

## Article

# Transcriptomic Insights into the Diversity and Evolution of Myxozoa (Cnidaria, Endocnidozoa) Toxin-like Proteins

Bin Xiao <sup>1,2,3</sup> , Qingxiang Guo <sup>1,2,3</sup> , Yanhua Zhai <sup>1,2,3</sup> and Zemao Gu <sup>1,2,3,\*</sup> 

<sup>1</sup> Department of Aquatic Animal Medicine, College of Fisheries, Huazhong Agricultural University, Wuhan 430070, China; ben.xiaobin@outlook.com (B.X.); qingxiang.guo@outlook.com (Q.G.); zhaiyh@mail.hzau.edu.cn (Y.Z.)

<sup>2</sup> Hubei Engineering Technology Research Center for Aquatic Animal Diseases Control and Prevention, Wuhan 430070, China

<sup>3</sup> Engineering Research Center of Green Development for Conventional Aquatic Biological Industry in the Yangtze River Economic Belt, Ministry of Education, Wuhan 430070, China

\* Correspondence: guzema@mail.hzau.edu.cn; Tel.: +86-027-8728-2114

**Abstract:** Myxozoa is a speciose group of endoparasitic cnidarians that can cause severe ecological and economic effects. Their cnidarian affinity is affirmed by genetic relatedness and the presence of nematocysts, historically called “polar capsules”. Previous studies have revealed the presence of toxin-like proteins in myxozoans; however, the diversity and evolution of venom in Myxozoa are not fully understood. Here, we performed a comparative analysis using the newly sequenced transcriptomes of five Myxobolidae species as well as some public datasets. Toxin mining revealed that myxozoans have lost most of their toxin families, while most species retained Kunitz, M12B, and CRISP, which may play a role in endoparasitism. The venom composition of Endocnidozoa (Myxozoa + *Polypodium*) differs from that of free-living cnidarians and may be influenced by ecological and environmental factors. Phylogenetic analyses showed that toxin families of myxozoans and free-living cnidarians were clustered into different clades. Selection analyses showed that purifying selection was the dominant evolutionary pressure in toxins, while they were still influenced by episodic adaptive selection. This suggests that the potency or specificity of a particular toxin or species might increase. Overall, our findings provide a more comprehensive framework for understanding the diversity and evolution of Myxozoa venoms.

**Keywords:** venomics; cnidaria; Myxobolidae; phylogenetics; selection analysis; adaptive evolution



**Citation:** Xiao, B.; Guo, Q.; Zhai, Y.; Gu, Z. Transcriptomic Insights into the Diversity and Evolution of Myxozoa (Cnidaria, Endocnidozoa) Toxin-like Proteins. *Mar. Drugs* **2022**, *20*, 291. <https://doi.org/10.3390/md20050291>

Academic Editor: Sebastien Dutertre

Received: 9 March 2022

Accepted: 22 April 2022

Published: 26 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cnidarians (anthozoans, medusozoans, and endocnidozoans) are among the earliest diverging extant venomous animals [1,2]. Their toxins are produced by the Golgi apparatus of stinging cells (cnidocytes), which are then housed in cells called nematocysts [3]. The venom of medically relevant species (box jellyfish), or easily collected species (sea anemones), has been extensively studied [4]. However, relatively few studies have been conducted within the parasitic cnidarians—Endocnidozoa. Considering that ecological factors have a profound influence on intra-species venom variation [5], this bias will not only compromise our understanding of venom diversity but also limit the exploration of venom function [6].

Myxozoans are obligate endoparasitic cnidarians that form Endocnidozoa with *Polypodium* hydriforme [7,8]. They have a complex life cycle that involves vertebrate and invertebrate hosts [9,10]. Their cnidarian affinity is confirmed by genetic relatedness and the presence of nematocysts known as polar capsules [11]. Unlike their free-living relatives, myxozoans do not use their nematocysts for feeding or defense, but rather discharge the tubules to anchor the spore to the host during infection [12]. It is unclear how this unique life history and the different utilization of nematocysts affect the venom composition of myxozoans.

Advances in omics technologies have enabled the investigation of the toxin components in several myxozoans at transcriptomic or proteomic level. Comparative analysis using the *Kudoa iwatai* genome and transcriptome did not detect venom proteins among the cnidarian-restricted genes [13]. This was further supported by the proteomic analysis of polar capsules isolated from *Ceratonova shasta*, which found minimal evidence of proteins similar to the toxins [14]. These studies suggested a hypothesis that myxozoans might lose their venom during the transition from a free-living lifestyle to parasitism. However, subsequent venom studies indicated that venom might be retained in myxozoans and used to infect hosts. For instance, transcriptomic analysis of *Myxobolus pendula* identified 49 putative toxin proteins [15]. Venomic analysis of some myxozoans provided evidence for the expression and translation of toxin homologs [6,16]. These studies have conducted pioneering works and provided valuable datasets. However, a comprehensive understanding of venom in the entire Myxozoa group is still lacking. Furthermore, previous studies have not fully explored the evolution of Myxozoa venoms. Unraveling the evolutionary patterns of venom in these dual-host parasitic groups may provide new insights into the mechanisms of venom diversification and molecular evolution in cnidarians.

Myxobolidae is a widespread and diverse component of Myxozoa. With some 1300 species currently described, they represent around 50% of the species diversity of myxozoans [17–19]. Certain Myxobolidae species cause emerging diseases. For example, *Myxobolus cerebralis* can cause the well-known whirling disease [20]. Some species of *Myxobolus* and *Thelohanellus* are also pathogenic to their hosts [21–23]. These diseases have impacted wild populations of some most iconic fish, resulting in substantial economic loss in aquaculture and fisheries [24]. As a result of these developments, Myxobolidae is now considered to be of ecological, economic, and medical significance, thus representing an excellent candidate for venom study.

Here, we conducted venom research on newly sequenced transcriptomes of five Myxobolidae species (*Thelohanellus kitauei*, *Myxobolus xiantaoensis*, *Myxobolus ampullicapsulatus*, *Myxobolus turpisrotundus*, and *Myxobolus honghuensis*) and some representative cnidarians. In the present work, we aimed to (i) reveal the distribution patterns of venom in the Myxozoa through a customized bioinformatics pipeline; (ii) compare the venom gene profiles between myxozoans and free-living cnidarians; (iii) explore the evolution of myxozoan venoms using phylogenetic and selection pressure analyses.

## 2. Results

### 2.1. Sequencing and De Novo Assembly of Transcriptome

A total of 67 gigabases (Gb) of compressed raw reads were generated for five myxozoans using paired-end sequencing. The assemblies ranged from 23,995 to 174,559 unigenes with an N50 range from 381 to 1368. Overall completeness was evaluated by CEGMA and BUSCO (Table 1).

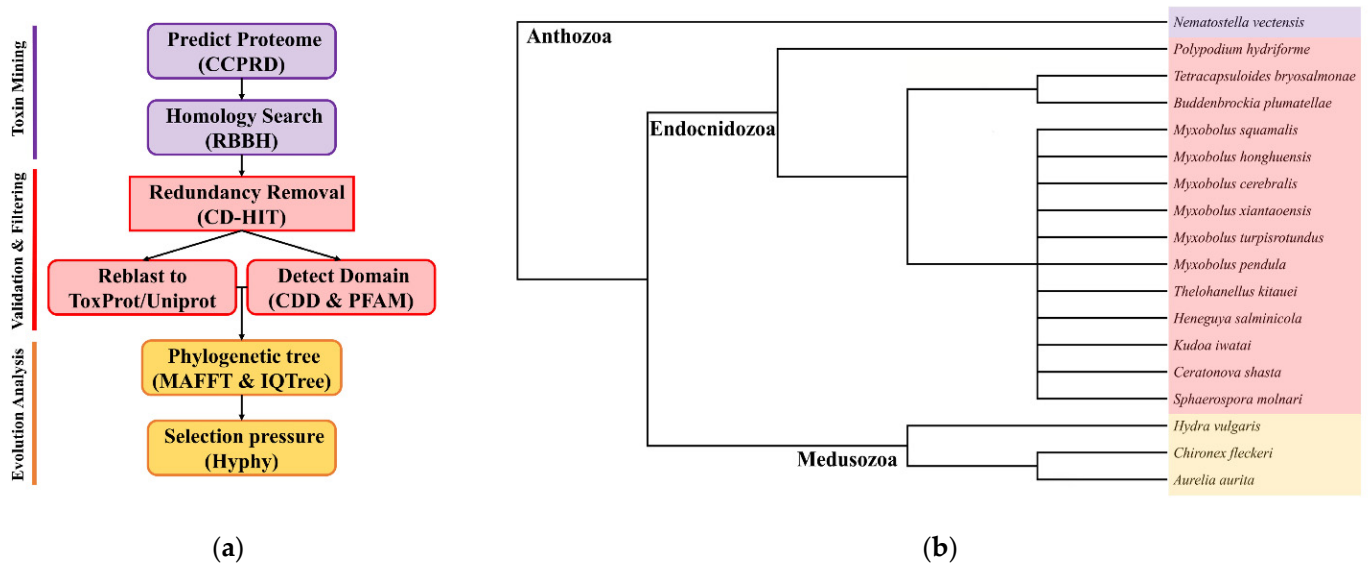
**Table 1.** Sequencing and assembly summary for five Myxobolidae transcriptomes.

