RESEARCH ARTICLE

A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*

Eisuke Hayakawa[1,2]*, **Hiroshi Watanabe**[2,3], **Gerben Menschaert**[4], **Thomas W. Holstein**[3], **Geert Baggerman**[5,6], **Liliane Schoofs**[1]

1 Research Group of Functional Genomics and Proteomics, KU Leuven, Leuven, Belgium, **2** Evolutionary Neurobiology Unit, Okinawa Institute of Science & Technology, Okinawa, Japan, **3** Centre for Organismal Studies (COS), Heidelberg University, Heidelberg, Germany, **4** Faculty of Bioscience Engineering, Laboratory for Bioinformatics and Computational Genomics, Ghent University, Ghent, Belgium, **5** CFP/Ceproma, University Antwerpen, Antwerpen, Belgium, **6** VITO, Applied Bio & molecular Systems (ABS), Mol, Belgium

* eisuke.hayakawa@oist.jp

Abstract

Neuropeptides are a class of bioactive peptides shown to be involved in various physiological processes, including metabolism, development, and reproduction. Although neuropeptide candidates have been predicted from genomic and transcriptomic data, comprehensive characterization of neuropeptide repertoires remains a challenge owing to their small size and variable sequences. *De novo* prediction of neuropeptides from genome or transcriptome data is difficult and usually only efficient for those peptides that have identified orthologs in other animal species. Recent peptidomics technology has enabled systematic structural identification of neuropeptides by using the combination of liquid chromatography and tandem mass spectrometry. However, reliable identification of naturally occurring peptides using a conventional tandem mass spectrometry approach, scanning spectra against a protein database, remains difficult because a large search space must be scanned due to the absence of a cleavage enzyme specification. We developed a pipeline consisting of *in silico* prediction of candidate neuropeptides followed by peptide-spectrum matching. This approach enables highly sensitive and reliable neuropeptide identification, as the search space for peptide-spectrum matching is highly reduced. *Nematostella vectensis* is a basal eumetazoan with one of the most ancient nervous systems. We scanned the *Nematostella* protein database for sequences displaying structural hallmarks typical of eumetazoan neuropeptide precursors, including amino- and carboxyterminal motifs and associated modifications. Peptide-spectrum matching was performed against a dataset of peptides that are cleaved *in silico* from these putative peptide precursors. The dozens of newly identified neuropeptides display structural similarities to bilaterian neuropeptides including tachykinin, myoinhibitory peptide, and neuromedin-U/pyrokinin, suggesting these neuropeptides occurred in the eumetazoan ancestor of all animal species.

Introduction

Neuropeptides are a highly diverse group of messenger molecules involved in neurotransmission. They are essential for many physiological processes, such as muscle contraction, food digestion, growth, development, and reproduction, as well as more complex behaviours, such as adaptation, learning and memory, and ageing [1]. A neuropeptide is usually encoded in a larger neuropeptide precursor gene, which also encodes an N-terminal signal peptide. The precursor is translated on the rough endoplasmic reticulum, and the signal peptide is removed by a signal peptidase. Afterwards, processing enzymes in the immature secretory granules typically produce one or more mature neuropeptides by processing the protein precursor at cleavage sites. A number of cleavage enzymes, so called prohormone convertases of the furin/subtilisin family with specific recognition patterns, have been identified [2,3]. After cleavage, the processed peptides can undergo various posttranslational modifications (PTMs) [4]. Especially, the N-terminal and/or C-terminal residues are often modified. A frequently observed modification is amidation at the C-terminus of mature neuropeptides, which is usually required for their biological activity. C-terminal amidation involves the enzymatic transformation of a glycine into an alpha-amide by peptidylglycine alpha-amidating monooxygenase [5]. First, peptidylglycine is transformed into peptidyl-alpha-hydroxyglycine in the presence of copper, ascorbate, and molecular oxygen and subsequently converted to peptide alpha-amide and glyoxylate. C-terminal amidation protects the mature peptide from enzymatic degradation by carboxypeptidases. Pyroglutamic acid is a posttranslational modification that is widely observed at N-termini of various neuropeptides and protects the peptide chain from enzymatic degradation by aminopeptidases. In addition to PTMs, many bioactive peptides contain a proline residue in the second or third position from the N-terminus. This feature also protects the peptide from peptidase activity at the N-terminus [6]. All these structural characteristics are widely observed in neuropeptide sequences across the Animal Kingdom [7].

Recent advances in mass spectrometry and liquid chromatography technology have led to the establishment of peptidomics, an efficient technology that combines liquid chromatography with tandem mass spectrometry to identify neuropeptides. Tandem mass spectrometry combined with peptide-spectrum matching tools enables systematic neuropeptide identification [8–10]. Unlike neuropeptide prediction tools that are based on sequence similarities, the peptidomics approach provides evidence for the *in vivo* occurrence of the mature peptides, and also shows the eventual presence of the peptide's PTMs [11,12]. However, identification of naturally occurring processed peptides by means of conventional peptide-spectrum matching tools remains difficult. Unlike proteomics, in which proteins are identified based on *in vitro* generated enzymatic peptide digests, peptidomics uses naturally occurring peptides already cleaved by processing enzymes. Because cleavage sites in a protein precursor of naturally occurring neuropeptides cannot be predicted with high accuracy [13], peptide-spectrum matching in classical peptidomics technology has to be performed without any enzyme specification. This drawback in peptidomics leads to a huge search space and often results in poor identification confidence values. In addition, all possible PTMs have to be taken into consideration, which further increases the search space for peptide-spectrum matching [14].

A commonly employed approach to identify neuropeptides is based on sequence similarities and has allowed the *in silico* prediction of putative peptide signatures in sequenced genomes and transcriptomes of several animal species [15–18]. Multiple reports have successfully tracked down the evolutionarily conserved neuropeptides within a phylum or between closely related phyla, by means of sequence similarity-based searches against protein databases [19]. This approach is very useful for searching peptide sequences that have been evolutionarily conserved between closely related species. However, when peptides have become

evolutionarily diverged, such as evolutionarily ancient organisms, the sequence similarity-based prediction of new neuropeptides may not always be successful. Moreover, within a particular peptide sequence, only a short motif required for the peptide's biological activity is conserved during evolution [18,20], and the non-peptide-coding region of the precursor is in general not conserved. These issues hamper reliable sequence similarity-based peptide prediction. This is especially true for neuropeptides and corresponding genes, based on homology searches, among evolutionarily distant species, thus necessitating experimental validation. To address these hurdles, we have developed an alternative strategy to identify neuropeptides with high sensitivity and confidence. The strategy involves the construction of a significantly reduced dataset that only comprises mature peptide sequences, cleaved from their predicted protein precursors *in silico*. The latter are extracted from the much larger protein database based on their typical hallmarks (signal peptide, specific cleavage sites, PTMs). Peptide-spectrum matching is performed against the significantly reduced dataset of peptide sequences. The narrowed search space for peptide-spectrum matching enables highly sensitive peptide identification.

Using this approach, we here present a systematic identification of neuropeptides from *Nematostella vectensis*. This sea anemone has an urbilaterian origin and is thus one of the most evolutionarily ancient animals with a nervous system. Urbilaterian origins of neuropeptide signaling systems are still debated. There are four known non-bilaterian metazoan phyla: two phyla that have nervous systems, the Ctenophora (comb jellies) and Cnidaria (e.g., sea anemones and jelly fish); and two phyla that lack nervous systems, the Placozoa (e.g., *Trichoplax*) and the Porifera (sponges). A variety of bioactive neuropeptides have been identified in the cnidarians *Renilla köllikeri* (class Anthozoa) and *Hydra magnipapillata* (class Hydrozoa) [21–27]. The cnidarian genome of *Nematostella* predicts the presence of putative neuropeptide precursors with following C-terminal motifs: RIamides, Rpamides, Rwamides, Lwamides, Itamide, Mtamide, Vramide, Rramide, Pgamides, Rgamides, Pvamides, and LVamide [28,29]. However, none of these neuropeptides appear to be orthologues of bilaterian neuropeptides. This complicates our understanding of the evolution of neuropeptides. To get deeper insight into the puzzle of neuropeptide evolution, identification of neuropeptides in urbilaterian animals is not only fascinating, but of utmost importance. We successfully identified 20 neuropeptides in *Nematostella*, many of which have not been predicted or annotated as neuropeptides before [28].

