Article

Multiomic Approach for Bioprospection: Investigation of Toxins and Peptides of Brazilian Sea Anemone *Bunodosoma caissarum*

Maria Eduarda Mazzi Esquinca [1], Claudia Neves Correa [1,2], Gabriel Marques de Barros [1,2], Horácio Montenegro [3] and Leandro Mantovani de Castro [1,2,*]

1 Department of Biological and Environmental Sciences, Bioscience Institute, Sao Paulo State University (UNESP), Sao Vicente 11330-900, SP, Brazil
2 Biodiversity of Coastal Environments Postgraduate Program, Department of Biological and Environmental Sciences, Bioscience Institute, Sao Paulo State University (UNESP), Sao Vicente 11330-900, SP, Brazil
3 NGS Soluções Genômicas, Piracicaba 13416-030, SP, Brazil
* Correspondence: leandro.mantovani@unesp.br

Abstract: Sea anemones are sessile invertebrates of the phylum Cnidaria and their survival and evolutive success are highly related to the ability to produce and quickly inoculate venom, with the presence of potent toxins. In this study, a multi-omics approach was applied to characterize the protein composition of the tentacles and mucus of *Bunodosoma caissarum*, a species of sea anemone from the Brazilian coast. The tentacles transcriptome resulted in 23,444 annotated genes, of which 1% showed similarity with toxins or proteins related to toxin activity. In the proteome analysis, 430 polypeptides were consistently identified: 316 of them were more abundant in the tentacles while 114 were enriched in the mucus. Tentacle proteins were mostly enzymes, followed by DNA- and RNA-associated proteins, while in the mucus most proteins were toxins. In addition, peptidomics allowed the identification of large and small fragments of mature toxins, neuropeptides, and intracellular peptides. In conclusion, integrated omics identified previously unknown or uncharacterized genes in addition to 23 toxin-like proteins of therapeutic potential, improving the understanding of tentacle and mucus composition of sea anemones.

Keywords: sea anemone; omics; toxins; intracellular peptides; neuropeptides

1. Introduction

Sea anemones are sessile invertebrates of the phylum Cnidaria that have inhabited the earth for over 700 million years. Their survival and evolutive success were highly related to the ability to produce and quickly inoculate venom. Cnidarians have a specialized cell responsible for venom production: the cnidocyte. This cell carries an organelle called cnida that penetrates the target to release the venom as soon as a movement is detected. Venom production is a powerful mechanism of defense, intra- and interspecific competition, and predative behavior [1,2]. A sea anemone's venom is enriched with biomolecules, such as polypeptides, proteins, polyamines, and salts, [3], and it represents an important weapon directly associated with their ecological success. Some of these biomolecules are potent toxins able to interact with sodium and potassium ion channels [4,5] by blocking or reducing their sensibility, leading to diverse effects on the prey, including the inability to move fast.

Biomolecules produced by marine organisms have been studied for several decades, and important drugs have been obtained to treat a variety of human diseases [6–10]. For example, Toxin ShK, a sequence of 34 amino acids, is a potent potassium channel inhibitor isolated from *Stichodactyla helianthus* [11]. Dalazatide, a compound derived from this toxin, is undergoing a clinical trial phase 1b due to its ability to inactivate Kv1.3 channels in human T lymphocytes, whose expression is upregulated in autoimmune diseases such as

psoriasis [12,13]. Another sea anemone toxin, BgK, acts on channels Kv1 and 3 in addition to Shaker-related Kv1 channels [14] and belongs to the same family of ShK.

Bunodosoma caissarum (*B. caissarum*) is a Brazilian sea anemone species that populates the intertidal zone of the southern coast. In the 1990s, *B. caissarum*'s extract was thoroughly studied by Malpezzi and Freitas research group, which published several findings related to its antimitotic, antitumor, cytotoxic, hemolytic, and neurotoxic activities [15–18], and described the first three toxins of this species [19]. However, access to different databases shows an absence of genome annotation for the genus of *B. caissarum*, and the number of deposited genomes for the phylum Cnidaria is still relatively low. This information is scarce when compared with the number of species estimated in the phylum—above 13,000 [20,21].

Recently, the advances in next-generation sequencing techniques, improvements in mass spectrometry instrumentation for protein sequencing, and also bioinformatic tools led to an increase in the number of large-scale molecular studies in anthozoans, using transcriptomic (RNA-seq) and/or proteomic approaches. *Exaiptasia diaphana, Nematostella vectensis, Edwardsiella lineata, Anthopleura elegantissima, Stichodactyla helianthus*, and *Stichodactyla haddoni* are some sea anemone species that have already been investigated [21–27]. These omics have provided a detailed overview of biomolecules, revealing a high complexity of the venom compared with results obtained in previous studies.

The understanding of the marine environment and its rich biodiversity still faces many obstacles for molecular, evolutionary, morphological, or even conservation studies. The lack of molecular information on marine fauna hinders the construction of phylogenies, the search for alternative drugs, and the analysis of the different adaptive mechanisms of these species. This study applied integrated omics approaches, including transcriptomics, proteomics, and peptidomics, to characterize the tentacles and mucus composition of *B. caissarum*.

2. Results

2.1. Tentacle Transcriptome

Sequencing data were deposited at the National Center for Biotechnology Information (Bioproject: PRJNA939171; Biosample: SAMN33456513). Transcriptome assembly yielded 186,978 transcripts from 166,364,140 paired reads. Protein prediction using TransDecoder resulted in 150,493 peptides from 75,867 complete proteins. Other metrics are represented in Table 1. Most of the genes annotated through BLASTx and BLASTp searches showed a high percentage of identity with the recently deposited draft genome of the sea anemone *Actinia tenebrosa* (*A. tenebrosa*) [28].

Table 1. Trinity and TransDecoder metrics.

	Trinity	**TransDecoder**
Transcripts	186,978	150,493
N50	1900	501
Genes	111,386	61,465

Sequences obtained by Trinity were searched against the Metazoa protein database. Annotation performed by Trinotate resulted in 23,444 annotated genes, categorized according to Gene Ontology terms. A total of 19,561 genes were assigned biological process classes, 19,933 were assigned cellular component classes, and 21,195 were assigned molecular function classes (Figure 1). In biological processes, the subcategories with the most genes were cellular and metabolic processes and biological regulation. In molecular function, genes related to binding, catalytic activity, and molecular transduction were more frequently found. Genes classified as subcategories of cellular components presented a homogeneous distribution. From this annotation, 1% of genes were identified as toxins (Figure 1). Initially, along with transcriptome sequences, 48 transcripts were identified with sequences similar

to toxins produced by Cnidarian species, corresponding to 11 families according to the UniProtKB database, with 3 of them previously described from *B. caissarum.*
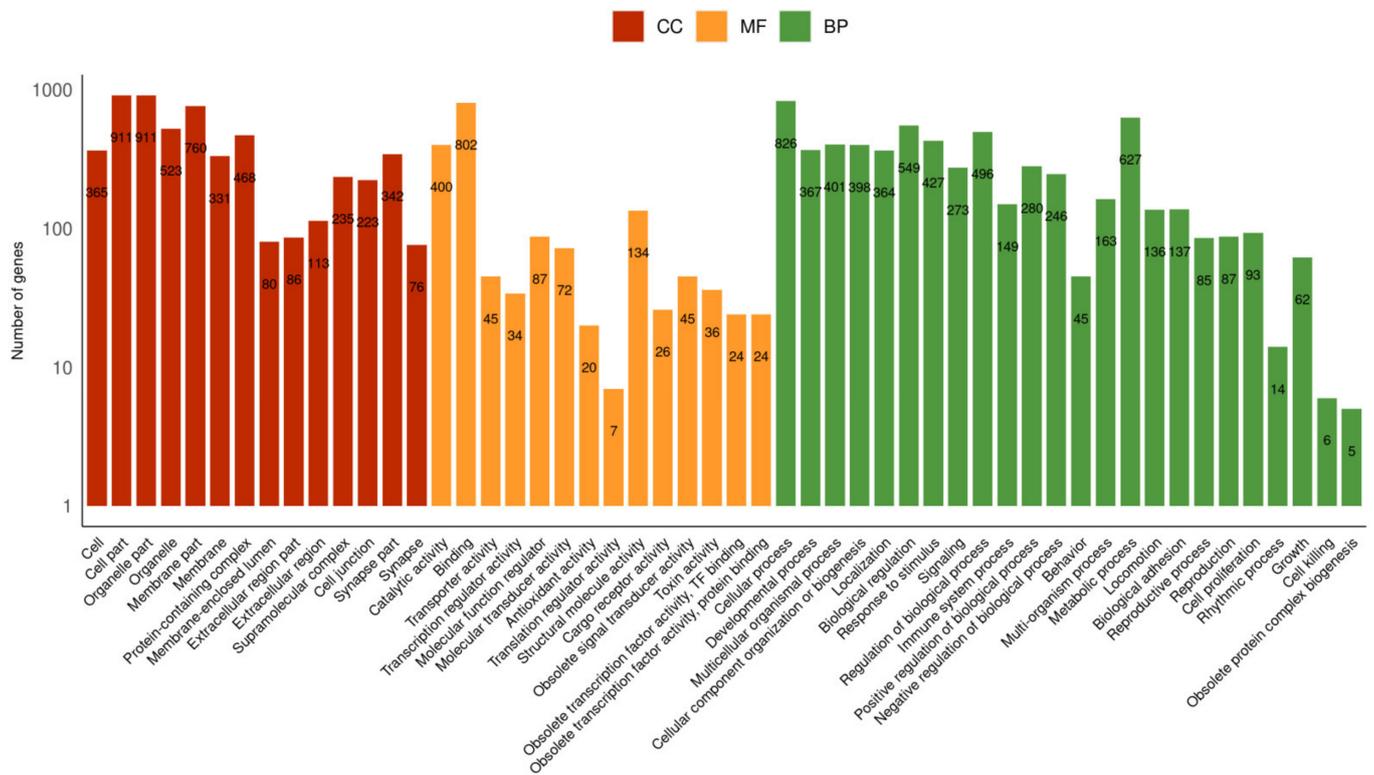

Figure 1. Number of annotated genes in subcategories according to Gene Ontology terms generated by Web Gene Ontology Annotation Plot (WEGO). The subcategories can be classified as cellular component (CC), molecular function (MF), and biological process (BP).

In addition, Trinity sequences obtained after the Trinotate annotation were searched against the Animal Toxin Annotation Project (Toxprot, accessed on 1 December 2021) using BLASTx local alignment tool with an e-value of 1E-04 Top BLAST hits were selected for each gene, and toxins were classified according to their families or main molecular function available on UniProtKB. A total of 2055 transcripts presented similarity with Toxprot sequences, distributed in 74 classes of toxins; nineteen of these classes with the highest number of transcripts are shown in Figure 2A. The category named 'Others' represents the toxin families with fewer transcripts found, such as actinoporins, sea anemone toxin 8, venom proteinase, and cysteine-type endopeptidase inhibitor (Figure 2A and Table S1). Of these 2055 genes, only 130 transcripts showed similarity with cnidaria toxins and venom-related proteins. These transcripts represent 50 different proteins present in the UniProtKB, 95% of which were from the following sea anemones: *Nematostella vectensis*, *A. tenebrosa*, *Stichodactyla gigantea*, and *Bunodosoma caissarum*. Sea anemone toxins were grouped into three main categories: neurotoxins, mixed-function enzymes, and membrane-active. Neurotoxins are molecules able to interact with ion channels, impacting the transmission of impulses, and they include families such as potassium channel toxins, sodium channel inhibitors, and small cysteine-rich proteins from the phylum Cnidaria [29,30]. Mixed-function enzymes display important cellular roles, not necessarily related to a toxic effect. Phospholipases A2 hydrolyze phospholipids, thus directly affecting membrane configuration. These enzymes are necessary for the cell because they may act as toxins [31]. Membrane-active toxins also interact with cellular membranes, for example, porins induce cell lysis by pore formation [32,33]. Transcripts related to two types of porins were identified: actinoporins and jellyfish toxins. The proteins with more genes associated were Nematocist express protein 6 (30 genes), toxin Bcs III 15.09 (11 genes), and *Urticina crassicornis'* phospholipase

A2 (6 genes). The unknown category gathers families whose functions are not clear or well described, such as acrohargins, sea anemone 8 toxin, EGF domain peptide, and toxins that were not characterized as any family member (Figure 2B).
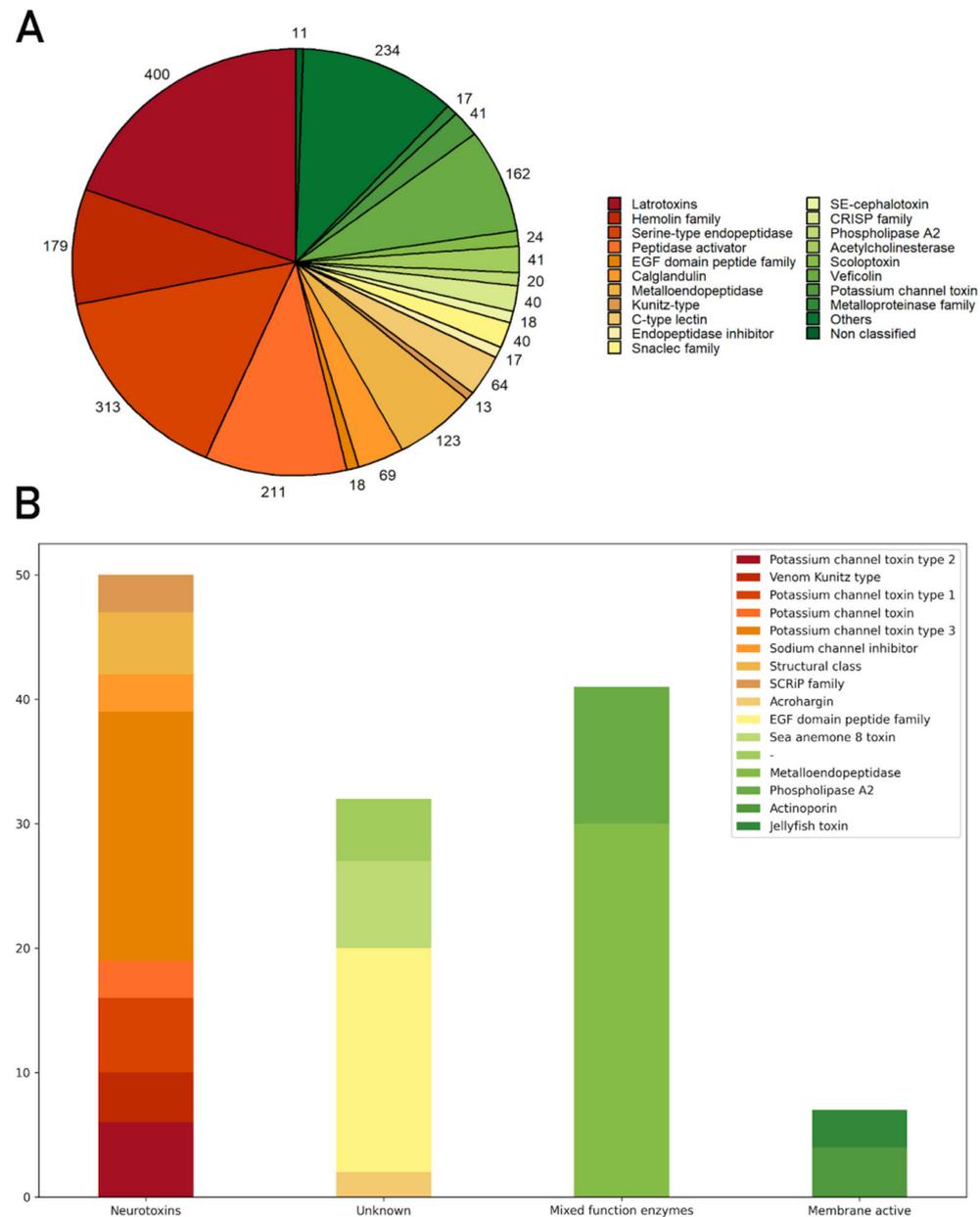

Figure 2. Resultant classes of toxins found in the transcriptome sequences related to the Toxprot database (**A**). Cnidaria toxin genes identified in BLASTp search of transcripts against the Toxprot database (https://www.uniprot.org/help/Toxins (accessed on 1 December 2021)) (**B**). In the graph, transcripts with similarity to toxins were grouped into four categories: neurotoxins, mixed-function enzymes, membrane-active, and unknown.