Species	Contigs	Unigene	Mean Length	N50	CEGMA	BUSCO
<i>T. kitauei</i>	90,775	88,081	590	1118	67.3%	57.3%
<i>M. xiantaoensis</i>	98,837	66,662	844	1368	77.8%	78.4%
<i>M. ampullicapsulatus</i>	276,231	174,559	380	381	71.8%	64.7%
<i>M. turpisrotundus</i>	171,878	153,649	453	759	62.1%	57.3%
<i>M. honghuensis</i>	27,502	23,995	6013	857	59.6%	59.6%

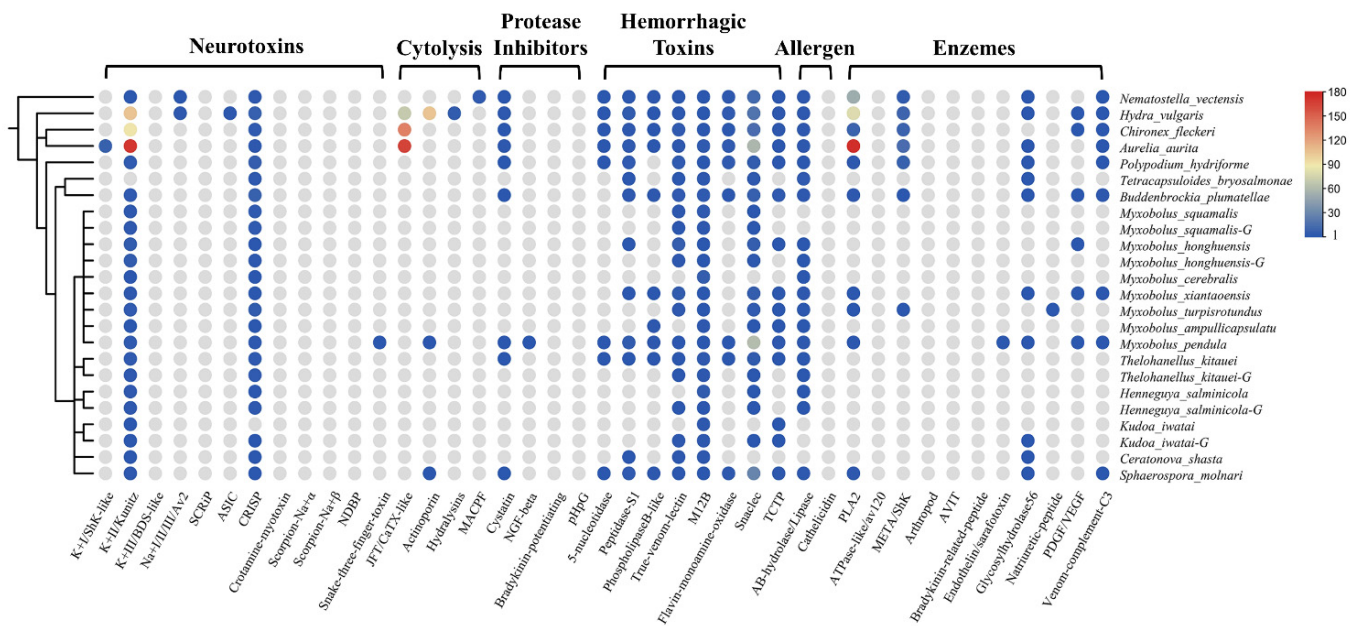
### 2.2. Identification of Toxin-like Proteins (TLPs) in Transcriptomes and Genomes

Using the customized bioinformatics pipeline depicted in Figure 1a, we found a diverse set of TLPs. Toxins were classified as neurotoxins, cytolytins, protease inhibitors, hemorrhagic toxins, allergens, and enzymes according to their biological functions. The presence and absence of toxin families are shown in Figure 2. We also calculated the proportion of different toxin types in all datasets (Table 2). In general, we found significantly

fewer TLPs in myxozoans (range 4 to 88) compared to free-living cnidarians (88, 419, 273, and 611 for *Nematostella vectensis*, *Hydra vulgaris*, *Chironex fleckeri*, and *Aurelia aurita*).



**Figure 1.** (a) The customized bioinformatics pipeline used in this study identified and filtered toxin-like proteins from the sequenced datasets and available omics data; (b) cladogram of species used in this study including 14 Endocnidozoa taxa. Adapted from Klompen et al. (2021) and Hartigan et al. (2021) [6,25].



**Figure 2.** The presence and absence of venom protein families in myxozoans and some cnidarians. A schematic tree of the species used in this project is shown on the left. The circles indicate the presence of a toxin family in that species, and the different colors represent the number of that family. The color gradient on the right corresponds to the numbers. Gray means absent.

**Table 2.** Composition of venom proteins identified in myxozoans and some cnidarians.

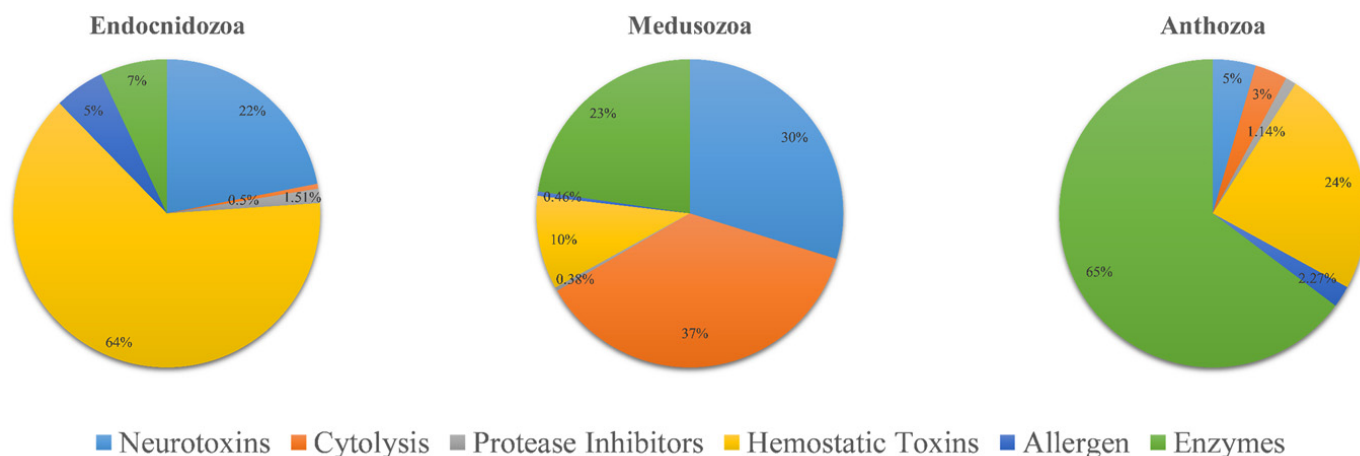
Class/Toxins	Neurotoxins	Cytolysins	Protease Inhibitors	Hemorrhagic Toxins	Allergen	Enzymes	ALL
<i>N. vectensis</i>	4/4.5%	3/3.4%	1/1.1%	21/23.9%	2/2.3%	57/64.8%	88
<i>H. vulgaris</i>	115/27.4%	179/42.7%	2/0.5%	31/7.4%	2/0.5%	90/21.5%	419
<i>C. fleckeri</i>	93/34.1%	136/49.8%	1/0.4%	25/9.2%	2/0.7%	16/5.9%	273
<i>A. aurita</i>	180/29.5%	165/27.0%	2/0.3%	72/11.8%	2/0.3%	190/31.1%	611
<i>P. hydriforme</i>	8/22.2%	0/0.0%	1/2.8%	16/44.4%	1/2.8%	10/27.8%	36
<i>T. bryosalmonae</i>	2/16.7%	0/0.0%	0/0.0%	6/50.0%	3/25.0%	1/8.3%	12
<i>B. plumatellae</i>	10/27.0%	0/0.0%	1/2.7%	20/54.1%	2/5.4%	4/10.8%	37
<i>M. squamalis</i>	5/45.5%	0/0.0%	0/0.0%	6/54.5%	0/0.0%	0/0.0%	11
<i>M. squamalis-G</i> <sup>1</sup>	5/55.6%	0/0.0%	0/0.0%	4/44.4%	0/0.0%	0/0.0%	9
<i>M. honghuensis</i>	5/26.3%	0/0.0%	0/0.0%	13/68.4%	1/5.3%	0/0.0%	19
<i>M. honghuensis-G</i> <sup>1</sup>	6/50.0%	0/0.0%	0/0.0%	5/41.7%	1/8.3%	0/0.0%	12
<i>M. cerebralis</i>	2/25.0%	0/0.0%	0/0.0%	5/62.5%	1/12.5%	0/0.0%	8
<i>M. xiantaoensis</i>	3/14.3%	0/0.0%	0/0.0%	14/66.7%	2/9.5%	2/9.5%	21
<i>M. turpisrotundus</i>	7/38.9%	0/0.0%	0/0.0%	7/38.9%	1/5.6%	3/16.7%	18
<i>M. ampullicapsulatu</i>	2/22.2%	0/0.0%	0/0.0%	6/66.7%	1/11.1%	0/0.0%	9
<i>M. pendula</i>	6/6.8%	1/1.1%	2/2.3%	74/84.1%	2/2.3%	3/3.4%	88
<i>T. kitauei</i>	6/30.0%	0/0.0%	1/5.0%	12/60.0%	1/5.0%	0/0.0%	20
<i>T. kitauei-G</i> <sup>1</sup>	5/50.0%	0/0.0%	0/0.0%	4/40.0%	1/10.0%	0/0.0%	10
<i>H. salminicola</i>	3/33.3%	0/0.0%	0/0.0%	5/55.6%	1/11.1%	0/0.0%	9
<i>H. salminicola-G</i> <sup>1</sup>	4/50.0%	0/0.0%	0/0.0%	3/37.5%	1/12.5%	0/0.0%	8
<i>K. iwatai</i>	1/25.0%	0/0.0%	0/0.0%	3/75.0%	0/0.0%	0/0.0%	4
<i>K. iwatai-G</i> <sup>1</sup>	2/22.2%	0/0.0%	0/0.0%	6/66.7%	0/0.0%	1/11.1%	9
<i>C. shasta</i>	2/28.6%	0/0.0%	0/0.0%	4/57.1%	0/0.0%	1/14.3%	7
<i>S. molnari</i>	3/6.0%	1/2.0%	1/2.0%	40/80.0%	2/4.0%	3/6.0%	50

<sup>1</sup> Represents the genome.

The venom system of cnidarians is mainly composed of three toxin types: enzymes, pore-forming toxins, and neurotoxins [2,26]. As expected, we could not find most of the above toxin families in myxozoans, such as type I and III K<sup>+</sup> channel inhibitors, any Na<sup>+</sup> channel inhibitors, SCRIP, ASIC, JFT, hydralysins, MACPF, or ATPase-like toxins. However, we found M12B in all species ( $n = 37$ ), which is classed as a hemorrhagic toxin. Most of the myxozoans, except *Tetracapsuloides bryosalmonae*, possessed the type II K<sup>+</sup> channel inhibitor Kunitz ( $n = 39$ ). Kunitz is a dual function toxin that inhibits serine protease and blocks potassium channels [27]. We also detected five PLA2 TLPs in *M. xiantaoensis*, *M. turpisrotundus*, *M. pendula*, *Sphaerospora molnari*, and *Buddenbrockia plumatellae*. Interestingly, we found two actinoporin TLPs in *M. pendula* and *S. molnari*, a family of venom proteins usually reported in anthozoans and hydrozoans [2].

