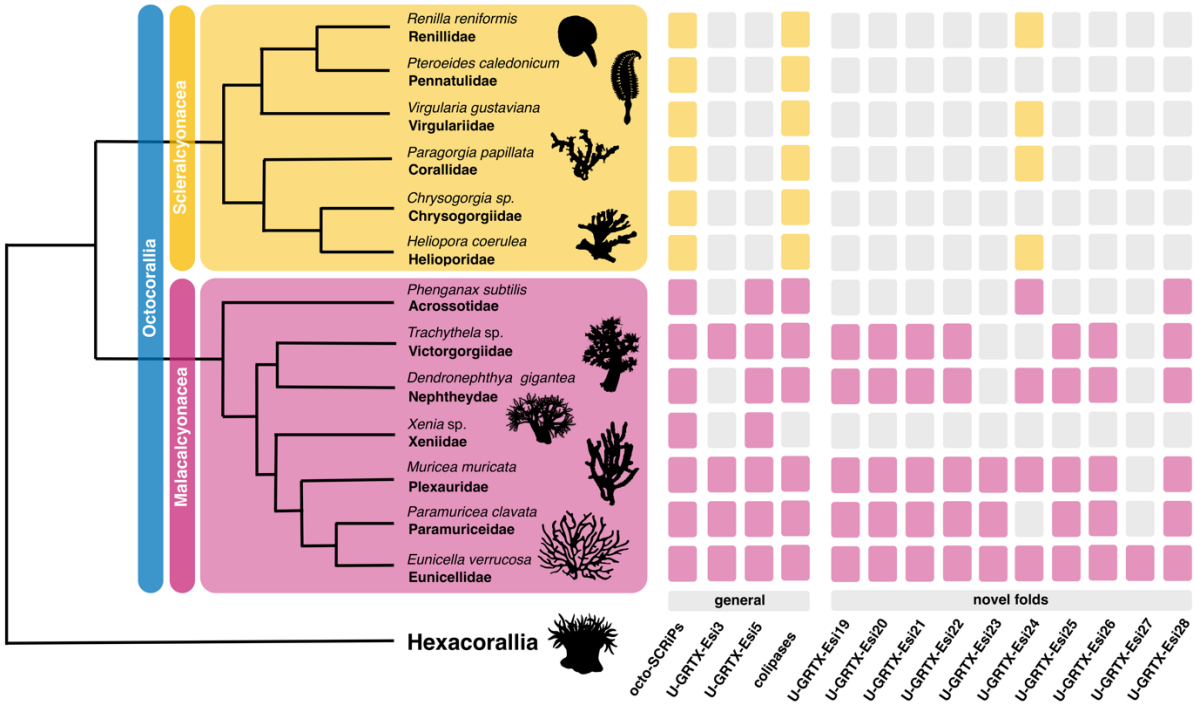


Figure 6. Distribution of orthologues of a subset of *E. singularis* toxins across the Octocorallia genomes here investigated. Phylogenetic relationships of the Octocorallia after [46].

In the nematocyst, 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 [103]. 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 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 (Fig. 6). In addition, a number of 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) (Fig. 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