2.2. Mucus and Tentacle Proteome

Proteomic analysis was performed on tentacles and mucus protein extracts, with a total of eight independent runs in the mass spectrometer, with six samples of the fraction above 10 KDa (three from tentacles and three from mucus), and two samples of the fraction below 10 KDa (two mixtures from all tentacles and all mucus samples) as shown in Figure S1.

A total of 1580 proteins were found in tentacles or mucus samples considering an FDR < 1% and using a database with protein sequences generated by Transdecoder from

the transcriptome experiment data of *B. caissarum* tentacles (Table S2). Of these proteins, 836 were found only in tentacles, 210 only in mucus, and 534 were common in both (Figure 3A). Moreover, 66.2% of them are known and approximately one-third (34.8%) are unknown or uncharacterized (Figure 3A). Known proteins were classified by family, class, or main molecular function according to UniProtKB and then grouped into six main categories: structural, enzymes, DNA- and RNA-associated proteins, transport and signal, protein metabolism, and toxin. Some proteins were associated with more than one category. Nonrelated proteins with these functions were grouped into others. Enzymes were the most abundant category, with 377 proteins classified as follows: transferases (45), hydrolases (44), kinases (30), metalloendopeptidases (28), and peptidases (25) (Figure 3A).

To quantitatively assess differences in protein profile composition between mucus and tentacle, the proteins found in at least two of three mucus or tentacle samples (a total of 430) were selected and compared (Figures S2 and S3). In the tentacle, 316 proteins found were present exclusively or in greater quantities than in the mucus, with almost 64% of them matching a known protein (Figure 3B). In the mucus, 114 proteins were present exclusively or in greater quantities than in the tentacle, with almost 55% of them known (Figure 3C). The known proteins were classified considering their main biological function or family according to the UniProtKB database. Known tentacle proteins were mostly enzymes (30.7%), followed by DNA- and RNA-related, while in mucus, known proteins were mostly toxins (29%) (Figure 3B,C). This quantitative analysis allows us to compare 24 transcripts with similarities to toxins, as shown in Figure 4.

2.3. Mucus and Tentacle Peptidome

Non-Reduced and Non-Alkylated Samples

A total of 681 peptides from tentacle and mucus non-reduced and non-alkylated samples were identified and quantified through the Transdecoder database; they were related to 84 transcripts and 72 described proteins (Tables S3 and S4). Peptides mass distribution showed a higher abundance of fragments between 1100 and 2100 Da (Figure 5A). Most peptides found were fragments from proteins, such as actin, tubulin, myosin, histones, microtubule-associated, and hemicentin. Quantitatively, most peptide fragments were more concentrated in the tentacle fraction than in the mucus. However, it is possible to notice that for the same precursor protein, some peptides were more present in one fraction than in the other (Figure S4). Furthermore, a transcript called TRINITY_DN1239, with a higher abundance of peptides (251 sequences, with a range of 8 at 40 amino acid residues) was observed in all peptidome and proteome samples. This transcript does not match any protein of the SWISS-PROT database. Meanwhile, in the BLASTp and BLASTx, this transcript is identified as antigen-like. Another transcript with no match to any annotated gene or protein, called TRINITY_DN942, was the second precursor with more peptides, totalizing 59 fragments. In addition, peptides related to transcripts with ShK and factor 5/8 C domains, which are common among toxins, were observed, but most of the sequences identified corresponded to the initial C-terminal region (Figure S4).

2.4. Peptidome with Reduced and Alkylated Cysteine Residues

In the initial peptidomic experiments from tentacle and mucus, a low identification of cysteine residues was observed, showing a technical limitation in the mass spectrometry fragmentation of peptides with the presence of disulfide bridges between cysteine residues. In order to solve this problem, two additional samples were prepared (one containing a mixture of tentacle samples and the other with mucus samples), with the following additional steps: cysteine reduction using dithiothreitol (DTT), alkylation with iodoacetamide (IAA) and without trypsin treatment. This procedure causes a modification in cysteine residues, preventing the formation of disulfide bonds and consequently facilitating their detection in mass spectrometry. Unlike proteomics, trypsin was not used here to preserve naturally-generated peptides.
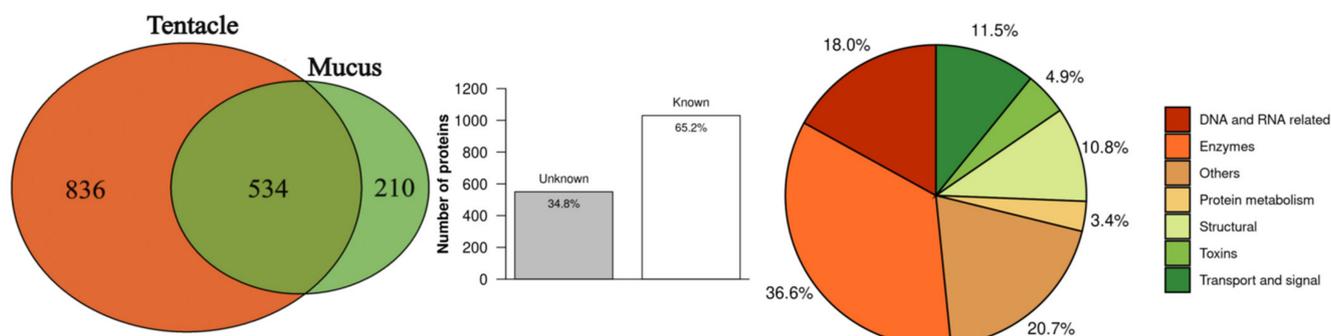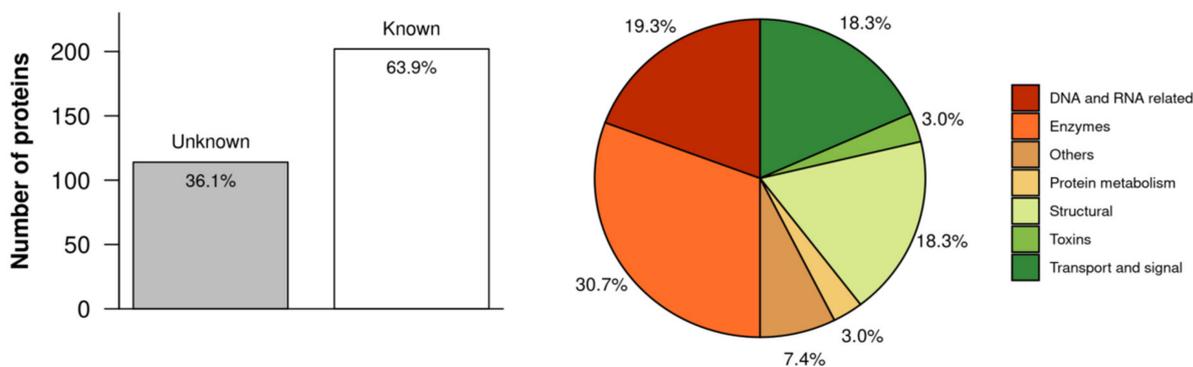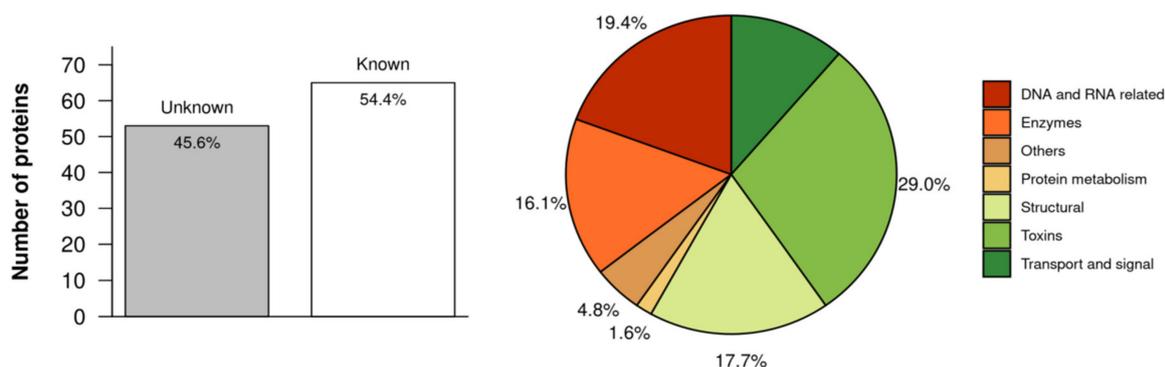
Figure 3. General profile of tentacle and mucus proteomics from *B. caissarum*. In (**A**) Venn diagram of proteins found in tentacle and mucus samples, percentage of known and unknown proteins, and protein classification into the following categories: structural, enzymes, DNA- and RNA-associated proteins, transport and signal, protein metabolism, and others. In (**B**) profile of enriched proteins in the tentacle. In (**C**) profile of enriched proteins in the mucus.

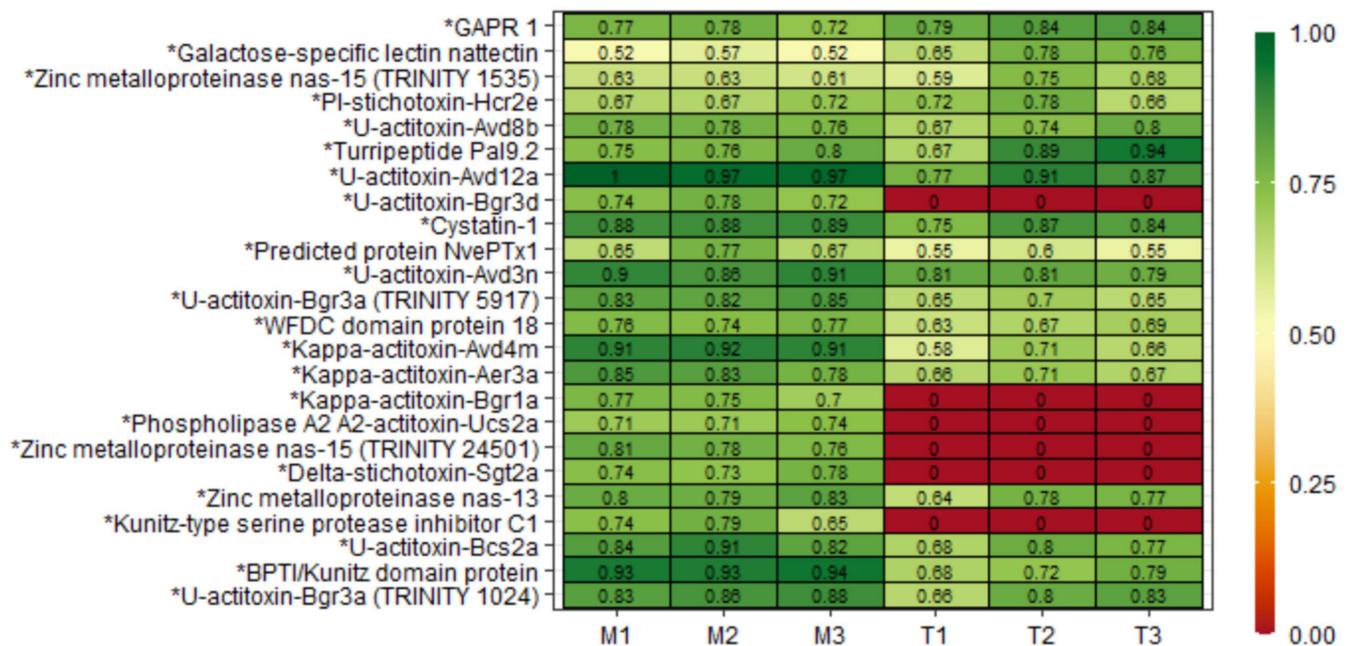|	M1	M2	M3	T1	T2	T3
*GAPR 1	0.77	0.78	0.72	0.79	0.84	0.84
*Galactose-specific lectin nattectin	0.52	0.57	0.52	0.65	0.78	0.76
*Zinc metalloproteinase nas-15 (TRINITY 1535)	0.63	0.63	0.61	0.59	0.75	0.68
*Pl-stichotoxin-Hcr2e	0.67	0.67	0.72	0.72	0.78	0.66
*U-actitoxin-Avd8b	0.78	0.78	0.76	0.67	0.74	0.8
*Turripeptide Pal9.2	0.75	0.76	0.8	0.67	0.89	0.94
*U-actitoxin-Avd12a	1	0.97	0.97	0.77	0.91	0.87
*U-actitoxin-Bgr3d	0.74	0.78	0.72	0	0	0
*Cystatin-1	0.88	0.88	0.89	0.75	0.87	0.84
*Predicted protein NvePTx1	0.65	0.77	0.67	0.55	0.6	0.55
*U-actitoxin-Avd3n	0.9	0.86	0.91	0.81	0.81	0.79
*U-actitoxin-Bgr3a (TRINITY 5917)	0.83	0.82	0.85	0.65	0.7	0.65
*WFDC domain protein 18	0.76	0.74	0.77	0.63	0.67	0.69
*Kappa-actitoxin-Avd4m	0.91	0.92	0.91	0.58	0.71	0.66
*Kappa-actitoxin-Aer3a	0.85	0.83	0.78	0.66	0.71	0.67
*Kappa-actitoxin-Bgr1a	0.77	0.75	0.7	0	0	0
*Phospholipase A2 A2-actitoxin-Ucs2a	0.71	0.71	0.74	0	0	0
*Zinc metalloproteinase nas-15 (TRINITY 24501)	0.81	0.78	0.76	0	0	0
*Delta-stichotoxin-Sgt2a	0.74	0.73	0.78	0	0	0
*Zinc metalloproteinase nas-13	0.8	0.79	0.83	0.64	0.78	0.77
*Kunitz-type serine protease inhibitor C1	0.74	0.79	0.65	0	0	0
*U-actitoxin-Bcs2a	0.84	0.91	0.82	0.68	0.8	0.77
*BPTI/Kunitz domain protein	0.93	0.93	0.94	0.68	0.72	0.79
*U-actitoxin-Bgr3a (TRINITY 1024)	0.83	0.86	0.88	0.66	0.8	0.83

Figure 4. Enriched toxins were identified in tentacle and mucus samples from *B. caissarum*. Selected proteins were found in at least two out of three samples. Zinc metalloproteinase nas-15 and U-actitoxin-Bgr3a presented similarities with more than one transcript. * Putative toxins identified through similarity with other described toxins in transcriptome and confirmed in proteome.

