Citation for published version (APA):
Comparative proteomics reveals recruitment patterns of some protein families in the venoms of Cnidaria

Adrian Jaimes-Becerra, Ray Chung, André C. Morandini, Andrew J. Weston, Gabriel Padilla, Ranko Gacesa, Malcolm Ward, Paul F. Long, Antonio C. Marques

PII: S0041-0101(17)30217-9
DOI: 10.1016/j.toxicon.2017.07.012
Reference: TOXCON 5671

To appear in: Toxicon

Received Date: 17 April 2017
Revised Date: 7 July 2017
Accepted Date: 10 July 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Comparative proteomics reveals recruitment patterns of some protein families in the venoms of Cnidaria

Adrian Jaimes-Becerra\textsuperscript{1*}, Ray Chung\textsuperscript{2+}, André C. Morandini\textsuperscript{1}, Andrew J. Weston\textsuperscript{3}, Gabriel Padilla\textsuperscript{4}, Ranko Gacesa\textsuperscript{5}, Malcolm Ward\textsuperscript{2+}, Paul F. Long\textsuperscript{5-7} and Antonio C. Marques\textsuperscript{1,8}

\textsuperscript{1}Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua Matão, Trav. 14, 101, 05508-090 São Paulo, SP, Brazil.

\textsuperscript{2}Proteomics Facility, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, 16 De Crespigny Park, London SE5 8AF, United Kingdom.

\textsuperscript{3}Mass Spectrometry Laboratory, UCL School of Pharmacy, 29/39 Brunswick Square, London WC1N 1AX, United Kingdom.

\textsuperscript{4}Instituto de Ciências Biomédicas, Universidade de São Paulo, Av. Professor Lineu Prestes 1374, 05508-000 Sao Paulo, SP, Brazil.

\textsuperscript{5}Faculty of Life Sciences & Medicine, King’s College London, 150 Stamford Street, London SE1 9NH, United Kingdom.

\textsuperscript{6}Brazil Institute, King’s College London, Strand, London WC2R 2LS, United Kingdom.

\textsuperscript{7}Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Av. Prof. Lineu Prestes, 580, B16, 05508-000 São Paulo, SP, Brazil.

\textsuperscript{8}Centro de Biologia Marinha, Universidade de São Paulo, Rodovia Manoel Hypólito do Rego, km. 131,5, 11600-000 São Sebastião, Brazil.

\*Corresponding author: Adrian Jaimes-Becerra, email: adrianjaimesb@usp.br
*Present addresses: Mr Ray Chung - Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, London, SE5 8AF, United Kingdom. Dr Malcolm Ward - Aulesa Biosciences Limited, 44, Elgar Drive, Shefford, Bedfordshire, SG17 5RZ, United Kingdom.
Abstract

Cnidarians are probably the oldest group of animals to be venomous, yet our current picture of cnidarian venom evolution is highly imbalanced due to limited taxon sampling. High-throughput tandem mass spectrometry was used to determine venom composition of the scyphozoan *Chrysaora lactea* and two cubozoans *Tamoya haplonema* and *Chiropsalmus quadrumanus*. Protein recruitment patterns were then compared against 5 other cnidarian venom proteomes taken from the literature. A total of 28 putative toxin protein families were identified, many for the first time in Cnidaria. Character mapping analysis revealed that 17 toxin protein families with predominantly cytolytic biological activities were likely recruited into the cnidarian venom proteome before the lineage split between Anthozoa and Medusozoa. Thereafter, venoms of Medusozoa and Anthozoa differed during subsequent divergence of cnidarian classes. Recruitment and loss of toxin protein families did not correlate with accepted phylogenetic patterns of Cnidaria. Selective pressures that drive toxin diversification independent of taxonomic positioning have yet to be identified in Cnidaria and now warrant experimental consideration.

**Keywords:** evolution; venom; Cnidaria; nematocysts; proteomics.
Introduction

Cnidaria is believed to be the most basal of the extant Metazoa to be venomous, having evolved since Neoproterozoic times, ~650 million years ago, long before the Cambrian radiation (Van Iten et al., 2014). Cnidaria is a diverse phylum comprising over 13,500 free living or parasitic marine, freshwater and terrestrial species (Daly et al., 2007 plus myxozoans by Okamura et al., 2015a). Cnidaria has two major subphyla: Anthozoa and Medusozoa. Anthozoa include sea anemones and both hard and soft corals (Bridge et al., 1992; Marques & Collins, 2004). Medusozoa comprise the classes Staurozoa (e.g. stalked jellyfish), Cubozoa (e.g. box jellyfish), Scyphozoa (e.g. ‘true’ jellyfish) and Hydrozoa (e.g. Hydra and relatives including several species of smaller jellyfish) (Marques & Collins, 2004; Collins et al., 2006; Van Iten et al., 2014). Recent molecular phylogenetic analyses have corroborated the cnidarian nature of Myxozoa, with strong support as a sister-group to Medusozoa (reviewed in Okamura et al. 2015b).