Materials and methods

Experimental design

A comprehensive and highly sensitive neuropeptide identification pipeline was designed for *Nematostella vectensis*, using a combination of *in silico* neuropeptide prediction combined with tandem mass spectrometry (MS/MS). Fig 1 shows an overview of the approach used. First, the protein database of *Nematostella* was processed with a software tool to extract potential neuropeptide precursor sequences, based on their structural hallmarks. These include cnidarian-specific amino- and carboxyterminal cleavage motifs that flank the peptides sequences and their associated posttranslational modifications (PTMs). Fragmentation spectra were acquired by liquid chromatography-matrix-assisted laser desorption/ionization (LC-MALDI) MS/MS. Unlike conventional peptide-spectrum matching for naturally occurring peptides (as in peptidomics) or for enzymatic (tryptic) digests (as in proteomics), we narrowed the search space by using a smaller target dataset for peptide-spectrum matching. This smaller dataset only contains peptide sequences that are cleaved *in silico* from extracted putative peptide precursors from the *Nematostella* protein database. In contrast to searches in a conventional
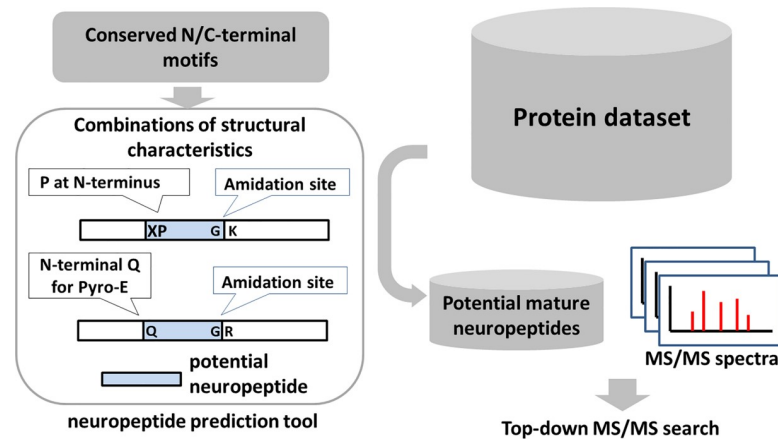
Fig 1. Schema of the neuropeptide identification strategy using a combination of peptide-spectrum matching against a dataset of *in silico* cleaved neuropeptide sequences extracted from putative neuropeptide precursors from the *Nematostella* protein database. Amino- and carboxyterminal motifs were used to scan the *Nematostella* protein database for neuropeptide precursor candidates, from which the peptides sequences were cleaved *in silico*. Peptide sequences were then exported into a target database for MS/MS spectral searching.

peptidomics workflow, which considers any amino acid residue as a potential cleavage site, the spectral matching search in the present method was performed directly against theoretical spectra of the mature forms of extracted neuropeptide sequences as in top-down proteomics.

Sample preparation

Adult polyps of *Nematostella vectensis* originally collected from the Rhode River in Maryland [30] were kept in 1/3 artificial seawater at 18˚C and fed three time per week with nauplius larvae of *Artemia salina*. The culture medium was changed once a week. Five adult polyps were homogenized with a probe sonicator in 1.5 mL methanol/water/formic acid (FA) solution (90:9:1) in 3 cycles (on for 5 s and off for 5 s) on ice. Large proteins were removed by centrifugation at $10,000 \times g$ for 15 min, and the supernatant was transferred to new tube. The sample was freeze-dried using a vacuum centrifuge (Speedvac concentrator SVC200H, Savant, USA) and stored at -80˚C for further treatment.

MALDI-MS/MS mass spectrometry

The sample was pre-fractionated on a C18 column (BEH C18 column, Waters, Milford, MA, USA) using mobile phase at high pH (MilliQ and acetonitrile with ammonium hydroxide (20 mM), pH 10) and five fractions were made during the gradient (B: 5 to 90% in 30 min) at a flow rate of 100 µL /min. Pre-fractionated samples were further separated on a C18 column at low pH. The eluent was (A) water containing 0.5% FA and (B) 90% acetonitrile in 0.5% aqueous formic acid. The column was first washed and equilibrated with eluent A and 5% of eluent B. After loading the sample, a linear gradient from 5% B to 60% B in 60 min at a flow rate of 100 µL /min was used as the mobile phase. Thirty fractions of 200 µL were collected from the beginning of the gradient using an automatic fraction collector. The resulting samples were dried in a Speedvac and stored at -80˚C until further analysis. Fractionated samples were resuspended in 1.5 µL water/ acetonitrile/FA (50/49.5/0.5 v/v). Subsequently, they were transferred onto a MALDI target plate (Bruker Daltonics, Bremen, Germany) and mixed with 1.5 mL of a saturated solution of CHCA in 50% acetonitrile containing 0.5% FA. After evaporation of the solvent, the MALDI target was introduced into the mass spectrometer ion source. Tandem mass spectrometry analysis was performed using the Ultraflex II instrument (Bruker

Daltonics, Bremen, Germany) in a positive ion, reflectron mode. The instrument was calibrated externally with a commercial peptide mixture (peptide calibration standard, Bruker Daltonics). All spectra were obtained using Flex Control software (Bruker Daltonics, Bremen, Germany). The plate was initially examined in MS1 mode and spectra were recorded within a mass range from m/z 500 to 4000. Subsequently, the peaks with S/N 10 were selected and used for the optimized LIFT method from the same target. All tandem mass spectra were processed by means of the FlexAnalysis software (Bruker Daltonics, Bremen, Germany), and m/z values and intensities of each peak were recorded in peak list files.

Mass spectrometry data analysis

An in-house software module was developed, enabling the prediction and extraction of potential neuropeptide sequences from *Nematostella* proteins as assembled from the genome at JGI (*Nematostella vectensis* version 1.0, all models) [31]. First, a protein was scanned for residues corresponding to the input C-terminal amino acid motif (e.g., GK or GR for C-terminal amidation). Second, an amino acid sequence corresponding to an N-terminal motif (e.g., XP or XXP; X is any amino acid)) was searched at the N-terminal side from the detected C-terminal motif within a given sequence length. In case multiple stretches matched N-terminal motif, they were considered potential N-termini of peptide sequences. All combinations of input amino acid motifs were applied. All peptide sequences that met the criteria were cleaved *in silico* and written in a FASTA formatted file that was subsequently used as a target database for peptide-spectrum matching.

To design the input motifs, we first assembled the sequences of all presently known cnidarian neuropeptides and their precursors from the UniProt sequence database and other public available resources (S1 Table) [23,25,32–39]. It is well known that bilaterian neuropeptide precursor proteins contain dibasic KR, KK, RK, and RR residues that flank the N- and C-termini of the peptides for cleavage of the peptides from their precursors [3]. We observed, however, that the already known cnidarian neuropeptide precursors rarely contain these dibasic residues (see S1 Table), which prompted us to search for cnidarian-specific input motifs. We observed that 52% of the cnidarian neuropeptide sequences had N-terminal prolines and 54% have N-terminal glutamines. In total, 84% had either an N-terminal proline or glutamine or both, indicating that these hallmarks can be used as input motifs to extract potential neuropeptide sequences from the *Nematostella* protein dataset. Currently known cnidarian neuropeptides also contain an amidation at their C-termini, which is a commonly observed PTM of neuropeptides across all animal species. Based on these observations, the following amino acid motifs were used to extract potential neuropeptide sequences from the *Nematostella* protein dataset:

C-terminal motif:

GK↓ or GR↓: glycine, required for amidation, before K or R

N-terminal motif:

↓XP: proline from the second position from N-terminus

↓XXP: proline from the third position from N-terminus

↓Q: glutamate for the formation of pyroglutamic acid

(↓ indicates the cleavage site)

C-terminal K or R of the extracted peptides were then removed and the saved sequences were stored as a FASTA formatted file.