After this procedure, 950 different peptides were identified: 459 were exclusive to the tentacle pool and 449 were exclusive to the mucus pool (Tables S5 and S6), with 51 common peptides between the samples (Figure S5). Furthermore, the tentacle contained 105 peptides with cysteine residues, while the mucus contained 108 of them. Most of these sequences corresponded to toxin-like sequence transcripts. Regarding the size of the detected peptides, there was an increase in the number of peptides in almost all mass ranges, with a range between 500 and 4300 Da (Figure 5A). Tables 2 and 3 show the peptide sequences that were detected with greater intensity. In mucus samples, three of the highest identified peptides matched *B. caissarum*'s toxin, U-AITX-Bcs2a, which was not observed in high expression in the tentacle. However, the tentacle showed the presence of another toxin, U-AITX-Aeq6a, also known as Acrorhagin II, which has an LD50 of 80 μg/kg in crabs [34].

2.5. Analysis of Cleavage Site in the Peptidome

In order to understand aspects related to proteolytic processing, the amino acid sequence of the precursor proteins and peptides generated were analyzed. Amino acids which preceded (upstream) and succeeded (downstream) peptide sequences, corresponding to the cleavage site (P1) were quantified in relation to the total number of peptides found in the procedures performed. In non-reduced and non-alkylated peptidome samples, the most frequent amino acid residue at the upstream position adjacent to the peptide sequences generated was tyrosine, followed by glutamic acid, phenylalanine, and methionine, showing a homogeneous distribution among these amino acids. Meanwhile, at the downstream position of cleavage the predominant amino acid residue present was serine, followed by leucine, and alanine. In reduced and alkylated peptidome samples, the most frequent amino acid in the upstream position was tyrosine, followed by methionine, arginine, phenylalanine, and glycine, while in the downstream position, they were serine, alanine, and glycine. Furthermore, many fragments in the alkylated and reduced samples were from the C-terminal region of the precursor proteins showing only the cleavage site corresponding to the upstream region adjacent to the N-terminal of the generated peptides (Figure 5B,C).
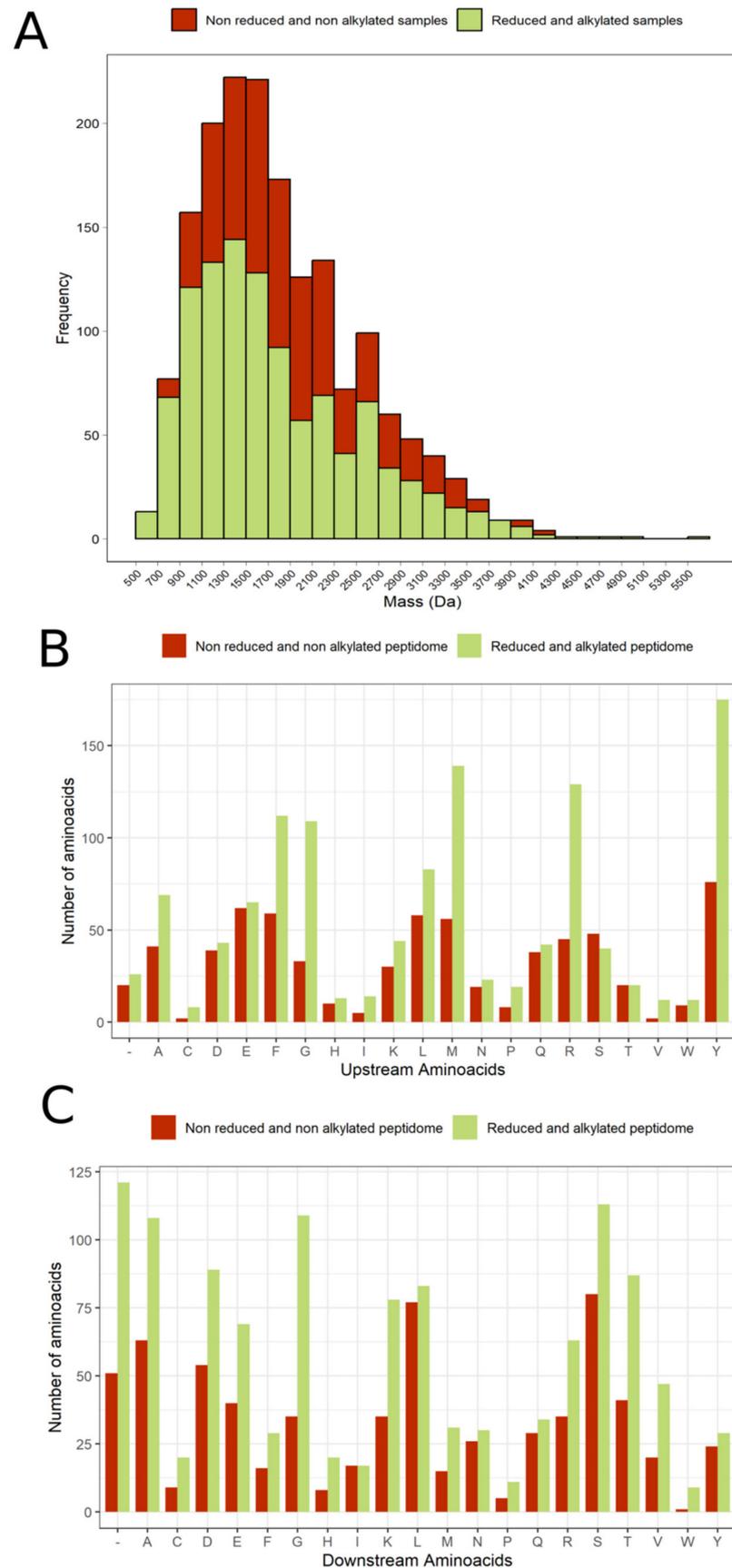
Figure 5. (**A**) Distribution of peptides obtained from non-reduced and non-alkylated samples (red) and reduced and alkylated samples (green). (**B**) Upstream amino acids. (**C**) Downstream amino acids.

Table 2. Most detected peptides of mucus samples in reduced and alkylated peptidome.

TRINITY	Name	Peptide	Mucus Sample Area	Tentacle Sample Area
TRINITY_DN1213	-	DCRGKHCQTGPFGD	9.06E+09	4.84E+08
TRINITY_DN58641	-	TLASSIQCVGKCKIK TSSGQCRTDLRCMLANKGAS	1.70E+09	3.88E+08
TRINITY_DN14686	U-actitoxin-Bcs2a	GLPCDCHGHTGTYWLNY YSKCPKGYGYTGRCRYLVG SCCYK	1.32E+09	-
TRINITY_DN3092	-	MATSCRKCKPGYGCWAV PCPKR	1.09E+09	-
TRINITY_DN557	Lamin-A	EELSFKRSVYDKE	5.92E+08	-
TRINITY_DN1213	-	DCRGKHCQTGPF	3.56E+08	1.38E+06
TRINITY_DN14686	U-actitoxin-Bcs2a	GLPCDCHGHTGTY	2.59E+08	-
TRINITY_DN3092	-	MATSCRKCKPGYGCWAV PCPKR	2.16E+08	-
TRINITY_DN3092	-	ATSCRKCKPGYGCWAVPCPKR	1.71E+08	-
TRINITY_DN14686	U-actitoxin-Bcs2a	WLNYYSKCPKGYGYT GRCRYLVGSCCYK	1.68E+08	-

Table 3. Most detected peptides of tentacle samples in reduced and alkylated peptidome.

TRINITY	Name	Peptide	Mucus Sample Area	Tentacle Sample Area
TRINITY_DN1213	-	DCRGKHCQTGPFGD	9.06E+09	4.84E+08
TRINITY_DN282	Histone H2B	LPGELAKHAVSEGTKAVTKYTSSK	-	3.99E+08
TRINITY_DN58641	-	TLASSIQCVGKCKIK TSSGQCRTDLRCMLANKGAS	1.7E+09	3.88E+08
TRINITY_DN3037	-	NPPYEEILEPAFFHIR	-	2.61E+08
TRINITY_DN10521	-	APPDTSILDKLGKL	7.26E+06	2.13E+08
TRINITY_DN12562	-	DEAGLLKYKTAA GAALVNERLKNLAERY	-	1.81E+08
TRINITY_DN3037	-	QPPFLGGPAYFHIR	-	1.52E+08
TRINITY_DN282	Histone H2B	LLLPGELAKHAVSEGT KAVTKYTSSK	8.66E+06	1.52E+08
TRINITY_DN25490	U-actitoxin-Aeq6a	KERCDLLGDPCVKG	-	1.51E+08
TRINITY_DN1828	-	AAAYVCDVARNLDCSAH	-	1.43E+08

3. Discussion

One of the most significant findings of this study was an increase in the repertoire of transcripts, proteins, and peptides identified in the tentacles and mucus of the sea anemone *B. caissarum*, highlighting the identification of new sequences similar to toxins, neuropeptides, and also peptide fragments of resident proteins in the cytosol, nucleus, and mitochondria, through the data integration obtained from a multi-omic approach. De novo transcriptomics allows the identification of low-expressed genes without an available genome, while proteomics and peptidomics confirm the transcriptome data in addition to providing information about post-translational modifications, such as N-terminus acetylation. Moreover, as shown in Figure 6, the integrated analysis of the data set from the omic techniques used here demonstrates an enrichment in the identification of molecules because

some of the transcripts were detected by proteomics, and others by peptidomics. These approaches together generate robust data on a poorly explored Brazilian sea anemone species, in which only nine toxin sequences have been characterized [6,19,35–37].
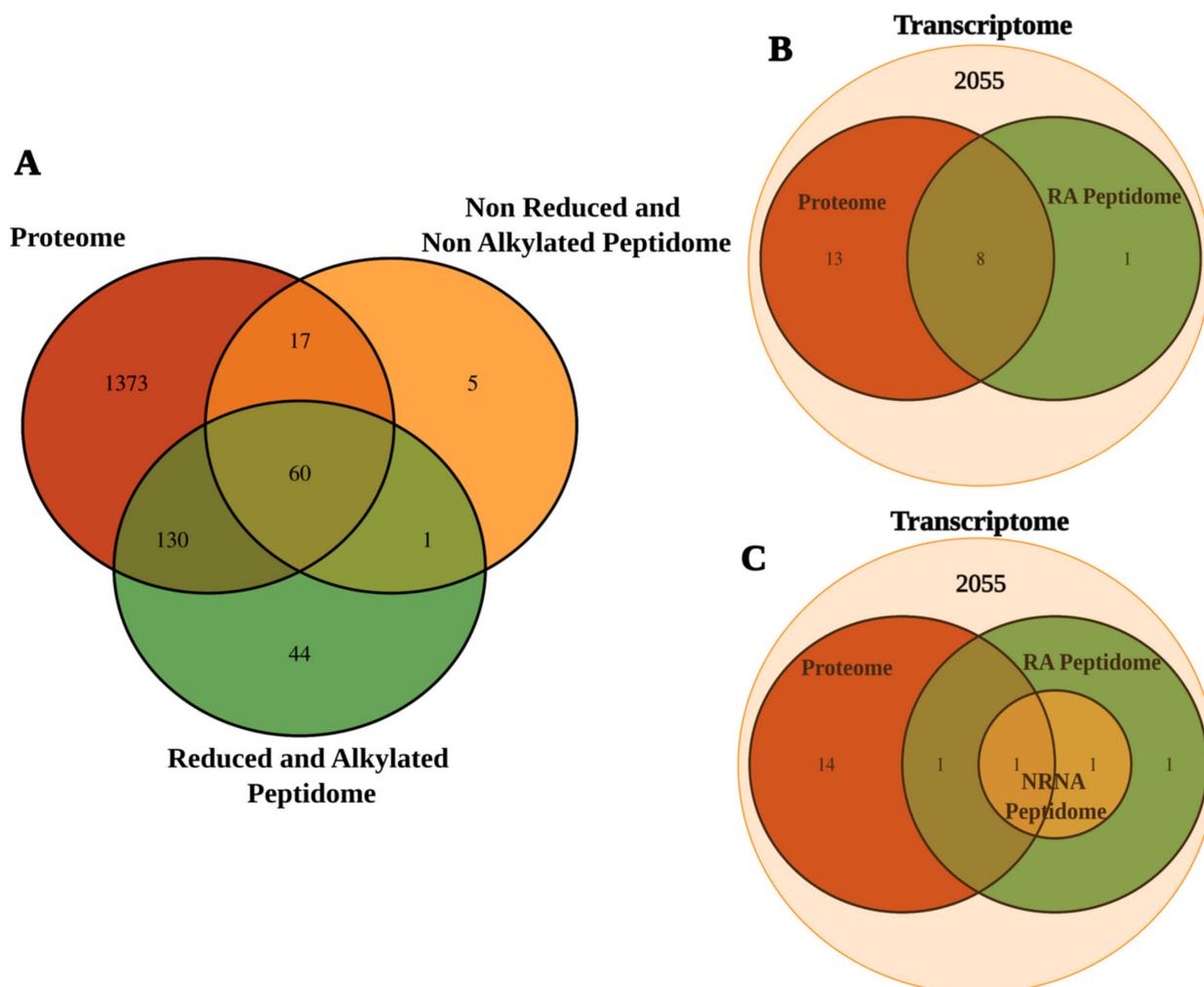

Figure 6. Integrated omics. Comparative analysis of transcripts confirmed by proteomics and peptidomics (**A**). Number of transcripts to the *Bunodosoma caissarum's* new toxins identified through similarity with other described toxins in transcriptome and confirmed in proteome peptidome (**B**). Number of transcripts to the putative toxins identified through similarity with other described toxins in transcriptome and confirmed in proteome or peptidome (**C**).