The most evident synapomorphy of Cnidaria is the presence of cnidae, organelles produced by the Golgi apparatus of specialised cells called cnidoblasts (Marques & Collins, 2004; Fautin, 2009; Beckmann & Özbek, 2012). Cnidae are found in various parts of the body of a cnidarian and are classified into three morphological types: nematocysts, spirocysts and ptychocysts (Östman, 2000; Özbek et al., 2009). The nematocysts discharge venom and are found in all cnidarians, but are morphologically and functionally heterogeneous (David et al., 2008; Fautin, 2009). In addition to prey capture and defence against predation, the venom of nematocysts may also mediate spatial intraspecific and interspecific competition (Bigger, 1980; Kass-Simon & Scappaticci, 2002).

There has been resurgence in interest surrounding the nature and evolutionary origins of cnidarian venom toxins, since the first application of high throughput tandem
mass spectrometry realised high sequence homology between cnidarian toxins and those
of other venomous animals (Weston et al., 2012, 2013). Many studies using genomic,
transcriptomic or proteomic approaches have also realised these astonishing similarities
(Balasubramanian et al., 2012; Brinkman et al., 2012, 2015; Li et al., 2012, 2014, 2016;
Gacesa et al., 2015; Jouiaei et al., 2015a; Macrander et al., 2015, 2016; Lewis et al.,
2016; Ponce et al., 2016, Huang et al., 2016), leading to the recognition that
understanding the mechanisms underpinning toxin diversification in Cnidaria could
provide a platform from which the evolution of this trait in higher animals might be
more fully explored (Starcevic & Long, 2013; Starcevic et al., 2015; Jouiaei et al.,
2015b). For this to be achieved, a comprehensive inventory of toxins must first be
undertaken and then mapped against different taxonomic levels from established
cnidarian phylogeny. To date, studies attempting to infer evolutionary aspects of toxin
recruitment in Cnidaria have suffered limited taxon sampling, but when taken together
these studies have demonstrated a degree of functional recruitment of certain toxin
protein families at different taxonomic levels (Rachamim et al., 2014; Brinkman et al.,
2015; Jouiaei et al., 2015b). Here, the number of venom proteomes is expanded and
used with data from the literature for character mapping analysis, to establish a more
complete venom assembly hypothesis between the major taxonomic classes of Cnidaria.
Material & Methods

Nematocyst proteomics: The scyphozoan *Chrysaora lactea* and two cubozoans *Tamoya haplonema* and *Chiropsalmus quadrumanus* (Figure 1) were collected with permission (SISBIO license 15031-2) on May 7th 2012 by bottom shrimp trawls (2 cm mesh size) dragged at 10 m depth along Enseada beach (Guarujá County, São Paulo State, ca. 23°43’20”S 43°23’40W). Animals were identified based on morphological characters (Morandini *et al.*, 2005; Morandini & Marques, 2010; Collins *et al.*, 2011) and intact nematocysts were isolated from excised tentacles as previously described (Weston *et al.*, 2013). To extract solubilised proteins, 1 mL of protein extraction buffer (50 mM TEAB, 0.04 % (w/v) SDS, Roche protease and phosphatase inhibitor cocktail) was added to freeze dried nematocysts. The reconstituted material was disrupted in a sonic bath (VWR, Lutterworth, UK) for 15 mins. The debris was removed by centrifugation (10,000 x g for 10 mins at 4 °C). The supernatant was decanted and the soluble protein concentration determined by Bradford assay. A volume equivalent to 15 µg of protein was made up to 15 µL in extraction buffer and added to 15 µL 2 x Laemmli sample buffer, heated for 10 mins at 95 °C and loaded onto a 4-12 % (w/v) NuPAGE gel (Life Technologies) and separated by 1D SDS-PAGE. Electrophoresis was performed in MES buffer (Life technologies) at 150 V for approximately 100 mins. The entire gel lane was then divided into 15 equal sections, excised and cut into 2 mm pieces. In-gel reduction, alkylation, and proteolytic digestion with trypsin were performed as follows: Cysteine residues were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide in 100 mM ammonium bicarbonate to form stable carbamidomethyl derivatives. Trypsin digestion was carried out overnight at 37 °C in 50 mM ammonium bicarbonate buffer and the supernatant was retained. Peptides were extracted from the gel pieces by two washes with 50 mM ammonium bicarbonate and
acetonitrile. Each wash involved shaking the gel pieces for 10 mins. The extracts were pooled with the initial digestion supernatant and then lyophilised. Lyophilised extract was reconstituted in 30 µL of 50 mM ammonium bicarbonate buffer for LC-MS/MS.

