bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Article

Transcriptomic analysis of four cerianthid (Cnidaria, Ceriantharia) venoms

Anna M. L. Klompen [1*], Jason Macrander [2,3], Adam M. Reitzel [2] , Sérgio N. Stampar [4]

[1] Department of Ecology and Evolutionary Biology, University of Kansas, 1200 Sunnyside Ave., Lawrence, KS, 66045, USA

[2] Department of Biological Sciences, University of North Carolina at Charlotte, 8812 Craver Rd, Charlotte, NC, 28262, USA

[3] Department of Biology, Florida Southern College, 111 Lake Hollingsworth Drive Lakeland, FL, 33801, USA

[4] Department of Biological Sciences, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), FCL, Assis, Brazil

* Correspondence: annaklompen@ku.edu

Abstract: Tube anemones, or cerianthids, are a phylogenetically informative group of cnidarians with complex life histories, including a pelagic larval stage and tube-dwelling adult stage, both known to utilize venom in stinging-cell rich tentacles. Cnidarians are an entirely venomous group that utilize their proteinaceous-dominated toxins to capture prey and defend against predators, in addition to several other ecological functions, including intraspecific interactions. At present there are no studies describing the venom for any species within cerianthids. Given their unique development, ecology, and distinct phylogenetic-placement within Cnidaria, our objective is to evaluate the venom-like gene diversity of four species of cerianthids from newly collected transcriptomic data. We identified 525 venom-like genes between all four species. The venom-gene profile for each species was dominated by enzymatic protein and peptide families, which is consistent with previous findings in other cnidarian venoms. However, we found few toxins that are typical of sea anemones and corals, and furthermore, three of the four species express toxin-like genes closely related to potent pore-forming toxins in box jellyfish. Our study is the first to provide a survey of the putative venom composition of cerianthids, and contributes to our general understanding of the diversity of cnidarian toxins.

Keywords: Cnidaria, Anthozoa, Ceriantharia, tube anemone, venom, transcriptome

1. Introduction

The phylum Cnidaria (sea anemones, corals, jellyfish, box jellies, hydroids/hydromedusae, etc.) is the earliest diverging venomous lineage (~ 600 million years) [1,2]. Cnidaria deliver their proteinaceous-dominant venom through organelles called nematocysts (a type of cnidae), housed in cells called nematocytes [3,4]. Venom from discharged nematocysts is used in prey capture and defense against predation, but cnidarians also use venom for a variety of other behaviors, such as intraspecific competition [5-7] and maternal care [8] (see review by [9]). This ecological diversity is complemented by the functional diversity of cnidarian venoms, which can include neurotoxic, cytotoxic, and enzymatic (e.g. phospholipase and metalloprotease) proteins and peptides, in addition to non-peptidic components [10,11]. For humans, stings from certain species can cause intense localized pain, scarring, induced anaphylaxis, and in the worst cases, cardiac and respiratory failure leading to death [12-15]. The venom of medically relevant species, such as the Portuguese Man-o-War (*Physalia physalis*) [16-18] and several species of box jellyfish ([19-22], reviewed in [23]), or easy to collect species, such as sea anemones [24,25], have been explored more extensively at a biochemical and pharmacological level [26]. However, these species represent a small fraction of the species diversity within the group, and only recently has the exploration the

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

venom composition for a wider number of cnidarians increased in an effort to characterize the evolution and ecological function of toxins within the group [27].

There is also a growing interest in cnidarian venoms as a potential resource for drug discovery, particularly the neurotoxin-rich venoms of sea anemones [28-30]. One of the best studied therapeutic proteins derived from a cnidarian toxin is an analogue of a potassium Kv1.3 channel blocker isolated from the sun sea anemone (*Stichodactyla helianthus*) called ShK [31], which completed Phase 1b trials for autoimmune diseases [32,33]. Because ShK-scaffolds are abundant in sea anemone venom peptides, characterizing the venoms from sea anemones (and cnidarians in general) could yield additional candidates for novel therapeutic compounds [30,34,35]. Kunitz-domain containing serine inhibitors, also found in sea anemone venoms, can also be used as potential therapeutic resources [25,36]. These cnidarian-derived neuropeptide inhibitors have potential applications as analgesics, antiepileptics, and other neuroprotective drugs [37].

While there has been a recent increase in transcriptomic and proteotranscriptomic analyses of cnidarian venoms (e.g. [7,8,22,38-54]), the phylum as a whole, which contains over 13,000 species, remains highly understudied. Cnidaria is split into three taxonomic groups: Anthozoa (sea anemones, corals, zoanthids, etc.), Medusozoa (jellyfish, box jellies, hydroids, siphonophores), and Endocnidozoa (*Polypodium* + myxozoans) [55,56]. Of the 7,142 animal toxins and venoms listed in Tox-Prot, a curated animal venom annotation database, only 273 are derived from cnidarians (as of May 2020, [57]), with that vast majority (>96%) are from anthozoans. Within that limited number there is even greater taxonomic bias; almost 90% of anthozoan toxins are from the Actinioidea superfamily of sea anemones [27,30], meaning less than 50 out of 1,100 known sea anemone species contribute to the database of annotated cnidarian toxins [54]. This taxon bias limits researcher's ability to discover novel therapeutic peptides and scaffolds from sea anemones, as well as limits to search for potential drug candidates in other anthozoan groups such as corals [58] and zoanthids [47-49].

One major hurdle to identifying the composition and comparative diversity of cnidarian toxins is their lack of a centralized venom system that can be easily isolated for study. This packaging of toxins into individual nematocysts scattered throughout the animal, impedes the ability to isolate crude venoms for downstream analysis, which is further exacerbated in smaller or rare species of cnidarians. There are several protocols for isolating venom from nematocysts (e.g. [59-62]), but these methods, as noted above, are typically restricted to larger or easy to obtain animals (e.g. corals and sea anemones, true jellies such as *Chrysaora* and *Cyanea*), species of medical relevance (e.g. *Physalia*, box jellies), or those that can be easily maintained in a lab (e.g. *Hydra* [63], *Nematostella* [64]). Next generation sequencing technologies provide a solution to this problem, and have greatly increased the ability of researchers to screen the diversity of putative venom-like genes for neglected or poorly studied venomous species, including cnidarians [65].

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

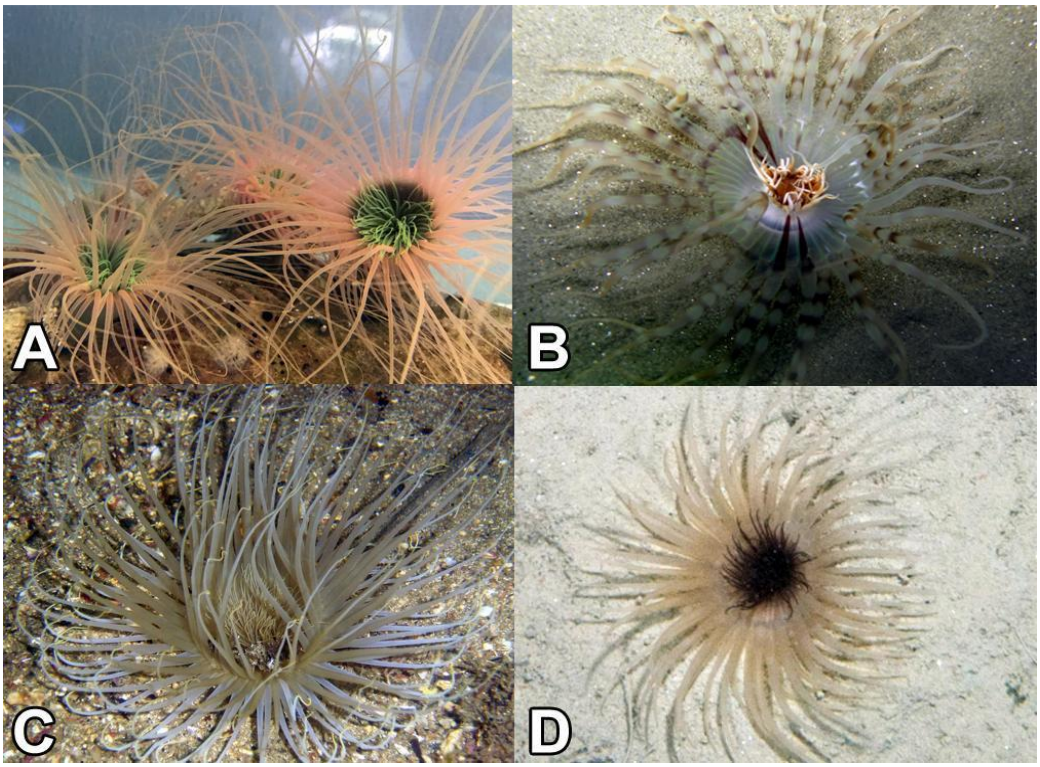

Figure 1. Ceriantharia species used in the current study. A) *Pachycerianthus cf. maua*; B) *Isarachnanthus nocturnus*; C) *Ceriantheomorphe brasiliensis* and D) *Pachycerianthus borealis*. Photos by Fisheries and Oceans Canada (Claude Nozères)).

One group of anthozoans whose venoms have yet to be explored are members of the subclass Ceriantharia, known as cerianthids (Phylum Cnidaria: Class Anthozoa) (Figure 1). Cerianthids are tube-dwelling anemones, so named because of their ability to create complex tubing from a specialized group of cnidae called ptychocysts [66]. Their phylogenetic placement within Cnidaria remains contentious, due to a combination of a lack of available sequence data and low species sampling [5,67,68]. Various studies place them as sister group to Hexacorallia, sister group to Octocorallia [69], or sister group to Hexacorallia + Octocorallia [70,71]. Although cerianthids are clearly members of Anthozoa, they have several features that are more similar to Medusozoa. For instance, cerianthids possess linear mitochondrial genomes, as in medusozoans, while all other anthozoans have circular mitochondrial genomes [71-73]. Also, unlike other anthozoans, cerianthids display a long-lived pelagic larval stage that superficially resembles a medusa [74]. It is unclear how this unique life history or their early diverging phylogenetic relationship to either, or both, of the major groups of anthozoans may be reflected in the venom composition of this group relative to other anthozoan venoms (or cnidarians more generally).

The aim of this project is to explore newly sequenced transcriptomes for four adult cerianthid species (*Ceriantheomorphe brasiliensis*, *Isarachnanthus nocturnus*, *Pachycerianthus borealis*, and *Pachycerianthus cf. maua*) and determine putative venom-like gene candidates across each using a customized annotation pipeline. This study is the first formal analysis of venom composition within this subclass Ceriantharia, and a targeted comparison of the venom gene profiles between cerianthids and other cnidarian species.

2. Results

2.1. Results for sequencing and de-novo transcriptome assembly of four cerianthids species

The number of paired end reads generated by Illumina HiSeq runranged from 27,865,720 to 36,520,791 across all taxa. The Trinity [75] assembly ranged from 92,757 to 158,663 unique

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

assembled transcripts with an N50 range from 1101 - 1282. Overall completeness evaluated in BUSCO ranged from 88.1% to 97.9% complete (Table 1).

Table 1. Sequencing and assembly parameters for various cerianthid transcriptomes.

Species	Reads (PE)	Transcripts	Genes	N50	BUSCO %
C. brasiliensis	34,877,883	131,550	110,524	1,276	95.4%
I. nocturnus	31,028,274	92,757	78,821	1170	89.2%
P. borealis	36,520,791	158,633	120,542	1,282	97.9%
P. maua.	27,865,720	179,576	145,788	1101	88.1%

2.2. Diversity and phylogenetic analysis of putative venom-like gene profiles for cerianthids species

Using the de-novo assemblies, we identified a diverse set of venom-like putative protein coding transcripts and peptides across the four cerianthids: 169, 69, 182, and 105 for *C. brasiliensis*, *I. nocturnus*, *P. borealis*, and *P. maua*, respectively. All toxins were categorized into families/scaffolds based on their highest Tox-Prot (i.e. UniProtKB/Swiss-Prot) BLAST hit [57], and categorized by biological function: neurotoxin, hemostatic and hemorrhagic toxins, membrane-active toxins, mixed function enzymes, protease inhibitors, allergen and innate immunity, and venom auxiliary proteins (modified from [49]). A summary of annotated contigs for each species is shown in Figure 2, Table 2. Below we provide short descriptions of select toxin groups and families represented by the identified toxins.

2.1.1. Neurotoxins

ShK-domain containing proteins and peptides are some of the most diverse toxins within the transcriptomes of the four species, which includes 15 cysteine-rich venom proteins, 27 ShK-domain containing toxins as identified from Pfam [76,77] (Supplemental Figures S1, S2), and a single sea anemone type 1 potassium channel toxin in *P. maua*. Interestingly, a single transcript in *P. borealis* that contains an ShK-domain had the closest match to propeptide 332-1 toxin from *Malo kingi*, a box jellyfish with a potent sting known to cause Irukandji syndrome [78]. Though the functions are highly variable and depend on the combination of present domains [30,79], ShK-domain toxins can cause paralysis due to potassium channel inhibition as well as induce hemolytic effects [80,81]. As noted above, these ShK toxins may also confer structural and/or functional properties of interest for pharmacological research.