521 of energy. In contrast, the other Malacalcyonacea included in our analysis (e.g., *Paramuricea*, *Trachytela*)
522 are suspension feeders that largely rely on zooplankton intake [104]. Novel toxins, such as U-GRTX-Esi26, U-
523 GRTX-Esi22, the acrorhagin-like U-GRTX-Esi21 and U-GRTX-Esi27 might thus constitute a predation-related
524 group of toxins. In particular, U-GRTX-Esi27 appears to be restricted to the genus *Eunicella* as a genus-
525 specific adaptation (Fig. 6).
526 The inference on the potential function of octocoral toxins is, however, heavily affected by the sparse
527 information available on the biology of these organisms. For example, some Octocorallia species possess
528 sweeper and thread-like tentacles, nematocysts-rich structures that have been hypothesized to mediate
529 interspecific competition, similarly to sea anemones acrorhagi [105]. The occurrence of these structures
530 across the Octocorallia radiation has been poorly investigated to date: this lack of comprehensive
531 information hampers the evaluation of a potential role of specific toxins in competition. Nevertheless,
532 although our comparative genomic approach has revealed a powerful tool to restrict the range of potential
533 functions, further functional characterization studies are needed to unequivocally establish the physiological
534 role of these toxins.
535 Interestingly, our proteomic analysis highlighted two cases of differential toxin maturation pathways in the
536 nematocyst vs whole body. The first concerns the two SCRiP isotoxins U-GRTX-Esi2a-b, derived from a
537 double-domain precursor, with both isoforms present in the WB-P while only U-GRTX-Esi2b represented in
538 the NEM-P. The second corresponds to the toxin PI-GRTX-Esi6 and the three isotoxins PI-GRTX-Esi7a-c,
539 encoded by a single transcript in which a Kunitz-like domain (retrieved both in the whole body and in the
540 nematocyst proteomes) is followed by three Kazal-type PI domains only identified in the whole-body
541 proteome. The contribution of alternative cleavage sites to venom diversity has been demonstrated in
542 conopeptides [106], but cell- or tissue-specific differential maturation patterns have not been so far
543 documented for venom toxins, suggesting that the mechanisms regulating venom toxin expression could be
544 more complex than previously assessed.
545 Through our multidisciplinary approach, we have demonstrated that Octocorallia produce a complex
546 neurotoxin-rich venom, showcasing an overall scarce resemblance to the hexacoral venom repertoire. This
547 distinctiveness is underscored by the presence of novel folds with a complex evolutionary history. These

548 findings call for extensive work, focusing both on the biology of this understudied group of organisms and
549 on the functional characterization of their toxins, to unravel the adaptive significance of Octocorallia arsenal.
550

551 **Data Availability**

552 The Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the
553 accession GKVI00000000. The version described in this paper is the first version, GKVI01000000. Proteomic
554 data have been submitted to PRIDE with reference 1-20240529-113849-2368759. Other data underlying this
555 article in are available in the online supplementary material.
556

557 **Acknowledgements**

558 This work has received funding from the European Union's Horizon 2020 research and innovation
559 programme under the Marie Skłodowska-Curie grant agreement No 748902. The MGX acknowledges
560 financial support from France Génomique National infrastructure, funded as part of "Investissement
561 d'Avenir" program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09). Mass
562 spectrometry experiments were carried out using the facilities of the Montpellier Proteomics Platform
563 (PPM, BioCampus Montpellier). Dr. Dany Dominguez-Perez is gratefully acknowledged for his help in data
564 handling and curation.
565

566 **References**

567 1. Van Iten H, Marques AC, Leme J, Pacheco MLA, Simões MG. 2014 Origin and early diversification of
568 the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's proterozoic-cambrian
569 history. *Palaeontology* 57, 677–690. (doi:10.1111/pala.12116)
570 2. McFadden CS, Quattrini AM, Brugler MR, Cowman PF, Dueñas LF, Kitahara M V., Paz-García DA,
571 Reimer JD, Rodríguez E. 2021 Phylogenomics, Origin, and Diversification of Anthozoans (Phylum
572 Cnidaria). *Syst Biol* **70**, 635–647. (doi:10.1093/sysbio/syaa103)
573 3. David CN, Özbek S, Adamczyk P, Meier S, Pauly B, Chapman J, Hwang JS, Gojobori T, Holstein TW.
574 2008 Evolution of complex structures: minicollagens shape the cnidarian nematocyst. *Trends in*

575 *Genetics* **24**, 431–438. (doi:10.1016/j.tig.2008.07.001)

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

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

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

583 7. Marques AC, Collins AG. 2004 Cladistic analysis of Medusozoa and cnidarian evolution. *Inv Biol* **123**,
584 23–42. (doi:10.1111/j.1744-7410.2004.tb00139.x)

585 8. Kayal E, Bentlage B, Sabrina Pankey M, Ohdera AH, Medina M, Plachetzki DC, Collins AG, Ryan JF.
586 2018 Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the
587 origins of key organismal traits. *BMC Evol Biol* **18**, 68. (doi:10.1186/s12862-018-1142-0)