**Data analysis:** Data analysis was performed as previously described (Weston et al., 2013; Gacesa et al., 2015) but with minor modifications. Briefly, a one search matching strategy of rawfile MS/MS data against the Tox-Prot UniProtKB/Swiss-Prot database (Jungo et al., 2012) using the MASCOT search engine was first executed (Perkins et al., 1999). Methionine oxidation, phosphorylation on S/T/Y, deamidation on N/D and carbamidomethyl cysteine were selected as fixed modifications. Digestion with trypsin allowed up to three missed cleavages. The data were searched with a parent ion tolerance of 5 ppm and a fragment ion tolerance of 0.5 Da. The MASCOT result files were next uploaded into Scaffold v4.3.4 (Proteome Software, Portland, Oregon, USA) (Searle, 2010) and spectra corresponding to likely venom toxin peptides were manually validated for unbroken series of overlapping b-type and y-type sequence specific fragments ions, where losses consistent with the sequence were acceptable. Validated spectra (Figures S1-S3) corresponding to peptides with predicted venom toxin functions were next distinguished from peptides with likely other non-toxic physiological functions using ‘ToxClassifier’ (Gacesa et al., 2016). This is a suite of machine learning based classifiers that provide consistent discrimination of toxins from non-toxin peptide sequences with > 99 % accuracy by performing BLAST and HMM comparisons against the Tox-Prot UniProtKB/Swiss-Prot (Jungo et al., 2012), UniProt Trembl (The UniProt consortium, 2017) and NR (NCBI Resource Coordinators, 2016) databases.

**Character mapping analysis:** In addition to the data acquired in this study, putative toxins from other cnidarians described in the literature were also included to enhance
the dataset. These putative toxins were from the anthozoans *Anemonia viridis* (Rachamim *et al*., 2014) and *Acropora digitifera* (Gacesa *et al*., 2015), the hydrozoans *Olindias sambaquiensis* (Weston *et al*., 2013) and *Hydra magnipapillata* (Rachamim *et al*., 2014) and, the scyphozoan *Aurelia aurita* (Rachamim *et al*., 2014). The putative toxins from the combined data set were assigned to venom toxin protein families using established KEGG ontology. Data were coded in a matrix as presence (1) or absence (0) of each toxin protein family in each species. Reconstruction of ancestral states at different nodes on an accepted taxonomic tree of Cnidaria (Marques & Collins, 2004; Collins *et al*., 2006) was performed using Mesquite version 3.04 (Maddison & Maddison, 2015) with the parsimony criterion for the model unordered. In addition, the matrix of presences and absences of toxin protein families was used to infer a phylogenetic pattern based on the parsimony criterion.

**Results**

Comparative proteomics of toxin protein families: The putative toxin proteomes of nematocysts for the 3 species experimentally acquired in this study are given in Table 1. The toxin protein families from 5 species taken from the literature are given in Table S1. A total of 28 toxin protein families were identified from the nematocyst proteomes of the 8 species studied and are shown in Figure 2. Nine (~33 %) out of the 28 toxin protein families were shared by all the four classes of cnidarians. These 9 protein families were conotoxins O, CRISP, latrotoxin, lipase, metalloproteinase, phospholipases A₂ (PLA₂), phospholipases D, CS αβ potassium channel blocker, and CS αβ sodium channel inhibitor. Nineteen protein toxin families were not distributed across all classes (Figure 2). These included three families of pore forming toxins, which were the jellyfish toxin family-like proteins (JFTs) found to be restricted to the sister classes
Cubozoa and Scyphozoa; the actinoporins found in the classes Anthozoa, Hydrozoa and Cubozoa, and latarcins found in the classes Anthozoa and Scyphozoa. The ficolins and snaclec belong to the lectin families of toxins and were limited to the Scyphozoa Anthozoa, and Hydrozoa. Peptides with similarity to three families of neurotoxins were also taxonomically restricted (Figure 2), these were the kunitz type family detected in Anthozoa and Scyphozoa, the calcium channel blocker Huwentoxin-1 reported here for the first time but solely in medusozoans, and snake three finger found in all classes except Cubozoa. Likewise, peptidase S1, flavin amino-oxidase and glycosyl hydrolase 56 families were identified in all classes except Cubozoa. Complement C3 family-like proteins were identified in the Hydrozoa and Scyphozoa. MAC-PF family-like proteins were identified in the Hydrozoa and Cubozoa. The presence of translationally controlled tumour like proteins (TCTP) was identified in the venom proteome from both Anthozoa and Hydrozoa.
Evolution of the cnidarian venom arsenal: Recruitment patterns of putative toxin protein families (Figure 3 and Table S3) were inferred using a presence and absence matrix (Table S2). This recruitment pattern indicated that venom of Medusozoa and Anthozoa ancestors might have been composed of at least seventeen types of protein toxin families (Figure 3 and Figure S4i). After separation of the ancestral lineage into Anthozoa and Medusozoa, some putative toxins were lost (or not expressed) in some clades. For example, the TCTP family was not present in the Cubozoa and Scyphozoa. Similarly, the actinoporin toxin protein family was lost from Scyphozoa. Unlike the other classes, the species of Cubozoa examined demonstrated large losses. Nine toxin protein families might have been recruited by a single clade after the split Anthozoa-Medusozoa (Figure 3 and Figure S4ii). Three families of cytolytic toxins (MAC-PF, ficolin lectin and JFTs) appear to have been recruited into Medusozoa after the basal diversification event into the venom of Hydrozoa, Scyphozoa, and Cubozoa (Figure 3). Equally, two families of neurotoxins ShK-like potassium channel and sea anemone sodium channel modulator appear to have been recruited into Anthozoa only. The latarcin and kunitz-type toxin protein families, might have been recruited independently (i.e., by convergence) into the venom of Anthozoa/Scyphozoa (Figure 3 and Figure S4iii). Phylogenetic analysis of the presence and absence matrix gave a topology of (Cubozoa(Anthozoa(Hydrozoa,Scyphozoa))), which disagreed with the more accepted phylogeny of Cnidaria (Anthozoa(Hydrozoa(Cubozoa,Scyphozoa))).