Turripeptides are ion channel blockers described from turrid gastropods, relatives of cone snails, but they have also been predicted or isolated from three species of zoanthid [47-49], a box jellyfish [22], a true jellyfish [82], and a stalked jellyfish [83], as well as bloodworms and marine annelids [81]. These toxin peptides contain a kazal domain with a conserved cysteine framework (C-C-C-C-C-C), and modulate ion channels, resulting in paralysis [84,85]. Four transcripts from cerianthids were shown to have similar cysteine patterns architecture, but have longer predicted protein sequences than the typical turripeptide sequences of <100 amino acids and four additional conserved cysteines upstream from the kazal domain (Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

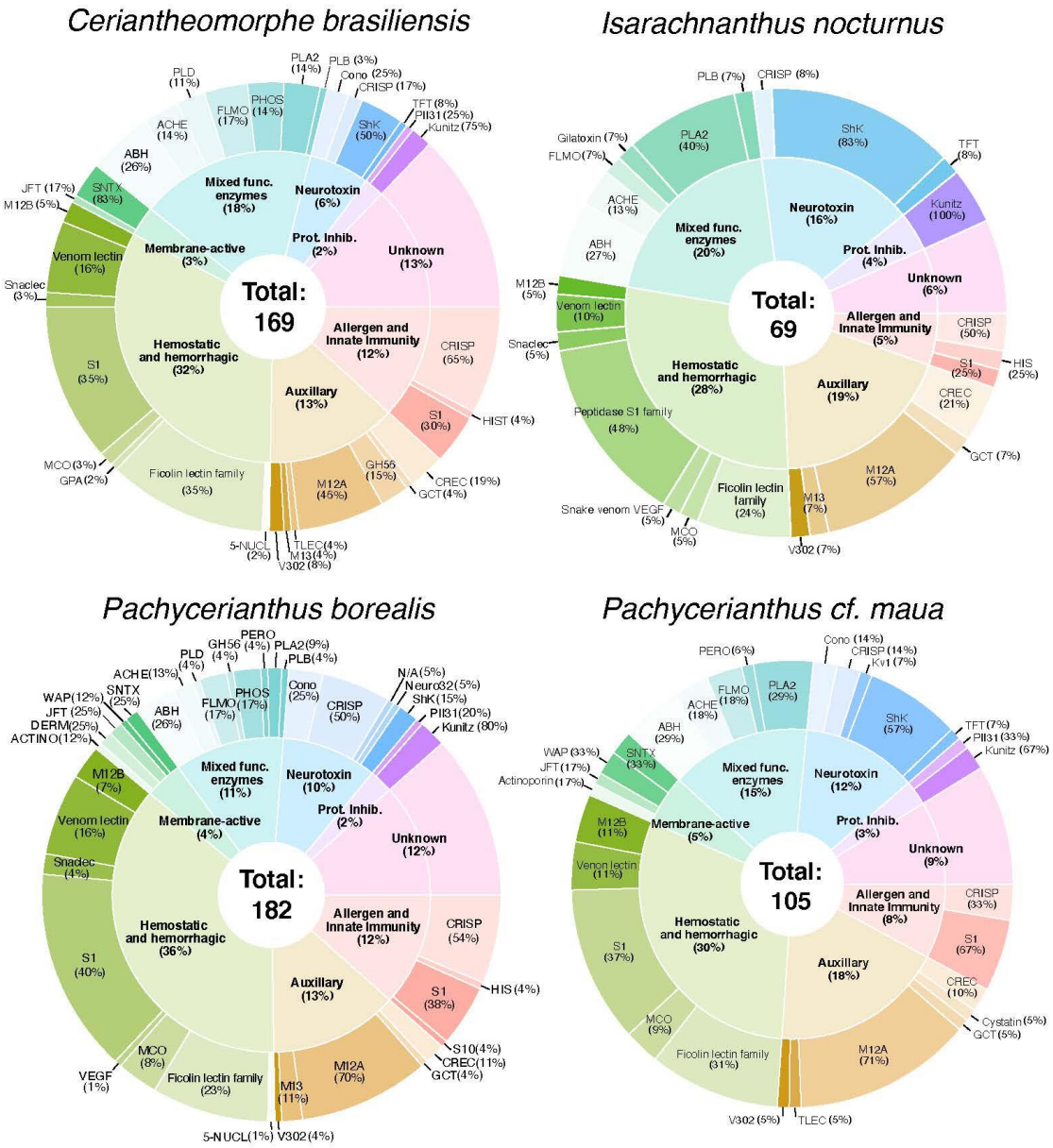

Figure 2. Number of venom-like genes identified for four cerianthid species. Inner circle: biological function and overall percentage of each over the total venom-like gene profile in each species. Outer circle: Venom-like genes families within each biological function category and overall percentage of that family within each category. ABH = AB hydrolase superfamily; ACHE = Acetylcholinesterase; ACTINO = Actinoporin-like; Cono = Conopeptide P–like superfamily; DERM = Dermatopontin; FLMO = lavin monoamine oxidase; GCT = Glutaminyl–peptide cyclotransferase; GH56 = Glycoside hydrolase 56; GPA = Glycoprotein hormones subunit alpha; HIS = Histidine acid phosphatase; JFT = Jellyfish Toxin; Kunitz = Venom Kunitz–type; Kv1 = Sea anemone type 1 potassium channel toxin ; M12A = Peptidase M12; MCO= Multicopper oxidase; M12B = Venom metalloproteinase (M12B); M13 = Peptidase M13; Neuro32 = Neurotoxin 32 Family; PII31 = Protease inhibitor I31; PHOS = Nucleotide pyrophosphatase/phosphodiesterase; PLA2 = Phospholipase A2; PLB = Phospholipase B-like; PLD = Arthropod phospholipase D; PERO = Peroxiredoxin ; SNTX = SNTX/VTX toxin; S1,S10 = Peptidase S1,S10; Venom Lectin = True venom lectin ; TLEC = Techylectin–like; TFT = Snake three–finger toxin; VEGF = Venom vascular endothelial growth factor; V302 = Venom protein 302; WAP = Snake waprin; 5-NUCL = 5'–nucleotidase.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

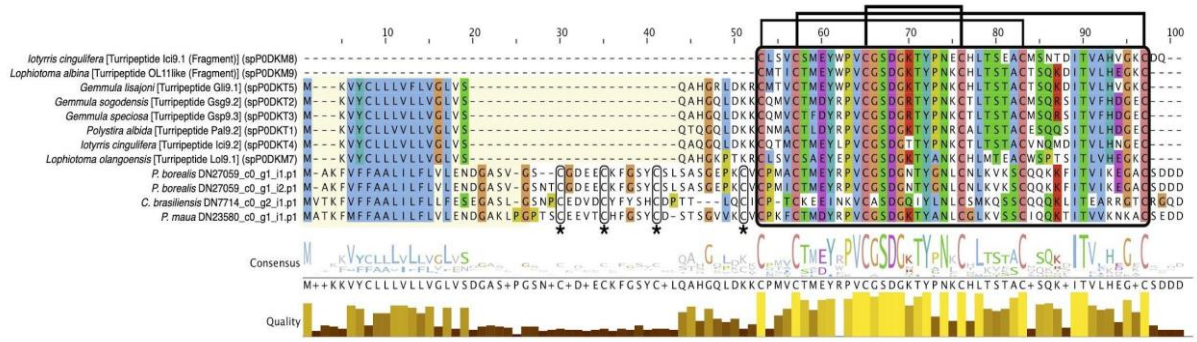

Figure 3. Multiple sequence alignment of candidate turripeptide-like sequences for cerianthid toxins and representatives from cone snails created using L-INS-I algorithm via MAFFT [153], viewed using Jalview [155] with Clustal color scheme. Kazal domain (in black box) and conserved cysteine patterning shown (bridging) are highlighted. The yellow box indicates the predicted signal peptide sequences as indicated by Signalp [147]. The stars and corresponding smaller black boxes indicate the four cysteine residues that are present in the cerianthid sequences preceding the kazal domain.

Three sequences, one each from *C. brasiliensis*, *I. nocturnus*, and *P. maua*, closely matched to three-finger toxins (TFTs), snake-derived toxins that display a wide diversity of functions such as neurotoxicity, acetylcholinesterase inhibitors, cytotoxins (cardiotoxins), platelet aggregation inhibitors, coagulation factor inhibitors, heparin binders, and K+ channel, and integral-receptor ligands [86]. A recent proteomic study found that the orange cup coral *Tubastrea coccinea* contains a putative TFT toxin [83], in addition to a predicted TFT in *P. varibilis* [47]. The TFT toxins in cerianthids and *P. varibilis* cluster as sister to bucandin, a TFT isolated from Malayan krait (*Bungarus candidus*) [87]. However, the bootstrap support throughout the phylogeny is generally low (<70%) (Supplemental Figures S4).

2.2.2. Hemostatic and hemorrhagic toxins

Hemostatic and hemorrhagic toxins are the most diverse type of toxins in all four cerianthid species (Figure 2). They generally interfere with hemostasis through various pathways, either individually or synergistically with other toxins. This group includes a variety of C-type lectin-containing toxins (C-type lectin lectoxin, galactose specific lectin, and snake c-type lectin (snaclec)), and are associated with blood coagulation, inflammation, myotoxicity, and homeostasis interference [87,88]. They have been reported in a variety of animal venoms, including, crustaceans, blood feeding insects, caterpillars, leeches, bloodworms, snakes, and stonefish [88], as well as cnidarian species [38,43,44,47,49]. We found 34 total toxins between the four species that match to a C-type lectin domain.

One of the most numerous groups of venom-like genes within this class are putative veficolin-like toxins (total 30), which are, comparatively, highly abundant in *P. borealis* (9 sequences) and *C. brasiliensis* (14 sequences). This toxin was described from the Komodo dragon (*Varanus komodoensis*), and is suggested to interfere with blood coagulation and/or platelet aggregation based on the similarity to ryncolin toxins [90]. Ryncolin toxins are represented in all cerianthid assemblies in relatively high abundance with 25 total sequences, originally described from the dog-faced water snake (*Cerberus rynchops*). Six sequences from the transcriptome of the zoanthid *Palythoa caribaeorum* (categorized in our study under allergen and innate immunity) [48] and three peptides from the proteome of the scyphozoan *Nemopilema nomurai* (as *Stomolophus meleagris*) [38] also belong in this group, suggesting ryncolin-like toxins may play be present across cnidarians.

We also found numerous venom prothrombin activators in two different groups: Factor 5/8 C-domain and trypsin domain. These types of toxins are well known from snake venoms, and cause hemostatic impairment by proteolytic cleavage of prothrombin to thrombin [91]. Putative

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

transcripts have been found in relatively high abundance in the mat anemone *Zoanthus natalensis* [49] as well as in the transcriptomes of *P. caribaeorum* [48] and sea anemone *Anthopleura dowii* [53]. They have also been found in a transcriptomic analysis of the sea anemone *Stichodactyla haddoni* venom, but no peptides were detected using mass spectrometry [46], suggesting that additional proteomic experiments will be needed to confirm the presence of these prothrombin activators (and other toxin groups) in cerianthid venoms.

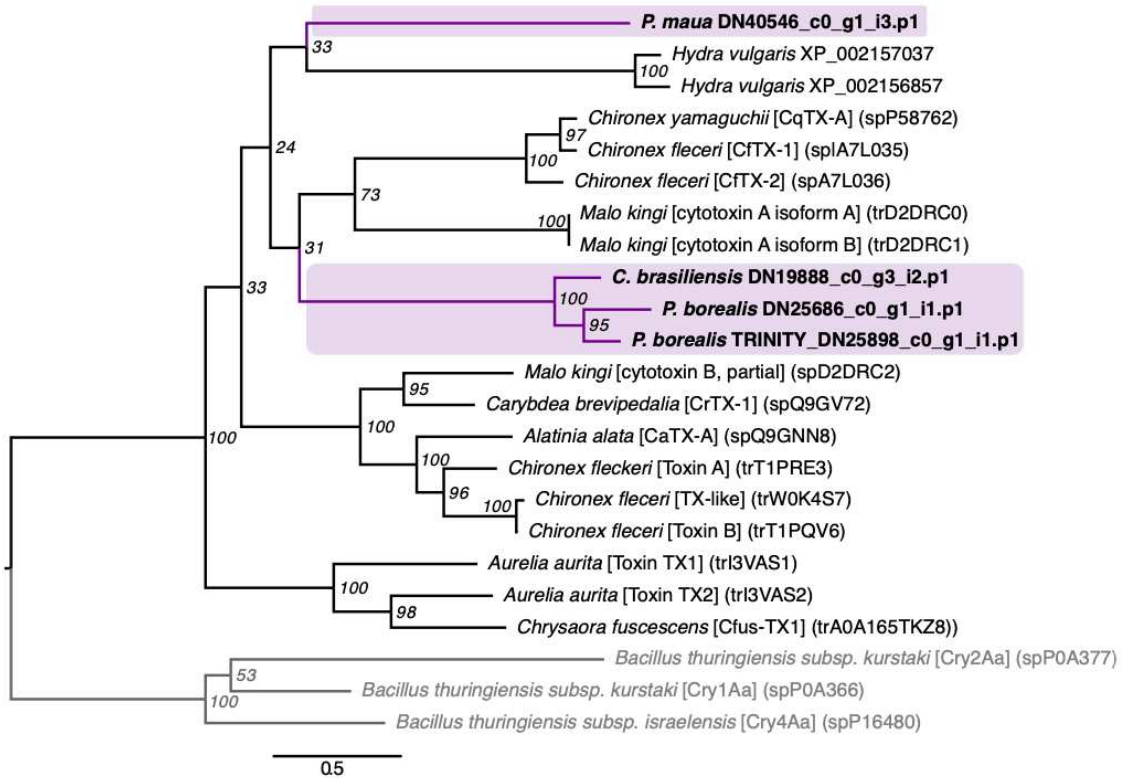

Figure 4. Phylogenetic tree of jellyfish toxin (or CaTX/CrTX) sequences. Phylogeny was constructed using RAxML with the PROTGAMMAWAG option [154]. Bootstrap support based on 500 rapid bootstrap replicates, and all support values are shown. Putative genes outlined in purple are from cerianthids sequences. Sequences in gray are bacterial pore-forming toxins that have closest structural homology to this toxin family [14], used to root the tree.

2.2.3. Membrane-active toxins, protease inhibitors

Jellyfish toxins (or CaTX/CrTX) are one of the most potent toxin families from cnidarians, initially isolated from several species of box jellyfish possessing stings that are dangerous to humans [20]. Two members within this family, CfTX-1 and CfTX-2 from the Australian box jellyfish (*Chironex fleckeri*), are highly cardiotoxic, and their stings are associated with cardiac failure [41]. Four sequence from cerianthids, two from *P. borealis* and one each from *C. brasiliensis* and *P. maua*, appear to belong in this family based on strong phylogenetic evidence, although the transcript from *P. maua* clustered with toxins from the hydroid *Hydra vulgaris* [92], which have yet to be functionally analyzed (Figure 4).