For comparison, another set of potential neuropeptide sequences was created in a similar way, but now making use of the processing sites commonly found in bilaterian neuropeptide precursors, such as those for prohormone convertases (PC1-3) and for furin [40]. These include the dibasic amino acid sites (KK, RR, KR, RK), as well as sites containing monobasic amino acid residues separated by 2, 3, or 6 other residues (e.g. RXXR, KXXK) [2]. Extraction of

potential neuropeptide sequences was thus achieved by applying the following cleavage pattern according to Falth *et al.*: **(K/R) Xm(K/R)** $\downarrow X_k$**(K/R) Xn(K/R)** $\downarrow$. where m and n = 0, 2, 4, 6; X can be any amino acid residue and k = 6–40 [41]. Residues in bold were subsequently removed and the sequence X_k was stored and saved as FASTA formatted file. To create decoy sequences, the amino acid sequences of the original protein dataset were randomly rearranged and subjected to the aforementioned peptide prediction and extraction processes.

The FASTA files used in the present study is available online.

The software was coded in C++ using Visual Studio 2008 (Microsoft) and Boost library. It is available online (https://sourceforge.net/projects/enpg/) with the source code under the MIT license.

The obtained tandem mass spectra were searched against three datasets for comparison: (i) a database holding all *Nematostella* proteins (ii) a database containing predicted neuropeptide sequences by applying cnidarian neuropeptide precursor hallmarks and (iii) a database containing predicted neuropeptide sequence by applying common bilaterian neuropeptide precursor processing motifs (**(K/R) Xm(K/R)** $\downarrow X_k$**(K/R) Xn(K/R)** $\downarrow$). The Mascot search engine version 2.3 (Matrix science, London, UK) was used. Peptide-spectrum matching was performed, allowing C-terminal amidation of glycine extended peptides, pyroglutamic acid (Q and E), and oxidation (M), all being common posttranslational modifications of neuropeptides. Searches against the reduced dataset containing the predicted peptide sequences were performed with "No Cleavage" setting, in which cleavage of input sequences was not considered. Precursor masses were matched to the theoretical masses of intact peptide sequences. Mass tolerance for precursor and fragment ions were set to 0.4 and 0.8 Da. Peptides were first tentatively identified with a Mascot expect value less than 0.05; then, they were further confirmed by manual verification of the product ions assigned. Decoy search was conducted by using the aforementioned decoy FASTA file. The presence of an N-terminal signal peptide in the protein precursor of a mass spectrometry-identified peptide was examined using SignalP (version 5) [42]. The output files of SignalP can be found in the supporting materials.

Expression analysis

Partial ORF sequences for the identified peptides were obtained from the National Center for Biotechnology Information (NCBI) trace archive of *Nematostella vectensis* (data generated by the Joint Genome Institute) and from *Stellabase* (http://cnidarians.bu.edu/stellabase/index.cgi). Gene-specific primers were designed based on the ORF sequences. The primer sequences are available in S3 Table. PCR products for the neuropeptide-encoding genes were subcloned into pGEM-T (Promega) and sequenced. Riboprobes were synthesized and purified as described previously [43]. *In situ* hybridization was performed as previously described [44], with the following modifications: specimens were fixed with 4% paraformaldehyde/PBS + 0.1% Tween 20 (PBST) for 1 h, washed with methanol 3 times and stored at -20°C. Hybridization of 0.4- to 1.1-kb digoxygenin (DIG)-labeled antisense RNA probes was carried out using hybridization solution containing 1% SDS at 50–65°C for at least 22 h. For post-hybridization washes, specimens were washed by serial dilutions (75%, 50%, and 25%) of hybridization solution with 2× SSC at 55°C. After DIG-labeled probe was visualized using BM purple (Roche), specimens were washed with PBST.

Results

Peptide identification by MALDI MS/MS

The tandem mass spectra were acquired from off-line LC-MALDI MS/MS analysis of a peptide extract of *Nematostella vectensis*. Peptide identification was carried out using the Mascot

Table 1. Sequences of detected peptides and their Mascot E-values.

Peptide name	ID	Peptide sequence	Mascot E-value		
			Peptide database Cnidarian-specific motifs	Protein database No motif	Peptide database-bilaterian neuropeptide processing motif
HIRamide	1	APPLDLSGPAYFHIRa	**1.9E-02**	4.2E+00	nd
	2	GPPYIDLTEPSFFHIRa	**8.8E-09**	**2.1E-06**	nd
	3	NPPIDLGPAYFHIRa	**7.3E-07**	**1.6E-04**	**5.9E-07**
	4	pQPPIDLSPAAYFHIRa	**1.1E-04**	**2.5E-02**	**7.7E-05**
	5	pQPPLDLGPAYFHIRa	**5.0E-07**	**1.1E-04**	nd
	6	pQPPYLDLGEPSFFHIRa	**4.5E-04**	1.0E-01	**3.6E-04**
	7	pQPPYLDLTPAYFHIRa	**2.5E-04**	5.5E-02	nd
	8	pQPPYLDLTPSYFHIRa	**5.4E-06**	**1.2E-03**	nd
	9	pQPPMIDLSEPAFFHIRa	**1.2E-06**	**2.7E-04**	nd
	10	pQQPPMIDLSEPAFFHIRa	**1.7E-03**	3.5E-01	nd
PRGamide	11	GPRGGRATEFGPRGa	**1.2E-04**	**2.8E-02**	nd
	12	GPRGGREVNLEGPRGa	**6.7E-03**	1.7E+00	nd
QWamide	13	IPPQGFRFNQWa	**3.4E-02**	8.2E+00	nd
	14	IPPQGLRFNQWa	**2.9E-03**	6.7E-01	nd
	15	IPPQGLRFSQWa	**2.8E-03**	6.4E-01	nd
RFamide	16	MPEQDANPQTRFDa	**2.3E-05**	**5.3E-03**	nd
	17	pQGRFGREDQGRFa	**3.2E-03**	6.3E-01	nd
RPamide	18	FPPGFHRPa	**2.0E-03**	4.9E-01	nd
RHamide	19	GPPMIKIPVRHa	**2.0E-03**	5.0E-01	nd
	20	GPPMoIKIPVRHa	**7.7E-04**	2.5E+01	nd

E-values lower than the threshold (0.05) are indicated in bold. Mascot E-values in the first column result from peptide spectrum matching (PSM) against the smaller dataset of peptides extracted *in silico*. Mascot E-values in the second column result from PSM against the *Nematostella* protein database. Mascot e-values in the third column result from PSM against a database of peptide sequences that were extracted from the *Nematostella* protein database using the most common neuropeptide processing motif, which is based on the presence of dibasic cleavage sites as substrates for prohormone convertases and carboxypeptidase E in bilaterian neuropeptide precursors. C-terminal amidation, oxidation, and N-terminal pyroglutamic acid are indicated as "a", "o," and "p," respectively. PC: protein convertase; nd: not detected.

peptide-spectrum matching tool, considering pyroglutamic acid formation, C-terminal amidation of glycine extended peptides, and methionine oxidation. In order to reduce the search space for peptide spectrum matching, we employed a reduced dataset that only comprised potential neuropeptide sequences that were extracted and cleaved from the *Nematostella* protein dataset using cnidarian-specific amino acid motifs as input queries *in silico*. This way, 400 to 800 peptides were extracted, efficiently reducing the size of the search space. This dataset was then used as a target sequence dataset for peptide-spectrum matching, and the search was done in the top-down fashion with "no-cleavage" setting. Twenty unique peptides were identified with a Mascot E-value < 0.05. Table 1 shows the identified peptide sequences. The representative fragment spectra used for neuropeptide identifications are shown in Fig 2. The amino acid sequences and the details of their precursor protein coding genes can be found in Fig 3 and S2 Table.