Our survey also included toxin families from snakes, scorpions, spiders, cone snails, and insects. The most diverse type was snakec (  $n = 124$  ) with *M. pendula* and *S. molnari* contributing 62 and 28 respectively. Nearly all species had true venom lectin ( $n = 28$ ) and CRISP ( $n = 39$ ). Interestingly, we found many TCTP in myxozoans, which was an important part of the sea anemone defense system [28]. We also detected snake three-finger toxin (TFT) in *M. pendula*. TFT is a non-enzymatic peptide with multiple biological activities found in all snakes [29].

The venom composition is shown in Figure 3. Overall, the pattern of TLPs in Endocnidozoa was similar. Hemorrhagic toxins were the most diverse family of toxins; they accounted for more than 63% of TLPs. The next most common were neurotoxins (21.9%). Enzymes, allergens, protease inhibitors, and cytolysins were also found in the endocnidozoans, but they accounted for only a small fraction of TLPs. In *N. vectensis*, we found that the most diverse toxin family was enzymes, which constituted more than 64.7% of TLPs. For medusozoans, cytolysins were found to be the most diverse toxin family (36.8%).



**Figure 3.** Comparative analysis of putative toxins identified from transcriptomic datasets. Transcripts were identified as coding for potential venoms: Anthozoa (*N. vectensis*,  $n = 88$ ), Medusozoa (*H. vulgaris*, *C. fleckeri*, *A. aurita*,  $n = 1303$ ), and Endocnidozoa (*P. hydriforme* + some myxozoans,  $n = 397$ ).

Our analysis was mainly based on transcriptomes, but we also added five genomes (*M. squamalis*, *T. kitauei*, *H. salminicola*, *K. iwatai*, and *M. honghuensis* [23,30,31]) to the project to detect the extent of trait differences. When comparing the results of genomic and transcriptomic analyses, we found that most of the species had more TLPs in the transcriptomes, such as *M. squamalis* (11 vs. 9), *M. honghuensis* (19 vs. 12), *T. kitauei* (20 vs. 10), and *H. salminicola* (9 vs. 8). However, there were more TLPs found in the *K. iwatai* genome ( $n = 9$  vs. 4 in transcriptome).

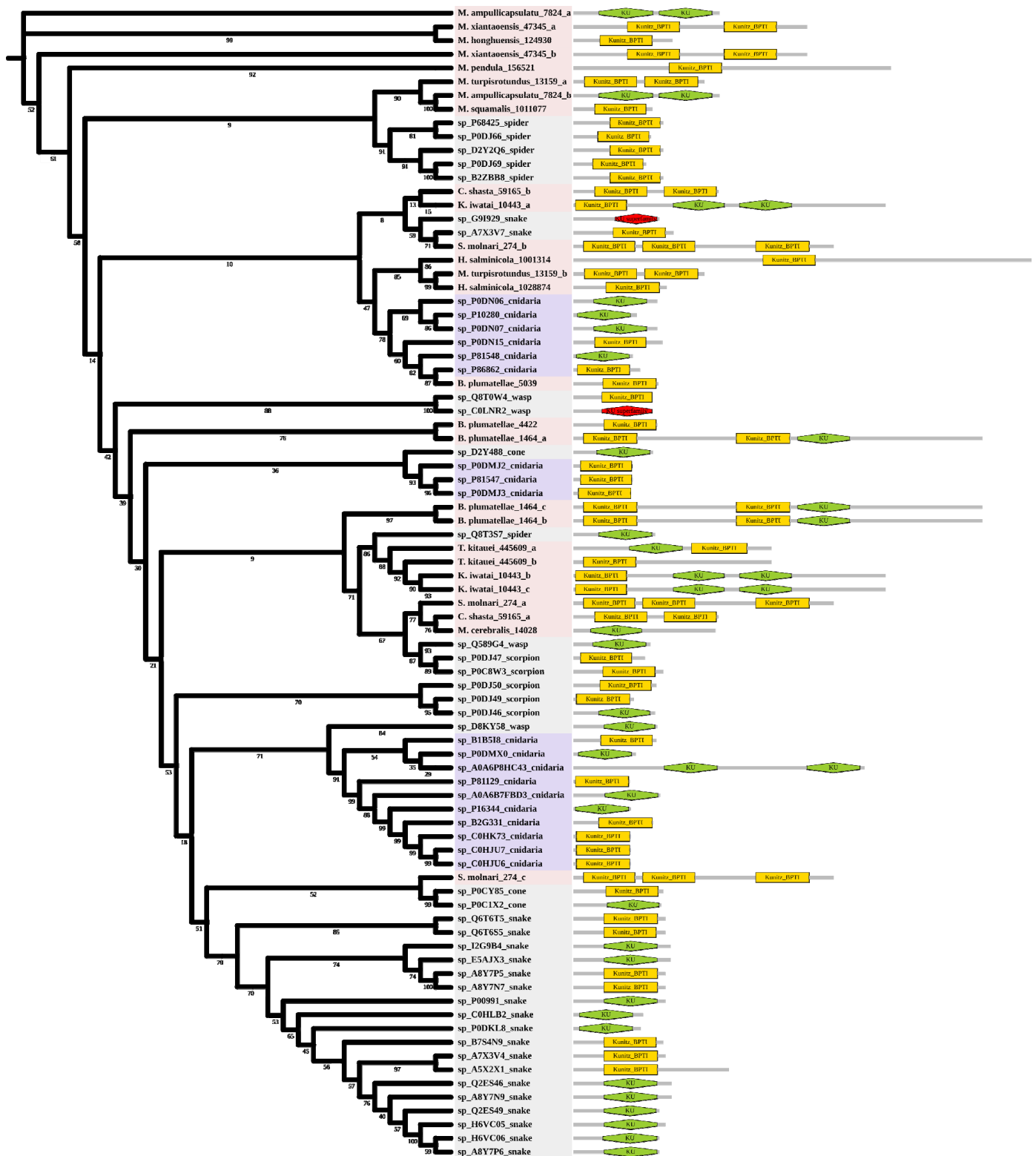
### 2.3. Phylogenetic Analysis of Toxin Families

We have developed a phylogenetic framework to reveal the evolutionary patterns and interspecies differences of Myxozoa venoms. Seven toxin families were selected for in-depth analysis, including Kunitz, CRISP, actinoporin, peptidase S1, true venom lectin, M12B and PLA2. These families were selected because they were found in most species or had essential biological functions. These both were frequently studied in the cnidarian literature and other venomous lineages, allowing easier comparison.

#### 2.3.1. Neurotoxins

Kunitz is a peptide with a molecular weight of about 6 kDa and usually has a dual function. Six myxozoan clades were supported. Some myxozoan sequences clustered with non-cnidarian species include the sequence of *S. molnari* clustered with the snake (A7X3V7), which is a serine protease inhibitor found in *Philodryas olfersii* [32]. The branch containing *T. kitauei*, *K. iwatai*, and spider (Q8T3S7) was supported by a relatively high bootstrap (86). Q8T3S7 was a putative insecticidal toxin that could inhibit trypsin [33]. We found a sequence of *S. molnari* clustered with cone snails (P0CY85 P0C1X2). They both specifically block voltage-activated potassium channels [34,35]. The free-living cnidarians formed three clades and had little intermixing with myxozoans. In contrast to other venomous taxa, domain duplication existed in both free-living cnidarians and myxozoans. We found two KU domains in *Actinia tenebrosa* (A0A6P8HC43), and three KU/Kunitz-BPTI domains in most myxozoans (Figure 4), which meant the duplication occurred twice in myxozoan at least.

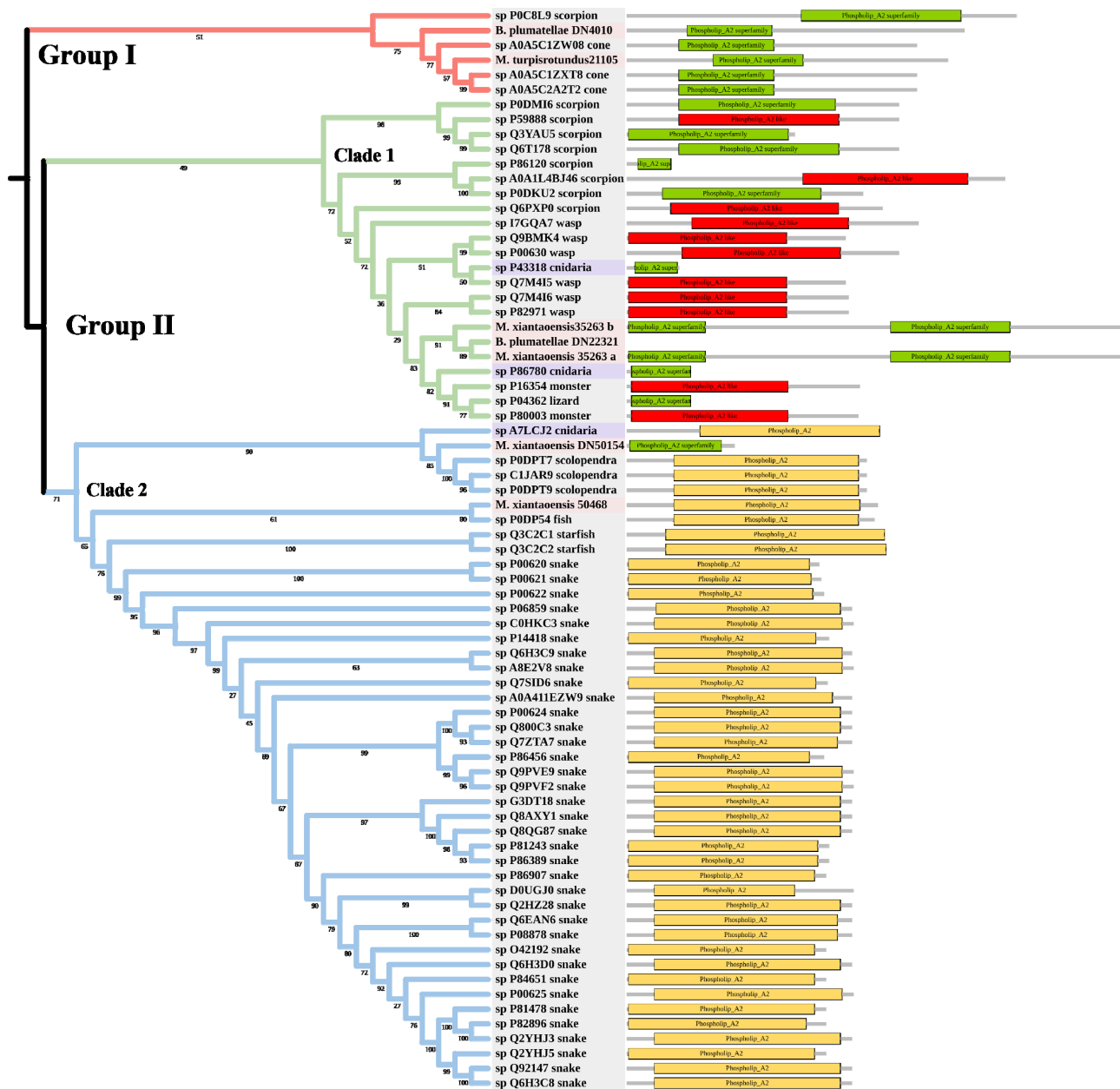
CRISP is a single-polypeptide protein with a molecular weight of 20 to 30 kDa [36]. In reptilian venom, CRISP blocks nucleotide-gated and inhibits potassium channels [37,38]. Two myxozoan clades were supported but did not cluster with other venomous taxa (Supplementary Materials Figure S1).