Originally derived from sea anemones, actinoporins are conserved 20kDa pore-forming toxins that exhibit cytolytic and hemolytic effects [93]. Actinoporin-like sequences have also been isolated from both molluscs [94] and chordates [95], and shown to be toxic to a wide variety of vertebrate and invertebrate species [96,97]. Two actinoporin sequences similar to DELTA-thalatoxin-Avl2a were found in *P. borealis* and *P. maua*, though both were phylogenetically closer to actinoporin-like sequences found in venomous gastropods and a putative actinoporin from *P. varibilis* [47] (Figure 5).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

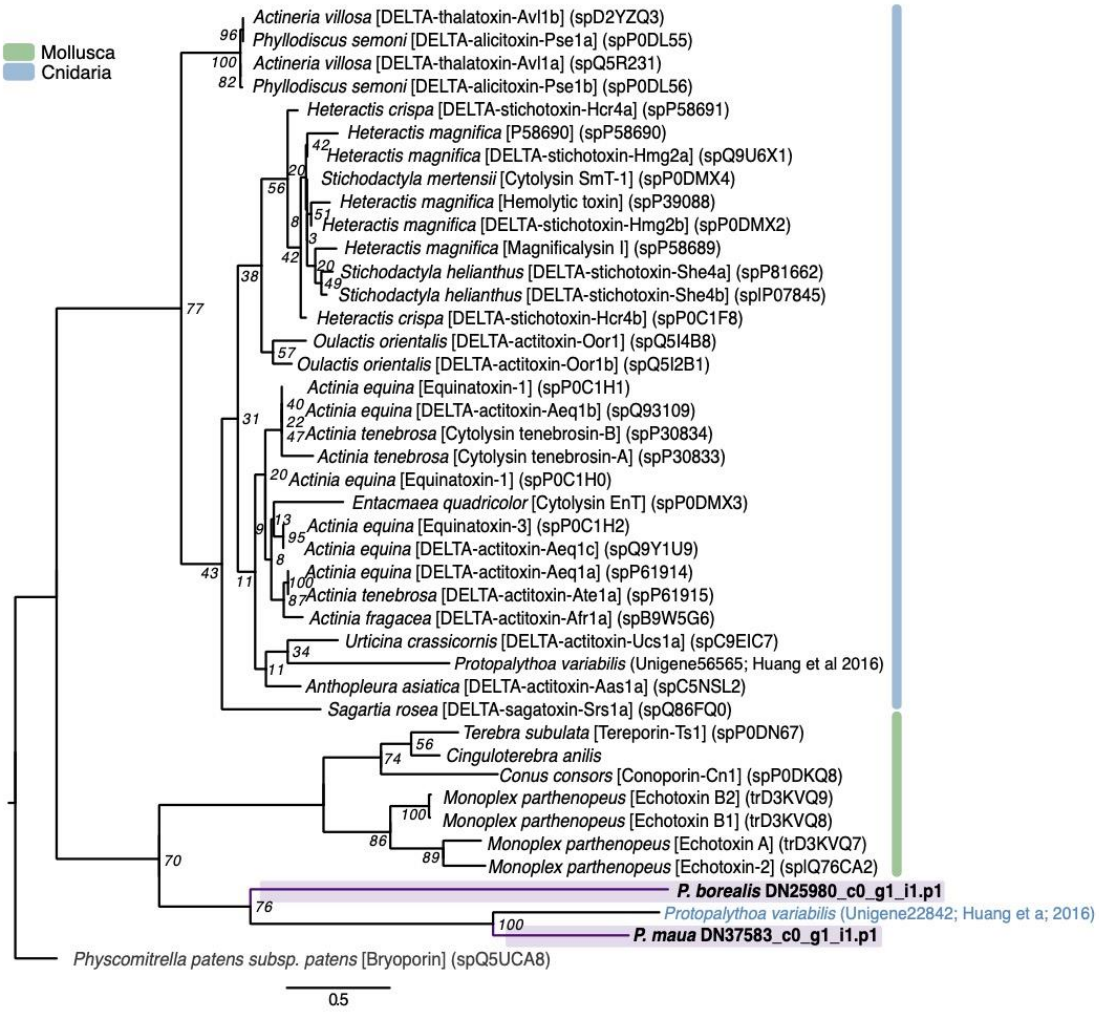

Figure 5. Phylogenetic tree of actinoporin and actinoporin-like sequences. Phylogeny was constructed using RAxML with the PROTGAMMAWAG option [154]. Bootstrap support based on 500 rapid bootstrap replicates, and all support values are shown. Putative genes outlined in purple are from cerianthids sequences. Sequences in gray are non-venomous representatives, and other colors outlined in the key are venom-like genes from other animal classes. Phylogeny modified from von [81]. Tree is rooted with actinoporin-like sequence from a moss (*Physcomitrella patens subsp. patens*).

SNTX-like transcripts include stonutoxin and neoverrucotoxin, non-enzymatic proteins found in a diversity of scorpaeniform fish and monotremes mammals [98,99]. In fish, these toxins cause lethal hemolysis and disrupt circulatory and neuromuscular systems [100,101]. *P. borealis*, *C. brasiliensis*, and *P. maua* express 9 SNTX-like transcripts, all of which phylogenetically cluster together in a group with two SNTX-like genes from non-venomous fish that is sister to a clade of SNTX genes from highly toxic stonefish (Figure 6).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

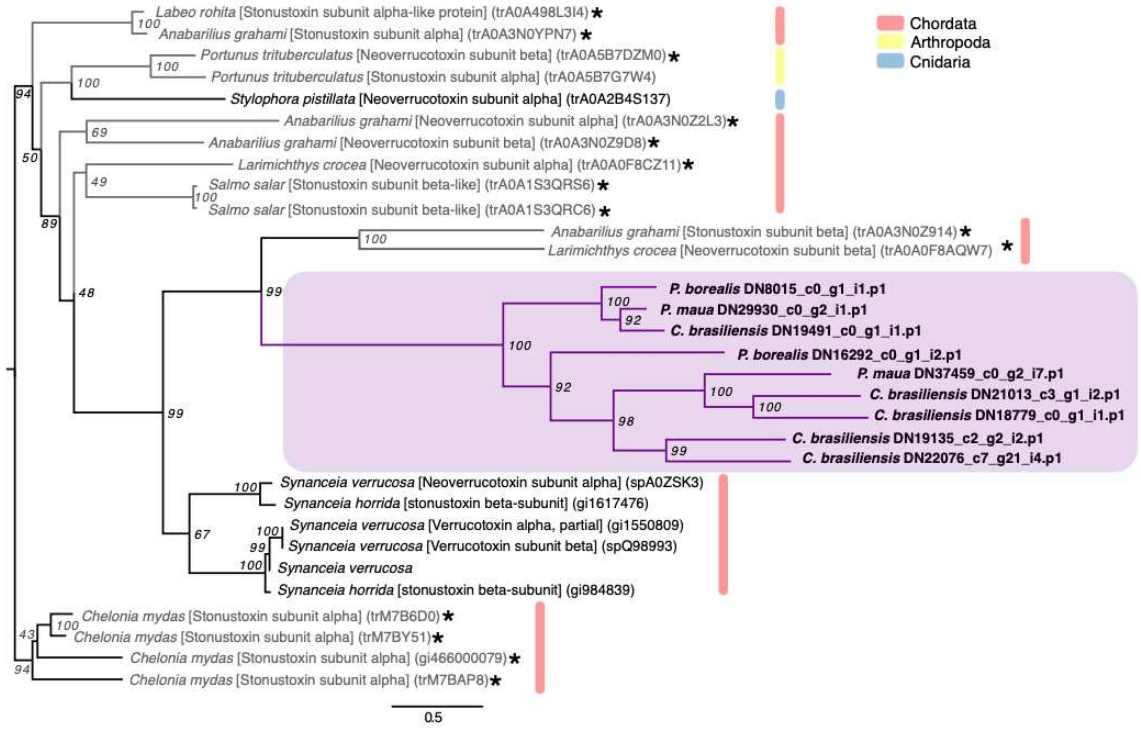

Figure 6. Phylogenetic tree of SNTX-like family sequences. Phylogeny was constructed using RAxML with the PROTGAMMAWAG option [154]. Bootstrap support based on 500 rapid bootstrap replicates, and all support values are shown. Putative genes outlined in purple are from cerianthids sequences. Sequences in gray and starred are non-venomous representatives, and other colors are from other animal classes. Phylogeny modified from [81]. Tree is rooted with sequences from green sea turtle (*Chelonia mydas*).

Waprins are membrane-active toxins derived from snakes that act as antimicrobial proteins, which are used by venomous animals as a defense against microbial infections of their venom glands [102,103]. One sequence of a waprin-like toxin from *P. borealis* and two from *P. maua* were identified in the cerianthids.

2.2.4. Mixed function enzymes

Phospholipases hydrolyze phospholipids to fatty acids and lysophospholipids, which in venoms induced hemolysis [104,105], as well as tissue necrosis, inflammation, blood coagulation inhibition, and neuromuscular transmission blockage [88,105]. These lipases are found in many animal venoms, including cephalopods, insects, spiders, scorpions, and reptiles [88]. Phospholipase A2 (PLA2) is a common and often abundant enzyme in cnidarians venom that aids in prey capture and digestion, and appears to have antimicrobial activity [106]. PLA2 are the most diverse of the enzymatic toxins detected in cerianthids, with 18 total sequences. Of these, 16 phylogenetically form a cluster that includes a putative PLA2 from *P. variabilis* [47] and conodipine-M alpha chain toxin, which was derived from the Magician's cone snail (*Conus magus*) and inhibits the binding of isradipine to L-type calcium channels [107] (Figure 7). The other two genes from *C. brasiliensis* and *I. nocturnus* cluster with a PLA2 from the broadclub cuttlefish (*Sepia latimanus*). We additionally found three phospholipase-B toxins within *P. borealis*, *C. brasiliensis*, and *I. nocturnus* and five phospholipase-D toxins, four in *C. brasiliensis* and a single transcript in *P. borealis*. Phospholipase-D in particular is thought to contribute to the dermonecrotic effects of brown spider venoms [108].

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

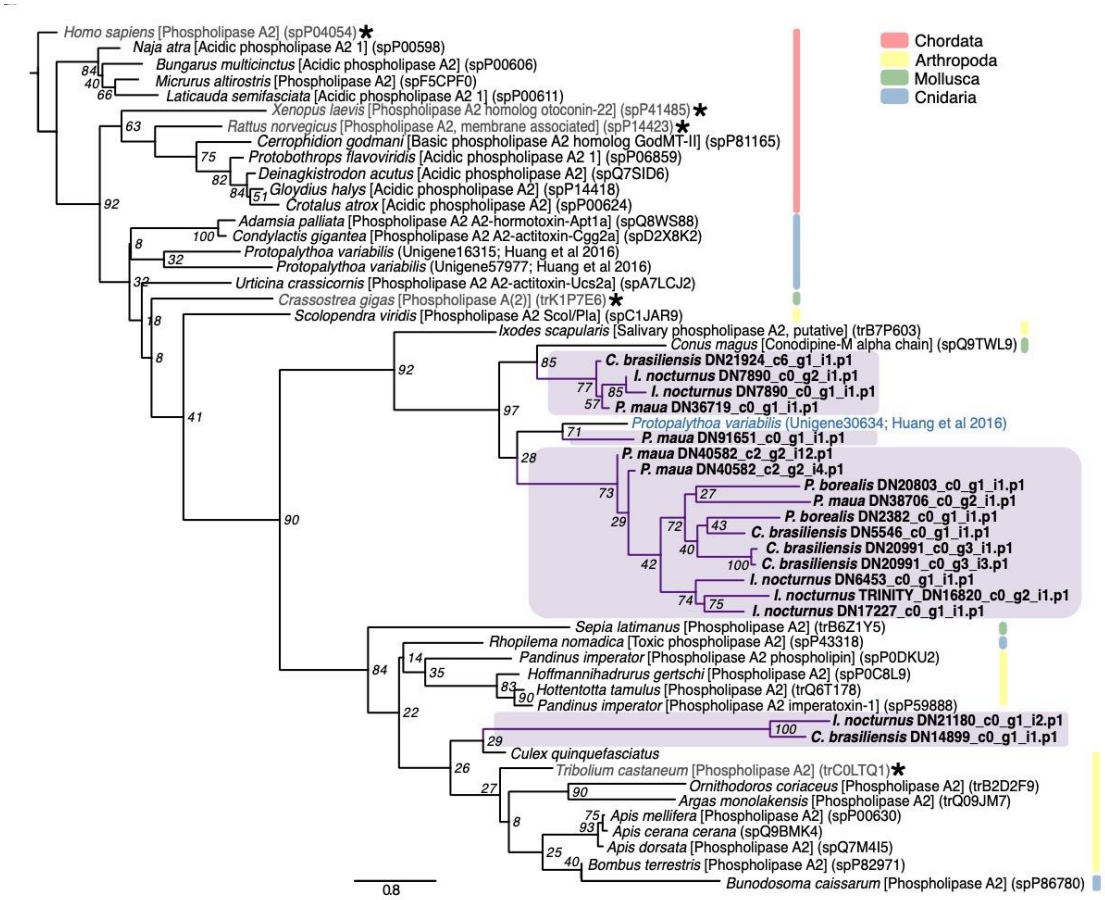

Figure 7. Phylogenetic tree of phospholipase A2 family sequences. Phylogeny was constructed using RAxML with the PROTGAMMAWAG option [154]. Bootstrap support based on 500 rapid bootstrap replicates, and all support values are shown. Putative genes outlined in purple are from cerianthids sequences. Sequences in gray and starred are non-venomous representatives, and other colors are from other animal classes. Phylogeny modified from [81].

2.2.5. Protease inhibitors

Kunitz-domain peptides both block ion channels and inhibit proteases, which can cause blood coagulation, fibrinolysis, and inflammation [109]. In sea anemones, kunitz-containing peptides are typically classified as type II potassium channel toxins, which cause paralysis by blocking potassium channels [25]. All four species have at least one kunitz-type serine protease inhibitor (total 11 across all four species), and *P. maua* specifically has a transcript that matches the sea anemone specific kunitz-containing toxin U-actitoxin-Avd3m, which, based on sequence similarity to other known toxins, may display hemolytic activity as well as potassium channel inhibition.