Discussion

The putative toxin component of nematocyst proteomes for 3 out of the 8 species examined (Chrysaora lactea, Tamoya haplonema, and Chiropsalmus quadrumanus) are described in this study for the first time (Table 1). Venom data from 3 other species
(Anemonia viridis, Hydra magnipapillata, and Aurelia aurita) were published elsewhere (Rachamim et al., 2014) and reassessed in this study. These data were combined with our own previously published putative nematocyst toxin proteomes from Acropora digitifera (Gacesa et al., 2015) and Olindias sambaquensis (Weston et al., 2013). Altogether, the data from this study has supported previous research that Anthozoa and Medusozoa have complex venom composition comprising multiple protein families (Rachamim et al., 2014; Jouiaei et al., 2015c) (Figures 2 and 3). We highlight that, although transcriptomes and proteomes from other species of cnidarians have also been published (Moran et al., 2013; Jouiaei et al., 2015a, Ponce et al., 2016, Macrander et al., 2016), our analysis focused on those species for which we had access to raw proteomics MS/MS data which could be analysed using identical bioinformatics methods, ensuring results were fully comparable. Our study was conservative, being restricted to putative toxin annotation in the expressed proteome and did not include a study of transcriptomes. This was because not all the transcripts that contributed to transcriptome diversity would equally be likely to be translated (if at all) and have contributed to protein diversity. Hence, correlation between sequences annotated as putative toxins in the transcriptome and proteome would not have been straightforward given the difficulty in differentiating sequences with toxic and other physiological functions. Future work to overcome this impediment will require the acquisition of genome sequence onto which other sequence data can be mapped (Gacesa et al., 2015).