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

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

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

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

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

598 14. Madio B, King GF, Undheim EAB. 2019 Sea anemone toxins: A structural overview. *Mar Drugs* **17**,
599 325. (doi:10.3390/md17060325)

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

16. Prentis PJ, Pavasovic A, Norton RS. 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**, 1–5. (doi:10.1371/journal.pone.0031902)
22. Raimundo I, Silva SG, 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*, pp. 123–195.
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)

629 27. Renzi JJ, Shaver EC, Burkepile DE, Silliman BR. 2022 The role of predators in coral disease dynamics.
630 *Coral Reefs* **41**, 405–422. (doi:10.1007/s00338-022-02219-w)
631 28. Radwan FFY, Aboul-Dahab HM, Burnett JW. 2002 Some toxicological characteristics of three
632 venomous soft corals from the Red Sea. *Comp Biochem Physiol C* **132**, 25–35. (doi:10.1016/S1532-
633 0456(02)00045-5)
634 29. Jouiaei M *et al.* 2015 Firing the sting: Chemically induced discharge of cnidae reveals novel proteins
635 and peptides from box jellyfish (*Chironex fleckeri*) Venom. *Toxins* **7**, 936–950.
636 (doi:10.3390/toxins7030936)
637 30. Grabherr MG *et al.* 2011 Full-length transcriptome assembly from RNA-Seq data without a reference
638 genome. *Nat Biotechnol* **29**, 644–652. (doi:10.1038/nbt.1883)
639 31. Wagner GP, Kin K, Lynch VJ. 2012 Measurement of mRNA abundance using RNA-seq data: RPKM
640 measure is inconsistent among samples. *Theor Biosci* **131**, 281–285. (doi:10.1007/s12064-012-0162-
641 3)
642 32. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G. 2003 PEAKS: Powerful software
643 for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectr* **17**,
644 2337–2342. (doi:10.1002/rcm.1196)
645 33. Zhang J *et al.* 2012 PEAKS DB: de novo sequencing assisted database search for sensitive and
646 accurate peptide identification. *Mol Cell Proteomics* **11**, M111.010587.
647 (doi:10.1074/mcp.M111.010587)
648 34. Han Y, Ma B, Zhang K. 2005 SPIDER: software for protein identification from sequence tags with de
649 novo sequencing error. *J Bioinform Comput Biol* **3**, 697–716. (doi:10.1142/s0219720005001247)
650 35. Degueldre M et al. 2017 Diversity in sequences, post-translational modifications and expected
651 pharmacological activities of toxins from four *Conus* species revealed by the combination of cutting-
652 edge proteomics, transcriptomics and bioinformatics. *Toxicon* **130**, 116–125.
653 (doi.org/10.1016/j.toxicon.2017.02.014)
654 36. Huang Y, Niu B, Gao Y, Fu L, Li W. 2010 CD-HIT Suite: A web server for clustering and comparing
655 biological sequences. *Bioinformatics* **26**, 680–682. (doi:10.1093/bioinformatics/btq003)