Three cerianthids, *P. borealis*, *C. brasiliensis*, and *P. maua* each contain a single transcript that corresponds to a ctenitoxin. Ctenotoxins are thyroglobulin type-1 protease inhibitors originally derived from the Brazilian spider (*Phoneutria nigriventer*), which inhibits cysteine proteases, aspartic proteases and metalloproteases [110].

2.2.6. Allergen and innate immunity

Several components from cnidarian stings have been known to cause immunological responses [14,111]. One common domain of these toxins is the CAP domain, which includes cysteine-rich secretory proteins (CRISPs), antigen 5 (Ag5), and pathogenesis-related 1 (Pr-1)

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

proteins [112]. These are found in many venomous taxa such as cephalopods, bloodworms, fireworms, scorpions, spiders, and reptiles [81,88,113], and are commonly found in cnidarians [22,43]. Function appears to vary by taxonomic group; in snakes, CAP proteins act as ion channel blockers and inhibit smooth muscle contraction [114], in cone snails as proteolytic compounds [115], and in hymenopterans as allergens [116]. The majority of CAP-domain cerianthid transcripts belong to a group called venom allergen proteins (total 31), though this is mainly driven by the number of genes present in *P. borealis* (12 sequences) and *C. brasiliensis* (14 sequences). Both species also have an additional CAP-domain (CRISP/Allergen/Pr-1) toxin. Multiple venom allergen proteins were also reported in the venom of the Pacific sea nettle (*Chrysaora fuscescens*) [43].

2.2.7. Venom auxiliary proteins

Venom auxiliary proteins are secreted in the venom gland to facilitate proper processing and stabilization. They can also work synergistically with other venom components to facilitate the spread of toxins after envenomation. One example is venom protein 302, originally derived from the scorpion *Lychas mucronatus* [117]. Each cerianthid has a putative single venom protein 302 match, two in the case of *C. brasiliensis*, but (weak) phylogenetic signals suggests that the cerianthids proteins are more closely related to an insulin-like growth factor-binding (IGLFP) protein from hexacorallian *S. pistillata* [118] (Supplementary Figure S9). Two venom protein 302 proteins were also identified in *P. variabilis* [47], and these zoanthid toxins formed a clade that is a sister group to non-venomous IGLFP-domain containing proteins in our study (Supplementary Figure S9). Venom 302-like peptides have been identified in *Z. natalensis* [49] and the proteomes of *N. nomurai* [38] and the cubozoan *C. fleckeri* [22]

Auxiliary proteins can also include various proteases that may facilitate diffusion of neurotoxins by breaking down the extracellular matrix in prey, and display cytolytic, gelatinolytic, caseinolytic, and fibrinolytic functions in cnidarians [119]. The most diverse auxiliary proteins in the four cerianthid transcriptomes match to astacin-like metalloproteases (M12A) with a total of 52 sequences between the four cerianthids. This includes transcripts with a close match to nematocyst expressed protein 6 (NEP-6), an astacin family metalloprotease previously reported from the starlet sea anemone *Nematostella vectensis* [120].

Additional metalloproteases, including neprilysin-like toxins (peptidase_M13_N domain), also found in the venom of *Cyanea capillata* [41], and glutaminyl-peptide cyclotransferases (peptidase_M28 domain) were also expressed within each species. Metalloprotease M12B containing domain proteases (zinc metalloproteinase-disintegrin and coagulation factor X-activating enzyme heavy chain) are also found in all four cerianthid species (13 total), but are categorized as hemostatic and hemorrhagic toxins (Section 2.2.2.), since, in snake venoms, these toxins disrupt capillary activity [120]. M12B metalloproteases have also been found in the venoms of *N. nomurai* [38] and the hydrozoan *Olindias sambaquiensis* [122].

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

326 **Table 2.** Toxin families identified for each cerianthid species.

Toxin Family ID	Pfam Domain	Cebr	Isn	Pasb	Pasm
Neurotoxin (%)		**7.1**	**17.4**	**11.0**	**13.3**
332-1 propeptide toxin	ShK	0	0	1	0
Cysteine-rich venom protein	CAP	2	1	10	2
ShK-domain	ShK	6	10	3	8
Three-finger toxin	/	1	1	0	1
Turripeptide	Kazal_1	3	0	5	2
U-actitoxin-Avd9a	ShK	0	0	0	1
U33-theraphotoxin-Cg1b	/	0	0	1	0
Hemostatic and hemorrhagic toxin (%)		**37.3**	**30.4**	**41.2**	**33.3**
Beta-fibrinogenase mucrofibrase-3	Trysin	0	0	0	1
Blarina Toxin	Trysin	3	0	1	0
C-type lectin lectoxin	Lectin_C	6	2	3	1
Coagualtion factor X	Trypin	1	2	2	0
Coagulation factor V	F5_F8_type_C	2	1	6	3
Coagulation factor X-activating enzyme heavy chain	Pep_M12B_propep/Reprolysin	1	0	1	0
Galactose-specific lectin	Lectin_C	4	0	9	3
Ryncolin	Fibrinogen_C	8	3	8	6
Snaclec	Lectin_C	2	1	3	0
Snake venom 5'-nucleotidase	5_nucleotid_C	1	0	1	0
Snake venom serine proteinase	Trypsin	0	0	0	1
Snake venom VEGF	PDGF	0	1	1	0
Thrombin-like enzyme	Trypsin	1	0	3	0
Thyrostimulin	DAN	1	0	0	0
Veficolin-1	Collagen	14	2	9	5
Venom peptide isomerase heavy chain	Trypsin	2	0	1	0
Venom prothrombin activator (F5/F8 type C)	F5_F8_type_C	6	3	15	4
Venom prothrombin activator (Trypsin)	Trypsin	9	5	8	7
Zinc metalloproteinase-disintegrin	Pep_M12B_propep/Reprolysin	2	1	4	4
Membrane-active (%)		**3.6**	**0**	**4.4**	**5.7**
DELTA-thalatoxin-Avl2a	MAPF	0	0	1	1
Jellyfish Toxin	/	1	0	2	1
Millepora cytotoxin	DERM	0	0	2	0
Stonutoxin/Neoverrucotoxin	/	5	0	2	2
Waprin	WAP	0	0	1	2
Mixed function enzyme (%)		**20.7**	**21.7**	**12.6**	**16.2**
Acetylcholinesterase	COesterase	5	2	3	3
Gilatoxin	Trypsin	0	1	0	0
L-amino-acid oxidase	Amino_oxidase	6	1	4	3
Peroxiredoxin	AhpC-TSA	0	0	1	1
Phospholipase-A2/Conodpine	Phospholip_A2	5	6	2	5
Phospholipase-B	Phospholip_B	1	1	1	0
Phospholipase-D	/	4	0	1	0
Putative endothelial lipase	Lipase	5	1	3	2
Putative lysosomal acid lipase/cholesteryl ester hydrolase	Abhydro_lipase/Abhydrolase_1	4	3	3	3
Trehalase	Trehalase	0	0	1	0
Venom phosphodiesterase	Phosphodiest	5	0	4	0
Protease inhibitor (%)		**2.4**	**4.3**	**2.7**	**2.9**
Kunitz-type serine protease inhibitor	Knuitz_BPTI	3	3	4	1
U-actitoxin-Avd3m	Knuitz_BPTI	0	0	0	1
U24-ctenitoxin-Pn1a	Thyroglobin_1	1	0	1	1
Allergen and innate immunity (%)		**12.7**	**2.2**	**13.2**	**5.4**
CRISP/Allergen/PR-1	CAP	1	0	1	0
Venom allergen	CAP	14	2	12	3
Venom phosphatase	His_Phos_2	1	1	1	0
Venom protease	Trysin	1	0	3	3
Venom serine carboxypeptidase	Peptidase_S10	0	0	1	0
Venom serine protease	Trysin	6	1	6	3
Techylectin-like	Fibrinogen_C	1	0	0	1

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Auxiliary protein (%)		**14.8**	**20.2**	**14.8**	**19.0**
Astacin-like metalloprotease toxin	Astacin	6	5	8	9
Cystatin	Cystatin	0	0	0	1
Glutaminyl-peptide cyclotransferase	Peptidase_M28	1	1	1	1
Hyaluronidase	Glyco_hydro_56	4	0	0	0
Nematocyst expressed protein	Astacin	6	3	11	6
Neprilysin	Peptidase_M13_N	1	1	3	0
Reticulocalbin	EF-hand_7	5	3	3	2
Venom protein 302	IGFBP	2	1	1	1
TOTAL		**169**	**69**	**182**	**105**
Unknown		*25*	*5*	*24*	*11*

3. Discussion

In this study we assembled de-novo transcriptomes of four members of Ceriantharia: *C. brasiliensis*, *I. nocturnus*, *P. borealis*, and *P. maua*, with BUSCO scores between 88.1-97.9% completeness (Table 1). From these transcriptomes, we identified a total of 525 venom-like genes between all four species using our customized bioinformatic pipeline, which are sorted into 135 clusters (124 orthologous clusters and 12 single-copy gene clusters) (Supplementary Figure S11). The venom-like gene profiles of the four cerianthids are similar in composition and generalized biological function, though the annotated number of toxin-like genes within each species is highly variable (69-182). Our four cerianthid toxin profiles are similar to previous transcriptome-based venom profiles for cnidarians, including the prevalence of ShK-domain containing toxins (e.g. [22,38,46,54]). While each species has a diversity of toxins within each of the seven functional categories, all toxin profiles were dominated by hemostatic and hemorrhagic toxins (30.4%-40.3%), mixed function enzymes (12.4-21.7%) and auxiliary venom proteins/peptides (14.5%-20.3%) followed by neurotoxins (7.2-17.4%), allergen and innate immunity toxins (2.2-12.9%), protease inhibitors (2.4-4.3%), and membrane-active toxins (0-5.7%). It should be noted that many of these toxins may have alternative or additional molecular functions, and the presented categorization only represents broad patterns based on previous studies on animal venoms. There was also a significant proportion of "unknown" toxins from each species within each transcriptome assembly (Table 2, Figure 2). Given that this is the first survey of putative toxins in this subclass within an already understudied group, it is unclear if these unknowns are potential novel venom-like transcripts or potential artifacts of assembly and annotation.

Some of the most common families we identified are common in anthozoan venoms, including PLA2, metalloproteases, serine proteases, and kunitz-domain protease inhibitors [11,43,51]. Several of the less common venom-gene families identified in cerianthids have also been identified in the transcriptomes of colonial zoanthids [47-49], another understudied group of anthozoans, including turripeptides, three-finger toxins, and venom protein 302 toxins (Figure 4; Supplementary Figure S4,S9), as well as snake venom VEGF toxins. However, the phylogenetic evidence for the majority of these candidate toxins is weak due to clustering with non-venomous taxa and/or low bootstrap scores. As mentioned above, several of these toxin groups have been identified in other cnidarian groups, including turripeptides [22,82,83] and venom protein 302 [22,38]. It is unclear if the similarities of these less common toxin families between zoanthid and cerianthid toxins are due to shared biology/evolutionary history or an artifact of the relatively limited dataset for cnidarians.

While membrane-active or pore-forming toxins are common in most cnidarian venoms [123], we had not expected to capture putative toxins in the jellyfish toxin family (also called CaTX/CrTX toxin family) in three of the four cerianthids species, given that these toxins are primarily found in medically relevant cubozoan venoms (Figure 6). In an ecological context, these highly potent toxins likely allow box jellyfish to capture fish [124,125]; while the diet of cerianthids remains fairly ambiguous, it is unlikely they capture fish as prey. Toxins from this family have

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

366 previously been identified in other anthozoan species through genomic and transcriptomic studies
367 [40,51,126], but to the best of our knowledge, these toxins have never been detected through
368 proteomic methods in anthozoans [40]. Given that these toxins are present in multiple cerianthids
369 (including two paralogs within *P. borealis*), these toxins are good candidates for proteomic analysis
370 and potentially functional characterization.

371 Because cerianthids group within the class Anthozoa, it is interesting that several toxins
372 groups commonly reported in anthozoans were absent from all four cerianthid species. For
373 example, we expected to find a diverse set of low molecular weight neurotoxins, such as sea
374 anemone sodium (Na+) channel toxins, potassium (K+) channel toxins, small cysteine-rich peptides
375 (SCRiPs), sodium-selection acid-sensing ion channel (ASICs) inhibitors, and nonselective cation
376 channel (TRPV1) inhibitors [11,25,30,127]. However, the four cerianthids transcriptomes contained
377 relatively low numbers of neurotoxins in general, and only a single transcript from *P. maua* closely
378 matched a sea anemone type I K+ channel toxin (Table 1). Additionally, actinoporin-like sequences
379 are often found in sea anemones and other organisms [93,123], but only two actinoporin-like
380 sequences were found in *P. borealis* and *P. maua*, despite often being found in sea anemones. We
381 also found no evidence of small cysteine-rich peptides (SCRiPS), neurotoxins with eight conserved
382 cysteine residues that cause paralysis in zebrafish (*Danio rerio*) [128], which were initially reported
383 in the corals *Orbicella faveolata* (as *Montastraea faveolata*), *Montipora capitata*, and *Acropora millepora*
384 [129]. The vast majority of candidate toxins containing ShK domains did not have a close match to
385 any toxin in the Tox-Prot database, but in 22 sequences we could confidently determine the six
386 cysteine residue patterns characteristic of ShK domains (Supplementary Figure S3). The exponential
387 increase in ShK domain peptides found in anthozoans prompted a recent sequence-function study
388 of the superfamily [130], and cerianthid ShK-domain toxins may represent additional structural
389 scaffolds with novel function for further study.

390 In general, our findings contrast with the previously observed pattern that anthozoan
391 venoms are typically neurotoxin-rich while medusozoan venoms are dominantly enzymatic. The
392 venoms of anthozoans and medusozoans have been broadly reported to be distinct, with
393 hydrozoans, scyphozoan, and cubozoan venoms being dominated by larger cytolytic proteins and
394 anthozoans by low molecular weight neuropeptides [26,40,83]. However, this pattern is based on
395 highly biased taxonomic data, as mentioned above [27]. Even though a greater diversity of
396 enzymatic-like genes is present within the four cerianthid transcriptomes, it is possible the level of
397 protein expression could shift towards a smaller subset of toxins dominating the venom
398 composition, and therefore overall venom function. For example, it has been shown in *S. haddoni*
399 that even when more enzymatic toxin-like sequences are present in the transcriptome, the
400 expression of neurotoxins is greater overall in milked venom (i.e. the proteomic level) [46]. Thus,
401 future quantitative gene expression and proteomic studies are needed to provide a more holistic
402 understanding of both single toxin and whole venom function in these species.