Comparative venom proteomic analysis from different Cnidaria classes

Our comparative proteomics data of putative venom toxins indicated that nearly half of the protein toxin families were distributed across all of the cnidarian classes studied (Figure 2). The biological activities of some of these toxin families are of note,
for example, PLA$_2$ toxins have thus far only been identified with haemolytic activity in cnidarians (Hessinger & Lenhoff, 1976; Grotendorst & Hessinger, 2000; Anderluh & Maček, 2002; Talvinen & Nevalainen, 2002; Nevalainen et al., 2004; Razpotnik et al., 2010), although neurotoxic and myotoxic activities as well as non-toxic physiological functions have also been widely reported in other venomous animals (Fry et al., 2009; Six & Dennis, 2000). Likewise, phospholipase D family proteins isolated from cnidarian venoms have been reported to exhibit necrotic activity (Burke, 2002; Uri et al., 2005), with homologs also recently identified in the giant jellyfish _Cyanea capillata_ (Liu et al., 2015). Most of the metalloproteinases identified in this study belonged to the zinc metalloproteinase family. This family of toxins is an important component found in the venoms of many terrestrial animals such as centipedes, snakes and ticks (Fry et al., 2009; Undheim et al., 2014), with diverse biological activities culminating in haemorrhage and tissue necrosis in the target prey following envenomation (Fox & Serrano, 2005; da Silveira et al., 2007). Transcriptomic and proteomic studies have identified zinc metalloprotease in venoms of the scyphozoans _Stomolophus meleagris_, _Cyanea capillata_, and _Cyanea nozakii_ (Li et al., 2014, 2016; Liu et al., 2015), the cubozoan _Chironex fleckeri_ (Brinkman et al., 2015; Jouiae et al., 2015a) and the anthozoan _Anthopleura elegantissima_ (Macrander et al., 2015). A study of metalloproteases from the scyphozoan _Nemopilema nomurai_, _Rhopilema esculenta_, _Cyanea nozakii_, and _Aurelia aurita_ confirmed the necrotic toxicity of these enzymes (Lee et al., 2011). Both sodium and potassium ion channel inhibitors were identified in representatives of all of the classes examined. These two types of neurotoxins have been widely studied in Anthozoa, especially sea anemones and comprise the largest number of toxins so far recorded in public databases for Cnidaria (Moran et al., 2009; Šuput, 2009; Turk & Kem, 2009; Frazão et al., 2012; Jouiae et al., 2015c; Macrander et al.,
256 2015; Mariottini et al., 2015). Neurotoxic effects have been identified in scyphozoans
257 such as Cyanea nozakii (Feng et al., 2010), Cyanea capillata (Helmholz et al., 2012),
258 and Pelagia noctiluca (Pang et al., 1993; Morabito et al., 2012) and, in cubozoans such
259 as Carukia barnesi (Winkel et al., 2005) and Malo kingi (Gershwin, 2007). In this study,
260 we identified two putative types of sodium and potassium ion channel inhibitors
261 (Figure 2, Table S2). ShK-like potassium channel and sea anemone sodium channel
262 modulator were only found in a single Anthozoan species, Anemonia viridis (a sea
263 anemone). It should be noted that this was the only species of sea anemone analysed in
264 this study. Both sodium and potassium putative ion channel inhibitors have been found
265 exclusively in sea anemones (Moran et al., 2009; Diochot and Lazdunski, 2009). CS \( \alpha \beta \)
266 potassium channel blocker and CS \( \alpha \beta \) sodium channel inhibitor were found in all of the
267 species of cnidarians analysed including another anthozoan, Acropora digitifera and
268 have sequence homology to sodium and potassium channel blockers of scorpions. This
269 observation might highlight a rare example of mechanistic convergence whereby
270 sodium and potassium ion channel inhibitors appeared on two separate occasions within
271 the cnidarians. Convergent evolution of these toxins in scorpions and sea anemones has
272 been previously reported and although these toxin protein families are structurally
273 different, functional mapping studies have shown similarities in the binding cores
274 (Gasparini et al., 2004). Further species sampling is required to substitute these
275 observations which are based here on a single MS/MS event in the anthozoa Acropora
276 digitifera, the hydrozoa Hydra magnipapillata, the scyphozoan Aurelia aurita and the
277 cubozoans Chiropsalmus quadrumanus and Tamoya haplonema. It should also be noted
278 that the names given to each putative ion channel inhibitors were used to distinguish
279 between the two different possible origins of the putative sodium and potassium ion
280 channel inhibitors identified in this study. Another family of neurotoxins were the
CRISP type toxins, which again were found in all classes of cnidarians studied herein. This toxin protein family has widely been reported in cnidarian venoms (Brinkman et al., 2015; Ponce et al., 2016; Lewis et al., 2016) and attributed many biological functions.

Just over half of the toxins protein families identified in this study were restricted to certain cnidarian classes only (Figure 2). Hyaluronidase-like proteins were found in all classes of cnidarians except Cubozoa, but these proteins are common and have non-toxic physiological function in many non-venomous animals. It is feasible that such proteins are likely recruited into venoms not as toxins, but as adjuvants to increase tissue permeability (Kemparaju & Girish, 2006; Fry et al., 2009). Non-toxic peptides and proteins present in nematocysts that may function in toxin maturation, toxin trafficking and delivery, or as self-defence mechanisms against the biological activities of the venom have received little study and perhaps warrant closer inspection. Likewise, the peptidase S1 family was also detected in all cnidarian classes studied except Cubozoa. This family is part of the group of serine protease inhibitors that is widely distributed in other marine venomous animals including marine cone snails and cephalopods (Mourão & Schwartz, 2013), as well as terrestrial reptiles (Fry et al., 2009). Recently, serine protease homologs were identified in the transcriptome of the sea anemone Anthopleura elegantissima (Macrander et al., 2015). However, the biological activity of the S1 peptidase family of toxins has yet to be confirmed in cnidarians. It is unclear why proteins commonly associated with innate immune responses are also apparently widely distributed in cnidarian venoms. For example, MAC-PF-like toxins have also been identified in sea anemones (Nagai et al., 2002b; Oshiro et al., 2004) and were recently annotated in the transcriptomes and proteomes of Hydrozoa and Scyphozoa (Rachamim et al., 2014). Likewise, the actinoporins are pore-
forming toxins were found in the proteomes of Anthozoa and Cubozoa classes. These
cytolysins have also been identified in transcriptome sequences and biological activity
recorded in nematocyst venom extracts of various Hydra species (Hydrozoa) (Hwang et
al., 2007; Glasser et al., 2014).