3.1. Considerations about the Mucus Collection Protocol and Omic Techniques

An important point to be mentioned here is the use of a protocol to release the content present in cnidae and its characterization through proteomics and peptidomics. Our analysis showed a marked enrichment of toxins, specific enzymes, and a few "housekeeping" proteins in the mucus samples in relation to the tentacles, allowing differentiation of proteins that could be related to components of the venom and those that are not. A similar procedure was performed in the sea anemone *Anthopleura dowii*, and enrichment with toxins and venom-related proteins was observed in the mucus. However, toxins were absent in tentacle proteomics, probably due to the discharging of cnidae after stress manipulation [38].

In the transcriptome data, many transcripts from the tentacle of *B. caissarum* showed a high percentage of identity with transcripts of closely related species, such as *A. tenebrosa* [28]. Nevertheless, few of these transcripts presented 100% identity. Moreover, when transcripts were related to toxins, the percentage of identity with sequences from other

species was decreased, showing differences between them. These data reinforce the need to use specific databases for further proteomics analyses to search for new polypeptide sequences with biological activity.

Additionally, for the identification of toxins or proteins related to the venom, the data obtained from the transcriptome were compared with the sequences deposited in the UniProt Animal Toxin Annotation Project (https://www.uniprot.org/help/Toxins (accessed on 1 December 2021)). The search was performed in all the databases and was restricted to only described toxins in the phylum Cnidaria. There were discrepancies among the results of Toxprot analysis as shown in Figure 2. When the search was completed in Toxprot, 2055 transcripts were identified as related to toxins, and the most abundant class was latrotoxins (19%), followed by serine-type endopeptidase (15%) and others (11%) (Figure 2A). When the analyses were carried out on the phylum Cnidaria, only 130 transcripts showed similarity with 50 toxin sequences present in the database. Thus, the search carried out in the Toxprot database must be analyzed carefully. Some of the toxin sequences may contain common motifs found in other proteins, as observed in latrotoxins. This class of toxins, from spiders of the genus *Latrodectus*, is involved in neurotransmitter liberation and has ankyrin repeats, which are present in many other intracellular proteins [39,40]. Madio et al.'s (2017) analysis of the venom of the sea anemone *Stichodactyla haddoni* found similar results, pointing to 52 families of toxins as false positives after analysis of proteomics data, including latrotoxins, which were classified, in fact, as proteins with ankyrin repeats [27]. In our study, approximately half of cnidarian transcripts identified in the transcriptome were confirmed in proteomics and peptidomics analysis. These data demonstrate that transcriptome analysis used alone can lead to false positive results.

Proteome showed a considerable proportion of unknown or uncharacterized proteins (34%), which were highly identifiable in peptidome and proteome. The mostly identified was TRINITY_DN1239, characterized as major antigen-like due to its similarity with an *A. tenebrosa's* protein. Considering protein roles, enzymatic activity was the main category among the 1580 proteins identified in all samples, and 316 highly expressed in the tentacle, while in mucus, toxin-like were more common among its 114 highest expressed proteins. According to Toxprot alignment with those proteins, the tentacle had approximately 1.9% of its proteins related to toxins, and mucus had 14%. The difference between these proportions corroborates mucus's ecological importance for sea anemones' survival and evolution [41,42].

Another aspect and the main differential applied here, compared with other integrated omic studies, was the inclusion of peptidome. Unlike proteome, where a set of peptides is generated in vitro through digestion with a specific enzyme, such as trypsin, for broad identification of protein sequences, the peptidome reflects the naturally generated peptides, adding information referent to proteolytic processing. Peptidome is restricted to molecules with less than 10 kDa, and most studies in different types of samples usually do not include the reduction and alkylation of cysteine residues in the protocol. However, cysteine residues are present in many described toxin sequences, so two protocols were performed for the peptidome to increase the number of sequenced fragments: with and without reduction and alkylation, both without trypsin. In this case, peptidomics allowed the identification of the larger and smaller fragments of the mature toxin. To exemplify, peptide fragments identified in proteomics and peptidomics for one toxin-like protein are shown in Figure 7. In this case, peptidomics allowed the identification of the larger fragment of the mature toxin, as well as smaller fragments.

3.2. Tentacle and Mucus Composition

The transcriptome carried out in this study aimed at obtaining a specific database, which could be used as a reference database for sequencing proteins and peptides, correlating with proteomics and peptidomics data obtained by mass spectrometry. In this context, and taking into account the diversity of genes and peptides found, the results will be

discussed below in the following subsections: toxins, cysteine-rich proteins, neuropeptides, and intracellular peptides (InPeps).
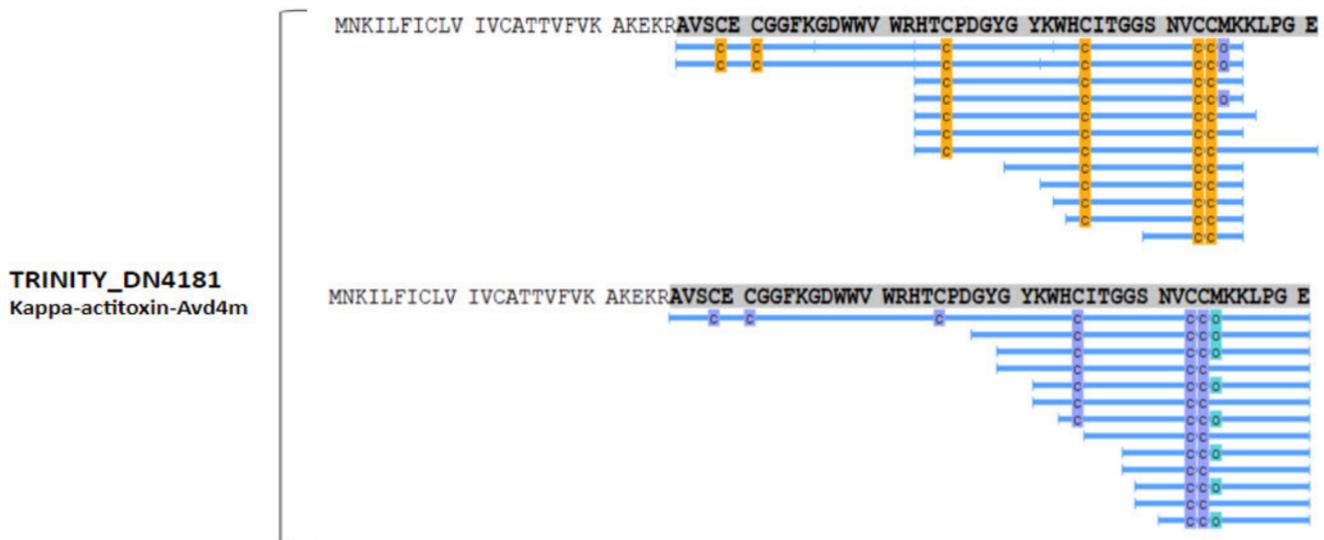

Figure 7. Peptide fragments identified in the proteomics and peptidomics experiments for Kappa-actitocin-Avd4m-like protein. In the sequences, differences in peptides with cysteine residues detected by proteomics (cysteines highlighted in yellow) compared with peptidomics (cysteines highlighted in purple) are observed. The letter *o*, highlighted in green, indicates methionine oxidation.

3.2.1. Toxins

Toxins detected in the annotated transcriptome and confirmed in the proteome, reduced and alkylated peptidome, and/or non-reduced and non-alkylated peptidome in more than one sample were selected and are listed in Table 4. Transcripts with toxin domains that presented sequences in the other omic approaches that were identified with a PFAM (Protein Families Database) search were selected and are presented in Table 5.

The most abundant toxin family identified was sea anemone type 3 (BDS) potassium channel toxin with 7 proteins codified by 9 transcripts (Table 4). This family's members share a defensin domain (PF07936.13). According to Beress et al. (1985), toxins from this family have antihypertensive and antiviral activity. U-AITX-Bcs2a, already characterized from *B. caissarum* [6], is one of the members. Two proteins matched the described toxins from *Bunodosoma granulifera*: U-AITX-Bgr3a and U-AITX-Bgr3d, which can selectively inhibit voltage-gated potassium channels type Kv11/KCNH/ERG according to UniProt database. Matches for Kappa-AITX-Avd4m and U-AITX-Ael2c, which is an APeTx-like peptide, were also found. In the reduced and alkylated peptidome, a toxin codified by TRINITY_DN4279 with a defensin domain and no similarity to any other described toxin is a candidate for the type 3 potassium channel family (Table 5).

The Kunitz domain is a highly conserved region common in toxins, characterized by serine protease inhibition, which is observed among sea anemone peptides. Four toxins belonging to the sea anemone potassium channel type 2 subfamily were identified: π-stichotoxin-Hcr2e, U-actitoxin-Avd3n, π-actitoxin-Aeq3a-like, and κπ-actitoxin-Avd3e-like (Table 4). π-SHTX-Hcr2e (also called InhVJ) inhibits trypsin and chymotrypsin, but not Kv channels. U-AITX-Avd3n or AsKC11 is an activator of G-protein-coupled inward-rectifier potassium channels, which regulates cellular excitability in the brain and cardiac cells. It also acts in voltage-gated potassium (Kv) 1.2 and dendrotoxin receptors [43,44]. Other proteins with the Kunitz domain were identified in proteomic data, such as Carboxypeptidase inhibitor SmCl-like from *A. tenebrosa* (Table 5).

Peptides with cysteine patterns are of increasing interest because of their higher stability, giving them resistance to proteolysis. The epidermal growth factor domain is a disulfide-rich structure highly abundant among the plant and animal kingdoms. Diverse

toxins from the EGF domain peptide family are deposited in the UniProt database from sea anemones *A. tenebrosa*, *Nematostella vectensis*, *Stichodactyla haddoni*, *Bunodosoma caissarum* and ants *Myrmecia gulosa* and *Manica rubida*. Recently, Eagles et al. (2022) described ant toxin MIITX2-Mg1a's activity as mimetic of EGF-like hormones, engendering a long-lasting hypersensitivity in mice [45]. One transcript confirmed in the proteome matched the described Gigantoxin 4 or U-AITX-avd12a from *Anemonia viridis* and a putative toxin with a similarity to neurogenic locus notch homolog protein 1 presented EGF-domain peptides.

Proteome and peptidome also contained turripeptides, members of the conopeptide P-like family, from *Polystira albida* and *Lophiotoma olangoensis*, two marine gastropods. Cone snails, as well as sea anemones, produce a wide range of biomolecules with high selectivity, turning them into potential pharmaceuticals. Turripeptides contain a Kazal-type domain, which is common in protease inhibitors, having six cysteine residues; turripeptides are also capable of ion channel regulation. These peptides were described in zoanthids, jellyfish, and marine annelids, but not in sea anemones [46–49]. Six transcripts have the Kazal-type domain, according to PFAM. Three were classified as putative toxins; one of them, identified as Agrin, is a protein that induces aggregation of nicotinic acetylcholine receptors [50]; one with fibrilin-2-like isoform X3; and one with neurogenic locus notch homolog protein 1-like (Table 5). Two new toxins with Kazal domain were identified by their similarity to π-AITX-Avd5a and turripeptides Pal 9.2 (Table 4), which were also found in the reduced and alkylated peptidome.

Phospholipases A2 cleaves glycerophospholipids, membrane components, generating lysophospholipids and fatty acids that compose the venom of snakes, scorpions, spiders, bees, corals, and sea anemones [51–56]. One phospholipase A2 was identified, but did not share higher similarity to the phospholipase A2, already characterized from *B. caissarum*'s, phospholipase A2 Bcs-2a. These sequences shared higher similarity to *A. tenebrosa*'s basic phospholipase A2 pseudexin A chain-like.

ShKT domain was described after the *Stichodactyla helianthus*, ShK, and is characteristically present among potassium channel inhibitors in sea anemones. The Actinaria family, which presents the ShKT domain, is potassium channel toxin type 1, and in the proteome, peptides with similarity to U-actitoxin-Avd8a-like, Kappa-actitoxin-Aer3a (AETX-K) and Kappa-actitoxin-Bgr1a (BgK) were identified. The U-AITX-Avd8a toxin belongs to the sea anemone 8 toxin family, whose functions remain poorly described. AETX-K acts in Shaker and Shaw Kv channels, as well as BgK, which inhibits Kv 1.1, 1.2, and 1.3 channels largely distributed in the central nervous system [14,57]. In addition to these described toxins, six transcripts were identified and confirmed in the proteome and in the reduced and alkylated peptidome with ShK domain, all of which had a similarity to uncharacterized proteins of *A. tenebrosa*. These sequences were classified as new putative toxins since they have the ShK domain and were identified in other sea anemone species.

Table 4. *Bunodosoma caissarum*'s new toxins identified through similarity with other described toxins in transcriptome and confirmed in proteome and/or reduced and alkylated peptidome. * These toxin names were suggested according to Oliveira et al. (2006) proposed nomenclature for sea anemone toxins [58].