403 Because the phylogenetic placement of Subclass Ceriantharia remains unclear, it is difficult
404 to interpret the evolutionary context of their venom profile within Anthozoa. For instance, if
405 Ceriantharia is sister to the Hexacorallia, that suggests that the expansion of neuropeptide toxins
406 occurred after the divergence of Ceriantharia, possibly through extensive gene duplications
407 [52,126,131]. Neurotoxins in sea anemones are important because they are sessile animals, and may
408 be critical to deterring predators [132]. Because cerianthids can fully contract into their tubes, they
409 have a distinct means of protecting themselves from predators in contrast to sea anemones which
410 cannot fully retract their bodies, which may ease the selective pressure to diversify or maintain
411 defensive toxins. If Ceriantharia is instead sister to Hexacorallia + Octcorallia, families such as the
412 jellyfish toxins may have been present in the last common ancestor and subsequently lost in the
413 other anthozoan lineages. Additionally, as noted above, cerianthids often have a long-lasting
414 pelagic larval stage. There is a general consensus that the composition and function of toxins

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reflects the ecological utility of that venom [133], thus, the increased time in the pelagic environment in the larval stage likely exposes cerianthids to different sets of potential predators and prey, resulting in different selection pressures driving venom composition and function. We can only speculate on the role of these various venom components and overall venom function in the ecological interactions of these animals until additional molecular studies are completed [27,134].

One interesting outcome is the difference in the number of venom-like putative protein coding transcripts found in *I. nocturnus* compared to the other three species (69 compared to 169, 182, 105). As this species is the only representative of the family Arachnactidae, this may be evidence of evolutionary difference compared to the family Cerianthidae, which is corroborated by morphology and traditionally accepted [73]. At the ecological level, the species *I. nocturnus*, as its name indicates, is nocturnal and thus increases its activity at night. This may indicate different needs in relation to predation and prey capture compared to species active during the day. For instance, species of the family Arachnactidae show considerable concentrations of green fluorescent protein [135], which can be an important mechanism of prey capture at night [136]. This may relax the selective pressures, or potentially the available metabolic energy, to sustain a large, complex toxin arsenal, and therefore result in the lower number of venom-like genes identified in our study.

While our findings suggest several interesting patterns about presence and absence of certain cerianthid venom components, there are some limitations to exploring the venom profiles of understudied species. Previous studies have shown that cnidarian transcriptomes often yield a larger diversity of putative toxin sequences than a combined transcriptome-proteome approach (e.g. [46,53,54,126]). This difference may be reflective of the state of the animal when collected; animals that have recently fired their stinging cells will likely express more venom-like genes as venom is being synthesized for developing nematocysts [46]. Consequently, animals that have not discharged their stinging cells recently may have a lower than expected expression of toxin-like sequences. There are also often issues using de-novo assemblies for venom gene discovery, including high false discovery rate or inability to annotate novel venom genes [137,138]. For instance, even though no membrane-active toxins were detected in *I. nocturnus*, it is unlikely that there are truly no toxins with this function, especially given their ubiquity in cnidarians [139]. Our study also focused on candidate transcripts that contained full ORFs (stop and start codon), which likely decreased the diversity of toxin-like gene candidates. The set of venom-like genes we present here are viewed as an initial step into exploring the diversity of the toxin peptides and proteins within a poorly studied cnidarian group.

We present the first sequence-based analysis of venom-like genes within the Subclass Ceriantharia. The four species of cerianthids expressed over 500 novel toxin-like genes that are functionally and structurally diverse. While the overall functional profiles are similar to other transcriptomic studies of cnidarians, many common anthozoan toxin families are not present in our study. This could have notable implications both for the evolution of venom genes with Anthozoan as well as ecological utility of candidate toxins within this specific anthozoan lineage. Furthermore, the addition set of ShK-domain, as well as kunitz-domain containing toxins, shows that cerinathid toxins provide potential candidates for therapeutic study. We hope that these new data will be utilized to further explore the diversity and function of these venom proteins and peptides.

4. Materials and Methods

4.1. Tissue collection, RNA extraction, next-gen sequencing, and transcriptome assembly

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Four species were used in the current study. The species C. brasiliensis and I. nocturnus were obtained in São Sebastião, São Paulo, Brazil while SCUBA Diving. The P. borealis specimen was purchased through Gulf Of Maine inc. (Pembroke, ME). The P. cf. maua specimen was purchased from an aquarium supplier and currently on exhibit at Discovery Place Science (Charlotte, NC). For each species, several (10+) tentacles were collected from each organism after acclimating them to aquariums for 48 hours or longer. Tissues were flash frozen in liquid nitrogen or stored in RNA later in -80°C. Total RNA was extracted using the RNAqueous Total RNA Isolation Kit from Thermo Fisher Scientific (Massachusetts, USA). RNA was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher). High throughput Sequencing was done on an Illumina HiSeq at the DHMRI (Kannapolis, NC). Total RNA was quantitated using the Quant-iT RiboGreen RNA Assay Kit (Thermo Fisher) and RNA integrity assessed using the Agilent Bioanalyzer. RNA sequencing libraries were generated using the Illumina TruSeq RNA Library Prep RNA Kit following the manufacturer's protocol and quantitated using qPCR and fragments visualized using an Agilent Bioanalyzer. Libraries were combined in equimolar amounts onto one flow cell for a 125 bp paired end sequencing run on the Illumina HiSeq2500. Overall quality of the sequencing run evaluated using FastQC [140]. Transcriptome assembly was done using the de novo assembly program Trinity v2.2 [74]. Transcriptome completeness was determined using the program BUSCO v3 [141].

4.2. Bioinformatic analysis and venom annotation

For the custom annotation pipeline, protein-coding regions were predicted from assembled transcriptomes using TransDecoder v5.5.0, minimum set to 50 (https://transdecoder.github.io) [142]. Using blastp from NCBI BLAST+ v.2.8.1 [143,144] with an e-value cutoff of 0.001, all transcripts were searched against 1) proteins and toxins from the Tox-prot animal venom annotation database ([57], downloaded March 2019), and 2) all cnidarian toxins and proteins from the Protein database on NCBI ("Cnidaria AND ((Toxin) OR (Venom))," downloaded March 2019). Additionally, predicated protein-coding regions were searched using hmmsearch with an e-value cutoff of 0.001 from HMMER 3.1b2 [145,146] against hidden markov model (HMM) profiles from alignments of 20 venom protein classes. HMM were modified from those used in a transcriptomic study on the venom of bloodworms [81] by supplementing several cnidarian specific toxins within respective venom protein families. Additionally, four cnidarian-specific pore-forming venom families were added based on annotations from VenomZone (venomzone.expasy.org, accessed March 2018): Actinoporin sea anemone subfamily, jellyfish toxin family, cnidaria small cysteine-rich protein (SCRiP) family and MACPF-domain toxins. The results from all three searches were combined and all complete coding sequences used for downstream analysis. Venoms are secreted proteins and peptides, thus signal peptides were predicted using the SignalP v5.0 server (https://services.healthtech.dtu.dk/service.php?SignalP-5.0) [147]. Redundant sequences were clustered using CD-HIT v.4.6.8 with a cutoff of 0.95 [148,149]. A reciprocal search using blastp was used with an e-value cutoff of 1e-5 against Tox-Prot animal venom database and the NCBI non-redundant protein sequences (nr) database (downloaded March 2019), as well as a hmmsearch search with an evalue cutoff of 1e-5 against Pfam (downloaded March 2019) [77].

The results were manually curated to confirm that BLASTp annotations from ToxProt matched the detected venom domain from Pfam [76,77]. In addition, several toxins were not identified from ToxProt that were from NCBI database (e.g. three-finger toxin W-IV-like (NCBI Reference Sequence: XP_015758456.1), 332-1 secreted propeptoide (GenBank: AKU77030.1). Candidates were considered "unknown" and not used for further analysis if there was no match to a protein from Tox-Prot, the best match from NCBI was an uncharacterized or predicated protein, and no toxin domain was detected. The final list of candidate toxins was classified into protein families, molecular function (based on annotation from UniProtKB/Swiss-Prot) [150], and putative biological function. The results were visualized using the PieDonut via the webr package v.0.1.2

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(https://cardiomoon.github.io/webr/) in R v3.6.2 [151] within Rstudio v1.0.153 [152] and final figures constructed in Inkscape v1.0beta2 (inkscape.org).

4.3. Phylogenetic analysis of select gene families

For select toxin families, gene trees were constructed using a representative set of venomous and non-venomous proteins for each protein family, modified from phylogenetic analyses in von [80] and [47]. Candidate cerianthids toxins and were aligned using the L-INS-I algorithm in MAFFT v7.312 [153]. Maximum likelihood phylogenies were constructed using RAxML v8.2.12 [154] under the PROTGAMMA + WAG model and branch support calculated using 500 rapid bootstrap replicates (-x). Trees were visualized using FigTree v1.4.4 (https://github.com/rambaut/figtree) and final figures constructed in Inkscape v1.0beta2 (inkscape.org).

4.4. Availability of supporting data

Raw reads used to construct the transcriptomes used in this analysis have been deposited under the SRA bioproject PRJNA633022, specifically SRR11802642 (*C. brasiliensis*), SRR11802641 (*I. nocturnus*), SRR11802643 (*P. borealis*), and SRR11802640 (*P. maua*) accessions.

Supplementary Materials: Figure S1: Bioinformatic pipeline for the annotation of venom-like genes for four cerianthid transcriptomes, Figure S2-S10: Phylogenetic relationships between several toxin gene families and putative cerianthid sequences, Figure S11: Orthologous gene clusters of the putative venom-like genes for all four cerianthids, Table S1: Annotation table for putative venom-like genes for four cerianthid species (Excel).

Author Contributions: SNS, JM and AMR obtained samples, JM and SNS extracted RNA, JM assembled transcriptomes, AMLK performed the analysis and annotation of toxins and wrote the initial draft. AMLK, SNS, JM and AMR contributed to the manuscript.

Funding: São Paulo Research Foundation FAPESP 2015/24408-4, 2017/50028-0 (SPRINT), 2019/03552-0, CNPq (PROTAX) 440539/2015-3 and CNPq (Research Productivity Scholarship) 301293/2019-8 to SNS. SPRINT award from UNC Charlotte to JM and AMR.