**Figure 4.** Phylogenetic tree of the Kunitz venom protein family. The maximum likelihood (ML) tree was constructed based on amino acid sequences using IQ-TREE version 2.1.1 software. The numbers below the nodes show bootstrap support values from 1000 replicates. ML phylogram was generated from amino-acid alignments of Kunitz homologs using the auto model. The location and length of conserved domains for each amino acid are shown on the right. Putative sequences outlined in pink, purple, and grey are from myxozoans, free-living cnidarians, and other toxic organisms.

### 2.3.2. Enzymes

PLA2 has been found in anthozoans, scyphozoans, hydrozoans, and cubozoans [39,40]. The toxic functions of PLA2 in cnidarian venoms include defense, prey digestion, and hemolytic activity [41,42]. The tree was divided into Group I and Group II (Figure 5). Two myxozoan sequences, three cone snail sequences, and one scorpion sequence formed Group I. Group II contained two main clades. Clade 1 included many venomous taxa, such as scorpions and wasps. The three myxozoan sequences were classed into Clade 1. Clade 2 mainly included snakes. There were only two myxozoan sequences in Clade 2: *S. molnari* clustered with scolopendras, and *M. pendula* clustered with a Blenniidae fish *Meiacanthus atrodorsalis*.



**Figure 5.** Phylogenetic tree of the PLA2 venom protein family. The maximum likelihood (ML) tree was constructed based on amino acid sequences using IQ-TREE version 2.1.1 software. The numbers below nodes show bootstrap support values from 1000 replicates. ML phylogram was generated from amino-acid alignments of PLA2 homologs using the auto model. The location and length of conserved domains for each amino acid are displayed on the right. Putative sequences outlined in pink, purple, and grey are from myxozoans, free-living cnidarians, and other toxic organisms.

### 2.3.3. Cytolysins

Actinoporins are  $\alpha$ -PFTs occurring in anthozoans and hydrozoans, with cardiovascular, respiratory, and cytotoxic effects [43]. We found two actinoporin-like sequences in *M. pendula* and *S. molnari* (Figure S2).

### 2.3.4. Hemorrhagic Toxins

Metalloproteinases were mainly found in snakes. Several biological effects are attributed to them including hemorrhage, hypotension, and necrosis [44]. Most myxozoan sequences were clustered in a large group and divided into three clades. A group of *K. iwatai*, *S. molnari*, and *M. pendula* was clustered. Additionally, we detected domain recruitment in the M12B tree (Figure 6).

Peptidase S1 is an enzyme that cleaves peptide bonds in proteins [45]. The phylogenetic tree supported three myxozoan clades (Figure S3). One group was clustered with snakes. While the other was clustered with arthropods. The last group only consisted of *M. pendula* and *S. molnari*, supporting a high bootstrap (100).

True venom lectins bind carbohydrates in snakes [46], exhibiting hemagglutination and platelet aggregation in other species [47]. Four myxozoan clades were supported, most of them belonging to Myxobolidae (Figure S4). Two sequences from *S. molnari* and *M. pendula* were clustered with toadfish (Q66S03) [48].

## 2.4. Selection Pressure Analysis

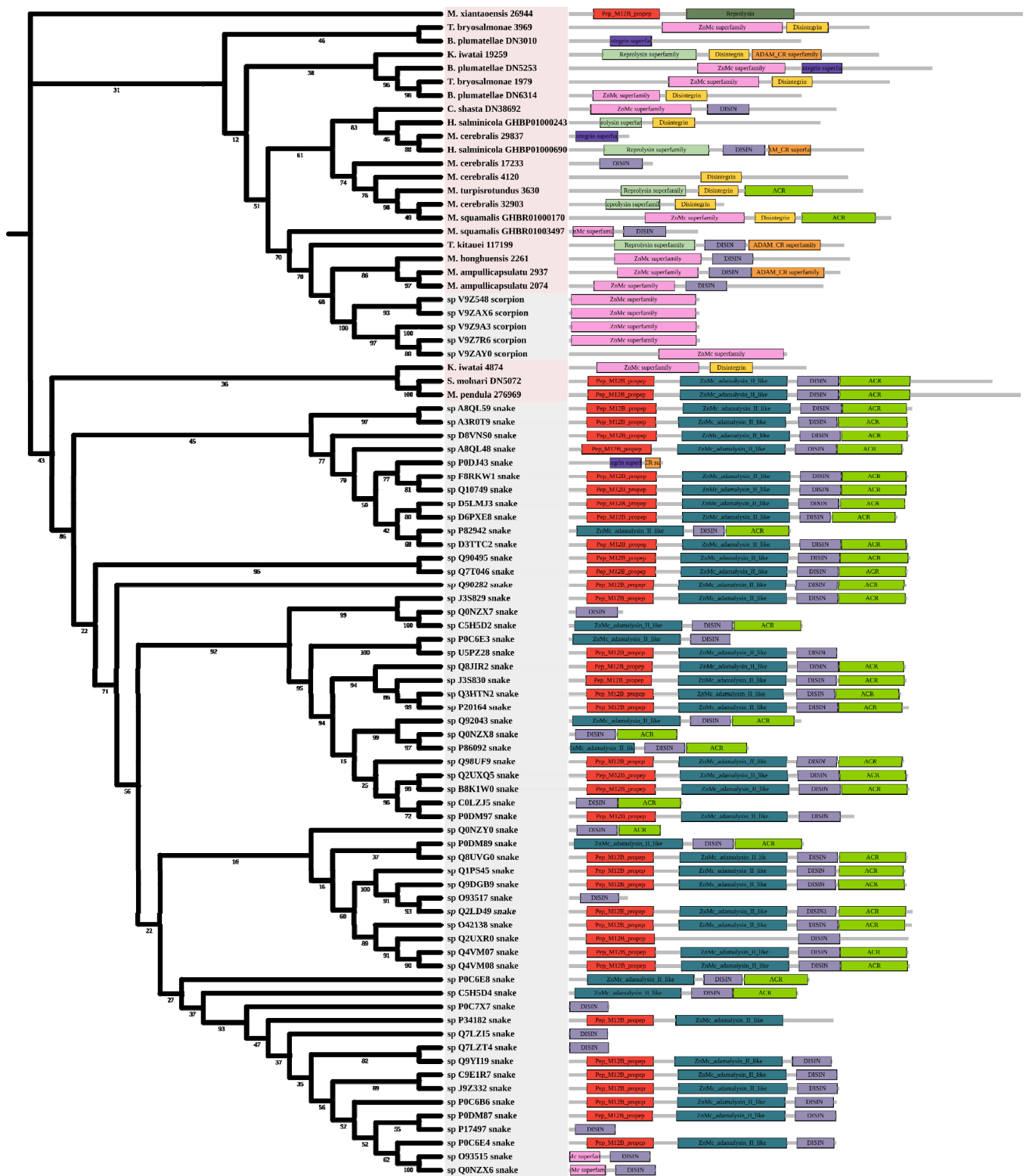
Considering the obvious differences in phylogeny between free-living cnidarians and myxozoans, we also analyzed the selection pressures to further explore the evolutionary patterns of Myxozoa venoms. Using the codeml program in PAML v4.9 [49] with NSITES = 0, we found the overall dN/dS ratio for Kunitz, M12B and CRISP was consistently 1 (Table 3). The dN/dS ratio is higher for M12B ( $\omega = 0.09904$ ) compared to CRISP ( $\omega = 0.09041$ ) and Kunitz ( $\omega = 0.05767$ ).

**Table 3.** Results of selection pressure analysis for toxin families.

Method/Family	Kunitz	CRISP	M12B
Codeml	0.05767	0.09041	0.09904
FEL	−32	−98	−29
MEME	1	3	4
aBSREL	0	2	1
BUSTED	found no evidence	found evidence	found no evidence
RELAX	K 1	K 1	K 1

Furthermore, we used gene-specific and branch-specific tests implemented through HyPhy [50]. First, we used the Fixed Effects Likelihood (FEL) [51] test to validate the dN/dS ratio. We found that 32 sites in Kunitz, 98 sites in CRISP, and 29 sites in M12B were under purifying selection. Using the Mixed Effects Model of Evolution (MEME) [52] test, we found evidence of episodic individual-sites positive selection across all toxin families. Four sites were found to be evolving under the influence of episodic diversifying selection in M12B. Three sites were found in CRISP and one site in Kunitz. Adaptive branch-site random effects likelihood (aBSREL) [53] found evidence for episodic positive selection across all toxin families except for Kunitz. We detected two branches in CRISP and one branch in M12B. We also used Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED) [54] to detect episodic gene-wide positive selection. The results showed that CRISP had undergone positive selection at least one site on at least one branch. Using the RELAX [55] test, we found the k value for Kunitz and CRISP was 0.88 and 0.99 respectively. Their selection intensification was not significant. While the k value of M12B was 1.07, the test of selection intensity was not significant.