Toxin	Trinity Code	Transcript (Identity)	E-Value	Toxin Family	Tentacle/Mucus	Omic Approach
* U-actitoxin-Bcs3d	TRINITY_DN72667	U-actitoxin-Bgr3d (42.3%)	4.4E−13	Sea anemone type 3 (BDS) potassium channel toxin	−/+	Proteome, RA peptidome
* U-actitoxin-Bcs3a	TRINITY_DN10015 TRINITY_DN5917 TRINITY_DN38124	U-actitoxin-Bgr3a (38.1%)	4.8E−06	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome, RA peptidome
* U-actitoxin-Bcs2c	TRINITY_DN42790	U-actitoxin-Ael2c (45.7%)	1.8E−09	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome

Table 4. *Cont.*

Toxin	Trinity Code	Transcript (Identity)	E-Value	Toxin Family	Tentacle/Mucus	Omic Approach
* K-actitoxin-Bcs4m	TRINITY_DN4181	Kappa-actitoxin-Avd4m (50%)	1.4E−16	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome, RA peptidome
U-actitoxin-Bcs2a	TRINITY_DN14686	U-actitoxin-Bcs2a (100%)	1.3E−30	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome, RA peptidome
Bcs Tx3	TRINITY_DN1024	type III potassium channel toxin protein, partial (96%)	1.4E−47	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome, RA peptidome
* K-actitoxin-Bcs2a	TRINITY_DN58216	Kappa-actitoxin-Ael2a (50%)	1.2E−08	Sea anemone type 3 (BDS) potassium channel toxin	−/+	RA peptidome
* KΠ-actitoxin-Bcs3e	TRINITY_DN13254	kappaPI-actitoxin-Avd3e-like (73.5%)	4.9E−90	Venom Kunitz type. Sea anemone type 2 potassium channel toxin.	+/+	Proteome
* Π-actitoxin-Bcs2e	TRINITY_DN217	PI-stichotoxin-Hcr2e (56.4%)	2.5E−28	Venom Kunitz type. Sea anemone type 2 potassium channel toxin.	+/−	Proteome
* U-actitoxin-Bcs3n	TRINITY_DN6572	U-actitoxin-Avd3n (74.7%)	1.6E−51	Venom Kunitz type. Sea anemone type 2 potassium channel toxin.	+/+	Proteome, RA peptidome
* Π-actitoxin-Bcs3a	TRINITY_DN10770	PI-actitoxin-Aeq3a-like (85.9%)	1.2E−59	Venom Kunitz type. Sea anemone type 2 potassium channel toxin.	−/+	Proteome
* U-actitoxin-Bcs12a	TRINITY_DN7989	U-actitoxin-Avd12a (74.4%)	9.4E−38	EGF-domain peptide	+/+	Proteome
Turripeptide Pal 9.2-like	TRINITY_DN4295	Turripeptide Pal 9.2 (38.2%)	1.2E−12	Kazal-type serine protease inhibitor domain	+/+	Proteome, RA peptidome
* Π-actitoxin-Bcs5a	TRINITY_DN713	PI-actitoxin-Avd5a (67.4%)	8E−24	Kazal-type serine protease inhibitor domain	+/+	Proteome
Basic phospholipase A2 pseudexin A chain-like	TRINITY_DN24551	Basic phospholipase A2 pseudexin A chain-like (72.8%)	7.8E−108	Phospholipase A2	−/+	Proteome
* K-actitoxin-Bcs1a	TRINITY_DN3227	Kappa-actitoxin-Bgr1a (89.2%)	1.4E−20	Sea anemone type 1 potassium channel toxin	−/+	Proteome
* K-actitoxin-Bcs3a	TRINITY_DN4127	Kappa-actitoxin-Aer3a (55.4%)	1.3E−28	Sea anemone type 1 potassium channel toxin	+/+	Proteome
* U-actitoxin-Bcs8a	TRINITY_DN1939	U-actitoxin-Avd8a-like (82.3%)	1.9E−51	Sea anemone 8 toxin	+/+	Proteome
* Δ-actitoxin-Bcs2a	TRINITY_DN24411	Delta-stichotoxin-Sgt2a (48.2%)	3.1E−15	Anemone neurotoxin	−/+	Proteome, RA peptidome
Zinc-metalloproteinase nas-13-like	TRINITY_DN24501	Zinc-metalloproteinase nas-13-like (84.9%)	3.5E−80	Astacin	−/+	Proteome
Zinc metalloproteinase nas-4-like	TRINITY_DN2272	Zinc metalloprotease nas-4-like isoform X2 (82.9%)	4E−269	Astacin	+/+	Proteome
Zinc metalloproteinase nas-15-like	TRINITY_DN1535	Zinc-metalloproteinase nas-15-like (68.2%)	1.2E−221	Astacin	+/+	Proteome
Cystatin	TRINITY_DN6968	Cystatin-like (57.2%)	7.3E−65	Cystatin domain	+/+	Proteome

RA peptidome = Reduced and alkylated peptidome; NRNA peptidome = Non-reduced and non-alkylated peptidome. Bgr: *Bunodosoma granuliferum*; Ae: *Actinia equina*; Av:*Anemonia viridis*; Bc: *Bunodosoma caissarum*; Hc: *Heteractis crispa*; Sg: *Stichodactyla gigantea*.

A similar candidate to Δ-stichotoxin-Sgt2a or gigantoxin II, which is a sodium channel toxin with an LD50 of 70 µg/kg in crabs, was also found in our proteomic data. This toxin is characterized as having a sea anemone neurotoxin domain. Another putative toxin with this domain (TRINITY_DN2326) was identified in the reduced and alkylated peptidome.

Table 5. *Bunodosoma caissarum*'s new toxins identified through similarity with other described toxins in transcriptome and confirmed in proteome and/or reduced and alkylated peptidome.

Trinity Code	Transcript (Identity)	E-Value	Toxin Family	Tentacle/Mucus	Omic Approach
TRINITY_DN4279	-	-	Sea anemone type 3 (BDS) potassium channel toxin	+/+	Proteome, RA peptidome
TRINITY_DN14881	Carboxypeptidase inhibitor SmCI-like (64.1%)	3E−32	Venom Kunitz type. Sea anemone type 2 potassium channel toxin	−/+	Proteome
TRINITY_DN4129	Neurogenic locus notch homolog protein 1 (37.1%)	7.9E−12	EGF-domain	+/+	Proteome
TRINITY_DN3459	Agrin-like (42.5%)	4.1E−71	Kazal-type serine protease inhibitor domain	−/+	Proteome
TRINITY_DN322	Fibrilin-2-like isoform X3 (69.5%)	4.3E−70	Kazal-type serine protease inhibitor domain	+/+	Proteome
TRINITY_DN14498	Neurogenic locus notch homolog protein 1-like isoform X6 (81.8%)	2.4E−55	Kazal-type serine protease inhibitor domain	+/+	Proteome
TRINITY_DN1312	Uncharacterized protein ZK643.6-like (47.2%)	6.4E−52	ShK domain-like	+/+	Proteome
TRINITY_DN22515	Uncharacterized protein LOC116291117 (72.1%)	3.5E−75	ShK domain-like	+/+	Proteome
TRINITY_DN10788	Uncharacterized protein LOC116293550 (67.9%)	2.2E−227	ShK domain-like	+/−	Proteome
TRINITY_DN10521	Uncharacterized protein LOC116287301 (52.2%)	1.9E−48	ShK domain-like	+/+	Proteome, RA peptidome, NRNA peptidome
TRINITY_DN4363	Uncharacterized protein LOC116301037 (80.4%)	0	ShK domain-like	+/−	Proteome
TRINITY_DN19291	Uncharacterized protein LOC116291117 (61.8%)	1.3E−135	ShK domain-like	+/+	Proteome
TRINITY_DN2326	mRNA, partial (76%)	5.9E−45	Anemone neurotoxin	−/+	RA peptidome
TRINITY_DN52114	Uncharacterized protein LOC116291022 (70.8%)	6.6E−102	F5/8 C domain	−/+	Proteome
TRINITY_DN16749	Uncharacterized protein LOC116297803 isoform X6 (67.5%)	1.4E−50	F5/8 C domain	−/+	Proteome
TRINITY_DN8245	Uncharacterized protein LOC116298543 (67.6%)	5.2E−70	F5/8 C domain	+/+	Proteome
TRINITY_DN37576	Uncharacterized protein LOC116297953 (65.3%)	6.2E−44	F5/8 C domain	−/+	RA peptidome, NRNA peptidome
TRINITY_DN10809	Perlucin-like protein (65.7%)	1.2E−98	Lectin C-type domain	+/+	Proteome

RA peptidome = Reduced and alkylated peptidome; NRNA peptidome = Non-reduced and non-alkylated peptidome.

C-type lectins (CTLs) are commonly found in snake venoms, and, recently, bioprospection of marine organisms also showed the presence of these lectins [59]. CTLs have a variety of functions, but the higher expression of receptors in this lectin type in infection situations suggests its role in immune responses [60]. C-type lectins are important for inflammation, phagocytosis, and antigen defense, due to their antimicrobial effects [61], including against viruses such as SARS-CoV-2 [62]. CTLs may inhibit entry, attachment, or replication of the antigen material by binding with glycoprotein from the viral envelope and preventing binding and recognition with receptors [59,63,64]. C-type lectins from diverse marine species are arousing interest among scientists due to their involvement with immune responses. Through domain detection with PFAM, a C-type lectin was classified as a putative toxin, and its presence was confirmed in the proteome (Table 5 and Figure 8).

Figure 8. New putative toxins found according to PFAM search of related toxin domains. The transcript sequence is tagged with a light green box representing the corresponding domain. Sequences found with the corresponding transcript in proteome, reduced and alkylated peptidome, and non-reduced and non-alkylated peptidome are represented by orange, green, and dark red boxes, respectively, below the transcript sequence. The alignment of putative toxin transcripts with the three best matches in NCBI with BLASTp is shown in Figure S6.

Another domain described in putative toxins is Factor V and Factor VIII, also called F 5/8C or discoidin domain, characterized by their role in the mammalian blood coagulation system. Proteins with this domain are usually responsible for cell recognition, adhesion and aggregation, and phospholipid binding, behaving as lectins according to Baumgartner et al. (1998) [65]. Madio et al. (2017) identified F 5/8C domain sequences in *Stichodactyla haddoni*'s transcriptome, although these transcripts were not confirmed in the proteome [27]. Here, three F 5/8C were identified in the proteome and one in the reduced and alkylated peptidome (Table 5 and Figure 8); it was only possible to identify them using PFAM because all of them matched uncharacterized proteins from *A. tenebrosa*.

Metalloproteases observed in omics data belong to the astacin (M12A), reprolysin (M12B), neprilysin (M13), leucine aminopeptidase (M17), and other families, and their major role is disrupting the extracellular matrix [66]. These proteins are important components of venom. From the astacin family, proteins similar to zinc metalloproteinase-nas from *Caenorhabditis elegans* were present. Astacins, considered as toxins, have diverse functions, such as processing biological peptides, regulating collagen fibers assembly and digestion, and they are critical for development and morphogenesis [67–69]. ShKT domain was also observed in *Hydractinia echinate* as well as in *C. elegans* [70]. Three transcripts with similarity to zinc metalloproteinase nas-13 and 15 were identified and confirmed in the proteome. Disintegrin and metalloproteinases belong to the reprolysin or M12B family and are involved in processes such as carcinogenesis, neoplastic progression, metastasis, extracellular matrix degradation, immune response evasion, and inflammation in mammals [71]. Two transcripts presented a similarity of around 40 percent with mammals but reached 80 percent identity to *A. tenebrosa*'s proteins. Peptidase M13 domain characterizes a group of metallopeptidases, represented by neprilysin, a protein involved in the inflammatory response. Two transcripts have the M13 domain, one of them matched endothelin-converting enzyme 1-like, which is a type II integral membrane protein and is important for muscle contraction and morphogenesis [66], and the other with membrane metallo-endopeptidase-like 1; both are described from *A. tenebrosa*. The peptidase M17 domain characterizes the leucine aminopeptidase family, whose major role is in the protein degradation process, releasing the N-terminal amino acid. Two transcripts contained the domain and were highly similar to putative aminopeptidases from *A. tenebrosa* (up to 90%). Other two metalloprotease domains were identified: M1 (aminopeptidase family) with one transcript and M2—peptidyl-dipeptidase A, which cleaves dipeptides from the C-terminal—with one. A serine peptidase domain was also identified in transcripts confirmed in the proteome, the subtilase domain. Two proteins with this domain matched *A. tenebrosa*'s proteins, and one of them was also identified in the reduced and alkylated peptidome (Table S5). Except for the astacin domain-containing proteins, all the other peptidase domains containing proteins were described as venom proteins (they are described as venom components and not as toxins).

3.2.2. Cysteine Rich Proteins

Cysteine-rich proteins found in proteomic data may be classified according to their domain: cysteine-rich secretory protein family (CAP) and scavenger receptor cysteine-rich domain (SRCR). The first family is distributed all over the animal kingdom, remarkably, in snake venoms. Four cysteine-rich secretory family proteins were identified in the proteome, two of them matched Golgi-associated plant pathogenesis-related protein 1-like, a C-type lectin, also observed in *Chrysaora fuscescens*' venom [72]. Those proteins presented higher enrichment in tentacle samples.

The scavenger receptor family is a superfamily of transmembrane cell glycoproteins related to the innate immune response, characterized by their ability to pattern recognition of diverse types of ligands. Neubauer et al. (2016) evaluated the SRCR domain proteins of six cnidaria species and observed their high importance for symbiosis regulation and immune modulation in the phylum [73]. In proteome samples, 13 proteins were identified with the scavenger receptor cysteine-rich domain and 5 were identified in the reduced and

alkylated peptidome (TRINITY_DN126, TRINITY_DN19873, TRINITY_DN21663, TRINITY_DN499, TRINITY_DN9220), only 2 of them in common with proteome. Five of these proteins matched the "deleted in malignant brain tumors 1 protein-like" of *A. tenebrosa*, which was named after the "deleted in malignant brain tumors" (DMBT) from *Homo sapiens*, a protein with 8 SRCR domain repeats. As other studies suggest, the innate immune gene families are expanding in invertebrates, and as observed by Neubauer et al. (2016), with anthozoan, it is not different [73]. Our results corroborate this hypothesis. The same study also relates the presence of proteins with diverse combinations of domains with SRCR, and in the reduced and alkylated peptidome, peptides of a transcript (TRINITY_DN499) with lectin C-type, EGF-like, and SRCR domains were identified.

3.2.3. Neuropeptides

Antho-RFamide neuropeptides occur in neurosecretory cells; these neuropeptides are related to photoreceptive organs in cnidarians and are classified as FMRF-amide-like peptides, or FLPs. This group of peptides controls several functions in cnidarians, such as reproduction, larval movement, metamorphosis, and muscle contraction, among others [74–76]. Katsukura et al. (2003) describe that RFamide neuropeptides inhibit the metamorphosis of *Hydractinia chinate*, while LWamide induce the process [77,78].

After cleavage, some neuropeptides undergo various post-translational modifications, especially at the N-terminal and/or C-terminal residues. A frequently observed modification is amidation at the C-terminus of neuropeptides, protecting them from degradation by carboxypeptidases. Pyroglutamic acid is a modification present at the N-terminus of several neuropeptides, which protects the peptide chain from enzymatic degradation by aminopeptidases. The LWamide and Antho-RFamide neuropeptides described in anemones have these modifications. As these are post-translational modifications, they are identified in mass spectrometry experiments [79,80].

Our results demonstrated the presence of Antho-RFamide neuropeptides, both in the transcriptome and in the tentacle peptidome. In the sequencing provided by the Transdecoder, it was possible to observe the presence of multiple copies of these sequences described in the literature. In the peptide extracts analyzed by mass spectrometry, 18 distinct fragments of Antho-RFamide were detected (Figure 9A). For neuropeptide LWamide, only the presence of the transcript was detected. (Figure 9B).