Acknowledgments: We are grateful to Elliot Provance, Drs Marymegan Daly, André C. Morandini, Paulyn Cartwright, Marcelo V. Kitahara, Joe Ryan, and Melissa Debiasse for assisting with obtaining materials and/or discussions about the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cartwright, P.; Collins, A. Fossils and phylogenies: integrating multiple lines of evidence to investigate the origin of early major metazoan lineages. *Integr. Comp. Biol.* **2007**, *47*, 744–51. https://doi.org/10.1093/icb/icm071.
2. Kayal, E.; Bentlage, B.; Sabrina Pankey, M.; Ohdera, A. H.; Medina, M.; Plachetzki, D. C.; Collins, A. G.; Ryan, J. F. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol. Biol.* **2018**, *18*. https://doi.org/10.1186/s12862-018-1142-0.
3. Mariscal, R.N. *Nematocysts.* In *Coelenterate Biology: Reviews and New Perspectives*; Muscatine, L., Lenhoff, H.M., Eds.; Academic Press: New York, USA, 1974; pp. 129–178.
4. Fautin, D. G. Structural diversity, systematics, and evolution of cnidae. *Toxicon* **2009**, *54*, 1054–1064. https://doi.org/10.1016/j.toxicon.2009.02.024.
5. Bigger, C. H. Interspecific and intraspecific acrorhagial aggressive behavior among sea anemones: a recognition of self and not-self. *Biol. Bull.* **1980**, *159*, 117-134.
6. Williams, R. B. Acrorhagi, catch tentacles and sweeper tentacles: a synopsis of 'aggression'of actiniarian and scleractinian Cnidaria. In *Coelenterate Biology: Recent Research on Cnidaria and Ctenophora;* Williams R.B., Cornelius P.F.S., Hughes R.G., Robson E.A., Eds. Developments in Hydrobiology Springer: Dordrecht, Netherlands, 1991; Volume 66, pp. 539-545.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7. Macrander, J.; Brugler, M. R.; Daly, M. A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps. *BMC Genomics* **2015**, *16*. https://doi.org/10.1186/s12864-015-1417-4.
8. Columbus-Shenkar, Y. Y.; Sachkova, M. Y.; Macrander, J.; Fridrich, A.; Modepalli, V.; Reitzel, A. M.; Sunagar, K.; Moran, Y. Dynamics of venom composition across a complex life cycle. *eLife* **2018**, *7*, e35014. https://doi.org/10.7554/eLife.35014.
9. Schendel; Rash; Jenner; Undheim. The diversity of venom: the importance of behavior and venom system morphology in understanding its ecology and evolution. *Toxins* **2019**, *11*, 666. https://doi.org/10.3390/toxins11110666.
10. Turk, T.; Kem, W. R. The phylum Cnidaria and investigations of its toxins and venoms until 1990. *Toxicon* **2009**, *54*, 1031–1037. https://doi.org/10.1016/j.toxicon.2009.06.031.
11. Jouiaei, M.; Yanagihara, A.; Madio, B.; Nevalainen, T.; Alewood, P.; Fry, B. Ancient venom systems: a review on Cnidaria toxins. *Toxins* **2015**, *7*, 2251–2271. https://doi.org/10.3390/toxins7062251.
12. Burnett, J. W. Treatment of atlantic cnidarian envenomations. *Toxicon* **2009**, *54*, 1201–1205. https://doi.org/10.1016/j.toxicon.2009.02.020.
13. Tibballs, J. Australian venomous jellyfish, envenomation syndromes, toxins and therapy. *Toxicon* **2006**, *48*, 830–859. https://doi.org/10.1016/j.toxicon.2006.07.020.
14. Tibballs, J.; A. Yanagihara, A.; C. Turner, H.; Winkel, K. Immunological and toxinological responses to jellyfish stings. *Inflamm. Allergy Drug Targets* **2011**, *10*, 438–446. https://doi.org/10.2174/187152811797200650.
15. Badré, S. Bioactive toxins from stinging jellyfish. *Toxicon* **2014**, *91*, 114–125. https://doi.org/10.1016/j.toxicon.2014.09.010.
16. Richard, J.; Portier, P. Fascicule 95: recherches sur la toxine des coelentéres et les phénomènes d'anaphylaxie. In *Résultats des campagnes scientifiques du prince de Monaco*. **1936.**
17. Lane, C. E.; Dodge, E. The toxicity of *Physalia* nematocysts. *Biol. Bull.* **1956**, *115*, 219-226.
18. Tamkun, M. M.; Hessinger, D. A. Isolation and partial characterization of a hemolytic and toxic protein from the nematocyst venom of the Portuguese man-of-war, *Physalia physalis*. *BBA Protein Struct.* **1981**, *667*, 87-98.
19. Barnes, J. H. *Chironex fleckeri* and *Chiropsalmus quadrigatus*: morphological distinctions. *North Queensland Naturalist* **1965**, 32, 13-22.
20. Nagai, H. Recent progress in jellyfish toxin study. *J. Health Sci.* **2003**, *49*, 337–340. https://doi.org/10.1248/jhs.49.337.
21. Brinkman, D. L.; Konstantakopoulos, N.; McInerney, B. V.; Mulvenna, J.; Seymour, J. E.; Isbister, G. K.; Hodgson, W. C. *Chironex fleckeri* (box jellyfish) venom proteins: expansion of cnidarian toxin family that elicits variable cytolytic and cardiovascular effects. *J. Biol. Chem.* **2014**, *289*, 4798–4812. https://doi.org/10.1074/jbc.M113.534149.
22. Brinkman, D. L.; Jia, X.; Potriquet, J.; Kumar, D.; Dash, D.; Kvaskoff, D.; Mulvenna, J. Transcriptome and venom proteome of the box jellyfish *Chironex fleckeri*. *BMC Genomics* **2015**, *16*, 407. https://doi.org/10.1186/s12864-015-1568-3.
23. Tibballs, J. Australian Chirodropid Cubozoan Jellyfish Envenomation. In *Clinical Toxinology in Australia, Europe, and Americas*; Vogel, C. W., Seifert, S. A., Tambourgi, D. V., Eds.; Springer: Dordrecht, Netherlands, 2018; pp 331–354. https://doi.org/10.1007/978-94-017-7438-3_78.
24. Goodwin, M. H.; Telford, M. The nematocyst toxin of *Metridium*. *Biol. Bull.* **1971**, *140*, 389–399. https://doi.org/10.2307/1540276.
25. Madio, B.; King, G. F.; Undheim, E. A. B. Sea anemone toxins: a structural overview. *Mar. Drugs* **2019**, *17*, 325. https://doi.org/10.3390/md17060325.
26. Hessinger, D. A. Nematocyst venoms and toxins. In *The biology of nematocysts;* Hessinger, D. A., Lenhoff., H. M., Eds.; Academic Press: California, USA, 1988; pp. 333-368.
27. Ashwood, L. M.; Norton, R. S.; Undheim, E. A. B.; Hurwood, D. A.; Prentis, P. J. Characterizing functional venom profiles of anthozoans and medusozoans within their ecological context. *Mar. Drugs* **2020**, *18*, 202. https://doi.org/10.3390/md18040202.
28. Rocha, J.; Peixe, L.; Gomes, N. C. M.; Calado, R. Cnidarians as a Source of New Marine Bioactive Compounds—An Overview of the Last Decade and Future Steps for Bioprospecting. *Mar. Drugs* **2011**, *9*, 1860–1886. https://doi.org/10.3390/md9101860.
29. Thangaraj, S; Bragadeeswaran, S.; Gokula, V. Bioactive compounds of sea anemones: a review. *Int J Pept Res Ther* **2019**, *25*, 1405–1416. https://doi.org/10.1007/s10989-018-9786-6.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

30. Prentis, P. J.; Pavasovic, A.; S. Norton, R. Sea Anemones: Quiet Achievers in the Field of Peptide Toxins. *Toxins* **2018**, *10*, 36. https://doi.org/10.3390/toxins10010036.
31. Castañeda, O.; Sotolongo, V.; Amor, A.M.; Stöcklin, R.; Anderson, A.J.; Harvey, A.L.; Engström, Å.; Wernstedt, C.; Karlsson, E. Characterization of a potassium channel toxin from the Caribbean Sea anemone *Stichodactyla helianthus*. *Toxicon* **1995**, *33*, 603-613.
32. Chandy, K. G.; Norton, R. S. Peptide blockers of Kv1. 3 channels in T cells as therapeutics for autoimmune disease. *Curr. Opin. Chem. Biol.* **2017**, *38*, 97-107. https://doi.org/10.1016/j.cbpa.2017.02.015.
33. Tarcha, E. J.; Olsen, C. M.; Probst, P.; Peckham, D.; Muñoz-Elías, E. J.; Kruger, J. G.; Iadonato, S. P. Safety and pharmacodynamics of dalazatide, a Kv1. 3 channel inhibitor, in the treatment of plaque psoriasis: A randomized phase 1b trial. *PloS One* **2017**, *12*, e0180762. https://doi.org/10.1371/journal.pone.0180762.
34. Beeton, C., W Pennington, M., & S Norton, R. Analogs of the sea anemone potassium channel blocker ShK for the treatment of autoimmune diseases. *Inflamm. Allergy Drug Targets* **2011**, *10*, 313-321. https://doi.org/10.2174/187152811797200641.
35. Norton, R. S., Pennington, M. W., & Beeton, C. Case study 2: transforming a toxin into a therapeutic: the sea anemone potassium channel blocker ShK toxin for treatment of autoimmune diseases. In *Venoms to drugs: venom as a source for the development of human* therapeutics; King GF, Eds.; Royal Society of Chemistry: Cambridge, UK, 2015; pp. 255–274. https://doi.org/10.1039/9781849737876-00255.
36. Chen, X.; Leahy, D.; Van Haeften, J.; Hartfield, P.; Prentis, P.J.; van der Burg, C.A.; Surm, J.M.; Pavasovic, A.; Madio, B.; Hamilton, B.R.; King, G.F. A versatile and robust serine protease inhibitor scaffold from *Actinia tenebrosa*. *Mar. Drugs* **2019**, *17*, 701. https://doi.org/10.3390/md17120701.
37. Liao, Q.; Feng, Y.; Yang, B.; Lee, S. M.-Y. Cnidarian peptide neurotoxins: a new source of various ion channel modulators or blockers against central nervous systems disease. *Drug Discov.Today* **2019**, *24*, 189–197. https://doi.org/10.1016/j.drudis.2018.08.011.
38. Li, R.; Yu, H.; Xue, W.; Yue, Y.; Liu, S.; Xing, R.; Li, P. Jellyfish venomics and venom gland transcriptomics analysis of *Stomolophus meleagris* to reveal the toxins associated with sting. *J. Proteomics* **2014**, *106*, 17–29. https://doi.org/10.1016/j.jprot.2014.04.011.
39. Li, R.; Yu, H.; Yue, Y.; Liu, S.; Xing, R.; Chen, X.; Li, P. combined proteomics and transcriptomics identifies sting-related toxins of jellyfish *Cyanea nozakii*. *J. Proteomics* **2016**, *148*, 57–64. https://doi.org/10.1016/j.jprot.2016.07.023.
40. Rachamim, T.; Morgenstern, D.; Aharonovich, D.; Brekhman, V.; Lotan, T.; Sher, D. the dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. *Mol. Biol. Evol.* **2015**, *32*, 740–753. https://doi.org/10.1093/molbev/msu335.
41. Liu, G.; Zhou, Y.; Liu, D.; Wang, Q.; Ruan, Z.; He, Q.; Zhang, L. global transcriptome analysis of the tentacle of the jellyfish *Cyanea capillata* using deep sequencing and expressed sequence tags: insight into the toxin- and degenerative disease-related transcripts. *PLoS One* **2015**, *10*, e0142680. https://doi.org/10.1371/journal.pone.0142680.
42. Macrander, J.; Broe, M.; Daly, M. Tissue-specific venom composition and differential gene expression in sea anemones. *Genome Biol. Evol.* **2016**, *8*, 2358–2375. https://doi.org/10.1093/gbe/evw155.
43. Ponce, D.; Brinkman, D.; Potriquet, J.; Mulvenna, J. tentacle transcriptome and venom proteome of the pacific sea nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa). *Toxins* **2016**, *8*, 102. https://doi.org/10.3390/toxins8040102.
44. Wang, C.; Wang, B.; Wang, B.; Wang, Q.; Liu, G.; Wang, T.; He, Q.; Zhang, L. Unique Diversity of Sting-Related Toxins Based on Transcriptomic and Proteomic Analysis of the Jellyfish *Cyanea Capillata* and *Nemopilema Nomurai* (Cnidaria: Scyphozoa). *J. Proteome Res.* **2018**, *18*, 436-448. https://doi.org/10.1021/acs.jproteome.8b00735.
45. Lewis Ames, C.; Ryan, J. F.; Bely, A. E.; Cartwright, P.; Collins, A. G. A New Transcriptome and Transcriptome Profiling of Adult and Larval Tissue in the Box Jellyfish *Alatina alata*: An Emerging Model for Studying Venom, Vision and Sex. *BMC Genomics* **2016**, *17*, 650. https://doi.org/10.1186/s12864-016-2944-3.
46. Madio, B.; Undheim, E. A. B.; King, G. F. Revisiting Venom of the Sea Anemone *Stichodactyla haddoni*: Omics Techniques Reveal the Complete Toxin Arsenal of a Well-Studied Sea Anemone Genus. *J. Proteomics* **2017**, *166*, 83–92. https://doi.org/10.1016/j.jprot.2017.07.007.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

47. Huang, C.; Morlighem, J.-É. R.; Zhou, H.; Lima, É. P.; Gomes, P. B.; Cai, J.; Lou, I.; Pérez, C. D.; Lee, S. M.; Rádis-Baptista, G. the transcriptome of the zoanthid *Protopalythoa variabilis* (Cnidaria, Anthozoa) predicts a basal repertoire of toxin-like and venom-auxiliary polypeptides. *Genome Biol. Evol.* **2016,** *8*, 3045–3064. https://doi.org/10.1093/gbe/evw204.
48. Liao, Q.; Li, S.; Siu, S. W. I.; Yang, B.; Huang, C.; Chan, J. Y.-W.; Morlighem, J.-É. R. L.; Wong, C. T. T.; Rádis-Baptista, G.; Lee, S. M.-Y. Novel kunitz-like peptides discovered in the zoanthid *Palythoa caribaeorum* through transcriptome sequencing. *J. Proteome Res.* **2018**, *17*, 891–902. https://doi.org/10.1021/acs.jproteome.7b00686.
49. Liao, Q.; Gong, G.; Poon, T. C. W.; Ang, I. L.; Lei, K. M. K.; Siu, S. W. I.; Wong, C. T. T.; Rádis-Baptista, G.; Lee, S. M.-Y. combined transcriptomic and proteomic analysis reveals a diversity of venom-related and toxin-like peptides expressed in the mat anemone *Zoanthus natalensis* (Cnidaria, Hexacorallia). *Arch. Toxicol.* **2019**, *93*, 1745–1767. https://doi.org/10.1007/s00204-019-02456-z.
50. Rivera-de-Torre, E.; Martínez-del-Pozo, Á.; Garb, J. E. *Stichodactyla Helianthus*' de novo transcriptome assembly: Discovery of a new actinoporin isoform. *Toxicon* **2018**, *150*, 105–114. https://doi.org/10.1016/j.toxicon.2018.05.014.
51. Surm, J. M.; Smith, H. L.; Madio, B.; Undheim, E. A. B.; King, G. F.; Hamilton, B. R.; Burg, C. A.; Pavasovic, A.; Prentis, P. J. A process of convergent amplification and tissue-specific expression dominates the evolution of toxin and toxin-like genes in sea anemones. *Mol. Ecol.* **2019**, *28*, 2272–2289. https://doi.org/10.1111/mec.15084.
52. Sachkova, M. Y.; Singer, S. A.; Macrander, J.; Reitzel, A. M.; Peigneur, S.; Tytgat, J.; Moran, Y. The birth and death of toxins with distinct functions: A case study in the sea anemone *Nematostella*. *Mol. Biol. Evol.* **2019**, *36*, 2001-2012. https://doi.org/10.1093/molbev/msz132.
53. Ramírez-Carreto, S.; Vera-Estrella, R.; Portillo-Bobadilla, T.; Licea-Navarro, A.; Bernaldez-Sarabia, J.; Rudiño-Piñera, E.; Verleyen, J. J.; Rodríguez, E.; Rodríguez-Almazán, C. transcriptomic and proteomic analysis of the tentacles and mucus of *Anthopleura Dowii* Verrill, 1869. *Mar. Drugs* **2019**, *17*, 436. https://doi.org/10.3390/md17080436.
54. Mitchell, M. L.; Tonkin-Hill, G. Q.; Morales, R. A. V.; Purcell, A. W.; Papenfuss, A. T.; Norton, R. S. Tentacle transcriptomes of the speckled anemone (Actiniaria: Actiniidae: *Oulactis* Sp.): Venom-related components and their domain structure. *Mar. Biotechnol.* **2020**, *22*, 207–219. https://doi.org/10.1007/s10126-020-09945-8.
55. Zapata, F.; Goetz, F. E.; Smith, S. A.; Howison, M.; Siebert, S.; Church, S. H.; Sanders, S. M.; Ames, C. L.; McFadden, C. S.; France, S. C.; Daly, M.; Collins, A. G.; Haddock, S. H. D.; Dunn, C. W.; Cartwright, P. Phylogenomic analyses support traditional relationships within Cnidaria. *PLoS One* **2015**, *10*, e0139068. https://doi.org/10.1371/journal.pone.0139068.
56. Chang, E. S.; Neuhof, M.; Rubinstein, N. D.; Diamant, A.; Philippe, H.; Huchon, D.; P. Cartwright. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proc. Natl. Acad. Sci.* **2015**, *112*, 14912-14917. https://doi.org/10.1073/pnas.1511468112.
57. Jungo, F.; Bougueleret, L.; Xenarios, I.; Poux, S. The UniProtKB/Swiss-Prot Tox-Prot program: a central hub of integrated venom protein data. *Toxicon* **2012**, *60*, 551-557. https://doi.org/10.1016/j.toxicon.2012.03.010.
58. Schmidt, C.; Daly, N.; Wilson, D. Coral venom toxins. *Front. Ecol. Evol.* **2019**, *7*, 320.
59. Weber, J., Klug, M., & Tardent, P. Some physical and chemical properties of purified nematocysts of *Hydra attenuata* Pall. (Hydrozoa, Cnidaria). *Comp. Biochem. Physiol. B, Biochem.* **1987**, *88*, 855-862. https://doi.org/10.1016/0305-0491(87)90255-0.
60. Marchini, B.; De, L. N.; Mazzei, M.; Mariottini, G. L. A fast centrifuge method for nematocyst isolation from *Pelagia noctiluca* Forskal (Cnidaria: Scyphozoa). *Rivista di biologia* **2004**, *97*, 505-515.
61. Carrette, T.; Seymour, J. Cardiotoxic effects of venoms from *Chironex fleckeri* and *Chiropsalmus sp.* on an invertebrate model. *J. Venom Anim. Toxins* **2004**, *12*, 245-254. https://doi.org/10.1016/j.toxicon.2004.04.008.
62. Feng, J.; Yu, H.; Li, C.; Xing, R.; Liu, S,; Wang, L.; Cai, S; Li, P. Isolation and characterization of lethal proteins in nematocyst venom of the jellyfish *Cyanea nozakii* Kishinouye. *Toxicon* **2010**, *55*, 118-125. https://doi.org/10.1016/j.toxicon.2009.07.008.
63. Klug, M.; Weber, J.; Tardent, P. Hemolytic and toxic properties of *Hydra attenuata* nematocysts. *Toxicon* **1989**, *27*, 325-339. https://doi.org/10.1016/0041-0101(89)90180-3.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