656 37. Oliveira JS, Fuentes-Silva D, King GF. 2012 Development of a rational nomenclature for naming
657 peptide and protein toxins from sea anemones. *Toxicon* **60**, 539–550.
658 (doi:10.1016/j.toxicon.2012.05.020)
659 38. Jumper J *et al.* 2021 Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–
660 589. (doi:10.1038/s41586-021-03819-2)
661 39. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022 ColabFold: making
662 protein folding accessible to all. *Nat Methods* **19**, 679–682. (doi:10.1038/s41592-022-01488-1)
663 40. Tunyasuvunakool K *et al.* 2021 Highly accurate protein structure prediction for the human
664 proteome. *Nature* **596**, 590–596. (doi:10.1038/s41586-021-03828-1)
665 41. Holm L. 2022 Dali server: structural unification of protein families. *Nucleic Acids Res* **50**, W210–
666 W215. (doi:10.1093/nar/gkac387)
667 42. Ayoub R, Lee Y. 2019 Rupee: A fast and accurate purely geometric protein structure search. *PLoS*
668 *One* **14**, e0213712. (doi:10.1371/journal.pone.0213712)
669 43. Webb B, Sali A. 2016 Comparative protein structure modeling using MODELLER. *Curr Protoc*
670 *Bioinformatics* **2016**, 5.6.1-5.6.37. (doi:10.1002/cpbi.3)
671 44. Jurrus E *et al.* 2018 Improvements to the APBS biomolecular solvation software suite. *Protein*
672 *Science* **27**, 112–128. (doi:10.1002/pro.3280)
673 45. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023 UCSF
674 ChimeraX: Tools for structure building and analysis. *Protein Science* **32**, e4792.
675 (doi:10.1002/pro.4792)
676 46. McFadden CS, Van Ofwegen LP, Quattrini AM. 2022 Revisionary systematics of Octocorallia
677 (Cnidaria: Anthozoa) guided by phylogenomics. *Bull Soc Syst Biol* **1**. (doi:10.18061/bssb.v1i3.8735)
678 47. Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF. 2006 Composition-based statistics and
679 translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biol* **4**, 41.
680 (doi:10.1186/1741-7007-4-41)
681 48. Reese MG. 1997 Improved splice site detection in Genie. *J Comput Biol* **4**, 311–323.
682 (doi:10.1089/cmb.1997.4.311)

683 49. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-
684 TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. 2020.
685 *Mol Biol Evol.* **37**, 1530-1534 (doi: 10.1093/molbev/msaa015).
686 50. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017 ModelFinder: Fast model
687 selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587–589.
688 (doi:10.1038/nmeth.4285)
689 51. Sunagawa S, DeSalvo MK, Voolstra CR, Reyes-Bermudez A, Medina M. 2009 Identification and gene
690 expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals.
691 *PLoS One* **4**, e4865. (doi:10.1371/journal.pone.0004865)
692 52. Logashina YA *et al.* 2017 New disulfide-stabilized fold provides sea anemone peptide to exhibit both
693 antimicrobial and TRPA1 potentiating properties. *Toxins* **9**, 154. (doi:10.3390/toxins9050154)
694 53. Barroso RA, Ramos L, Moreno H, Antunes A. 2024 Evolutionary Analysis of Cnidaria Small Cysteine-
695 Rich Proteins (SCRiPs), an Enigmatic Neurotoxin Family from Stony Corals and Sea Anemones
696 (Anthozoa: Hexacorallia). *Toxins* **16**, 75. (doi:10.3390/toxins16020075)
697 54. Pennington MW, Czerwinski A, Norton RS. 2018. Peptide therapeutics from venom: Current status
698 and potential. *Bioorg. Med. Chem.* 26:2738–2758.
699 55. Sintsova O *et al.* 2021 Sea anemone kunitz-type peptides demonstrate neuroprotective activity in
700 the 6-hydroxydopamine induced neurotoxicity model. *Biomedicines* **9**, 283.
701 (doi:10.3390/biomedicines9030283)
702 56. Peigneur S, Billen B, Derua R, Waelkens E, Debaveye S, Béress L, Tytgat J. 2011 A bifunctional sea
703 anemone peptide with Kunitz type protease and potassium channel inhibiting properties. *Biochem*
704 *Pharmacol* **82**, 81–90. (doi:10.1016/j.bcp.2011.03.023)
705 57. Nikolaev M V., Dorofeeva NA, Komarova MS, Korolkova Y V., Andreev YA, Mosharova I V., Grishin E
706 V., Tikhonov DB, Kozlov SA. 2017 TRPV1 activation power can switch an action mode for its
707 polypeptide ligands. *PLoS One* **12**, e0177077. (doi:10.1371/journal.pone.0177077)
708 58. García-Fernández R, Peigneur S, Pons T, Alvarez C, González L, Chávez MA, Tytgat J. 2016 The kunitz-
709 type protein ShPI-1 inhibits serine proteases and voltage-gated potassium channels. *Toxins* **8**, 110.