3.2.4. Intracellular Peptides (InPeps)

Previous studies conducted in rodent tissues [81], zebrafish [82], human cell lines [83], fungi [84], and plants [85], revealed the presence of a very large set of stable and bioactive intracellular peptides [86]. Some of these peptide fragments of intracellular proteins found altered in different biological models have shown a variety of biological and pharmacological activities, such as regulating signal transduction [87], regulating glucose uptake [88], altering intracellular calcium levels [89], potentiating interferon-gamma activity [90], and inducing cell death in various tumor cell lines [91,92].

Many intracellular peptides identified in this study were from histones. In the literature, the antimicrobial activity of histone fragments, both complete sequences and N-terminal [93] or C-terminal fragments [94], have been demonstrated in fish and invertebrates [95]. There was also a fragment of histone H2B type 1-H intracellular peptide (PepH) showing a protective effect against cell death in neuro2A cells [96]. In this context, histone fragments are one of the examples that show the importance of applying extraction methodologies that prioritize the fraction below 10 kDa. The exciting and largely recognized therapeutic potential of intracellular peptides [87] suggests that pharmacological characterization of intracellular peptides identified here should still be conducted.
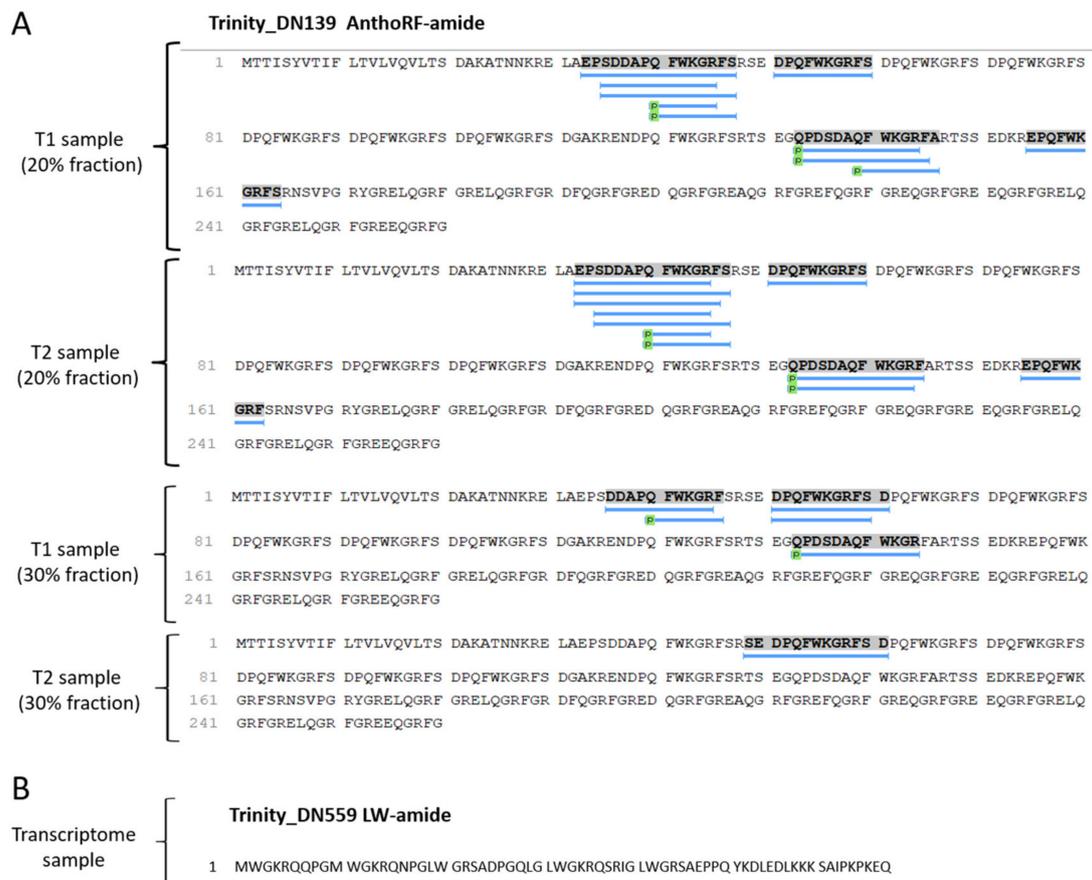
Figure 9. Neuropeptides identified in the tentacles of *B. caissarum*. In (**A**) peptides identified in the Antho-RFamide protein sequence. In (**B**) LWamide protein sequence.

4. Materials and Methods

4.1. Collection and Maintenance of Specimens

The specimens of *B. caissarum* were collected during the low tide period on Itaqui-tanduva beach, located in the State Park "Xixová Japuí", in São Vicente County, on the south coast of São Paulo State ($-23{,}9991794$, $-46{,}3915857$). The two collection events were conducted with an interval of 6 months and, together, totalized 15 specimens of *B. caissarum*. The sea anemones were manually collected off the rocky shore and inserted on a recipient with seawater, protease inhibitor, and aerator. The acclimatization period lasted 7 days with a photoperiod of 12 h, salinity of 33 ppm in artificial seawater, and 25 °C temperature. The specimens were fed two or three times a week with small fragments of squid. For mucus obtention, the specimens were inserted in hypertonic medium (salinity: 40 ppm) and physically stimulated with a spatula around the oral disc in three cycles of 3 min of stimulating activity and 7 min of rest. Three samples with the content of three sea anemones' mucus each were obtained and posteriorly prepared as described in proteome protocols [38].

4.2. Transcriptomic

4.2.1. RNA Extraction, mRNA Library Synthesis and Illumina Sequencing

For the transcriptomic approach, one specimen was inserted in ice for 5 min and the tentacles were removed with surgical microscissors and laboratory glass. After the extraction, the sample was immediately stored in liquid nitrogen for shipping to NGS Soluções Genomicas. After adding Rnase-free Dnase I in order to eliminate genomic DNA, RNA extraction was performed following the protocol based on the Rneasy Lipid Tissue Kit. RNA integrity was evaluated with Bioanalyzer. The samples presenting an optic density

ratio of 260/280 nm higher than 1,8 and RNA integrity higher than 7 were selected for the following experiments. The cDNA library was synthesized with ~1 ug of total RNA, following the recommendations of the RNA TrueSeq manufacturer. The resultant cDNA libraries were sequenced using an Illumina HiSeq 2500 platform. The Real-Time Analysis program (Illumina) performed the base calling of sequencing images, converting them into fastq sequences. Before assembly, raw reads were evaluated using FastQC.

4.2.2. Transcriptome Assembly, Annotation and Functional Enrichment

The resultant sequencing reads were submitted to de novo assembly using Trinity version 2.11.0, a pipeline for transcriptome assembly. Standard assembly parameters were used with a minimum length of 50 bp. Trinity filtered bases of low quality, removed Illumina adapters with Trimmomatic version 0.36, counted k-mers present in sequencing reads with Jellyfish version 2.3.0, and performed digital normalization to reduce data volume, without affecting the assembly's accuracy.

4.3. Proteomic and Peptidomic

Protein and Peptide Fractions Preparation

Protein and peptide fractions of tentacle and mucus from *B. caissarum* were prepared as previously described with some modifications [97]. The tentacles collected were resuspended in 10 mL of deionized water at 80 °C each, homogenized, and kept at the same temperature for 20 min. After this incubation, the samples were cooled at 4 °C on ice for 40 min and then acidified with HCl to a final concentration of 10 mM. The cooled samples were sonicated 3 times with 20 pulses of 1s at 4 Hz and the homogenates centrifuged at $15,000 \times g$ for 40 min at 4 °C. After this point, the supernatants were filtered through a Millipore membrane that allows the passage of molecules with a molecular weight of less than 10 Kda (Amicon Ultra, Millipore, Burlington, MA, USA).

The eluates not retained by the membrane (peptide fractions) were loaded onto C18 pre-columns (OASIS—Waters, UK), washed with deionized water, and eluted within 20 and 30% acetonitrile containing 0.15% trifluoroacetic acid (TFA). Sample volumes were concentrated to 5 μL in a vacuum centrifuge and stored at −80 °C.

The samples retained by the membrane (protein fraction, to 10 Kda) were processed for proteomic analysis. A solution of 8 M Urea was added to the incubation sample to a final concentration of 4 M, followed by the addition of dithiothreitol (DTT) to a final concentration of 5 mM. The mixture was incubated at 65 °C for 60 min. Iodoacetamide (IAA) was then added to a final concentration of 15 mM, and the samples were incubated for 60 min at room temperature in the dark. To quench the excess of IAA, DTT was added to a final concentration of 10 mM; proteins were digested with 1:50 trypsin:sample (porcine; Promega, Madison, WI, USA) overnight at 37 °C, acidified with formic acid, and desalted.

4.4. NanoLC and Mass Spectrometry

The peptide mixture was suspended in 0.1% formic acid and analyzed as follows. An UltiMate 3000 Basic Automated System (Thermo Fisher®, Waltham, MA, USA) was set up and connected online with a Fusion Lumos Orbitrap mass spectrometer (Thermo Fisher®) at the mass spectrometry facility RPT02H/Carlos Chagas Institute—Fiocruz, Paraná. The peptide mixture was chromatographically separated on a column (15 cm in length with an internal diameter of 75 μm), packed in-house with ReproSil-Pur C18-AQ 3 μm resin (Dr. Maisch GmbH HPLC) with a flow rate of 250 nL/min of 5% to 38% ACN in 0.1% formic acid on a 120 min gradient. The Fusion Lumos Orbitrap was placed in data-dependent acquisition (DDA) mode to automatically turn between full-scan MS and MS/MS acquisition with 60 s dynamic exclusion. Survey scans (300–1500 m/z) were acquired in the Orbitrap system with a resolution of 120,000 at m/z 200. The most intense ions captured in a 2 s cycle time were chosen, excluding those which were unassigned or had a 1+ charge state. The selected ions were then isolated in sequence and fragmented using HCD (higher-energy collisional dissociation) with normalized collision energy of

30%. The fragment ions were analyzed with a resolution of 50,000 at 200 m/z. The general mass spectrometric conditions were as follows: 2.3 kV spray voltage, no sheath or auxiliary gas flow, heated capillary temperature of 175 °C, predictive automatic gain control (AGC) enabled, and an S-lens RF level of 30%. Mass spectrometer scan functions and nLC solvent gradients were regulated using the Xcalibur 4.1 data system (Thermo Fisher®).

4.5. Protein and Peptide Identification

The raw data files (.raw) generated by the mass spectrometer were searched in the Transdecoder database built from the *B. caissarum* transcriptome, using the PEAKS Studio software (version 8.5; Bioinformatics Solution, Waterloo, ON, Canada) [98,99]. The research parameters used were as follows: no enzymatic specificity; precursor mass tolerance adjusted to ±30 ppm and fragmentation ion mass (tolerance ±0.5 Da); oxidized methionine (+15.99 Da) and acetylation (+42.01 Da) were defined as variable modifications. Carbamidomethylation (+57.02) was also added as a variable modification for alkylated and reduced samples. The identified peptides were then sorted by their mean local confidence to select the best spectra for annotation and filtered by FDR ≤1%.

4.6. Quantitative Data Analysis

Enriched proteins of tentacle and mucus were identified according to their sample area. Proteins that presented sample area as zero were replaced by a minimum value of 1E−02 because PEAKS show detected proteins with low signal as zero and when the protein was not present in the sample, it showed as "-". This symbol was replaced by zero and zero by 1E−02, followed by standard normalization. The mean value of each protein for tentacle and mucus samples was calculated, and proteins that were first replaced by 1E−02, were replaced with the mean value. RStudio ggplot package and Python matplotlib were used to plot data.

5. Conclusions

In summary, we report the first integrated omics analysis to characterize the protein composition of tentacles and mucus for the sea anemone species *B. caissarum*. The proteomics and peptidomics data obtained using the sequences generated from the tentacle transcriptome as a database proved to be a robust method in the identification of similar toxins but not completely identical to the sequences already characterized in the phylum of cnidarians, belonging to the following families: sea anemone type 3 (BDS) potassium channel toxin Venom Kunitz type, sea anemone type 2 potassium channel toxin, EGF-domain peptide, Kazal-type serine protease inhibitor domain, phospholipase A2, sea anemone type 1 potassium channel toxin, sea anemone 8 toxin, astacin and cystatin domain, among others. In addition, toxins were more present in the mucus than in the tentacle, showing that the stimulation protocol helps to distinguish which proteins are more related to the venom and which are more structural. Finally, the peptidome also revealed the presence of many fragments of intracellular proteins, neuropeptides, and toxin fragments of different sizes, showing proteolytic processing. This approach demonstrates the complexity of the venom of this species of a sea anemone.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/md21030197/s1, Figure S1. Experimental fluxogram; Figure S2: Tentacle enriched proteins; Figure S3. A. Mucus enriched known proteins. B. Mucus enriched unknown proteins codified by Trinity transcripts; Figure S4: Peptides identified in non reduced and non alkylated peptidome with its corresponding Trinity transcripts and matched known protein when identified; Figure S5: Peptides identified in reduced and alkylated peptidome with its corresponding Trinity transcripts and matched known protein when identified; Figure S6: Alignment of putative toxins transcripts with the three best matches in NCBI with BLASTp; Table S1: Transcripts identified as putative toxin related through BLASTp search of TransDecoder data against UniProt Animal Toxin Annotation Project (ToxProt); Table S2: Proteome data; Table S3: Non-reduced and non-alkylated peptidome (20% fraction); Table S4: Non-reduced and Non-alkylated peptidome (30% fraction);

Table S5: Mucus pool reduced and alkylated peptidome; Table S6: Tentacle pool reduced and alkylated peptidome; Table S7: Toxins, putative toxins and venom proteins identified in proteome and reduced and alkylated peptidome grouped by domains identified with PFAM search. Putative toxins are in yellow.