Assembly of the cnidarian venom proteome

To date, only one previous study has been published that used similar
approaches to those described here to investigate evolutionary aspects of toxin
recruitment in Cnidaria (Rachamim et al., 2014). In this previous study, the kunitz type
family of toxins was only found in Anthozoa. In the study present here, this family of
toxins was found in Anthozoa, but was also identified in Scyphozoa. In the study of
Rachamim et al., (2014), the PLA2 family of toxin proteins were only found in species
of Scyphozoa and Hydrozoa. In our analyses, PLA2 were present in all the Cnidaria
classes studied and hence, most likely arose as a recruitment event at the base ancestor
of the Cnidaria. It should be noted that PLA2 like proteins have also been identified in
recent studies of Anthozoan venoms (Macrander et al., 2015, 2016). Differences in the
recruitment patterns between studies might be explained because of the low number of
species sampled. The extent of comparison groups (all Cnidaria) in light of the
sparseness of data at terminals in the analysis is a concern, for example, no data is
currently available on the toxin compliment of venoms from the Staurozoa or Myxozoa
(Marques & Collins, 2004; Okamura et al., 2015a). Based on the data presented, many
of the neurotoxic and cytotoxic protein toxin families might have been recruited into the
venom proteome early in cnidarian evolution, before the first major radiation in this
phylum around 800 million years ago (Park et al., 2012; Van Iten et al., 2017).
The approaches used in this study were very conservative, with analyses based exclusively on putative toxin protein families found in each proteomic profile and not specific toxin peptides or proteins. These proteomic profiles can be considered phenotypes, or a "morphological representation" of the venom, allowing variation in the toxin complement to be evaluated. For example, in this study JFTs were found only in the venoms of Scyphozoa and Cubozoa. However, previous studies have demonstrated JFTs encoded in the genome and expressed in the proteome of Hydrozoa, and in the transcriptome of the anthozoan *Anemonia viridis* (Rachamim *et al*., 2014). Few reports in the literature have documented variation in toxin composition of venom at taxonomic level in the phylum Cnidaria (Orts *et al*., 2013; Rachamim *et al*., 2014), and certainly there have been no studies that have attempted to put into context what the biological consequences of venom variation might be (Gacesa *et al*., 2015; Knittell *et al*., 2016).

The difference between the two phylogenetic patterns (accepted vs inferred using the presence and absence matrix) found in this study could be due to various ecological factors that need to be investigated in future studies. But increased sampling and analysis at different taxonomic levels is a priority in order to identify the influence of history and ecology in the origin of these contrasting patterns.

**Conclusions**

Venom composition of Medusozoa and Anthozoa are different, with cytolytic toxin protein families slightly more abundant in Medusozoa. When only toxin protein family composition was used for phylogenetic inference, the resulting topology (Cubozoa(Anthozoa(Hydrozoa,Scyphozoa))) did not match the classic published phylogeny (Anthozoa(Hydrozoa(Cubozoa,Scyphozoa))). Understanding the functional context (environment versus morphological form) that may drive expression of toxins in
Cnidaria requires future experimentally consideration, including wider taxonomic sampling.

Acknowledgments

We are indebted to Dr David Morganstern, Prof Tamar Lotan and Prof Daniel Sher for making available their mass spectrometry data. We extend our thanks to the Santos Family for providing logistic support of shipboard operations. We are also grateful for onshore technical assistance of the staff at CEBIMar. This is a contribution to the NP-BioMar program at the Universidade de São Paulo.

Funding

This work was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo [grant numbers 2010/52324-6, 2010/06927-0, 2011/50242-5, 2013/50484-4, 2010/50174-7], the Conselho Nacional de Desenvolvimento Científico e Tecnológico [grant numbers 562143/2010-6, 477156/2011-8, 305805/2013-4, 445444/2014-2, 301039/2013-5], the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [grant number 236.507.518-52], the Universidade de São Paulo [grant number 13.1.1502.9.8] and the Medical Research Council [grant number G82144A].