64. Zenkert, C., Takahashi, T., Diesner, M. O., & Özbek, S. Morphological and molecular analysis of the *Nematostella vectensis* cnidome. *PloS One* **2011**, *6*, e22725. https://doi.org/10.1371/journal.pone.0022725.
65. Von Reumont, B. M. Studying smaller and neglected organisms in modern evolutionary venomics implementing RNASeq (transcriptomics)—A critical guide. *Toxins* **2018**, *10*, 29. https://doi.org/10.3390/toxins10070292.
66. Mariscal, R. N.; Conklin, E. J.; Bigger, C. H. The ptychocyst, a major new category of cnida used in tube construction by a cerianthid anemone. *Biol. Bull.* **1977**, *152*, 392-405. https://doi.org/10.2307/1540427.
67. Stampar, S. N.; Maronna, M. M.; Kitahara, M. V.; Reimer, J. D.; Beneti, J. S.; Morandini, A. C. Ceriantharia in current systematics: life cycles, morphology and genetics. In *The Cnidaria, past, present and future*; Goffredo S., Dubinsky Z., Eds.; Springer: Cham, Switzerland, 2016; pp. 61-72.
68. Quattrini, A.M.; Faircloth, B.C.; Dueñas, L.F.; Bridge, T.C.; Brugler, M.R.; Calixto-Botía, I.F.; DeLeo, D.M.; Foret, S.; Herrera, S.; Lee, S.M.; Miller, D.J. Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. *Mol. Ecol. Resour.* **2018**, *18*, 281-295. https://doi.org/10.1111/1755-0998.12736.
69. Daly, M.; Brugler, M. R.; Cartwright, P.; Collins, A. G.; Dawson, M. N.; Fautin, D. G.; France, S. C.; Mcfadden, C. S.; Opresko, D. M.; Rodriguez, E.; Romano, S. L.; Stake, J. L. The phylum Cnidaria: A review of phylogenetic patterns and diversity 300 years after Linnaeus*. *Zootaxa* **2007**, *1668*, 127–182. https://doi.org/10.11646/zootaxa.1668.1.11.
70. Stampar, S. N., Maronna, M. M., Kitahara, M. V., Reimer, J. D., & Morandini, A. C. Fast-evolving mitochondrial DNA in Ceriantharia: a reflection of Hexacorallia paraphyly?. *PLoS One* **2014**, *9*, e86612. https://doi.org/10.1371/journal.pone.0086612.
71. Stampar, S. N., Broe, M. B., Macrander, J., Reitzel, A. M., Brugler, M. R., & Daly, M. Linear mitochondrial genome in Anthozoa (Cnidaria): A case study in Ceriantharia. *Sci. Rep.* **2019**, *9*, 1-12. https://doi.org/10.1038/s41598-019-42621-z.
72. Bridge, D.; Cunningham, C. W.; DeSalle, R.; Buss, L. W. Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. *Mol. Biol. Evol.* **1995**, *12*, 679-689. https://doi.org/10.1093/oxfordjournals.molbev.a040246.
73. Kayal, E.; Roure, B.; Philippe, H.; Collins, A. G.; Lavrow, D. V. Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evol. Bio.l* **2013,** *13*, 5. https://doi.org/10.1186/1471-2148-13-5.
74. Stampar, S. N.; Morandini, A. C.; Branco, L. C.; Da Silveira, F. L.; Migotto, A. E. Drfting in the oceans: *Isarachnanthus nocturnus* (Cnidaria, Ceriantharia, Arachnactidae), an anthozoan with an extended planktonic stage. *Mar. Biol.* **2015**, *162*, 2161–2169. https://doi.org/10.1007/s00227-015-2747-0.
75. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; Chen, Z. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644. https://doi.org/10.1038/nbt.1883.
76. Bateman, A. The Pfam Protein Families Database. *Nucleic Acids Res.* **2004**, *32*, D138-D141. https://doi.org/10.1093/nar/gkh121.
77. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; Sonnhammer, E.L.L. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427-D432. https://doi.org/10.1093/nar/gky995.
78. Griselda, A. V. Molecular characterization of *Carukia barnesi* and *Malo* kingi, Cnidaria; Cubozoa; Carybdeidae. Doctoral dissertation, James Cook University, Queensland, Australia, 2009. Research Online at JCU. Available online: https://researchonline.jcu.edu.au/8218/ (accessed April 2020).
79. Rangaraju, S.; Khoo, K.K.; Feng, Z.P.; Crossley, G.; Nugent, D.; Khaytin, I.; Chi, V.; Pham, C.; Calabresi, P.; Pennington, M.W.; Norton, R.S. Potassium channel modulation by a toxin domain in matrix metalloprotease 23. *J. Biol. Chem.* **2010**, *285*, 9124-9136. https://doi.org/10.1074/jbc.M109.071266.
80. Krebs, H. C.; Habermehl, G. G. Isolation and structural determination of a hemolytic active peptide from the sea anemone *Metridium senile*. *Naturwissenschaften* **1987**, *74*, 395-396.
81. von Reumont, B. M.; Campbell, L. I.; Richter, S.; Hering, L.; Sykes, D.; Hetmank, J.; Jenner, R. A.; Bleidorn, C. A polychaete's powerful punch: venom gland transcriptomics of *Glycera* reveals a complex cocktail of toxin homologs. *Genome Biol. Evol.* **2014**, *6*, 2406–2423. https://doi.org/10.1093/gbe/evu190.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

82. Liang, H.; Jiang, G.; Wang, T.; Zhang, J.; Liu, W.; Xu, Z.; Zhang, J.; Xiao, L. an integrated transcriptomic and proteomic analysis reveals toxin arsenal of a novel antarctic jellyfish *Cyanea* Sp. *J. Proteomics* **2019**, *208*, 103483. https://doi.org/10.1016/j.jprot.2019.103483.
83. Jaimes-Becerra, A.; Gacesa, R.; Doonan, L. B.; Hartigan, A.; Marques, A. C.; Okamura, B.; Long, P. F. "Beyond primary sequence"—proteomic data reveal complex toxins in cnidarian venoms. *Integr. Comp. Biol.* **2019**, *59*, 777–785. https://doi.org/10.1093/icb/icz106.
84. Aguilar, M. B.; Flores-Torres, A.; Batista, C. V.; Falcón, A.; López-Vera, E.; de la Cotera, E. P. Structural characterization of five post-translationally modified isomorphs of a novel putative delta-conotoxin from the vermivorous snail *Conus delessertii* from the Mexican Caribbean sea. *Peptides* **2009**, *30*, 458–466. https://doi.org/10.1016/j.peptides.2008.12.005.
85. Gonzales, D. T. T.; Saloma, C. P. A bioinformatics survey for conotoxin-like sequences in three turrid snail venom duct transcriptomes. *Toxicon* **2014**, *92*, 66-74. https://doi.org/10.1016/j.toxicon.2014.10.003.
86. Kini, R. M. Evolution of three-finger toxins – a versatile mini protein scaffold. *Acta. Chim. Slov.* **2011**, *58*, 693–701.
87. Kuhn, P.; Deacon, A. M.; Comoso, S.; Rajaseger, G.; Kini, R. M.; Usón, I.; Kolatkar, P. R. The atomic resolution structure of bucandin, a novel toxin isolated from the malayan krait, determined by direct methods. *Acta. Crystallogr. D. Biol. Crystallogr.* **2000**, *56*, 1401–1407. https://doi.org/10.1107/S0907444900011501.
88. Fry, B. G.; Roelants, K.; Champagne, D. E.; Scheib, H.; Tyndall, J. D. A.; King, G. F.; Nevalainen, T. J.; Norman, J. A.; Lewis, R. J.; Norton, R. S.; Renjifo, C.; de la Vega, R. C. R. the toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genom. Hum. G.* **2009**, *10* , 483–511. https://doi.org/10.1146/annurev.genom.9.081307.164356.
89. Arlinghaus, F. T.; Eble, J. A. C-Type lectin-like proteins from snake venoms. *Toxicon* **2012**, *60*, 512–519. https://doi.org/10.1016/j.toxicon.2012.03.001.
90. OmPraba, G.; Chapeaurouge, A.; Doley, R.; Devi, K.R.; Padmanaban, P.; Venkatraman, C.; Velmurugan, D.; Lin, Q.; Kini, R.M. Identification of a novel family of snake venom proteins veficolins from *Cerberus rynchops* using a venom gland transcriptomics and proteomics approach. *J. Proteome Res.* **2010**, *9*, 1882-1893. https://doi.org/10.1021/pr901044x.
91. Kini, R.M. The intriguing world of prothrombin activators from snake venom. *Toxicon* **2005**, *45*, 1133-1145. https://doi.org/10.1016/j.toxicon.2005.02.019.
92. Balasubramanian, P. G.; Beckmann, A.; Warnken, U.; Schnölzer, M.; Schüler, A.; Bornberg-Bauer, E.; Holstein, T. W.; Özbek, S. Proteome of *Hydra* nematocyst. *J. Biol. Chem.* **2012**, *287*, 9672–9681. https://doi.org/10.1074/jbc.M111.328203.
93. Anderluh, G.; Maček, P. Dissecting the actinoporin pore-forming mechanism. *Structure* **2003**, *11*, 1312–1313. https://doi.org/10.1016/j.str.2003.10.007.
94. Violette, A.; Biass, D.; Dutertre, S.; Koua, D.; Piquemal, D.; Pierrat, F.; Stöcklin, R.; Favreau, P. Large-scale discovery of conopeptides and conoproteins in the injectable venom of a fish-hunting cone snail using a combined proteomic and transcriptomic approach. *J. Proteomics* **2012**, *75*, 5215–5225. https://doi.org/10.1016/j.jprot.2012.06.001.
95. Warren, W.C.; Hillier, L.W.; Graves, J.A.M.; Birney, E.; Ponting, C.P.; Grützner, F.; Belov, K.; Miller, W.; Clarke, L.; Chinwalla, A.T.; Yang, S.P *et al*. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **2008**, *453*, 175–183. https://doi.org/10.1038/nature06936.
96. García-Ortega, L.; Alegre-Cebollada, J.; García-Linares, S.; Bruix, M.; Martínez-del-Pozo, Á.; Gavilanes, J. G. the behavior of sea anemone actinoporins at the water–membrane interface. *BBA Biomembranes* **2011**, *1808*, 2275–2288. https://doi.org/10.1016/j.bbamem.2011.05.012.
97. Andrich, F.; Carnielli, J. B. T.; Cassoli, J. S.; Lautner, R. Q.; Santos, R. A. S.; Pimenta, A. M. C.; de Lima, M. E.; Figueiredo, S. G. A potent vasoactive cytolysin isolated from *Scorpaena Plumieri* scorpionfish venom. *Toxicon* **2010**, *56*, 487–496. https://doi.org/10.1016/j.toxicon.2010.05.003.
98. Kiriake, A.; Suzuki, Y.; Nagashima, Y.; Shiomi, K. proteinaceous toxins from three species of scorpaeniform fish (Lionfish *Pterois Lunulata*, Devil Stinger *Inimicus Japonicus* and Waspfish *Hypodytes Rubripinnis*): Close similarity in properties and primary structures to stonefish toxins. *Toxicon* **2013**, *70*, 184–193. https://doi.org/10.1016/j.toxicon.2013.04.021.
99. Whittington, C. M.; Papenfuss, A. T.; Locke, D. P.; Mardis, E. R.; Wilson, R. K.; Abubucker, S.; Mitreva, M.; Wong, E. S.; Hsu, A. L.; Kuchel, P. W.; Belov, K.; Warren, W. C. Novel venom gene discovery in the platypus. *Genome Biol.* **2010**, *11*, R95. https://doi.org/10.1186/gb-2010-11-9-r95.
100. Ueda, A.; Suzuki, M.; Honma, T.; Nagai, H.; Nagashima, Y.; Shiomi, K. Purification, properties and cdna cloning of neoverrucotoxin (NeoVTX), a hemolytic lethal factor from the stonefish *Synanceia*

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Verrucosa venom. *BBA Gen. Subjects* **2006**, *1760*, 1713–1722. https://doi.org/10.1016/j.bbagen.2006.08.017.