**Figure 6.** Phylogenetic tree of the M12B venom protein family. The maximum likelihood (ML) tree was constructed based on amino acid sequences using IQ-TREE version 2.1.1 software. The numbers below nodes show bootstrap support values from 1000 replicates. ML phylogram was generated from amino-acid alignments of M12B homologs using the auto model. The location and length of conserved domains for each amino acid are displayed on the right. Putative sequences outlined in pink, purple, and grey are from myxozoans, free-living cnidarians, and other toxic organisms.

### 3. Discussion

Here, we sequenced and assembled the transcriptomes of five Myxobolidaes species, with BUSCO scores ranging from 57–78%. However, we do not think this is a defect in the data and will not affect the validity of our results. First, the fast evolutionary rates of myxozoans have led to underperformance in both BUSCO and CEGMA [30]. Secondly, the transcriptome sequencing was deep and our samples contained spores from different stages [56].

Our results suggest that the venom composition of endocnidozoans differs from that of anthozoans and medusozoans. Hemorrhagic toxins were the most diverse type of toxin family in endocnidozoans. Anthozoans were rich in enzymes and medusozoans contained more cytolytic and neurotoxins. There are strong indications that various ecological and environmental factors (diet, gender, and geographical distribution) might influence venom composition in animals [57]. We speculate that the use of venom in myxozoans has changed during the transitions to parasitism. They did not use venom for hunting or defense in the same manner as their free-living relatives. A large number of hemorrhagic toxins may be used for host infection or immune escape. Further experimental exploration is needed to understand this question. In addition, we found there were few differences in the types and composition patterns of venom among myxozoans. We speculate that different myxozoans in different living environments may be similar in their utilization of the venom. Future research can expand the scope of species, and combine bioinformatics analysis with functional experiments to explore this issue, which may help explain some interesting biological phenomena in myxozoans.

We noted some differences in venom-mining results at the transcriptomic and genomic levels. Overall, we found more TLPs in transcriptomes, suggesting that gene duplication may occur in some toxin families. Some toxin families were found in transcriptomes but not in genomes, we speculate that this may be caused by alternative splicing, which has been proved to influence transcriptional and post-transcriptional regulation of venom genes [58]. Considering that our proposed set of toxin-like proteins is seen as a first step in exploring the diversity and evolution of an understudied group, future research can build on this foundation to examine the effects of alternative splicing on the myxozoan venom, and determine the mechanisms of diversification. Meanwhile, some toxin families were only found in the genome. This can be a result of selective loss or “decommissioning” of protein families from the venom [1]. We have detected many TLPs with different functions in myxozoans, but the protein expression level of TLPs may be inconsistent with the hints in genomes and transcriptomes. For example, it has been shown in the mixed myxosporean that even when more serine peptidase sequences are present in the transcriptome, the expression of neurotoxins is greater overall at the proteomic level [6]. Thus, future quantitative gene expression and proteomic studies are needed to provide a holistic understanding of myxozoan venoms.

Compared to free-living cnidarians, myxozoans have lost most of their toxin families. However, we can identify a small set of toxin families common to almost all myxozoans, such as Kunitz, CRISP, and M12B. These “core” toxin families may consist of proteins with an ancient and conserved role in nematocyst function [1], which may play a key role in the adaptive evolution of myxozoans. Previous studies support the idea that these protein families are closely related to parasitism. For example, Kunitz is identified in the transcriptome of *Echinococcus granulosus* and has been shown to interfere with host physiological processes during the initial stages of infection [59]. In addition, CRISP is thought to play an important role in the development and parasite adaptation of cestodes [60]. Metalloproteinases have been identified in ectoparasitic wasps with abilities to manipulate host development [61]. To gain a better understanding of these “core” toxin families, we focused on their evolutionary history in our subsequent studies.

Kunitz-type venom is present in proteins from most venomous taxa and plays an important role in envenomation [62]. In sea anemones, Kunitz is classified as a type II  $K^+$  channel inhibitor. It protects the organisms from endogenous proteases, causes paralysis

in prey, and defends against enemies [63,64]. Molecular phylogenetic analysis of Kunitz showed that myxozoans and free-living cnidarians are clustered into distinct clades. Some myxozoan sequences are grouped with other venomous animals. Our results support the hypothesis proposed by Hartigan et al. (2021) that the toxins of endoparasitic cnidarians are more divergent from homologs in free-living relatives [6]. This may be caused by the potential convergent evolution of venoms in diverse animal groups. Interestingly, we found that the Kunitz conserved domain is duplicated in myxozoans and free-living cnidarians, although not in other toxic organisms. Conserved domains represent the basic evolutionary units that form proteins. Domain duplication and reorganization are the most important forces driving protein evolution and increasing proteome complexity [65]. We also note that domain duplication differs in the two groups. Duplication occurred twice in myxozoans and only once in free-living cnidarians. We hypothesize that the Kunitz domain was first duplicated in their common ancestor, and there is a separate duplication for myxozoans after the divergence of endoparasitism and free-living. Further experiments are needed to explore the effect of duplication on the function of Kunitz.

Metalloproteases are important venom components of terrestrial animals [66]. They have also been detected in jellyfishes and sea urchins [67–69] with gelatinolytic, caseinolytic, and fibrinolytic activity [49]. The toxicity of CRISP in cnidarians is not fully understood, but it has been detected in anthozoans, scyphozoans, and hydrozoans [1]. M12B and CRISP shared similar phylogenetic results. Most of the myxozoans did not cluster with other venomous taxa. We detected domain recruitment in M12B and domain divergence in CRISP. In many venomous taxa, toxins evolve through a combination of gene recruitment, duplication, and neofunctionalization, leading to extended families of closely related yet often functionally distinct toxin isoforms [70]. In that case, myxozoan may have evolved new functions of M12B and CRISP to support the endoparasitic lifestyles.

The evolution of venom is usually accompanied by strong evidence of accelerated evolution and positive selection [71]. This is especially true in recently diverged venomous clades, such as snakes, scorpions, spiders, and cone snails [72–75]. In contrast, venom evolution is dominated by purifying selection in ancient venomous lineages (e.g., cnidarians, coleoids, and arthropods) [76]. In the present work, we found that the Myxozoa venoms were under purifying selection, which is consistent with Cnidaria [77]. Why do the evolutionary patterns of cnidarian venoms differ from that of other venomous organisms? This may be because venom production in cnidarians is associated with the development of nematocysts. This will pose genomic constraints on the evolution of venom in cnidarians [1]. However, we detected episodic positive selection on sites or branches in myxozoan toxin families using MEME and aBSREL, which suggests that weak diversity selection was involved in Myxozoa venoms. We also interpret this discrete positive selection as concomitant with increased potency and/or specificity of a particular toxin or species. Overall, these patterns suggest that myxozoan venoms have undergone a varied history of selection pressures, including widespread purifying selection and a few instances of gene-specific or lineage-specific episodic positive selection.

Collectively, we provide a comprehensive view of Myxozoa venoms. We detected relatively few toxin-like proteins in myxozoans, suggesting that the reduction in myxozoan genome size includes reduced toxin diversity [6]. The phylogenetics and selection pressure results reflect differences in venoms between myxozoans and free-living cnidarians, suggesting that toxins might have evolved to adapt to parasitic lifestyles. The set of toxin-like proteins we present is viewed as an initial step into exploring diversity and evolution within a poorly studied group of cnidarians. These new data will be used to further explore the diversification and molecular evolution of toxins encoded by these fascinating and ancient animals.

## 4. Materials and Methods

### 4.1. Collecting Sample

*M. honghuensis* was collected from infected allogynogenetic gibel carp *Carassius auratus gibelio* in Zoumaling Farm, Hubei Province, China on 29 July 2015. *M. ampullicapsulatus* was collected from infected allogynogenetic gibel carp *C. auratus gibelio* in Datonghu lake, Hubei Province, China on 12 May 2016. *M. xiantaoensis* was collected from infected yellow catfish *Tachysurus fulvidraco* in Xiantao, Hubei Province, China on 23 April 2016. *M. turpisrotundus* was collected from infected allogynogenetic gibel carp *C. auratus gibelio* in Wuhan, Hubei Province, China on 16 November 2015. *T. kitauei* was collected from infected common carp *Cyprinus carpio* in Liuerbao Town, Shenyang Province, China on 11 August 2015.

Fish were held on ice before being killed with an overdose of MS-222 (Sigma-Aldrich, Co., Ltd., St. Louis, MO, USA). From each fish, tissue containing one large cyst was homogenized by a manual glass tissue grinder and suspended in 0.1 M phosphate-buffered saline (PBS) with a pH 7.2 and then filtered through cotton gauze. Myxospores were separated from the filtrate by sucrose gradient centrifugation and Percoll gradient centrifugation, in turn. They were washed several times with distilled water and then examined microscopically to verify purity and identity. Purified myxospores were either placed into RNAlater (Sigma), frozen in liquid nitrogen, and finally stored at  $-80\text{ }^{\circ}\text{C}$ , or immediately sent to nematocyst isolation as described below. Myxozoan identification was performed based on morphology and 18S sequencing [78,79]. Maintenance and care of experimental animals complied with the National Institutes of Health Guide for the care and use of laboratory animals [80] and was approved by the animal care and use committee of Huazhong Agricultural University, China (HZAUI-2015-011).

### 4.2. Next-Generation Sequencing and Assembly

Total RNA was extracted from frozen preserved specimens using TRIzol 550 (Invitrogen, Carlsbad, CA, USA). The purity and integrity of RNA were assessed using NanoPhotometer<sup>®</sup> spectrophotometer (IMPLEN, Westlake Village, CA, USA), and RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA) respectively. Libraries were prepared from purified mRNA using NEBNext<sup>®</sup> Ultra RNA kit (New England Biolab (NEB), Ipswich, MA, USA) following the manufacturer's recommendations and sequenced as  $2 \times 125$  paired-end (PE) runs with the Illumina HiSeq 2500 for *M. honghuensis* and *T. kitauei*; as  $2 \times 150$  paired-end runs with the Illumina HiSeq 4000 for *M. ampullicapsulatus*, *M. turpisrotundus*, and *M. xiantaoensis*. Raw sequence data were cleaned and trimmed by removing adaptor and low-quality reads using Trimmomatic v0.33 [81]. Filtered reads were de novo assembled through Trinity v2.6.6 [82] and clustered using CD-HIT v4.6.8 [83] and TGICL v2.1 [84]. BUSCO v4.0 [85] and CEGMA [86] were used to evaluate the completeness of assemblies.