710 (doi:10.3390/toxins8040110)

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

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

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

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

722 63. Dai SX, Zhang A Di, Huang JF. 2012 Evolution, expansion and expression of the Kunitz/BPTI gene
723 family associated with long-term blood feeding in *Ixodes scapularis*. *BMC Evol Biol* **12**, 4.
724 (doi:10.1186/1471-2148-12-4)

725 64. Modica MV, Lombardo F, Franchini P, Oliverio M. 2015 The venomous cocktail of the vampire snail
726 Colubraria reticulata (Mollusca, Gastropoda). *BMC Genomics* **16**, 441. (doi:10.1186/s12864-015-
727 1648-4)

728 65. Mishra M. 2020 Evolutionary Aspects of the Structural Convergence and Functional Diversification
729 of Kunitz-Domain Inhibitors. *J Mol Evol* **88**, 537–548. (doi:10.1007/s00239-020-09959-9)

730 66. Chen Z *et al.* 2013 Genomic and Structural Characterization of Kunitz-Type Peptide LmKTT-1a
731 Highlights Diversity and Evolution of Scorpion Potassium Channel Toxins. *PLoS One* **8**, e60201.
732 (doi:10.1371/journal.pone.0060201)

733 67. Fry BG *et al.* 2009 The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal
734 Venoms. *Annu Rev Genomics Hum Genet* **10**, 483–511.
735 (doi:10.1146/annurev.genom.9.081307.164356)

736 68. Boisbouvier J, Albrand J-P, Blackledge M, Jaquinod M, Schweitz H, Lazdunski M, Marion D. 1998 A

737 structural homologue of colipase in black mamba venom revealed by NMR floating disulphide bridge
738 analysis. *J Mol Biol* **283**, 205–219. (doi:10.1006/jmbi.1998.2057)
739 69. Kaser A, Winklmayr M, Lepperdinger G, Kreil G. 2003 The AVIT protein family. *EMBO Rep* **4**, 469–
740 473. (doi:10.1038/sj.embor.embor830)
741 70. Negri L, Lattanzi R, Giannini E, Melchiorri P. 2007 Bv8/Prokineticin proteins and their receptors. *Life*
742 *Sci* **81**, 1103–1116. (doi:10.1016/j.lfs.2007.08.011)
743 71. Wen S *et al.* 2005 Discovery of an MIT-like atracotoxin family: Spider venom peptides that share
744 sequence homology but not pharmacological properties with AVIT family proteins. *Peptides* **26**,
745 2412–2426. (doi:10.1016/j.peptides.2005.05.012)
746 72. Hecht O, van Nuland NA, Schleinkofer K, Dingley AJ, Bruhn H, Leippe M, Grötzinger J. 2004 Solution
747 Structure of the Pore-forming Protein of *Entamoeba histolytica*. *J Biol Chem* **279**, 17834–17841.
748 (doi:10.1074/jbc.M312978200)
749 73. Tosi T, Nickerson NN, Mollica L, Jensen MR, Blackledge M, Baron B, England P, Pugsley AP, Dessen
750 A. 2011 Pilotin-secretin recognition in the type II secretion system of *Klebsiella oxytoca*. *Mol*
751 *Microbiol* **82**, 1422–1432. (doi:10.1111/j.1365-2958.2011.07896.x)
752 74. Korotkov K V., Hol WGJ. 2013 Crystal structure of the pilotin from the enterohemorrhagic
753 *Escherichia coli* type II secretion system. *J Struct Biol* **182**, 186–191. (doi:10.1016/j.jsb.2013.02.013)
754 75. Krishnarjuna B, Sunanda P, Villegas-Moreno J, Csoti A, A.V. Morales R, Wai DCC, Panyi G, Prentis P,
755 Norton RS. 2021 A disulfide-stabilised helical hairpin fold in acrorhagin I: An emerging structural
756 motif in peptide toxins. *J Struct Biol* **213**, 107692. (doi:10.1016/j.jsb.2020.107692)
757 76. Honma T, Minagawa S, Nagai H, Ishida M, Nagashima Y, Shiomi K. 2005 Novel peptide toxins from
758 acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon* **46**, 768–774.
759 (doi:10.1016/j.toxicon.2005.08.003)
760 77. Anderluh G, Lakey JH. 2008 Disparate proteins use similar architectures to damage membranes.
761 *Trends Biochem Sci* **33**, 482–490. (doi:10.1016/j.tibs.2008.07.004)
762 78. Morando MA, Venturella F, Sollazzo M, Monaca E, Sabbatella R, Vetri V, Passantino R, Pastore A,
763 Alfano C. 2022 Solution structure of recombinant Pvfp-5β reveals insights into mussel adhesion.