101. Yazawa, K.; Wang, J.-W.; Hao, L.-Y.; Onoue, Y.; Kameyama, M. Verrucotoxin, a stonefish venom, modulates calcium channel activity in guinea-pig ventricular myocytes: Verrucotoxin modulates cardiac calcium channels. *Br. J. Pharmacol.* **2009**, *151*, 1198–1203. https://doi.org/10.1038/sj.bjp.0707340.
102. Nair, D. G.; Fry, B. G.; Alewood, P.; Kumar, P. P.; Kini, R. M. antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins. *Biochem. J.* **2007**, *402*, 93–104. https://doi.org/10.1042/BJ20060318.
103. Liu, D.; Wang, Y.; Wei, L.; Ye, H.; Liu, H.; Wang, L.; Liu, R.; Li, D.; Lai, R. Snake venom-like waprin from the frog of *Ceratophrys calcarata* contains antimicrobial function. *Gene* **2013**, *514*, 99–104. https://doi.org/10.1016/j.gene.2012.11.007.
104. Hessinger, D. A.; Lenhoff, H. M. Mechanism of Hemolysis Induced by Nematocyst Venom: Roles of Phospholipase A and Direct Lytic Factor. *Arch. Biochem. Biophys.* **1976**, *173*, 603–613. https://doi.org/10.1016/0003-9861(76)90297-6.
105. Kini, R. M. Excitement ahead: Structure, function and mechanism of snake venom phospholipase a2 enzymes. *Toxicon* **2003**, *42*, 827–840. https://doi.org/10.1016/j.toxicon.2003.11.002.
106. Nevalainen, T. J.; Peuravuori, H. J.; Quinn, R. J.; Llewellyn, L. E.; Benzie, J. A. H.; Fenner, P. J.; Winkel, K. D. Phospholipase A2 in Cnidaria. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **2004**, *139*, 731–735. https://doi.org/10.1016/j.cbpc.2004.09.006.
107. McIntosh, J. M.; Ghomashchi, F.; Gelb, M. H.; Dooley, D. J.; Stoehr, S. J.; Giordani, A. B.; Naisbitt, S. R.; Olivera, B. M. Conodipine-M, a novel phospholipase a 2 isolated from the venom of the marine snail *Conus Magus*. *J. Biol. Chem.* **1995**, *270*, 3518–3526. https://doi.org/10.1074/jbc.270.8.3518.
108. Lee, S.; Lynch, K. R. Brown recluse spider (*Loxosceles Reclusa*) venom phospholipase D (PLD) generates lysophosphatidic acid (LPA). *Biochem. J.* **2005**, *391*, 317–323. https://doi.org/10.1042/BJ20050043.
109. Wan, H.; Lee, K. S.; Kim, B. Y.; Zou, F. M.; Yoon, H. J.; Je, Y. H.; Li, J.; Jin, B. R. A spider-derived kunitz-type serine protease inhibitor that acts as a plasmin inhibitor and an elastase inhibitor. *PLoS One* **2013**, *8*, e53343. https://doi.org/10.1371/journal.pone.0053343.
110. Mihelič, M.; Turk, D. Two decades of thyroglobulin type-1 domain research. *Biol. Chem.* **2007**, *388*, 1123–1130. https://doi.org/10.1515/BC.2007.155.
111. Horiike, T.; Nagai, H.; Kitani, S. identification of allergens in the box jellyfish *Chironex Yamaguchii* that cause sting dermatitis. *Int. Arch. Allergy Immunol.* **2015**, *167*, 73–82. https://doi.org/10.1159/000434721.
112. Gibbs, G. M.; Roelants, K.; O'Bryan, M. K. The CAP superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—roles in reproduction, cancer, and immune defense. *Endocr. Rev.* **2008**, *29*, 865–897. https://doi.org/10.1210/er.2008-0032.
113. Verdes, A.; Simpson, D.; Holford, M. Are fireworms venomous? evidence for the convergent evolution of toxin homologs in three species of fireworms (Annelida, Amphinomidae). *Genome Biol. Evol.* **2018**, *10*, 249–268. https://doi.org/10.1093/gbe/evx279.
114. Yamazaki, Y.; Koike, H.; Sugiyama, Y.; Motoyoshi, K.; Wada, T.; Hishinuma, S.; Mita, M.; Morita, T. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction: novel proteins in snake venoms. *Eur. J. Biochem.* **2002**, *269*, 2708–2715. https://doi.org/10.1046/j.1432-1033.2002.02940.x.
115. Milne, T. J.; Abbenante, G.; Tyndall, J. D. A.; Halliday, J.; Lewis, R. J. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J. Biol. Chem.* **2003**, *278*, 31105–31110. https://doi.org/10.1074/jbc.M304843200.
116. Fang, K. S.; Vitale, M.; Fehlner, P.; King, T. P. CDNA cloning and primary structure of a white-face hornet venom allergen, antigen 5. *Proc. Natl. Acad. Sci.* **1988**, *85*, 895–899. https://doi.org/10.1073/pnas.85.3.895.
117. Ruiming, Z.; Yibao, M.; Yawen, H.; Zhiyong, D.; Yingliang, W.; Zhijian, C.; Wenxin, L. Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. *BMC Genomics* **2010**, *11*, 452. https://doi.org/10.1186/1471-2164-11-452.
118. Ben-Ari, H.; Paz, M.; Sher, D. The chemical armament of reef-building corals: inter- and intra-specific variation and the identification of an unusual actinoporin in *Stylophora Pistilata*. *Sci. Rep.* **2018**, *8*, 251. https://doi.org/10.1038/s41598-017-18355-1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

119. Lee, H.; Jung, E.; Kang, C.; Yoon, W. D.; Kim, J.-S.; Kim, E. Scyphozoan jellyfish venom metalloproteinases and their role in the cytotoxicity. *Toxicon* **2011**, *58*, 277–284. https://doi.org/10.1016/j.toxicon.2011.06.007.
120. Moran, Y.; Praher, D.; Schlesinger, A.; Ayalon, A.; Tal, Y.; Technau, U. Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Mar. Biotechnol.* **2013**, *15*, 329–339. https://doi.org/10.1007/s10126-012-9491-y.
121. Fox, J. W.; Serrano, S. M. T. structural considerations of the snake venom metalloproteinases, key members of the M12 reprolysin family of metalloproteinases. *Toxicon* **2005**, *45*, 969–985. https://doi.org/10.1016/j.toxicon.2005.02.012.
122. Weston, A. J.; Chung, R.; Dunlap, W. C.; Morandini, A. C.; Marques, A. C.; Moura-da-Silva, A. M.; Ward, M.; Padilla, G.; da Silva, L. F.; Andreakis, N.; Long, P. F. Proteomic characterization of toxins isolated from nematocysts of the south atlantic jellyfish *Olindias Sambaquiensis*. *Toxicon* **2013**, *71*, 11–17. https://doi.org/10.1016/j.toxicon.2013.05.002.
123. Podobnik, M.; Anderluh, G. Pore-forming toxins in Cnidaria. *Semin. Cell Dev. Bio.* **2017**, *72*, 133–141. https://doi.org/10.1016/j.semcdb.2017.07.026.
124. Carrette, T.; Alderslade, P.; Seymour, J. Nematocyst ratio and prey in two australian cubomedusans, *Chironex fleckeri* and *Chiropsalmus* Sp. *Toxicon* **2002**, *40*, 1547–1551. https://doi.org/10.1016/S0041-0101(02)00168-X.
125. Courtney, R.; Sachlikidis, N.; Jones, R.; Seymour, J. Prey capture ecology of the cubozoan *Carukia barnesi*. *PLoS One* **2015**, *10*, e0124256. https://doi.org/10.1371/journal.pone.0124256.
126. Gacesa, R.; Chung, R.; Dunn, S. R.; Weston, A. J.; Jaimes-Becerra, A.; Marques, A. C.; Morandini, A. C.; Hranueli, D.; Starcevic, A.; Ward, M.; Long, P. F. Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora Digitifera* (Cnidaria: Anthozoa: Scleractinia). *BMC Genomics* **2015**, *16*. https://doi.org/10.1186/s12864-015-1976-4.
127. Frazão, B.; Vasconcelos, V.; Antunes, A. Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: An overview. *Mar. Drugs* **2012**, *10*, 1812–1851. https://doi.org/10.3390/md10081812.
128. Jouiaei, M.; Sunagar, K.; Federman Gross, A.; Scheib, H.; Alewood, P. F.; Moran, Y.; Fry, B. G. Evolution of an ancient venom: Recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. *Mol. Biol. Evol.* **2015**, *32*, 1598–1610. https://doi.org/10.1093/molbev/msv050.
129. Sunagawa, S.; DeSalvo, M. K.; Voolstra, C. R.; Reyes-Bermudez, A.; Medina, M. Identification and gene expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals. *PLoS One* **2009**, *4*, e4865. https://doi.org/10.1371/journal.pone.0004865.
130. Shafee, T.; Mitchell, M. L.; Norton, R. S. Mapping the chemical and sequence space of the ShKT superfamily. *Toxicon* **2019**, *165*, 95–102. https://doi.org/10.1016/j.toxicon.2019.04.008.
131. Surm, J. M.; Stewart, Z. K.; Papanicolaou, A.; Pavasovic, A.; Prentis, P. J. The draft genome of *Actinia Tenebrosa* reveals insights into toxin evolution. *Ecol. Evol.* **2019**, *9*, 11314–11328. https://doi.org/10.1002/ece3.5633.
132. Niermann, C. N.; Tate, T. G.; Suto, A. L.; Barajas, R.; White, H. A.; Guswiler, O. D.; Secor, S. M.; Rowe, A. H.; Rowe, M. P. Defensive venoms: Is pain sufficient for predator deterrence? *Toxins* **2020**, *12*, 260. https://doi.org/10.3390/toxins12040260.
133. Sunagar, K.; Morgenstern, D.; Reitzel, A. M.; Moran, Y. Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. *J. Proteomics* **2016**, *135*, 62–72. https://doi.org/10.1016/j.jprot.2015.09.015.
134. Arbuckle, K. From Molecules to macroevolution: Venom as a model system for evolutionary biology across levels of life. *Toxicon: X* **2020**, *6*, 100034. https://doi.org/10.1016/j.toxcx.2020.100034.
135. Stampar, S. N.; El Didi, S. O.; Paulay, G.; Berumen, M. L. A new species of Arachnanthus from the Red Sea (Cnidaria, Ceriantharia). *Zookeys* **2018**, *748*, 1–10. https://doi.org/10.3897/zookeys.748.22914.
136. Haddock, S. H.; Dunn, C. W. Fluorescent proteins function as a prey attractant: experimental evidence from the hydromedusa Olindias formosus and other marine organisms. *Biology open* **2015**, *4*, 1094-1104. https://doi.org/10.1242/bio.012138.
137. Macrander, J.; Broe, M.; Daly, M. Multi-copy venom genes hidden in de novo transcriptome assemblies, a cautionary tale with the snakelocks sea anemone *Anemonia sulcata* (Pennant, 1977). *Toxicon* **2015**, *108*, 184–188. https://doi.org/10.1016/j.toxicon.2015.09.038.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.18.159541; this version posted June 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

138. Holding, M.; Margres, M.; Mason, A.; Parkinson, C.; Rokyta, D. Evaluating the performance of de novo assembly methods for venom-gland transcriptomics. *Toxins* **2018**, *10*, 249. https://doi.org/10.3390/toxins10060249.
139. Mariottini, G.; Pane, L. Cytotoxic and cytolytic cnidarian venoms. A review on health implications and possible therapeutic applications. *Toxins* **2013**, *6*, 108–151. https://doi.org/10.3390/toxins6010108.
140. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. **2010**. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
141. Waterhouse, R. M.; Seppey, M.; Simão, F. A.; Manni, M.; Ioannidis, P.; Klioutchnikov, G.; Kriventseva, E. V.; Zdobnov, E. M. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Bio. Evol.* **2018**, *35*, 543–548. https://doi.org/10.1093/molbev/msx319.
142. Haas, B. J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P. D.; Bowden, J.; Couger, M. B.; Eccles, D.; Li, B.; Lieber, M.; MacManes, M. D.; Ott, M.; Orvis, J.; Pochet, N.; Strozzi, F.; Weeks, N.; Westerman, R.; William, T.; Dewey, C. N.; Henschel, R.; LeDuc, R. D.; Friedman, N.; Regev, A. De novo transcript sequence reconstruction from rna-seq using the trinity platform for reference generation and analysis. *Nat Protoc* **2013**, *8*, 1494–1512. https://doi.org/10.1038/nprot.2013.084.
143. Altschul, S. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.
144. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T. L. BLAST+: Architecture and applications. *BMC Bioinformatics* **2009**, *10*, 421. https://doi.org/10.1186/1471-2105-10-421.
145. Finn, R. D.; Clements, J.; Eddy, S. R. HMMER Web server: Interactive sequence similarity searching. *Nucleic Acids Research* **2011**, *39* (suppl), W29–W37. https://doi.org/10.1093/nar/gkr367.
146. Potter, S. C.; Luciani, A.; Eddy, S. R.; Park, Y.; Lopez, R.; Finn, R. D. HMMER web server: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W200–W204. https://doi.org/10.1093/nar/gky448.
147. Almagro Armenteros, J. J.; Tsirigos, K. D.; Sønderby, C. K.; Petersen, T. N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. https://doi.org/10.1038/s41587-019-0036-z.
148. Li, W.; Godzik, A. Cd-Hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.
149. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565.
150. The UniProt Consortium, UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515. https://doi.org/10.1093/nar/gky1049
151. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2019. https://www.R-project.org/.
152. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, 2016. http://www.rstudio.com/.
153. Katoh, K.; Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. https://doi.org/10.1093/molbev/mst010.
154. Stamatakis, A. RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 2014, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
155. Waterhouse, A.M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189-1191. https://doi.org/10.1093/bioinformatics/btp033.
156. Xu, L.; Dong, Z.; Fang, L.; Luo, Y.; Wei, Z.; Guo, H.; Zhang, G.; Gu, Y.Q.; Coleman-Derr, D.; Xia, Q.; Wang, Y. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* **2019** *47*, W52-W58. https://doi.org/10.1093/nar/gkz333.