### 4.3. Data Decontamination and Construction Predict Proteomes

Here we used our recently developed method to construct a clean, efficient, and comprehensive protein reference that we denote as the customized comprehensive proteomic reference database (CCPRD) (for details, see [56]). Briefly, nonredundant host databases containing proteins or nucleotide sequences from the host and nonredundant closely related databases containing proteins or nucleotide sequences from species closely related to myxozoans were constructed. These four databases were then blasted against the transcriptome assemblies using TBLASTN or TBLASTX in BLAST+ v2.4.0 [87] (e-value  $1 \times 10^{-5}$ ). Only transcriptome sequences that exclusively matched host databases were removed. The retained transcriptome sequences were further blasted against a bacterial database using BLASTX (e-value  $1 \times 10^{-10}$ ). The positive hits were blasted to the NCBI nonredundant protein database (NR) using BLASTX (e-value  $1 \times 10^{-5}$ ). Only sequences that were annotated as "bacteria" were removed. For assessment and visualization of contamination in genomes, taxonomic assignment of each contig was carried out using Blobtools v1.0 [88]. Sequences strongly matched to Chordata and Proteobacteria were excluded. The decontam-

ination process was carried out conservatively to prevent possible over-decontamination that could result in a loss of a large portion of actually expressed genes and proteins.

For RNA-seq data, coding sequences were firstly predicted by TransDecoder [89] and GeneMarkS-T [90]. Next, transcriptome homology-based predictions were c by a customized Perl script Hercules (<https://github.com/qingxiangguo/hercules>, accessed on 5 April 2020). Furthermore, those transcripts that were translated neither by de novo nor homology-based method were translated into amino acid sequences using the Transeq script from the EMBOSS [91]. Finally, all the proteins (30 amino acids) resulting from genomic and RNA-seq data were processed by CD-HIT with a threshold of 100%, to collapse the group into a nonredundant data set, leading to the final predicted proteomes. Besides the data we sequenced, 18 publicly available transcriptomes and genomes were used in this project (for details, see Table S1).

#### 4.4. Identification Toxin in Predict Proteomes

We used a pipeline aimed at conservatively identifying toxins: a group of toxin sequences from cnidarian and other venomous animals was generated as seed for reciprocal best BLAST hits (RBBH) analysis to predict proteomes (e-value  $1 \times 10^{-5}$ ). All the candidate toxin proteins identified by RBBH were validated by blasting against the Tox-Prot UniProtKB/Swiss-Prot [92] and only proteins with the best match to toxins were retained (e-value  $1 \times 10^{-5}$ ). PFAM annotation by HMMER [93] was conducted for those putative toxins. They were considered positive only if the presence of the toxin motif or family domain (partial or complete) could be confirmed in protein sequences. Results were manually curated to confirm that toxin-like sequences matched the detected venom domain from PFAM. CDD annotations were added (See Table S2 for detailed results). The results were visualized using Microsoft Excel, and final figures were constructed in Inkscape v1.0.

#### 4.5. Alignment and Phylogenetic Analysis

We used TBtools v1.098669 [94] to extract the domain of toxin protein sequences for conducting phylogenetic trees. Other sequences were obtained from VenomZone (<https://venomzone.expasy.org>, accessed on 10 May 2020). Retrieved amino acid sequences were aligned with MAFFT v7.205 [95] by the L-INS-i method. Multiple alignments were trimmed using Gblocks v0.91b [96]. Phylogenetic trees were constructed using a maximum likelihood (ML) approach by IQ-TREE v2.1.1 [97] and node support was obtained with 1000 ultrafast bootstrap. Searches with multiple 500 initial parsimony trees were performed. Bootstrap support values were assigned on the best maximum likelihood tree branches. The tree was collapsed and formatted using iTOL v4 [98].

#### 4.6. Selection Analysis

We first extracted the coding sequences of the different toxin families using TBLASTN, and then used MEGA-X [99] for sequence alignment and stop codon movement. Codon alignments were obtained by matching the aligned amino acids to their DNA sequences using pal2nal v14 [100] on the online server. The overall synonymous to nonsynonymous substitutions (dN/dS) ratio values were evaluated using codeml. Then, the package HyPhy was used to conduct several positive selection analyses (FEL, MEME, Absrel, BUSTED, and RELAX). Tree topologies were modified to be consistent with the ML tree for all test runs through Datamonkey [101]. Values of dN/dS 0, =0, and 0 indicate negative selection, neutral evolution, and positive selection, respectively.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/md20050291/s1>. Figure S1: Phylogenetic tree of the CRISP venom protein family, Figure S2: Phylogenetic tree of the actinoporin venom protein family, Figure S3: Phylogenetic tree of the peptidase S1 venom protein family, Figure S4: Phylogenetic tree of the true venom lectin venom protein family, Table S1: Summary of transcriptomes and genomes used in this project, Table S2: Annotation table for putative toxin-like proteins identified for five Myxobolidae species, File S1: Toxin families used as seeds for RBBH in this project.

**Author Contributions:** Conceptualization, B.X., Q.G. and Z.G.; Methodology, B.X., Q.G. and Z.G.; Investigation, B.X. and Q.G.; Resources, Y.Z. and Z.G.; Writing—Original Draft, B.X. and Z.G.; Writing—review editing, B.X., Q.G., Y.Z. and Z.G.; Visualization, B.X., Q.G. and Z.G.; Funding Acquisition, Z.G.; Supervision, Z.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Nature Science Foundation of China (Grant No. 32070431), China Agriculture Research System of MOF and MARA (CARS-46), and Featuring Talents Cultivation Project (Grant No. 4611300108).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data generated and analyzed during this study are included in the published article, its additional files, and publicly available repositories. The raw reads of *M. honghuensis* transcriptome sequencing and PacBio genome sequencing have been deposited at the NCBI Short Read Archive (SRA) with the Bioproject accession numbers PRJNA779260, PRJNA778632, and PRJNA779846. The raw reads of other myxozoan transcriptome sequencing were also deposited under the Bioproject accession numbers in GenBank: *M. ampullicapsulatus* (PRJNA786097), *M. turpisrotundus* (PRJNA786098), *M. xiantaoensis* (PRJNA786274), *T. kitauei* (PRJNA786278, Illumina Hiseq 2500), and *T. kitauei* (PRJNA786276, Illumina Hiseq 4000).

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

## References

- 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. [[CrossRef](#)]
- Jouiaei, M.; Yanagihara, A.A.; Madio, B.; Nevalainen, T.J.; Alewood, P.F.; Fry, B.G. Ancient venom systems: A review on cnidaria toxins. *Toxins* **2015**, *7*, 2251–2271. [[CrossRef](#)] [[PubMed](#)]
- Fautin, D.G. Structural diversity, systematics, and evolution of cnidae. *Toxicon* **2009**, *54*, 1054–1064. [[CrossRef](#)] [[PubMed](#)]
- Klompen, A.M.L.; Macrander, J.; Reitzel, A.M.; Stampar, S.N. Ancient venom systems: A review on cnidaria toxins. *Mar. Drugs* **2020**, *18*, 2251–2271.
- O’Hara, E.; Wilson, D.; Seymour, J. The influence of ecological factors on cnidarian venoms. *Toxicon* **2021**, *9*, 100067. [[CrossRef](#)]
- Hartigan, A.; Jaimes-Becerra, A.; Okamura, B.; Doonan, L.B.; Ward, M.; Marques, A.C.; Long, P.F. Recruitment of toxin-like proteins with ancestral venom function supports endoparasitic lifestyles of Myxozoa. *PeerJ* **2021**, *9*, e11208. [[CrossRef](#)] [[PubMed](#)]
- Chang, E.S.; Neuhof, M.; Rubinstein, N.D.; Diamant, A.; Philippe, H.; Huchon, D.; Cartwright, P. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14912–14917. [[CrossRef](#)] [[PubMed](#)]
- 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*, 68. [[CrossRef](#)]
- Kent, M.L.; Andree, K.B.; Bartholomew, J.L.; El-Matbouli, M.; Desser, S.S.; Devlin, R.H.; Feist, S.W.; Hedrick, R.P.; Hoffmann, R.W.; Khattria, J.; et al. Recent advances in our knowledge of the Myxozoa. *J. Eukaryot. Microbiol.* **2001**, *48*, 395–413. [[CrossRef](#)]
- Lom, J.; Dyková, I. Myxozoan genera: Definition and notes on taxonomy, life-cycle terminology and pathogenic species. *Folia Parasitol.* **2006**, *53*, 1–36. [[CrossRef](#)]
- Americus, B.; Lotan, T.; Bartholomew, J.L.; Atkinson, S.D. A comparison of the structure and function of nematocysts in free-living and parasitic cnidarians (Myxozoa). *Int. J. Parasitol.* **2020**, *50*, 763–769. [[CrossRef](#)] [[PubMed](#)]
- Cannon, Q.; Wagner, E. Comparison of discharge mechanisms of cnidarian cnidae and myxozoan polar capsules. *Rev. Fish. Sci.* **2010**, *11*, 185–219. [[CrossRef](#)]
- Shpirer, E.; Diamant, A.; Cartwright, P.; Huchon, D. A genome wide survey reveals multiple nematocyst-specific genes in Myxozoa. *BMC Evol. Biol.* **2018**, *18*, 138. [[CrossRef](#)] [[PubMed](#)]
- Piriatskiy, G.; Atkinson, S.D.; Park, S.; Morgenstern, D.; Brekhman, V.; Yossifon, G.; Bartholomew, J.L.; Lotan, T. Functional and proteomic analysis of *Ceratonova shasta* (Cnidaria: Myxozoa) polar capsules reveals adaptations to parasitism. *Sci. Rep.* **2017**, *7*, 9010. [[CrossRef](#)]
- Foxx, J.; Ringuette, M.; Desser, S.S.; Siddall, M.E. In silico hybridization enables transcriptomic illumination of the nature and evolution of Myxozoa. *BMC Genom.* **2015**, *16*, 840. [[CrossRef](#)] [[PubMed](#)]
- Americus, B.; Hams, N.; Klompen, A.M.L.; Alama-Bermejo, G.; Lotan, T.; Bartholomew, J.L.; Atkinson, S.D. The cnidarian parasite *Ceratonova shasta* utilizes inherited and recruited venom-like compounds during infection. *PeerJ* **2021**, *9*, e12606. [[CrossRef](#)] [[PubMed](#)]