We compared the performance of our peptide identification approach, using peptide-spectrum matching against a reduced cnidarian-dedicated dataset, with the conventional peptidomics approach that instead uses the entire protein dataset.

As shown in Table 1, the e-values in the spectral search against the reduced dataset of predicted *Nematostella* peptides (first column) significantly improved compared to those of the search against the entire *Nematostella* protein dataset (second column). Second, spectral searching against the reduced dataset yielded more than twice as many peptide identifications.
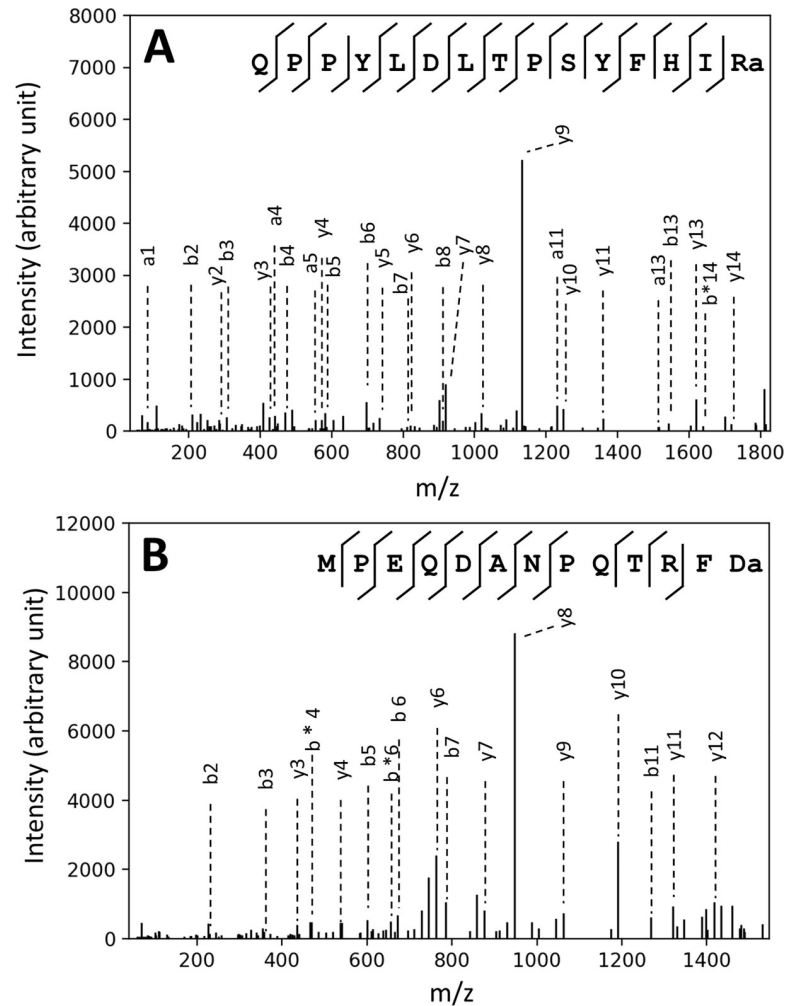
Fig 2. Representative fragment spectra of identified peptides. Fragmentation spectra of the peptide "QPPYLDLTPSYFHIRa" (A) and "MPEQDANPQTRFDa" (B). The dotted lines indicate fragment ions assigned. Ion labeled with * means loss of NH3.

Indeed, the E-values of the peptide hits in the second column (search against entire *Nematostella* protein dataset) were not only distributed in the higher range; many of them were also similar to the E-values of the decoy search result (S1 Fig), which evidently complicates the peptide identification process. In contrast, the E-values of the peptide hits in the first column (search against the reduced dataset of predicted candidate *Nematostella* peptide sequences) were distributed in a lower range, showing clear separation from the range of E-values of the decoy search result (S1 Fig).

By comparison, scanning the much larger search space of the entire *Nematostella* protein dataset, without any cleavage site specification (which is the conventional peptidomics search), revealed about 100,000 to 200,000 peptide sequences that matched the molecular masses of the mass spectrometry-detected peptides in the sample. In this third peptide spectrum matching search, we therefore used the most common cleavage rule of neuropeptide precursors, instead of the 'no cleavage setting'. This includes sites containing 2 basic amino acid residues (KK, RR, KR, RK), as well as sites containing pairs of basic amino acid residues separated by 2, 3, or 6 other residues (e.g. RXXR, KXXK), followed by removal of C-terminal basic residues [2].

HIRamide precursor

```
1
MRLYLFVPVF ALVLAVEGAS DEKRDSKQPP IDLSPAAYFH IRGKRTYNAP PLDLSGPAYF HIRGKRTAKQ PPYLDLGEPS
                                4 p        a          1         a            6 p
81
FFHIRGKRTE GPPYIDLTEP SFFHIRGKRS SEQPPLDLGP AYFHIRGKRT KNPPIDLGPA YFHIRGKRLS GEQPPYLDLT
    a          2          a        5 p          a           3           a          7 p
161
PAYFHIRGKR TQQPPMIDLS EPAFFHIRGR RAVEQPPYLD LTPSYFHIRG KRTEYPPFLE LGQPSYFHIR GRRAEKTTKD
    a          9 p         a          8 p          a
                10 p
```

PRGamide precursor

```
1
MAYFKWTLCA FVLAVLCLST NYVAGEKNEE KSEKSTLLQE LSKKDFDKEL AADEAIVLEL LKEKADDLDE ESLLDLLAPR
81
GGRDAPRGGR SLLDAPRGGR SLLDLLAAPR GGRDAPRGGR SLTELLNAPR RGRSLADLLD APRGGRSLAE LLDAPRGGRS
          a                     a                    a                              a
161
IEAPRGGRSL AELLEAPRGG RSLIELLEAP RGGRSVEAPR GGRSILELLT APRGGRSAPR GGRSAEKSDA IHKEKKAPRG
          a                    a                    a                    a             a
241
GRKRRSLPEE EMDGPRGGRS AVSGRSSEDG PGGRAVYGP GGRSFDGSR GGRSYDGPRG GRSMENGPRG GRATEFGPRG
                                          a                               11
321
GRSYEGSRGG REIDGPRGGR SFKEGPRGGR AVEGPRGGRD LYEEGPRGGR EVYLEGPRGG RDLVLEGPRG GRDLYEEGPR
                               a                     a                     a        12
401
GGREVNLEGP RGGRDLYEEG PRGGRSYDGP RGGRSAVYGP RGGRDANDGP RGGRDANDGP RGGRDANDGP RGGRDANEGP
         a                                                      a
481
RGGRDANEGP RGGRDANKGP  RGGRNTNTVG AQQADARQGQ RGGRSTGMGP RGGRSVSSEK DNRH
         a
```

QWamide precursor

```
1
MLKLITLAKS AQFKLTKRSR IYIKASFYNH AWCLHRLLWL SSYCRKSKFR QGLTDQPMDE SKEIPPQGLR FSQWGKKEIP
                                                                   15         a       13
81
PQGFRFNQWG KKEIPPQGLR FNQWGKKEIP PQGFRFSQWG KKEIPPQGFR FNQWGKKEIP PQGLRFNQWG KKEIPPQGLR
        a          14          a                   13          a          14         14
161
FNQWGKKEIP PQGLRFSQWG KREIPPQGLR FNQWGKKEIP PQGLRFNQWG KKEIPPQGLR FNQWGKKEIP PQGLRFNQWG
a       15          a          14          a        14          a        14          a
241
KKEIPPQGLR FNQWGKKEIP PQGLRFNQWG KKEIPPQGLR FSQWGKKEIP PQGLRFNQWG KKEIPPQGLR FNQWGKRKLI
         14          a        14          a        15          a        14          a
321
NQVMI
```