References


Starcevic A, Moura-da-Silva AM, Cullum J, Hranueli D, Long PF. 2015. Combinations of long peptide sequence blocks can be used to describe toxin diversification in venomous animals. Toxicon 95:84–92 DOI:10.1016/j.toxicon.2015.01.005.


Figure legends

**Figure 1:** A) *Chrysaora lactea*, B) *Tamoya haplonema* and C) *Chiropsalmus quadrumanus*. Medusa adult stages of Cnidaria from which the venom proteomes of isolated nematocysts were acquired for this study. (Photos courtesy of Dr Alvaro Migotto, Centro de Biologia Marinha, Universidade de São Paulo São Sebastião, Brasil).

**Figure 2. Comparison of Cnidarian venom composition.** Venn diagram showing the number of putative toxin protein families shared among the soluble nematocyst proteomes of the four classes of cnidarians studied (Note that the protein families marked with an asterisk are described here for the first time).

**Figure 3. Recruitment patterns of putative toxin protein families into Cnidaria venom, based on a established cnidarian phylogenies** (Marques & Collins, 2004; Collins et al., 2006). Solid black rectangles represent recruitment events. Dotted rectangles represent absence of toxin families. White rectangles represent multiple recruitments of toxin families. The numbers above of the lines represents the toxin families: 1. actinoporins; 2. complement C3; 3. conotoxins O; 4. Conotoxins T; 5. CRISP; 6. ficolin lectin; 7. flavin monoamine oxidase; 8. Jellyfish toxin; 9. kunitz-type; 10. latrotoxin; 11. MAC-PF; 12. metalloproteinase; 13. natriuretic peptide; 14. peptidase S1; 15. phospholipase A2; 16. phospholipase B; 17. phospholipase D; 18. ShK-like potassium channel; 19. snaclec; 20. snake three finger; 21. Sea anemone sodium channel modulator; 22. TCTP; 23. glycosyl hydrolase 56*; 24. huwentoxin-1*; 25. latarcin*; 26. lipase*; 27. CS αβ potassium channel blocker*; 28. CS αβ sodium channel inhibitor*. The proteins families marked with asterisk (*) have never previously been recorded in Cnidaria.
Table 1: Predicted venom proteomes of potential toxins isolated from nematocysts. A) *Chiropsalmus quadrumanus*, B) *Tamoya haplonema* and C) *Chrysaora lactea*. Peptide fragments used for putative toxin annotation are given with validated spectra in Figures S1-S3.