17. Szekely, C.; Ghosh, S.; Borzak, R.; Goswami, U.; Molnar, K.; Cech, G. The occurrence of known Myxobolus and Thelohanellus species (Myxozoa, Myxosporidia) from Indian major carps with the description of *Myxobolus bandyopadhyayi* n. sp. in West Bengal. *Int. J. Parasitol. Parasites Wildl.* **2021**, *16*, 18–25. [[CrossRef](#)] [[PubMed](#)]
18. Eiras, J.C.; Cruz, C.F.; Saraiva, A.; Adriano, E.A. Synopsis of the species of Myxobolus (Cnidaria, Myxozoa, Myxosporidia) described between 2014 and 2020. *Folia Parasitol.* **2021**, *68*. [[CrossRef](#)] [[PubMed](#)]
19. Eiras, J.C.; Adriano, E.A. A checklist of new species of Henneguya Thelohan, 1892 (Myxozoa: Myxosporidia, Myxobolidae) described between 2002 and 2012. *Syst. Parasitol.* **2012**, *83*, 95–104. [[CrossRef](#)] [[PubMed](#)]
20. Hoffman, G. *Myxobolus cerebralis*, a worldwide cause of salmonid whirling disease. *J. Aquat. Anim. Health* **1990**, *2*, 30–37. [[CrossRef](#)]
21. Liu, Y.; Whipps, C.M.; Gu, Z.M.; Zeng, C.; Huang, M.J. *Myxobolus honghuensis* n. sp. (Myxosporidia: Bivalvulida) parasitizing the pharynx of allogynogenetic gibel carp *Carassius auratus gibelio* (Bloch) from Honghu Lake, China. *Parasitol. Res.* **2012**, *110*, 1331–1336. [[CrossRef](#)] [[PubMed](#)]
22. Zhang, J.; Wang, J.; Li, A.; Gong, X. Infection of *Myxobolus turpisrotundus* sp. n. in allogynogenetic gibel carp, *Carassius auratus gibelio* (Bloch), with revision of *Myxobolus rotundus* (sl) Nemeček reported from *C. auratus auratus* (L.). *J. Fish. Dis.* **2010**, *33*, 625–638. [[CrossRef](#)] [[PubMed](#)]
23. Yang, Y.; Xiong, J.; Zhou, Z.; Huo, F.; Miao, W.; Ran, C.; Liu, Y.; Zhang, J.; Feng, J.; Wang, M. The genome of the myxosporidian *Thelohanellus kitauei* shows adaptations to nutrient acquisition within its fish host. *Genome Biol. Evol.* **2014**, *6*, 3182–3198. [[CrossRef](#)] [[PubMed](#)]
24. Okamura, B.; Gruhl, A.; Reft, A.J. Cnidarian origins of the Myxozoa. In *Myxozoan Evolution, Ecology and Development*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 45–68.
25. Klompen, A.M.L.; Kayal, E.; Collins, A.G.; Cartwright, P. Phylogenetic and Selection Analysis of an Expanded Family of Putatively Pore-Forming Jellyfish Toxins (Cnidaria: Medusozoa). *Genome Biol. Evol.* **2021**, *13*, evab081. [[CrossRef](#)]
26. D’Ambra, I.; Lauritano, C. A review of toxins from cnidaria. *Mar. Drugs* **2020**, *18*, 507. [[CrossRef](#)] [[PubMed](#)]
27. Liao, Q.; Li, S.; Siu, S.W.I.; Yang, B.; Huang, C.; Chan, J.Y.; Morlighem, J.R.L.; Wong, C.T.T.; Radis-Baptista, G.; Lee, S.M. Novel Kunitz-like peptides discovered in the Zoanthid *Palythoa caribaeorum* through transcriptome sequencing. *J. Proteome Res.* **2018**, *17*, 891–902. [[CrossRef](#)]
28. Nicosia, A.; Bennici, C.; Biondo, G.; Costa, S.; Di Natale, M.; Masullo, T.; Monastero, C.; Ragusa, M.A.; Tagliavia, M.; Cuttitta, A. Characterization of translationally controlled tumour protein from the sea anemone *Anemonia viridis* and transcriptome wide identification of cnidarian homologues. *Genes* **2018**, *9*, 30. [[CrossRef](#)]
29. Kini, R.M.; Doley, R. Structure, function and evolution of three-finger toxins: Mini proteins with multiple targets. *Toxicon* **2010**, *56*, 855–867. [[CrossRef](#)]
30. Yahalomi, D.; Atkinson, S.D.; Neuhof, M.; Chang, E.S.; Philippe, H.; Cartwright, P.; Bartholomew, J.L.; Huchon, D. A cnidarian parasite of salmon (Myxozoa: Henneguya) lacks a mitochondrial genome. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 5358–5363. [[CrossRef](#)]
31. Guo, Q.; Atkinson, S.D.; Xiao, B.; Zhai, Y.; Bartholomew, J.L.; Gu, Z. A myxozoan genome reveals mosaic evolution in a parasitic cnidarian. *BMC Biol.* **2022**, *20*, 51. [[CrossRef](#)]
32. Fry, B.G.; Scheib, H.; van der Weerd, L.; Young, B.; McNaughtan, J.; Ramjan, S.F.; Vidal, N.; Poelmann, R.E.; Norman, J.A. Evolution of an arsenal: Structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol. Cell Proteom.* **2008**, *7*, 215–246. [[CrossRef](#)]
33. Chung, E.-H.; Lee, K.-S.; Han, J.-H.; Je, Y.-H.; Chang, J.-H.; Roh, J.-Y. Molecular cloning of two cDNAs encoding an insecticidal toxin from the spider, *Araneus ventricosus*, and construction of a recombinant baculovirus expressing a spider toxin. *Int. J. Ind. Entomol.* **2002**, *4*, 43–49.
34. Hu, H.; Bandyopadhyay, P.K.; Olivera, B.M.; Yandell, M. Characterization of the *Conus bullatus* genome and its venom-duct transcriptome. *BMC Genom.* **2011**, *12*, 60. [[CrossRef](#)] [[PubMed](#)]
35. Bayrhuber, M.; Graf, R.; Ferber, M.; Zweckstetter, M.; Imperial, J.; Garrett, J.E.; Olivera, B.M.; Terlau, H.; Becker, S. Production of recombinant Conkunitzin-S1 in *Escherichia coli*. *Protein Expr. Purif.* **2006**, *47*, 640–644. [[CrossRef](#)]
36. Pollastri, M.P.; Adade, C.M.; Carvalho, A.L.O.; Tomaz, M.A.; Costa, T.F.R.; Godinho, J.L.; Melo, P.A.; Lima, A.P.C.A.; Rodrigues, J.C.F.; Zingali, R.B.; et al. Crovirin, a snake venom cysteine-rich secretory protein (CRISP) with promising activity against trypanosomes and leishmania. *PLoS Negl. Trop. Dis.* **2014**, *8*, e3252.
37. Yamazaki, Y.; Brown, R.L.; Morita, T. Purification and cloning of toxins from elapid venoms that target cyclic nucleotide-gated ion channels. *Biochemistry* **2002**, *41*, 11331–11337. [[CrossRef](#)]
38. Brown, R.L.; Haley, T.L.; West, K.A.; Crabb, J.W. Pseudochetoxin: A peptide blocker of cyclic nucleotide-gated ion channels. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 754–759. [[CrossRef](#)] [[PubMed](#)]
39. Nevalainen, T.J.; Peuravuori, H.J.; Quinn, R.J.; Llewellyn, L.E.; Benzie, J.A.; Fenner, P.J.; Winkel, K.D. Phospholipase A2 in cnidaria. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2004**, *139*, 731–735. [[CrossRef](#)] [[PubMed](#)]
40. Mariottini, G.L.; Pane, L. Cytotoxic and cytolytic cnidarian venoms. A review on health implications and possible therapeutic applications. *Toxins* **2013**, *6*, 108–151. [[CrossRef](#)] [[PubMed](#)]
41. Talvinen, K.A.; Nevalainen, T.J. Cloning of a novel phospholipase A2 from the cnidarian *Adamsia carciniopados*. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **2002**, *132*, 571–578. [[CrossRef](#)]