RFamide precursor

```
1
MAPRPGTLLL LGILIQVLIC TAKSTYKKEI ADLLDDNKDT PQFWKGRFSD PQFWKGRFSD PQFWKGRFSD PQFWKGRFSD
81
PQFWKGRFSD PQFWKGRFSD PQFWKGRFSD PQFWKGRFAD ELLNGGHHKH HHEEGEWKRT AGPGRFGRED QGRFGREDQG
                                                                              17 p
161
RFGREDQGRF GREDQGRFGR EDQGRFGRED QRFGREDQG RFGREDQGRF GREDQGRFGR EDQGRFGRED QGRFGREDQG
a       17 p          17 p          a       17 p          a        17 p          17 p
241
RFGREDQGRF GREDQGRFGR EDQGRFGRED QGRFGREDQG RFGREDQGRF GREDQGRFGR EDQGRFGRED QGRFGREDQG
a       17 p          17 p          a       17 p          17 p          a       17 p
321
RFGRNKIARV IDLDQGRFGR TMDTATKKDT VASMPEQDAN PQTRFDGKKR QAAKERSIEK KSTISSDAKA SDAKQS
a                                    16
```

RPamide precursor

```
1
MKVLCVTLVL MALVCSVQSR NLRHKDYERE LQDYRRDFAR ALSMVEQDAF LPKPRPGRRE QDSSNYEFPP GFHRPGKKRS
                                                                  18         a
81
FKEEK
```

RHamide precursor

```
1
MKITPVMSCL VLASALLVTA ECYRITDPGD LEKPQENDEP GPPMIKIPVR HGKREFNDDL EDTTLLHYPF RVARKRELFN
                                           19                 a
                                          20 o
81
AGGPPPIYLP VRPGKRGYYR DNIQETRQYL DPRPGKRSRN H
```

a Amidation
p Pyroglutamic acid
o Oxidation

Fig 3. Primary structures of neuropeptide precursor proteins. The location of the detected neuropeptides is indicated by full lines. Numbers correspond to the ID in Table 1. Predicted neuropeptides that were not detected in this study are indicated by dotted lines. Signal peptides predicted by SignalP are highlighted in red.

https://doi.org/10.1371/journal.pone.0215185.g003

In this approach proposed by Falth *et al*, the search space is also narrowed, preserving only potential neuropeptide structures that are cleaved through the most common cleavage patterns in bilaterian neuropeptide precursors. [41]. However, the third column in Table 1 shows that only 3 peptides could be identified this way. This poor peptide identification indicates that bilaterian-typical dibasic prohormone convertase cleavage sites are not commonly utilized for neuropeptide precursor processing in *Nematostella*, which is in accordance with neuropeptide precursor sequences identified in other cnidarian species [33,36,45–47].

Novel neuropeptides in *Nematostella vectensis*

Most of the peptides identified in the present study have never been reported so far. Their protein-coding genes have not been annotated as neuropeptide precursor genes, except for the RFamide precursor. Fig 3 shows the structures of the newly identified peptide precursors containing the mass spectrometry-identified peptides. All precursors start with a signal peptide at their N-terminus, typical for proteins destined for secretion, except for the QWamide precursor, which, however, contains several hydrophobic amino acid residues at the N-terminus. Overall, the identified peptide precursors contain multiple copies of the mature peptide sequences sharing the same C-terminal motifs. It should be noted that the precursors also contain multiple other potential neuropeptide sequences that display the same C-terminal motif as the identified peptides, as well as the structural motifs that we used for the *in silico* peptide extraction. For instance, the precursor protein of HIRamide peptides contains an additional putative peptide that shows the same HIRamide motif at C-terminus, together with the amidation motif and a proline at the second and third position from the N-terminus. It remains elusive why these peptides were not detected in our analysis. Either these peptides are not produced, or their presence was below the detection limit of the mass spectrometer.

The HIRamide peptide precursor contains 9 structurally related peptide sequences and a typical signal peptide sequence at the N-terminus. The HIRamide peptides share sequence similarities with arthropod tachykinin peptides at their C-termini (Fig 4A). Tachykinins and related peptides are well studied neuropeptides that have been evolutionarily conserved in both protostome (nematodes, arthropods, annelids, and mollusks) and deuterostome animals (echinodermates and chordates). Based on their C-termini, tachykinin and related peptides can be further classified into two subfamilies: the "-FXGLMa" and "-GFXGXRa" subfamilies. *Nematostella* HIR peptides display more similarities to the "GFXGXRa" subfamily, as indicated by the aromatic phenylalanine or tyrosine residues at the fifth position from the C-terminus and by the C-terminal arginine residue. Arthropod tachykinins and Nematostella HIRamides share the PXXFYXXRamide motif.

The RFamide neuropeptide family is widespread among bilaterian organisms and RFamides have previously been reported in cnidarians including the sea anemone. In this study, an additional member of the RFamide family, MPEQDANPQTRFDa, was identified within the RFamide precursor protein [29]. This peptide has an additional asparagine residue at the C-terminus, in contrast to other peptides contained in this precursor, which all display the carboxyterminal RFamide motif.

Two peptides, GPRGGRATEFGPRGamide and GPRGGREVNLEGPRG, share the C-terminal motif GPRGamide. The corresponding peptide precursor contains many peptide copies, which all share this motif. These cnidarian peptides seem to display sequence similarities with protostomian pyrokinins that are characterized by the C-terminal FXPRLamide sequence (Fig 4B) [65,66].

Three peptides derived from a single precursor protein display either the C-terminal sequence motif FSQWamide or FNQWamide (Fig 4C). This motif aligns with the motif that
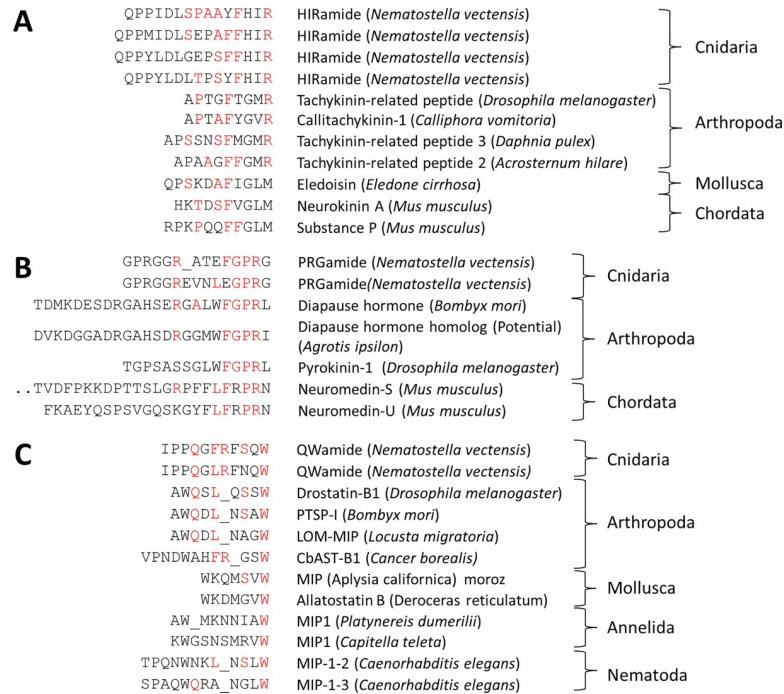
A

```
          QPPIDLSPAAYFHIR    HIRamide (Nematostella vectensis)
          QPPMIDLSEPAFFHIR   HIRamide (Nematostella vectensis)        ─┐
          QPPYLDLGEPSFFHIR   HIRamide (Nematostella vectensis)         ├ Cnidaria
          QPPYLDLTPSYFHIR    HIRamide (Nematostella vectensis)        ─┘
                APTGFTGMR    Tachykinin-related peptide (Drosophila melanogaster) ─┐
                APTAFYGVR    Callitachykinin-1 (Calliphora vomitoria)              ├ Arthropoda
                APSSNSFMGMR  Tachykinin-related peptide 3 (Daphnia pulex)          │
                APAAGFFGMR   Tachykinin-related peptide 2 (Acrosternum hilare)    ─┘
                QPSKDAFIGLM  Eledoisin (Eledone cirrhosa)             ─ Mollusca
                HKTDSFVGLM   Neurokinin A (Mus musculus)              ─ Chordata
                RPKPQQFFGLM  Substance P (Mus musculus)
```