<table>
<thead>
<tr>
<th>Toxin with closest homology</th>
<th>Predicted toxin protein family</th>
<th>Uniprot accession number</th>
<th>Example of animal species with closest homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) <em>Chiropsalmus quadrumanus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha-latroinsectotoxin-Lt1a</td>
<td>Latrotoxin</td>
<td>Q02989</td>
<td><em>Latrodectus tredecimguttatus</em> (European black widow spider)</td>
</tr>
<tr>
<td>Conotoxin Bu2</td>
<td>Conotoxin O1</td>
<td>P0CY61</td>
<td><em>Conus bullatus</em> (Bubble cone snail)</td>
</tr>
<tr>
<td>Echotoxin-2</td>
<td>Actinoporin</td>
<td>Q76CA2</td>
<td><em>Monoplex parthenopeus</em> (Giant triton sea snail)</td>
</tr>
<tr>
<td>Hainantoxin-XVIII-5</td>
<td>Putative ion channel inhibitor</td>
<td>D2Y2N9</td>
<td><em>Haplopelma hainanum</em> (Chinese bird spider)</td>
</tr>
<tr>
<td>Neurotoxin LmNaTx1</td>
<td>CS αβ sodium channel inhibitor</td>
<td>D9U297</td>
<td><em>Lychas mucronatus</em> (Chinese swimming scorpion)</td>
</tr>
<tr>
<td>Toxin CfTX-2</td>
<td>Jellyfish toxin</td>
<td>A7L036</td>
<td><em>Chironex fleckeri</em> (Sea wasp)</td>
</tr>
<tr>
<td>B) <em>Tamoya haplonema</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha-latroinsectotoxin-Lt1a</td>
<td>Latrotoxin</td>
<td>Q02989</td>
<td><em>Latrodectus tredecimguttatus</em> (European black widow spider)</td>
</tr>
<tr>
<td>Conotoxin Lt5.9</td>
<td>Conotoxin T</td>
<td>Q1A3Q7</td>
<td><em>Conus litteratus</em> (Lettered cone snail)</td>
</tr>
<tr>
<td>DELTA-alicitoxin-Pse2b</td>
<td>MACPF</td>
<td>P58912</td>
<td><em>Phyllodiscus semoni</em> (Wasp sea anemone)</td>
</tr>
<tr>
<td>Disintegrin acostatin-alpha</td>
<td>Disintegrin</td>
<td>Q805F7</td>
<td><em>Agkistrodon contortrix contortrix</em> (northern copperhead pit viper)</td>
</tr>
<tr>
<td>Echotoxin-2</td>
<td>Actinoporin</td>
<td>Q76CA2</td>
<td><em>Monoplex parthenopeus</em> (Giant triton sea snail)</td>
</tr>
<tr>
<td>Equinatoxin-3</td>
<td>Actinoporin</td>
<td>P0C1H2</td>
<td><em>Actinia equina</em> (Beadlet sea anemone)</td>
</tr>
<tr>
<td>Im-conomarphin</td>
<td>Conotoxin A</td>
<td>P0CH39</td>
<td>Conus imperialis (Imperial cone snail)</td>
</tr>
<tr>
<td>Maximins 3/H2</td>
<td>Bombinin</td>
<td>P83082</td>
<td>Bombina maxima (Yunnan firebelly toad)</td>
</tr>
<tr>
<td>Phospholipase A2 3</td>
<td>Phospholipase A2</td>
<td>P21792</td>
<td>Micrurus nigrocinctus (Central American coral snake)</td>
</tr>
<tr>
<td>Phospholipase D LISicTox-alphaIII2</td>
<td>Arthropod phospholipase D CS aβ potassium channel blocker</td>
<td>Q1KY79</td>
<td>Loxosceles laeta (Chilean recluse spider)</td>
</tr>
<tr>
<td>Potassium channel toxin alpha-KTx Tx773</td>
<td>Long chain scorpion toxin</td>
<td>Q0GY43</td>
<td>Tityus discrepans (Venezuelan scorpion)</td>
</tr>
<tr>
<td>Potassium channel toxin TdiKIK</td>
<td>Venom metalloproteinase (M12B)</td>
<td>Q9W7S2</td>
<td>Deinagkistrodon acutus (Sharp-nosed pit viper)</td>
</tr>
<tr>
<td>Snake venom metalloproteinase aculysin-1</td>
<td>Spider toxin Tx2</td>
<td>P85276</td>
<td>Ctenus ornatus (Brazilian spider)</td>
</tr>
<tr>
<td>U5-ctenitoxin-Co1a</td>
<td>CRISP</td>
<td>A9QQ26</td>
<td>Lycosa singoriensis (Chinese wolf spider)</td>
</tr>
<tr>
<td>Venom allergen 5</td>
<td>Lipase</td>
<td>B2D0I5</td>
<td>Apis mellifera (European honey bee)</td>
</tr>
<tr>
<td>Venom carboxylesterase-6</td>
<td>NGF-beta</td>
<td>Q2XXL6</td>
<td>Azemiops feae (Black-headed Burmese viper)</td>
</tr>
<tr>
<td>Venom nerve growth factor 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C) Chrysaora lactea

<p>| CfTX-2 | Jellyfish toxin | A7L036 | Chironex fleckeri (Sea wasp) |
| Cathelicidin-related peptide Na.CRAMP | Cathelicidin | B6S2X0 | Naja atra (Chinese cobra) |
| Conotoxin Bu2 | Conotoxin O1 | P0CY61 | Conus bullatus (Bubble cone snail) |
| Cysteine-rich venom protein LIO1 | CRISP | Q2XXQ0 | Erythrolamprus poecilogyrus (Water snake) |
| L-amino-acid oxidase | Flavin monoamine oxidase | P0DI84 | Vipera ammodytes (Sand viper) |
| M-zodatoxin-Lt4a | Latarcin | Q1ELU5 | Lachesana tarabaevi (Ant spider) |
| Snake venom serine protease KN2 | Peptidase S1 | Q71QJ0 | Trimeresurus stejnegeri (Chinese green tree viper) |</p>
<table>
<thead>
<tr>
<th>Venom peptide</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>U16-lycotoxin-Ls1a</td>
<td><em>Lycosa singoriensis</em> (Chinese wolf spider)</td>
</tr>
<tr>
<td>Venom peptide Ocy2</td>
<td><em>Opisthacanthus cayaporum</em> (South American scorpion)</td>
</tr>
</tbody>
</table>
• Early diverging metazoans offer a phylogenetic anchor to study evolution of the venom trait.

• Venom proteomes of the scyphozoan *Chrysaora lactea* and two cubozoans *Tamoya haplonema* and *Chiropsalmus quadrumanus* are presented.

• Toxin recruitment and retention patterns do not always correlate with accepted phylogeny.

• Factors that drive toxin diversification independent of phylogeny merit closer scrutiny.