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

References

1. Bakus, G.J.; Targett, N.M.; Schulte, B. Chemical ecology of marine organisms: An overview. *J. Chem. Ecol.* **1986**, *12*, 951–987. [CrossRef] [PubMed]
2. Coll, J.; La Barre, S.; Sammarco, P.; Williams, W.; Bakus, G. Chemical Defences in Soft Corals (Coelenterata: Octocorallia) of the Great Barrier Reef: A Study of Comparative Toxicities. *Mar. Ecol. Prog. Ser.* **1982**, *8*, 271–278. [CrossRef]
3. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: The evolutionary novelty of venoms. *Trends Ecol. Evol.* **2013**, *28*, 219–229. [CrossRef] [PubMed]
4. Moran, Y.; Gordon, D.; Gurevitz, M. Sea anemone toxins affecting voltage-gated sodium channels—Molecular and evolutionary features. *Toxicon* **2009**, *54*, 1089–1101. [CrossRef]
5. Diochot, S.; Lazdunski, M. Sea anemone toxins affecting potassium channels. In *Marine Toxins as Research Tools. Progress in Molecular and Subcellular Biology*; Fusetani, N., Kem, W., Eds.; Springer: Berlin, Germany, 2009; Volume 46, pp. 99–122.
6. Oliveira, J.S.; Zaharenko, A.J.; Ferreira, W.A.; Konno, K.; Shida, C.S.; Richardson, M.; Lúcio, A.D.; Beirão, P.S.L.; de Freitas, J.C. BcIV, a new paralyzing peptide obtained from the venom of the sea anemone Bunodosoma caissarum. A comparison with the Na+ channel toxin BcIII. *Biochim. Biophys. Acta (BBA)—Proteins Proteom.* **2006**, *1764*, 1592–1600. [CrossRef]
7. Chun, H.G.; Davies, B.; Hoth, D.; Suffness, M.; Plowman, J.; Flora, K.; Grieshaber, C.; Leyland-Jones, B.; Didemnin, B. The first marine compound entering clinical trials as an antineoplastic agent. *Investig. New Drugs* **1986**, *4*, 279–284. [CrossRef]
8. Schweikart, K.; Guo, L.; Shuler, Z.; Abrams, R.; Chiao, E.T.; Kolaja, K.L.; Davis, M. The effects of jaspamide on human cardiomyocyte function and cardiac ion channel activity. *Toxicol. Vitr.* **2012**, *27*, 745–751. [CrossRef]
9. Tarcha, E.J.; Olsen, C.M.; Probst, P.; Peckham, D.; Muñoz-Elías, E.J.; Kruger, J.G.; Iadonato, S.P. Safety and pharmacodynamics of dalazatide, a Kv1.3 channel inhibitor, in the treatment of plaque psoriasis: A randomized phase 1b trial. *PLoS ONE* **2017**, *12*, e0180762. [CrossRef]
10. Rinehart, K.L.; Holt, T.G.; Fregeau, N.L.; Keifer, P.A.; Wilson, G.R.; Perun, T.J., Jr.; Sakai, R.; Thompson, A.G.; Stroh, J.G.; Shield, L.S.; et al. Bioactive Compounds from Aquatic and Terrestrial Sources. *J. Nat. Prod.* **1990**, *53*, 771–792. [CrossRef]
11. Castañeda, O.; Sotolongo, V.; Amor, A.M.; Stöcklin, R.; Anderson, A.J.; Harvey, A.L.; Engström, Å.; Wernstedt, C.; Karlsson, E. Characterization of a potassium channel toxin from the Caribbean sea anemone Stichodactyla helianthus. *Toxicon* **1995**, *33*, 603–613. [CrossRef] [PubMed]
12. Beeton, C.; Wulff, H.; Standifer, N.E.; Azam, P.; Mullen, K.M.; Pennington, M.W.; Kolski-Andreaco, A.; Wei, E.; Grino, A.; Counts, D.R.; et al. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17414–17419. [CrossRef] [PubMed]
13. Devarajan, P.; Chen, Z. Autoimmune effector memory T cells: The bad and the good. *Immunol. Res.* **2013**, *57*, 12–22. [CrossRef] [PubMed]
14. Cotton, J.; Crest, M.; Bouet, F.; Alessandri, N.; Gola, M.; Forest, E.; Karlsson, E.; Castaneda, O.; Harvey, A.L.; Vita, C.; et al. A Potassium-Channel Toxin from the Sea Anemone Bunodosoma Granulifera, An Inhibitor for Kv1 Channels—Revision of the

Amino Acid Sequence, Disulfide-Bridge Assignment, Chemical Synthesis, and Biological Activity. *Eur. J. Biochem.* **1997**, *244*, 192–202. [CrossRef] [PubMed]

15. Malpezzi, E.L.; Freitas, J.C. Antimitotic effect of an extract of the sea anemone *Bunodosoma caissarum* on sea urchin egg development. *Braz. J. Med. Biol. Res.* **1990**, *23*, 811–814.

16. Malpezzi, E.L.; Freitas, J.C. Hemolytic activity of the nematocyst venom from the sea anemone *Bunodosoma caissarum*. *Braz. J. Med. Biol. Res.* **1991**, *24*, 1245–1249.

17. Malpezzi, E.L.A.; Matsui, D.H.; Groote, S.C.T.S.; Freitas, J.C.; Santelli, G.M.; Fernandas, J.B. Antitumoral activity in an organic extract of the sea anemone *Bunodosoma caissarum*. *Toxicon* **1993**, *33*, 291.

18. Malpezzi, E.L.A. Pharmacologically Active Substances of the Sea Anemone *Bunodosoma caissarum*, Corrêa, 1964 (Cnidaria, Anthozoa, Actiniidae). *J. Venom. Anim. Toxins Incl. Trop. Dis.* **1996**, *2*, 60. [CrossRef]

19. Malpezzi, E.L.; de Freitas, J.; Muramoto, K.; Kamiya, H. Characterization of peptides in sea anemone venom collected by a novel procedure. *Toxicon* **1993**, *31*, 853–864. [CrossRef]

20. Appeltans, W.; Ahyong, S.T.; Anderson, G.; Angel, M.V.; Artois, T.; Bailly, N.; Bamber, R.; Barber, A.; Bartsch, I.; Berta, A.; et al. The Magnitude of Global Marine Species Diversity. *Curr. Biol.* **2012**, *22*, 2189–2202. [CrossRef]

21. Klompen, A.; Macrander, J.; Reitzel, A.; Stampar, S. Transcriptomic Analysis of Four Cerianthid (Cnidaria, Ceriantharia) Venoms. *Mar. Drugs* **2020**, *18*, 413. [CrossRef]

22. Fu, J.; He, Y.; Peng, C.; Tang, T.; Jin, A.; Liao, Y.; Shi, Q.; Gao, B. Transcriptome Sequencing of the Pale Anemones (Exaiptasia diaphana) Revealed Functional Peptide Gene Resources of Sea Anemone. *Front. Mar. Sci.* **2022**, *9*, 856501. [CrossRef]

23. Warner, J.F.; Röttinger, E. Transcriptomic Analysis in the Sea Anemone Nematostella vectensis. In *Developmental Biology of the Sea Urchin and Other Marine Invertebrates*; Carroll, D.J., Stricker, S.A., Eds.; Methods in Molecular Biology; Springer Protocols: New York, NY, USA, 2020; Volume 2219, pp. 231–240.

24. Stefanik, D.J.; Lubinski, T.J.; Granger, B.R.; Byrd, A.L.; Reitzel, A.M.; DeFilippo, L.; Lorenc, A.; Finnerty, J.R. Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, *Edwardsiella lineata*, a parasitic cnidarian. *BMC Genom.* **2014**, *15*, 71. [CrossRef] [PubMed]

25. Macrander, J.; Brugler, M.R.; Daly, M. A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive Anthopleura elegantissima polyps. *BMC Genom.* **2015**, *16*, 221. [CrossRef] [PubMed]

26. Rivera-De-Torre, E.; Martínez-Del-Pozo, Á.; Garb, J.E. Stichodactyla helianthus' de novo transcriptome assembly: Discovery of a new actinoporin isoform. *Toxicon* **2018**, *150*, 105–114. [CrossRef]

27. Madio, B.; Undheim, E.A.B.; King, G.F. Revisiting venom of the sea anemone Stichodactyla haddoni: Omics techniques reveal the complete toxin arsenal of a well-studied sea anemone genus. *J. Proteom.* **2017**, *166*, 83–92. [CrossRef]

28. Surm, J.M.; Stewart, Z.K.; Papanicolaou, A.; Pavasovic, A.; Prentis, P.J. The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution. *Ecol. Evol.* **2019**, *9*, 11314–11328. [CrossRef]

29. Liao, Q.; Feng, Y.; Yang, B.; Lee, S.M.-Y. Cnidarian peptide neurotoxins: A new source of various ion channel modulators or blockers against central nervous systems disease. *Drug Discov. Today* **2018**, *24*, 189–197. [CrossRef]

30. Linial, M.; Rappoport, N.; Ofer, D. Overlooked Short Toxin-Like Proteins: A Shortcut to Drug Design. *Toxins* **2017**, *9*, 350. [CrossRef] [PubMed]

31. Vernon, L.P.; Bell, J.D. Membrane structure, toxins and phospholipase A2 activity. *Pharmacol. Ther.* **1992**, *54*, 269–295. [CrossRef]

32. Rojko, N.; Dalla Serra, M.; Maček, P.; Anderluh, G. Pore formation by actinoporins, cytolysins from sea anemones. *Biochimica Biophysica Acta (BBA)—Biomembr.* **2016**, *1858*, 446–456. [CrossRef]

33. Podobnik, M.; Anderluh, G. Pore-forming toxins in Cnidaria. *Semin. Cell Dev. Biol.* **2017**, *72*, 133–141. [CrossRef] [PubMed]

34. Honma, T.; Minagawa, S.; Nagai, H.; Ishida, M.; Nagashima, Y.; Shiomi, K. Novel peptide toxins from acrorhagi, aggressive organs of the sea anemone Actinia equina. *Toxicon* **2005**, *46*, 768–774. [CrossRef] [PubMed]

35. Martins, R.D.; Alves, R.S.; Martins, A.M.; Barbosa, P.S.F.; Evangelista, J.S.; Evangelista, J.J.F.; Ximenes, R.M.; Toyama, M.H.; Toyama, D.O.; Souza, A.J.F.; et al. Purification and characterization of the biological effects of phospholipase A2 from sea anemone Bunodosoma caissarum. *Toxicon* **2009**, *54*, 413–420. [CrossRef] [PubMed]

36. Orts, D.J.B.; Peigneur, S.; Madio, B.; Cassoli, J.S.; Montandon, G.G.; Pimenta, A.M.C.; Bicudo, J.E.P.W.; Freitas, J.C.; Zaharenko, A.J.; Tytgat, J. Biochemical and Electrophysiological Characterization of Two Sea Anemone Type 1 Potassium Toxins from a Geographically Distant Population of *Bunodosoma caissarum*. *Mar. Drugs* **2013**, *11*, 655–679. [CrossRef] [PubMed]

37. Orts, D.J.B.; Moran, Y.; Cologna, C.T.; Peigneur, S.; Madio, B.; Praher, D.; Quinton, L.; de Pauw, E.; Bicudo, J.E.P.W.; Tytgat, J.; et al. BcsTx3 is a founder of a novel sea anemone toxin family of potassium channel blocker. *FEBS J.* **2013**, *280*, 4839–4852. [CrossRef]

38. Ramírez-Carreto, S.; Vera-Estrella, R.; Portillo-Bobadilla, T.; Licea-Navarro, A.; Bernaldez-Sarabia, J.; Rudiño-Piñera, E.; Verleyen, J.J.; Rodríguez, E.; Rodríguez-Almazán, C. Transcriptomic and Proteomic Analysis of the Tentacles and Mucus of Anthopleura dowii Verrill, 1869. *Mar. Drugs* **2019**, *17*, 436. [CrossRef]

39. Kiyatkin, N.; Dulubova, I.; Grishin, E. Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of ankyrin-like repeats. *Eur. J. Biochem.* **1993**, *213*, 121–127. [CrossRef]

40. Holz, G.G.; Habener, J.F. Black widow spider α-latrotoxin: A presynaptic neurotoxin that shares structural homology with the glucagon-like peptide-1 family of insulin secretagogic hormones. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **1998**, *121*, 177–184. [CrossRef]

41. Stabili, L.; Schirosi, R.; Parisi, M.G.; Piraino, S.; Cammarata, M. The Mucus of Actinia equina (Anthozoa, Cnidaria): An Unexplored Resource for Potential Applicative Purposes. *Mar. Drugs* **2015**, *13*, 5276–5296. [CrossRef]

42. Savoca, S.; Di Fresco, D.; Alesci, A.; Capillo, G.; Spanò, N. Mucus secretions in Cnidarian, an ecological, adaptive and evolutive tool. *Adv. Oceanogr. Limnol.* **2022**, *13*, 1–24. [CrossRef]

43. Schweitz, H.; Bruhn, T.; Guillemare, E.; Moinier, D.; Lancelin, J.-M.; Béress, L.; Lazdunski, M. Kalicludines and kaliseptine. *J. Biol. Chem.* **1995**, *270*, 25121–25126. [CrossRef]

44. An, D.; Pinheiro-Junior, E.L.; Béress, L.; Gladkikh, I.; Leychenko, E.; Undheim, E.A.B.; Peigneur, S.; Tytgat, J. AsKC11, a Kunitz Peptide from *Anemonia sulcata*, Is a Novel Activator of G Protein-Coupled Inward-Rectifier Potassium Channels. *Mar. Drugs* **2022**, *20*, 140. [CrossRef] [PubMed]

45. Eagles, D.A.; Saez, N.J.; Krishnarjuna, B.; Bradford, J.J.; Chin, Y.K.-Y.; Starobova, H.; Mueller, A.; Reichelt, M.E.; Undheim, E.A.B.; Norton, R.S.; et al. A peptide toxin in ant venom mimics vertebrate EGF-like hormones to cause long-lasting hypersensitivity in mammals. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2112630119. [CrossRef] [PubMed]

46. Huang, C.; Morlighem, J.R.; Zhou, H.; Lima, É.P.; Gomes, P.B.; Cai, J.; Lou, I.; Pérez, C.D.; Lee, S.M.; Rádis-Baptista, G. The Transcriptome of the Zoanthid *Protopalythoa variabilis* (Cnidaria, Anthozoa) Predicts a Basal Repertoire of Toxin-like and Venom-Auxiliary Polypeptides. *Genome Biol. Evol.* **2016**, *8*, 3045–3064. [CrossRef]

47. Liao, Q.; Gong, G.; Poon, T.C.W.; Ang, I.L.; Lei, K.M.K.; Siu, S.W.I.; Wong, C.T.T.; Rádis-Baptista, G.; Lee, S.M.-Y. Combined transcriptomic and proteomic analysis reveals a diversity of venom-related and toxin-like peptides expressed in the mat anemone *Zoanthus natalensis* (Cnidaria, Hexacorallia). *Arch. Toxicol.* **2019**, *93*, 1745–1767. [CrossRef]

48. Liang, H.; Jiang, G.; Wang, T.; Zhang, J.; Liu, W.; Xu, Z.; Zhang, J.; Xiao, L. An integrated transcriptomic and proteomic analysis reveals toxin arsenal of a novel Antarctic jellyfish *Cyanea* sp. *J. Proteom.* **2019**, *208*, 103483. [CrossRef] [PubMed]

49. Jaimes-Becerra, A.; Gacesa, R.; Doonan, L.B.; Hartigan, A.; Marques, A.C.; Okamura, B.; Long, P.F. "Beyond primary sequence"—Proteomic data reveal complex toxins in cnidarian venoms. *Integr. Comp. Biol.* **2019**, *59*, 777–785. [CrossRef]

50. Patthy, L.; Nikolics, K. Functions of agrin and agrin-related proteins. *Trends Neurosci.* **1993**, *16*, 76–81. [CrossRef]

51. Lomonte, B.; Križaj, I. Snake venom phospholipase A2 toxins. In *Handbook of Venoms and Toxins of Reptiles*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2021; pp. 389–412.