42. 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. [[CrossRef](#)]
43. Fedorov, S.; Dyshlovoy, S.; Monastyrnaya, M.; Shubina, L.; Leychenko, E.; Kozlovskaya, E.; Jin, J.O.; Kwak, J.Y.; Bode, A.M.; Dong, Z.; et al. The anticancer effects of actinoporin RTX-A from the sea anemone *Heteractis crispa* (= *Radianthus macrodactylus*). *Toxicon* **2010**, *55*, 811–817. [[CrossRef](#)] [[PubMed](#)]
44. Sajevic, T.; Leonardi, A.; Krizaj, I. Haemostatically active proteins in snake venoms. *Toxicon* **2011**, *57*, 627–645. [[CrossRef](#)] [[PubMed](#)]
45. Serrano, S.M.; Maroun, R.C. Snake venom serine proteinases: Sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon* **2005**, *45*, 1115–1132. [[CrossRef](#)] [[PubMed](#)]
46. Sartim, M.A.; Sampaio, S.V. Snake venom galactoside-binding lectins: A structural and functional overview. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2015**, *21*, 35. [[CrossRef](#)] [[PubMed](#)]
47. Earl, S.T.; Robson, J.; Trabi, M.; de Jersey, J.; Masci, P.P.; Lavin, M.F. Characterisation of a mannose-binding C-type lectin from *Oxyuranus scutellatus* snake venom. *Biochimie* **2011**, *93*, 519–527. [[CrossRef](#)]
48. Lopes-Ferreira, M.; Magalhaes, G.S.; Fernandez, J.H.; Junqueira-de-Azevedo Ide, L.; Le Ho, P.; Lima, C.; Valente, R.H.; Moura-da-Silva, A.M. Structural and biological characterization of Nattectin, a new C-type lectin from the venomous fish *Thalassophryne nattereri*. *Biochimie* **2011**, *93*, 971–980. [[CrossRef](#)]
49. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)] [[PubMed](#)]
50. Kosakovsky Pond, S.L.; Poon, A.F.Y.; Velazquez, R.; Weaver, S.; Hepler, N.L.; Murrell, B.; Shank, S.D.; Magalis, B.R.; Bouvier, D.; Nekrutenko, A.; et al. HyPhy 2.5-A customizable platform for evolutionary hypothesis testing using phylogenies. *Mol. Biol. Evol.* **2020**, *37*, 295–299. [[CrossRef](#)]
51. Kosakovsky Pond, S.L.; Frost, S.D. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **2005**, *22*, 1208–1222. [[CrossRef](#)] [[PubMed](#)]
52. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond, S.L. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **2012**, *8*, e1002764. [[CrossRef](#)]
53. Smith, M.D.; Wertheim, J.O.; Weaver, S.; Murrell, B.; Scheffler, K.; Kosakovsky Pond, S.L. Less Is More: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **2015**, *32*, 1342–1353. [[CrossRef](#)]
54. Murrell, B.; Weaver, S.; Smith, M.D.; Wertheim, J.O.; Murrell, S.; Aylward, A.; Eren, K.; Pollner, T.; Martin, D.P.; Smith, D.M.; et al. Gene-wide identification of episodic selection. *Mol. Biol. Evol.* **2015**, *32*, 1365–1371. [[CrossRef](#)] [[PubMed](#)]
55. Wertheim, J.O.; Murrell, B.; Smith, M.D.; Kosakovsky Pond, S.L.; Scheffler, K. RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **2015**, *32*, 820–832. [[CrossRef](#)] [[PubMed](#)]
56. Guo, Q.; Li, D.; Zhai, Y.; Gu, Z. CCPRD: A novel analytical framework for the comprehensive proteomic reference database construction of nonmodel organisms. *ACS Omega* **2020**, *5*, 15370–15384. [[CrossRef](#)]
57. 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. Proteom.* **2016**, *135*, 62–72. [[CrossRef](#)] [[PubMed](#)]
58. Zancolli, G.; Casewell, N.R. Venom systems as models for studying the origin and regulation of evolutionary novelties. *Mol. Biol. Evol.* **2020**, *37*, 2777–2790. [[CrossRef](#)] [[PubMed](#)]
59. Gonzalez, S.; Flo, M.; Margenat, M.; Duran, R.; Gonzalez-Sapienza, G.; Grana, M.; Parkinson, J.; Maizels, R.M.; Salinas, G.; Alvarez, B.; et al. A family of diverse Kunitz inhibitors from *Echinococcus granulosus* potentially involved in host-parasite cross-talk. *PLoS ONE* **2009**, *4*, e7009. [[CrossRef](#)] [[PubMed](#)]
60. Britos, L.; Lalanne, A.I.; Castillo, E.; Cota, G.; Senorale, M.; Marin, M. *Mesocestoides corti* (syn. *vogae*, cestoda): Characterization of genes encoding cysteine-rich secreted proteins (CRISP). *Exp. Parasitol.* **2007**, *116*, 95–102. [[CrossRef](#)]
61. Price, D.R.; Bell, H.A.; Hinchliffe, G.; Fitches, E.; Weaver, R.; Gatehouse, J.A. A venom metalloproteinase from the parasitic wasp *Eulophus pennicornis* is toxic towards its host, tomato moth (*Lacanobia oleraceae*). *Insect Mol. Biol.* **2009**, *18*, 195–202. [[CrossRef](#)] [[PubMed](#)]
62. Ranasinghe, S.; McManus, D.P. Structure and function of invertebrate Kunitz serine protease inhibitors. *Dev. Comp. Immunol.* **2013**, *39*, 219–227. [[CrossRef](#)] [[PubMed](#)]
63. Huang, C.; Morlighem, J.R.; Zhou, H.; Lima, E.P.; Gomes, P.B.; Cai, J.; Lou, I.; Perez, C.D.; Lee, S.M.; Radis-Baptista, G. The transcriptome of the zoanthid *Protopalathoa variabilis* (Cnidaria, Anthozoa) predicts a basal repertoire of toxin-like and venom-auxiliary polypeptides. *Genome Biol. Evol.* **2016**, *8*, 3045–3064. [[CrossRef](#)] [[PubMed](#)]
64. Nicosia, A.; Mikov, A.; Cammarata, M.; Colombo, P.; Andreev, Y.; Kozlov, S.; Cuttitta, A. The *Anemonia viridis* venom: Coupling biochemical purification and RNA-Seq for translational research. *Mar. Drugs* **2018**, *16*, 407. [[CrossRef](#)]
65. Vogel, C.; Teichmann, S.A.; Pereira-Leal, J. The relationship between domain duplication and recombination. *J. Mol. Biol.* **2005**, *346*, 355–365. [[CrossRef](#)] [[PubMed](#)]
66. Undheim, E.A.; Jones, A.; Clauser, K.R.; Holland, J.W.; Pineda, S.S.; King, G.F.; Fry, B.G. Clawing through evolution: Toxin diversification and convergence in the ancient lineage Chilopoda (centipedes). *Mol. Biol. Evol.* **2014**, *31*, 2124–2148. [[CrossRef](#)] [[PubMed](#)]
67. 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. Proteom.* **2014**, *106*, 17–29. [[CrossRef](#)] [[PubMed](#)]



68. Jouiaei, M.; Casewell, N.R.; Yanagihara, A.A.; Nouwens, A.; Cribb, B.W.; Whitehead, D.; Jackson, T.N.; Ali, S.A.; Wagstaff, S.C.; Koludarov, I.; et al. Firing the sting: Chemically induced discharge of cnidae reveals novel proteins and peptides from box jellyfish (*Chironex fleckeri*) venom. *Toxins* **2015**, *7*, 936–950. [[CrossRef](#)] [[PubMed](#)]
69. da Silva, G.F.; Reuille, R.L.; Ming, L.-J.; Livingston, B.T. Overexpression and mechanistic characterization of blastula protease 10, a metalloprotease involved in sea urchin embryogenesis and development. *J. Biol. Chem.* **2006**, *281*, 10737–10744. [[CrossRef](#)] [[PubMed](#)]
70. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et al. The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genom. Hum. Genet.* **2009**, *10*, 483–511. [[CrossRef](#)] [[PubMed](#)]
71. Kordiš, D.; Gubenšek, F. Adaptive evolution of animal toxin multigene families. *Gene* **2000**, *261*, 43–52. [[CrossRef](#)]
72. Chang, D.; Duda, T.F., Jr. Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol.* **2012**, *29*, 2019–2029. [[CrossRef](#)] [[PubMed](#)]
73. Binford, G.J.; Bodner, M.R.; Cordes, M.H.; Baldwin, K.L.; Rynerson, M.R.; Burns, S.N.; Zobel-Thropp, P.A. Molecular evolution, functional variation, and proposed nomenclature of the gene family that includes sphingomyelinase D in sicariid spider venoms. *Mol. Biol. Evol.* **2009**, *26*, 547–566. [[CrossRef](#)] [[PubMed](#)]
74. Weinberger, H.; Moran, Y.; Gordon, D.; Turkov, M.; Kahn, R.; Gurevitz, M. Positions under positive selection-key for selectivity and potency of scorpion alpha-toxins. *Mol. Biol. Evol.* **2010**, *27*, 1025–1034. [[CrossRef](#)]
75. Fry, B.G.; Wuster, W.; Kini, R.M.; Brusica, V.; Khan, A.; Venkataraman, D.; Rooney, A.P. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. *J. Mol. Evol.* **2003**, *57*, 110–129. [[CrossRef](#)] [[PubMed](#)]
76. Sunagar, K.; Moran, Y. The rise and fall of an evolutionary innovation: Contrasting strategies of venom evolution in ancient and young animals. *PLoS Genet.* **2015**, *11*, e1005596. [[CrossRef](#)] [[PubMed](#)]
77. Surm, J.M.; Smith, H.L.; Madio, B.; Undheim, E.A.B.; King, G.F.; Hamilton, B.R.; van der 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. [[CrossRef](#)] [[PubMed](#)]
78. Naldoni, J.; Zatti, S.A.; da Silva, M.R.M.; Maia, A.A.M.; Adriano, E.A. Morphological, ultrastructural, and phylogenetic analysis of two novel *Myxobolus* species (Cnidaria: Myxosporaea) parasitizing bryconid fish from Sao Francisco River, Brazil. *Parasitol. Int.* **2019**, *71*, 27–36. [[CrossRef](#)]
79. Lom, J.; Arthur, J. A guideline for the preparation of species descriptions in Myxosporaea. *J. Fish. Dis.* **1989**, *12*, 151–156. [[CrossRef](#)]
80. National Research Council. *Guide for the Care and Use of Laboratory Animals*; National Academies Press: Washington, DC, USA, 2010.
81. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
82. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; A Thompson, D.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
83. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinform.* **2006**, *22*, 1658–1659. [[CrossRef](#)] [[PubMed](#)]
84. Pertea, G.; Huang, X.; Liang, F.; Antonescu, V.; Sultana, R.; Karamycheva, S.; Lee, Y.; White, J.; Cheung, F.; Parvizi, B.; et al. TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinform.* **2003**, *19*, 651–652. [[CrossRef](#)]
85. Manni, M.; Berkeley, M.R.; Seppey, M.; Simao, F.A.; Zdobnov, E.M. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647–4654. [[CrossRef](#)] [[PubMed](#)]
86. Parra, G.; Bradnam, K.; Korf, I. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **2007**, *23*, 1061–1067. [[CrossRef](#)] [[PubMed](#)]
87. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
88. Laetsch, D.R.; Blaxter, M.L. BlobTools: Interrogation of genome assemblies. *F1000Research* **2017**, *6*, 1287. [[CrossRef](#)]
89. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)]
90. Tang, S.; Lomsadze, A.; Borodovsky, M. Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res.* **2015**, *43*, e78. [[CrossRef](#)] [[PubMed](#)]
91. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **2000**, *16*, 276–277. [[CrossRef](#)]
92. 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. [[CrossRef](#)]
93. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* **2011**, *39* (Suppl. 2), W29–W37. [[CrossRef](#)] [[PubMed](#)]

94. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
95. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
96. Castresana, J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Mol. Biol. Evol.* **2000**, *17*, 540–552. [[CrossRef](#)]
97. Minh, B.Q.; A Schmidt, H.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)] [[PubMed](#)]
98. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)] [[PubMed](#)]
99. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K.; Battistuzzi, F.U. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
100. Suyama, M.; Torrents, D.; Bork, P. PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **2006**, *34*, W609–W612. [[CrossRef](#)] [[PubMed](#)]
101. Weaver, S.; Shank, S.D.; Spielman, S.; Li, M.; Muse, S.V.; Pond, S.L.K. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [[CrossRef](#)] [[PubMed](#)]