764 *Commun Biol* **5**, 1–10. (doi:10.1038/s42003-022-03699-w)

765 79. Yeung SYM, Thompson D, Wang Z, Fedida D, Robertson B. 2005 Modulation of Kv3 subfamily
766 potassium currents by the sea anemone toxin BDS: Significance for CNS and biophysical studies. *J*
767 *Neurosci* **25**, 8735–8745. (doi:10.1523/JNEUROSCI.2119-05.2005)

768 80. Sher D, Fishman Y, Melamed-Book N, Zhang M, Zlotkin E. 2008 Osmotically driven prey
769 disintegration in the gastrovascular cavity of the green hydra by a pore-forming protein. *The FASEB*
770 *Journal* **22**, 207–214. (doi:10.1096/fj.07-9133com)

771 81. Macrander J, Daly M. 2016 Evolution of the cytolytic pore-forming proteins (Actinoporins) in sea
772 anemones. *Toxins* **8**, 1–16. (doi:10.3390/toxins8120368)

773 82. Darmoise A, Maschmeyer P, Winau F. 2010 The Immunological Functions of Saposins. *Adv Immunol*
774 **105**, 25–62. (doi:10.1016/S0065-2776(10)05002-9)

775 83. Ahn VE, Leyko P, Alattia J-R, Chen L, Privé GG. 2006 Crystal structures of saposins A and C. *Protein*
776 *Sci* **15**, 1849–1857. (doi:10.1110/ps.062256606)

777 84. Munford RS, Sheppard PO, O'Hara PJ. 1995 Saposin-like proteins (SAPLIP) carry out diverse functions
778 on a common backbone structure. *J Lipid Res* **36**, 1653–1663.

779 85. Bruhn H. 2005 A short guided tour through functional and structural features of saposin-like
780 proteins. *Biochem J* **389**, 249–257. (doi:10.1042/BJ20050051)

781 86. Herbst R, Marciano-Cabral F, Leippe M. 2004 Antimicrobial and pore-forming peptides of free-living
782 and potentially highly pathogenic *Naegleria fowleri* are released from the same precursor molecule.
783 *Journal of Biological Chemistry* **279**, 25955–25958. (doi:10.1074/jbc.M401965200)

784 87. Hill CH, Read RJ, Deane JE. 2015 Structure of human saposin A at lysosomal pH. *Acta Crystallogr.*
785 *Sect F* **71**, 895–900. (doi:10.1107/S2053230X15008584)

786 88. Hawkins CA, De Alba E, Tjandra N. 2005 Solution structure of human saposin C in a detergent
787 environment. *J Mol Biol* **346**, 1381–1392. (doi:10.1016/j.jmb.2004.12.045)

788 89. Alattia JR, Shaw JE, Yip CM, Privé GG. 2006 Direct Visualization of Saposin Remodelling of Lipid
789 Bilayers. *J Mol Biol* **362**, 943–953. (doi:10.1016/j.jmb.2006.08.009)

790 90. Popovic K, Holyoake J, Pomès R, Privé GG. 2012 Structure of saposin A lipoprotein discs. *PNAS* **109**,

791 2908–2912. (doi:10.1073/pnas.1115743109)

792 91. Leippe M, Bruhn H, Hecht O, Grotzinger J. 2005 Ancient weapons: the three-dimensional structure
793 of amoebapore A. *Trends Parasitol* **21**, 5–7. (doi:10.1016/j.pt.2004.10.009)