52. Krayem, N.; Gargouri, Y. Scorpion venom phospholipases A2: A minireview. *Toxicon* **2020**, *184*, 48–54. [CrossRef]

53. Atakuziev, B.U.; Nuritova, F.; Usmanov, P.B. Phospholipase A2 from the venom of the spider Eresus niger. *Chem. Nat. Compd.* **1991**, *27*, 487–489. [CrossRef]

54. Palm, N.W.; Rosenstein, R.K.; Yu, S.; Schenten, D.D.; Florsheim, E.; Medzhitov, R. Bee Venom Phospholipase A2 Induces a Primary Type 2 Response that Is Dependent on the Receptor ST2 and Confers Protective Immunity. *Immunity* **2013**, *39*, 976–985. [CrossRef]

55. Nevalainen, T.J.; Peuravuori, H.J.; Quinn, R.J.; Llewellyn, L.E.; Benzie, J.A.H.; Fenner, P.J.; Winkel, K.D. Phospholipase A2 in cnidaria. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **2004**, *139*, 731–735. [CrossRef] [PubMed]

56. Razpotnik, A.; Križaj, I.; Šribar, J.; Kordiš, D.; Maček, P.; Frangež, R.; Kem, W.R.; Turk, T. A new phospholipase A2 isolated from the sea anemone *Urticina crassicornis*—Its primary structure and phylogenetic classification. *FEBS J.* **2010**, *277*, 2641–2653. [CrossRef] [PubMed]

57. Hasegawa, Y.; Honma, T.; Nagai, H.; Ishida, M.; Nagashima, Y.; Shiomi, K. Isolation and cDNA cloning of a potassium channel peptide toxin from the sea anemone *Anemonia erythraea*. *Toxicon* **2006**, *48*, 536–542. [CrossRef] [PubMed]

58. Oliveira, J.S.; Fuentes-Silva, D.; King, G.F. Development of a rational nomenclature for naming peptide and protein toxins from sea anemones. *Toxicon* **2012**, *60*, 539–550. [CrossRef]

59. Ahmmed, M.K.; Bhowmik, S.; Giteru, S.G.; Zilani, N.H.; Adadi, P.; Islam, S.S.; Kanwugu, O.N.; Haq, M.; Ahmmed, F.; Ng, C.C.W.; et al. An Update of Lectins from Marine Organisms: Characterization, Extraction Methodology, and Potential Biofunctional Applications. *Mar. Drugs* **2022**, *20*, 430. [CrossRef]

60. Ojeda, N.; Salazar, C.; Cárdenas, C.; Marshall, S.H. Expression of DC-SIGN-like C-Type Lectin Receptors in Salmo salar. *Dev. Comp. Immunol.* **2020**, *113*, 103806. [CrossRef]

61. Schröder, H.C.; Ushijima, H.; Krasko, A.; Gamulin, V.; Thakur, N.L.; Diehl-Seifert, B.; Müller, I.M.; Müller, W.E.G. Emergence and Disappearance of an Immune Molecule, an Antimicrobial Lectin, in Basal Metazoa. *J. Biol. Chem.* **2003**, *278*, 32810–32817. [CrossRef]

62. Kim, C.-H. SARS-CoV-2 evolutionary adaptation toward host entry and recognition of receptor o-acetyl sialylation in virus–host interaction. *Int. J. Mol. Sci.* **2020**, *21*, 4549. [CrossRef]

63. Rahimi, N. C-Type lectin CD209L/L-SIGN and CD209/DC-SIGN: Cell adhesion molecules turned to pathogen recognition receptors. *Biology* **2020**, *10*, 1. [CrossRef]

64. El-Maradny, Y.A.; El-Fakharany, E.M.; Abu-Serie, M.M.; Hashish, M.H.; Selim, H.S. Lectins purified from medicinal and edible mushrooms: Insights into their antiviral activity against pathogenic viruses. *Int. J. Biol. Macromol.* **2021**, *179*, 239–258. [CrossRef]

65. Baumgartner, S.; Hofmann, K.; Bucher, P.; Chiquet-Ehrismann, R. The discoidin domain family revisited: New members from prokaryotes and a homology-based fold prediction. *Protein Sci.* **1998**, *7*, 1626–1631. [CrossRef] [PubMed]

66. Sarras, M.P., Jr.; Yan, L.; Leontovich, A.; Zhang, J.S. Structure, expression, and developmental function of early divergent forms of metalloproteinases in Hydra. *Cell Res.* **2002**, *12*, 163–176. [CrossRef]

67. Dumermuth, E.; Sterchi, E.; Jiang, W.; Wolz, R.; Bond, J.; Flannery, A.; Beynon, R. The astacin family of metalloendopeptidases. *J. Biol. Chem.* **1991**, *266*, 21381–21385. [CrossRef] [PubMed]

68. Reddi, A.H. BMP-1: Resurrection as Procollagen C-Proteinase. *Science* **1996**, *271*, 463. [CrossRef]
69. Trevisan-Silva, D.; Gremski, L.H.; Chaim, O.M.; da Silveira, R.B.; Meissner, G.O.; Mangili, O.C.; Barbaro, K.C.; Gremski, W.; Veiga, S.S.; Senff-Ribeiro, A. Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the brown spider (genus *Loxosceles*). *Biochimie* **2010**, *92*, 21–32. [CrossRef] [PubMed]
70. Möhrlen, F.; Maniura, M.; Plickert, G.; Frohme, M.; Frank, U. Evolution of astacin-like metalloproteases in animals and their function in development. *Evol. Dev.* **2006**, *8*, 223–231. [CrossRef]
71. Zadka, L.; Kulus, M.J.; Piatek, K. ADAM protein family—Its role in tumorigenesis, mechanisms of chemoresistance and potential as diagnostic and prognostic factors. *Neoplasma* **2018**, *65*, 823–839. [CrossRef]
72. Ponce, D.; Brinkman, D.L.; Potriquet, J.; Mulvenna, J. Tentacle Transcriptome and Venom Proteome of the Pacific Sea Nettle, *Chrysaora fuscescens* (Cnidaria: Scyphozoa). *Toxins* **2016**, *8*, 102. [CrossRef]
73. Neubauer, E.F.; Poole, A.Z.; Weis, V.M.; Davy, S.K. The scavenger receptor repertoire in six cnidarian species and its putative role in cnidarian-dinoflagellate symbiosis. *PeerJ* **2016**, *4*, e2692. [CrossRef]
74. McFarlane, I.D.; Graff, D.; Grimmelikhuijzen, C.J.P. Excitatory Actions of Antho-Rfamide, an Anthozoan Neuropeptide, on Muscles and Conducting Systems in the Sea Anemone *Calliactis parasitica*. *J. Exp. Biol.* **1987**, *133*, 157–168. [CrossRef]
75. Tremblay, M.; Henry, J.; Anctil, M. Spawning and gamete follicle rupture in the cnidarian *Renilla koellikeri*: Effects of putative neurohormones. *Gen. Comp. Endocrinol.* **2004**, *137*, 9–18. [CrossRef] [PubMed]
76. Takahashi, T. Comparative Aspects of Structure and Function of Cnidarian Neuropeptides. *Front. Endocrinol.* **2020**, *11*, 339. [CrossRef] [PubMed]
77. Katsukura, Y.; David, C.N.; Grimmelikhuijzen, C.J.P.; Sugiyama, T. Inhibition of metamorphosis by RFamide neuropeptides in planula larvae of Hydractinia echinata. *Dev. Genes Evol.* **2003**, *213*, 579–586. [CrossRef]
78. Leitz, T.; Morand, K.; Mann, M. Metamorphosin A: A Novel Peptide Controlling Development of the Lower Metazoan Hydractinia echinata (Coelenterata, Hydrozoa). *Dev. Biol.* **1994**, *163*, 440–446. [CrossRef] [PubMed]
79. Hayakawa, E.; Watanabe, H.; Menschaert, G.; Holstein, T.W.; Baggerman, G.; Schoofs, L. A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *PLoS ONE* **2019**, *14*, e0215185. [CrossRef]
80. Koch, T.L.; Grimmelikhuijzen, C.J.P. A comparative genomics study of neuropeptide genes in the cnidarian subclasses Hexacorallia and Ceriantharia. *BMC Genom.* **2020**, *21*, 666. [CrossRef]
81. Fricker, L.D. Analysis of mouse brain peptides using mass spectrometry-based peptidomics: Implications for novel functions ranging from non-classical neuropeptides to microproteins. *Mol. Biosyst.* **2010**, *6*, 1355–1365. [CrossRef]
82. Teixeira, C.M.M.; Correa, C.N.; Iwai, L.K.; Ferro, E.S.; de Castro, L.M. Characterization of Intracellular Peptides from Zebrafish (*Danio rerio*) Brain. *Zebrafish* **2019**, *16*, 240–251. [CrossRef]
83. Gelman, J.S.; Sironi, J.; Castro, L.M.; Ferro, E.S.; Fricker, L.D. Peptidomic Analysis of Human Cell Lines. *J. Proteome Res.* **2011**, *10*, 1583–1592. [CrossRef]
84. Dasgupta, S.; Yang, C.; Castro, L.M.; Tashima, A.K.; Ferro, E.S.; Moir, R.D.; Willis, I.M.; Fricker, L.D. Analysis of the Yeast Peptidome and Comparison with the Human Peptidome. *PLoS ONE* **2016**, *11*, e0163312. [CrossRef] [PubMed]
85. Kmiec, B.; Branca, R.M.; Berkowitz, O.; Li, L.; Wang, Y.; Murcha, M.W.; Whelan, J.; Lehtiö, J.; Glaser, E.; Teixeira, P.F. Accumulation of endogenous peptides triggers a pathogen stress response in *Arabidopsis thaliana*. *Plant J.* **2018**, *96*, 705–715. [CrossRef] [PubMed]
86. De Araujo, C.B.; Heimann, A.S.; Remer, R.A.; Russo, L.C.; Colquhoun, A.; Forti, F.L.; Ferro, E.S. Intracellular Peptides in Cell Biology and Pharmacology. *Biomolecules* **2019**, *9*, 150. [CrossRef]
87. Cunha, F.M.; Berti, D.A.; Ferreira, Z.S.; Klitzke, C.F.; Markus, R.P.; Ferro, E.S. Faculty Opinions recommendation of Intracellular peptides as natural regulators of cell signaling. *J. Biol. Chem.* **2008**, *283*, 24448–24459. [CrossRef]
88. Berti, D.A.; Russo, L.C.; Castro, L.M.; Cruz, L.; Gozzo, F.C.; Heimann, J.C.; Lima, F.B.; Oliveira, A.C.; Andreotti, S.; Prada, P.O.; et al. Identification of intracellular peptides in rat adipose tissue: Insights into insulin resistance. *Proteomics* **2012**, *12*, 2668–2681. [CrossRef] [PubMed]
89. Russo, L.C.; Asega, A.F.; Castro, L.M.; Negraes, P.D.; Cruz, L.; Gozzo, F.C.; Ulrich, H.; Camargo, A.C.; Rioli, V.; Ferro, E.S. Natural intracellular peptides can modulate the interactions of mouse brain proteins and thimet oligopeptidase with 14-3-3epsilon and calmodulin. *Proteomics* **2012**, *12*, 2641–2655. [CrossRef]
90. Monte-Silva, E.R.C.R.C.; Russo, L.C.; Castro, L.M.; Gozzo, F.C.; de Araujo, C.B.; Peron, J.P.S.; Rioli, V.; Ferro, E.S. EL28 is a novel intracellular peptide that activates immune proteasome and CD8+ T-cell response. *J. Proteom.* **2016**, *16*, S1874–S3919.
91. De Araujo, C.B.; Russo, L.C.; Castro, L.M.; Forti, F.L.; do Monte, E.R.; Rioli, V.; Gozzo, F.C.; Colquhoun, A.; Ferro, E.S. A novel intracellular peptide derived from g1/s cyclin d2 induces cell death. *J. Biol. Chem.* **2014**, *289*, 16711–16726. [CrossRef]
92. Russo, L.C.; Araujo, C.B.; Iwai, L.K.; Ferro, E.S.; Forti, F.L. A Cyclin D2-derived peptide acts on specific cell cycle phases by activating ERK1/2 to cause the death of breast cancer cells. *J. Proteom.* **2016**, *151*, 24–32. [CrossRef] [PubMed]
93. Patrzykat, A.; Zhang, L.; Mendoza, V.; Iwama, G.K.; Hancock, R.E.W. Synergy of Histone-Derived Peptides of Coho Salmon with Lysozyme and Flounder Pleurocidin. *Antimicrob. Agents Chemother.* **2001**, *45*, 1337–1342. [CrossRef]
94. Fernandes, J.M.O.; Molle, G.M.; Kemp, G.D.; Smith, V.J. Isolation and characterisation of oncorhyncin II, a histone H1-derived antimicrobial peptide from skin secretions of rainbow trout Oncorhynchus mykiss. *Dev. Comp. Immunol.* **2004**, *28*, 127–138. [CrossRef]

95. Smith, V.J.; Desbois, A.P.; Dyrynda, E.A. Conventional and Unconventional Antimicrobials from Fish, Marine Invertebrates and Micro-algae. *Mar. Drugs* **2010**, *8*, 1213–1262. [CrossRef] [PubMed]

96. Café-Mendes, C.; Ferro, E.; Torrão, A.; Crunfli, F.; Rioli, V.; Schmitt, A.; Falkai, P.; Britto, L.; Turck, C.; Martins-De-Souza, D. Peptidomic analysis of the anterior temporal lobe and corpus callosum from schizophrenia patients. *J. Proteom.* **2017**, *151*, 97–105. [CrossRef] [PubMed]

97. Correa, C.N.; Fiametti, L.O.; Mazzi Esquinca, M.E.; Mantovani de Castro, M. Sample Preparation and Relative Quantitation using Reductive Methylation of Amines for Peptidomics Studies. *J. Vis. Exp.* **2021**, *177*, e62971.

98. Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2337–2342. [CrossRef]

99. Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G.A.; Ma, B. PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification. *Mol. Cell. Proteom.* **2012**, *11*, M111.010587. [CrossRef]