B

```
      GPRGGR_ATEFGPRG        PRGamide (Nematostella vectensis)        ─┐
      GPRGGREVNLEGPRG        PRGamide(Nematostella vectensis)          ├ Cnidaria
   TDMKDESDRGAHSERGALWFGPRL  Diapause hormone (Bombyx mori)          ─┐
                                                                       │
   DVKDGGADRGAHSDRGGMWFGPRI  Diapause hormone homolog (Potential)      ├ Arthropoda
                             (Agrotis ipsilon)                         │
        TGPSASSGLWFGPRL      Pyrokinin-1 (Drosophila melanogaster)   ─┘
  ..TVDFPKKDPTTSLGRPFFLFRPRN Neuromedin-S (Mus musculus)             ─┐
    FKAEYQSPSVGQSKGYFLFRPRN  Neuromedin-U (Mus musculus)              ─┘ Chordata
```

C

```
          IPPQGFRFSQW        QWamide (Nematostella vectensis)         ─┐
          IPPQGLRFNQW        QWamide (Nematostella vectensis)          ─ Cnidaria
          AWQSL_QSSW         Drostatin-B1 (Drosophila melanogaster)  ─┐
          AWQDL_NSAW         PTSP-I (Bombyx mori)                      ├ Arthropoda
          AWQDL_NAGW         LOM-MIP (Locusta migratoria)              │
        VPNDWAHFR_GSW        CbAST-B1 (Cancer borealis)              ─┘
            WKQMSVW           MIP (Aplysia californica) moroz         ─┐
            WKDMGVW           Allatostatin B (Deroceras reticulatum)   ─ Mollusca
          AW_MKNNIAW         MIP1 (Platynereis dumerilii)            ─ Annelida
          KWGSNSMRVW         MIP1 (Capitella teleta)
        TPQNWNKL_NSLW        MIP-1-2 (Caenorhabditis elegans)        ─┐
        SPAQWQRA_NGLW        MIP-1-3 (Caenorhabditis elegans)         ─┘ Nematoda
```

Fig 4. Structural similarities of identified neuropeptides in *Nematostella vectensis* and other species. A: HIRamides and Tachykinin related peptides[18,48–52]. B: PRGamides and PRXamide related peptides [53–57]. C: QWamides, myoinhibitory peptide (MIP) and allatostatin type B [58–64]. Conserved amino acid residues are shown in red.

https://doi.org/10.1371/journal.pone.0215185.g004

typifies bilaterian myoinhibitory peptide/allatostatin type B peptide family, which is a neuro-peptide family widespread among protostomes with various functions such as inhibitory effect on muscles, juvenile hormone synthesis, mating behavior, sleep and learning. [60,67–70].

In addition to the peptides described above, we discovered two groups of peptides (RPa-mide and RHamide) that originated from two precursor proteins. So far, we have not found any sequence similarities to known neuropeptide families (Table 1).

Expression patterns of `HIRamide`, `PRGamide`, and `VRHamide` neuropeptides

Our whole mounts in situ hybridization (WISH) analyses confirmed that the peptide genes encoding the identified peptides `HIRamide`, `PRGamide`, and `VRHamide` are exclusively expressed in neurons. As shown in Fig 5, the expression of all peptide genes as detected by WISH was observed in specific cells housed mainly in the endodermal layer at the juvenile polyp stages. Careful observation of the morphology of the positively stained cells showed round-shaped cell bodies with neurite-like processes (Fig 5, lower panel). This indicates that the genes encoding the detected peptides are expressed in neurons, and that the newly identi-fied peptides are true neuropeptides. The genes encoding HIRamide and PRGamide neuro-peptides were strongly expressed in neuronal subsets around the mouth opening (Fig 5A), indicating that these neurons develop at the oral side to form region-specific neural network. The tissue around the mouth of cnidarian polyps has been shown to express a number of neu-ronal markers including RFamides, which are evolutionarily conserved in metazoans (Fig 5, upper panel) [28], and to develop a regionalized nervous system, which is known as the oral nervous system (*Nematostella*) or the nerve ring (*Hydrozoa*) [71–73]. The expression pattern
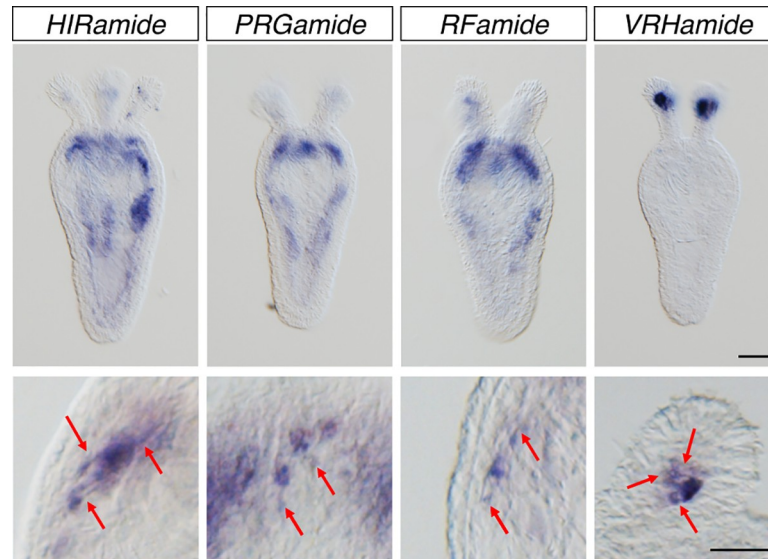
Fig 5. WISH staining of juvenile polyps of *Nematostella vectensis* (8 days post fertilization). The figure shows localized expression of *HIRamide*, *PRGamide RFamide*, and *VRHamide* genes at low (upper panels) and high magnification (lower panels). Scale bars, 100 μm (upper) and 50 μm (lower). Neural processes are indicated by red arrows.

https://doi.org/10.1371/journal.pone.0215185.g005

of the VRHamide-encoding gene was in sharp contrast with that of HIRamide, PRGamide, and RFamide-encoding genes. Expression of this gene could not be detected at the oral tissue. Instead, VRHamides were strongly and exclusively detected in neurons located at the most distal region of the tentacle endoderm (Fig 5, upper panel). This unexpected and interesting expression pattern suggests a specific function of VRHamides in development and/or in contractility of the tentacles.

Discussion

We studied the neuropeptidome of *Nematostella vectensis*, a sea anemone that belongs to the Cnidaria phylum, possessing one of most basal nervous systems in the Animal Kingdom. Recent advances in mass spectrometry technologies have greatly enhanced the identification of neuropeptidomes from biological matrices [14,74–76]. However, an effective bioinformatics solution for neuropeptide identification is still missing and leaves the interpretation of tandem mass spectral data difficult. We overcame this difficulty by performing peptide spectrum matching against a narrowed database comprising *in silico* cleaved putative neuropeptide sequences. We successfully identified 20 peptides encoded in six neuropeptide precursor genes from *Nematostella* and experimentally confirmed that at least four precursor genes are expressed in neurons exhibiting distinct neurophysiological activities. Our approach can be easily incorporated with recent advanced technologies in peptidomics, and such a combination will greatly improve the analysis of the neuropeptidome in a broad range of biological samples.

Strategies for the improvement of peptide identification

Peptide-spectrum matching has become the main tool for identifying naturally occurring peptides. Unlike classical methods, such as Edman degradation, this approach takes advantage of the high sensitivity of modern mass spectrometry instruments and high throughput peptide search engines, therefore enabling comprehensive neuropeptide characterization. One of the

main difficulties in peptidomics is that the peptide-spectrum matching tool needs to used with the "no enzyme" setting, which results in an enormous search space. Consequently, the peptide-spectrum matching result is usually worse in comparison with a peptide search in a proteomics setting.