794 92. Mysliwy J, Dingley AJ, Stanisak M, Jung S, Lorenzen I, Roeder T, Leippe M. 2010 Caenopore-5: The
795 three-dimensional structure of an antimicrobial protein from. *Dev Comp Immunol* **34**, 323-30 (doi:
796 10.1016/j.dci.2009.11.00

797 93. Liepinsh E, Andersson M, Ruysschaert JM, Otting G. 1997 Saposin fold revealed by the NMR
798 structure of NK-lysin. *Nat Struct Biol* **4**, 793–795. (doi:10.1038/nsb1097-793)

799 94. Anderson DH, Sawaya MR, Cascio D, Ernst W, Modlin R, Krensky A, Eisenberg D. 2003 Granulysin
800 Crystal Structure and a Structure-derived Lytic Mechanism. *J Mol Biol* **325**, 355–365.
801 (doi:10.1016/S0022-2836(02)01234-2)

802 95. Chen L, Dürr KL, Gouaux E. 2014 X-ray structures of AMPA receptor-cone snail toxin complexes
803 illuminate activation mechanism. *Science* **345**, 1021–1026. (doi:10.1126/science.1258409)

804 96. Hackney CM *et al.* 2023 A previously unrecognized superfamily of macro-conotoxins includes an
805 inhibitor of the sensory neuron calcium channel Cav2.3. *PLoS Biol* **21**, e3002217.
806 (doi:10.1371/journal.pbio.3002217)

807 97. Olmeda B, García-Álvarez B, Pérez-Gil J. 2013 Structure-function correlations of pulmonary
808 surfactant protein SP-B and the saposin-like family of proteins. *EBJ* **42**, 209–222.
809 (doi:10.1007/s00249-012-0858-9)

810 98. Olmeda B, García-Álvarez B, Gómez MJ, Martínez-Calle M, Cruz A, Pérez-Gil J. 2015 A model for the
811 structure and mechanism of action of pulmonary surfactant protein B. *FASEB Journal* **29**, 4236–
812 4247. (doi:10.1096/fj.15-273458)

813 99. Garrido-Arandia M, Cuevas-Zuviría B, Díaz-Perales A, Pacios LF. 2018 A comparative study of human
814 saposins. *Molecules* **23**, 422. (doi:10.3390/molecules23020422)

815 100. Yin W *et al.* 2016 Evolutionary trajectories of snake genes and genomes revealed by comparative
816 analyses of five-pacer viper. *Nat Commun* **7**, 13107. (doi:10.1038/ncomms13107)

817 101. Maxwell M, Undheim EAB, Mobli M. 2018 Secreted cysteine-rich repeat proteins 'SCREPs': A novel

818 multi-domain architecture. *Front Pharmacol* **9**, 1333. (doi:10.3389/fphar.2018.01333)
819 102. Gerdol M, Cervelli M, Mariottini P, Oliverio M, Dutertre S, Modica MV. 2019 A recurrent motif:
820 Diversity and evolution of ShKT domain containing proteins in the vampire snail *Cumia reticulata*.
821 *Toxins* **11**. (doi:10.3390/toxins11020106)
822 103. Berking S, Herrmann K. 2006 Formation and discharge of nematocysts is controlled by a proton
823 gradient across the cyst membrane. *Helgol Mar Res* **60**, 180–188. (doi:10.1007/s10152-005-0019-
824 y)
825 104. Fabricius KE, Dommisse M. 2000. Depletion of suspended particulated matter over coastal reef
826 communities dominated by zooxanthellate soft corals. *MEPS* **196**, 157-167
827 (doi:10.3354/meps196157)
828 105. Sebens KP, Miles JS. 1988 Sweeper Tentacles in a Gorgonian Octocoral: Morphological
829 Modifications for Interference Competition. *Biol Bull* **175**, 378–387. (doi:10.2307/1541729)
830 106. Dutertre S, Jin AH, Kaas Q, Jones A, Alewood PF, Lewis RJ. 2013 Deep venomics reveals the
831 mechanism for expanded peptide diversity in cone snail venom. *Mol Cell Proteomics* **12**, 312–329.
832 (doi:10.1074/mcp.M112.021469)
833