The peptide characterization strategy employed in this study is a combination of predicting potential neuropeptide sequences and peptide-spectrum matching. This approach is effective in narrowing the search space and thereby significantly increases the sensitivity of peptide identification. Fälth *et al.* employed a similar approach to narrow search space in order to improve peptide-spectrum matching [41]. They used the most common processing rule of neuropeptide precursors to extract potential neuropeptide sequences from a large protein database, which significantly improved peptide identification rate. We tested their approach by extracting peptide sequences from the *Nematostella* protein dataset using the most common cleavage sites containing basic residues, but failed to identify the majority of *Nematostella* neuropeptides identified by the present strategy. This is explained by the fact that the most common basic amino acid cleavage site in bilaterian neuropeptide precursors seems to be rare in cnidarian (*Nematostella*) neuropeptide precursors.

Southey *et al.* developed a dedicated tool to predict basic cleavage sites in neuropeptide precursor genes [77]. The prediction engine, NEUROPRED, needs to be trained with the information of known neuropeptides and their precursors; hence many identified peptides and information of processing sites are required. Compared to the prediction with simple motif matching, the prediction by NEUROPRED using logistic regression modeling is more reliable, but, on the other hand, rather strict and potentially discards many forms of cleaved peptides, especially if the large amount of training data is not available. These approaches are very effective for the prediction of neuropeptides in bilaterian phyla where a lot of sequence information is already available. In contrast, the number of known cnidarian neuropeptides is limited compared to the massive amounts of peptide sequences in bilaterian phyla. Cnidarian neuropeptides display, however, very clear N/C-terminal motifs. The limited number of peptide sequences and highly conserved motifs in cnidarian neuropeptides make regression modeling unnecessary and less effective for peptide identification in this animal group.

Unlike bilaterian animals, non-basic cleavage sites are commonly observed in cnidarian neuropeptide precursors [33,34,37,39,45–47,78]. In fact, most of the peptides identified in the present study were generated via cleavages at non-basic residues. Therefore, the approach we employed in the present study (i.e., the prediction of putative neuropeptide sequences based on other structural hallmarks, such as the N-terminal pyroglutamation or Proline at the second or third position from the N-terminus and the C-terminal amidation) is more comprehensive and effective to preserve as many potential neuropeptide sequences as possible in the target database for peptide-spectrum matching. Nevertheless, the approach also has a drawback in that it only allows identification of neuropeptides that display these specific features. Non-amidated peptides or peptides that are not N-terminally protected are not contained in this reduced database.

Noteworthy, prohormone convertases distinct from the furin/subtilisin family likely exist in other bilaterians as has been suggested in several peptidomic and non-peptidomic studies. Hence, our method may eventually reveal novel peptides in bilaterian species [13,79–83].

Neuropeptide evolution

Many pioneer studies in the late 80s and early 90s already pointed to homologies between protostomian and deuterostomian neuropeptides based on neuropeptide sequence identifications in vertebrate, insect, molluscan, and cnidarian species [20,36,84–88]. However, until today,

knowledge on signaling molecules, including neuropeptides, in the more ancient eumetazoan animals remained scarce. A previous study showed a highly conserved set of genes in the *Nematostella vectensis* genome, including molecules involved in neurotransmission [28,89]. Anctil *et al*. has reported potential neuropeptide coding genes by means of homology searching [29]. Although these predicted precursor proteins contain short repeats of neuropeptide-like motifs, it is difficult to prove the peptide identities as the corresponding precursors were predicted solely based on their protein sequence without any experimental evidence that the peptides are processed. Some of the neuropeptide precursor proteins predicted by Anctil showed a significant degree of structural similarities to non-peptide precursors, leaving doubts on their identity as true neuropeptide precursors. In the present study, we were unable to identify any of those predicted peptides, neither by conventional "no-enzyme" searches against the entire *Nematostella* protein database, nor by searches against the reduced dataset of predicted peptides. Homology-based searches based on short neuropeptide sequences remain a difficult task. More importantly, the homology-based prediction of neuropeptide precursor genes does not provide evidence for the occurrence of the predicted neuropeptides *in vivo*. Therefore, peptide identification approaches that are based on conventional biochemical purification methods and powered by upcoming bioinformatic technologies are necessary to empirically prove the predicted peptides. They will be important to identify not only evolutionary conserved peptides, but also evolutionarily derived or species-specific peptides that correspond to the diverged physiological traits.

Our data have also important implications on our insight into the evolution of neuropeptides. Since the neuropeptide repertoire in ancient eumetazoans is largely unknown, the origin of most of neuropeptides found in bilaterian animals remains difficult to reveal. Jekely (2013) performed a similarity-based clustering analysis of genes encoding neuropeptides and neuropeptide GPCRs across metazoan phyla and concluded that the last common ancestor of eumetazoans had various small amidated peptides including RFamide, RYamide, and Wamide) [90]. Ancestral bilaterian neuropeptide-receptor families include GnRH, vasopressin, GnIH/SIFamide, CRF/diuretic hormone, calcitonin/DH31, NPY/NPF, neuromedin-U/pyrokinin, CCK/sulfakinin, galanin/allatostatin-A, and orexin/allatotropin [90,91]. It has been suggested that these neuropeptide families may have originated concomitantly with the origin of a complex bilaterian body plan, with control of food intake and digestion, excretory and circulatory systems, light-controlled reproduction, a centralized nervous system, complex reproductive behavior, and learning.

Thus far, the urbilaterian origin of neuropeptide signaling pathways remains largely elusive. Nevertheless, the present study strongly suggests that the extent of the conservation of some neuropeptide families may even be deeper than previously proposed. Indeed, we were able to identify neuropeptides structurally related to the myoinhibitory peptide/allatostatin type B, tachykinin, and neuromedin-U/pyrokinin families in the cnidarian species, *Nematostella*, suggesting that these neuropeptide families were already present in the common ancestors of all eumetazoan species. Neuropeptides modulate various biological processes by signaling through G protein coupled receptors (GPCRs). The *Nematostella* genome contains at least 79 GPCR coding genes structurally related to known neuropeptide receptors in bilaterians [29]. However, at this moment, in any of the four non-bilaterian phyla–Porifera, Placozoa, Ctenophora and Cnidaria–functional studies showing which neuropeptides signal through which GPCRs are still lacking, but at least they are now on the horizon for cnidarians.

Taken together, there are still many missing pieces in the 'jigsaw puzzle' of neuropeptide evolution, but we anticipate that the discovery of the receptors, downstream targets, and functions of cnidarian representatives of ancient eumetazoan neuropeptide families will provide important new insights into the evolution of neuropeptide functions in the Animal Kingdom.

Supporting information

S1 File. Output files of SignalP. Original output files of SignalP analysis on identified neuro-peptide precursor proteins.
(ZIP)

S1 Fig. The distribution of e-values of PSMs. Each circle indicates the e-value of the top-scoring hit in the dataset. A: Search against the predicted neuropeptide dataset. B: decoy of the predicted neuropeptides dataset. C: the whole protein models. D: decoy of the whole protein models.
(DOCX)

S1 Table. The structures of the previously identified cnidarian amidated neuropeptides.
(TSV)

S2 Table. Details of the neuropeptide precursor genes.
(XLSX)

S3 Table. Primer and cDNA sequences of neuropeptides.
(XLSX)

Author Contributions

Conceptualization: Eisuke Hayakawa, Hiroshi Watanabe, Gerben Menschaert, Geert Baggerman, Liliane Schoofs.

Data curation: Gerben Menschaert.

Funding acquisition: Thomas W. Holstein, Liliane Schoofs.

Investigation: Eisuke Hayakawa, Hiroshi Watanabe, Geert Baggerman.

Methodology: Eisuke Hayakawa, Gerben Menschaert, Thomas W. Holstein.

Project administration: Eisuke Hayakawa, Thomas W. Holstein.

Resources: Liliane Schoofs.

Software: Eisuke Hayakawa.

Supervision: Eisuke Hayakawa, Thomas W. Holstein, Geert Baggerman, Liliane Schoofs.

Validation: Eisuke Hayakawa, Hiroshi Watanabe, Liliane Schoofs.

Visualization: Hiroshi Watanabe.

Writing – original draft: Eisuke Hayakawa, Hiroshi Watanabe, Gerben Menschaert, Thomas W. Holstein, Geert Baggerman, Liliane Schoofs.

Writing – review & editing: Eisuke Hayakawa, Hiroshi Watanabe, Gerben Menschaert, Thomas W. Holstein, Geert Baggerman, Liliane Schoofs.

References

1. Van Bael S, Watteyne J, Boonen K, De Haes W, Menschaert G, Ringstad N, et al. Mass spectrometric evidence for neuropeptide-amidating enzymes in *Caenorhabditis elegans*. J Biol Chem. 2018; 293: 6052–6063. https://doi.org/10.1074/jbc.RA117.000731 PMID: 29487130

2. Fricker LD. Neuropeptide-processing enzymes: applications for drug discovery. AAPS J. 2005; 7: E449–55. https://doi.org/10.1208/aapsj070244 PMID: 16353923

3. Hook V, Funkelstein L, Lu D, Bark S, Wegrzyn J, Hwang S-R. Proteases for processing proneuropep-tides into peptide neurotransmitters and hormones. Annu Rev Pharmacol Toxicol. 2008; 48: 393–423. https://doi.org/10.1146/annurev.pharmtox.48.113006.094812 PMID: 18184105

4. Hökfelt T, Broberger C, Xu ZQ, Sergeyev V, Ubink R, Diez M. Neuropeptides—an overview. Neuro-pharmacology. 2000; 39: 1337–56. https://doi.org/10.1016/s0028-3908(00)00010-1 PMID: 10818251

5. Emeson RB. Hypothalamic peptidyl-glycine alpha-amidating monooxygenase: preliminary characteri-zation. J Neurosci. 1984; 4: 2604–13. https://doi.org/10.1523/JNEUROSCI.04-10-02604.1984 PMID: 6491725

6. Harada M, Fukasawa KM, Fukasawa K, Nagatsu T. Inhibitory action of proline-containing peptides on Xaa-Pro-dipeptidylaminopeptidase. Biochim Biophys Acta. 1982; 705: 288–90. https://doi.org/10.1016/0167-4838(82)90191-1 PMID: 6126216

7. Conlon JM, Larhammar D. The evolution of neuroendocrine peptides. Gen Comp Endocrinol. 2005; 142: 53–9. https://doi.org/10.1016/j.ygcen.2004.11.016 PMID: 15862548

8. Baggerman G, Cerstiaens A, De Loof A, Schoofs L. Peptidomics of the larval *Drosophila melanogaster* central nervous system. J Biol Chem. 2002; 277: 40368–74. https://doi.org/10.1074/jbc.M206257200 PMID: 12171930

9. Clynen E, Baggerman G, Veelaert D, Cerstiaens A, Van der Horst D, Harthoorn L, et al. Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, *Locusta migratoria*. Eur J Biochem. 2001; 268: 1929–39. https://doi.org/10.1046/j.1432-1327.2001.02067.x PMID: 11277915

10. Fricker LD, Lim J, Pan H, Che F-Y. Peptidomics: identification and quantification of endogenous pep-tides in neuroendocrine tissues. Mass Spectrom Rev. 2006; 25: 327–44. https://doi.org/10.1002/mas.20079 PMID: 16404746

11. De Haes W, Van Sinay E, Detienne G, Temmerman L, Schoofs L, Boonen K. Functional neuropeptido-mics in invertebrates. Biochim Biophys Acta. 2014; https://doi.org/10.1016/j.bbapap.2014.12.011 PMID: 25528324

12. Dallas DC, Guerrero A, Parker EA, Robinson RC, Gan J, German JB, et al. Current peptidomics: Appli-cations, purification, identification, quantification and functional analysis. Proteomics. 2014; https://doi.org/10.1002/pmic.201400310 PMID: 25429922

13. Pauls D, Chen J, Reiher W, Vanselow JT, Schlosser A, Kahnt J, et al. Peptidomics and processing of regulatory peptides in the fruit fly *Drosophila*. EuPA Open Proteomics. 2014; 3: 114–127. https://doi.org/10.1016/j.euprot.2014.02.007

14. Hayakawa E, Menschaert G, De Bock P-J, Luyten W, Gevaert K, Baggerman G, et al. Improving the identification rate of endogenous peptides using electron transfer dissociation and collision-induced dis-sociation. J Proteome Res. 2013; 12: 5410–21. https://doi.org/10.1021/pr400446z PMID: 24032530

15. Hummon AB, Richmond TA, Verleyen P, Baggerman G, Huybrechts J, Ewing MA, et al. From the genome to the proteome: uncovering peptides in the *Apis* brain. Science. 2006; 314: 647–9. https://doi.org/10.1126/science.1124128 PMID: 17068263

16. Christie AE. Prediction of the peptidomes of *Tigriopus californicus* and *Lepeophtheirus salmonis* (Cope-poda, Crustacea). Gen Comp Endocrinol. 2014; 201: 87–106. https://doi.org/10.1016/j.ygcen.2014.02.015 PMID: 24613138

17. Hauser F, Neupert S, Williamson M, Predel R, Tanaka Y, Grimmelikhuijzen CJP. Genomics and pepti-domics of neuropeptides and protein hormones present in the parasitic wasp *Nasonia vitripennis*. J Pro-teome Res. 2010; 9: 5296–310. https://doi.org/10.1021/pr100570j PMID: 20695486

18. Dircksen H, Neupert S, Predel R, Verleyen P, Huybrechts J, Strauss J, et al. Genomics, transcrip-tomics, and peptidomics of *Daphnia pulex* neuropeptides and protein hormones. J Proteome Res. 2011; 10: 4478–504. https://doi.org/10.1021/pr200284e PMID: 21830762

19. Jékely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. Proc Natl Acad Sci U S A. 2013; 110: 8702–7. https://doi.org/10.1073/pnas.1221833110 PMID: 23637342

20. De Loof A, Schoofs L. Homologies between the amino acid sequences of some vertebrate peptide hor-mones and peptides isolated from invertebrate sources. Comp Biochem Physiol B. 1990; 95: 459–68. https://doi.org/10.1016/0305-0491(90)90003-c PMID: 2184989

21. Anctil M, Grimmelikhuijzen CJP. Excitatory action of the native neuropertide Antho-RFamide on mus-cles in the pennatulid *Renilla köllikeri*. Gen Pharmacol Vasc Syst. 1989; 20: 381–384. https://doi.org/10.1016/0306-3623(89)90277-2

22. Fujisawa T, Hayakawa E. Peptide signaling in *Hydra*. Int J Dev Biol. 2012; 56: 543–50. https://doi.org/10.1387/ijdb.113477tf PMID: 22689368

23. Grimmelikhuijzen CJ, Graff D. Isolation of pyroGlu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide from sea anemones. Proc Natl Acad Sci U S A. 1986; 83: 9817–21. https://doi.org/10.1073/pnas.83.24.9817 PMID: 2879288

24. McFarlane ID, Anderson PA, Grimmelikhuijzen CJ. Effects of three anthozoan neuropeptides, Antho-RWamide I, Antho-RWamide II and Antho-RFamide, on slow muscles from sea anemones. J Exp Biol. 1991; 156: 419–31. Available: http://www.ncbi.nlm.nih.gov/pubmed/1675657 PMID: 1675657

25. Takahashi T, Muneoka Y, Lohmann J, Lopez de Haro MS, Solleder G, Bosch TC, et al. Systematic isolation of peptide signal molecules regulating development in hydra: LWamide and PW families. Proc Natl Acad Sci U S A. 1997; 94: 1241–6. https://doi.org/10.1073/pnas.94.4.1241 PMID: 9037037

26. Takahashi T, Hayakawa E, Koizumi O, Fujisawa T. Neuropeptides and their functions in *Hydra*. Acta Biol Hung. 2008; 59 Suppl: 227–35. https://doi.org/10.1556/ABiol.59.2008.Suppl.32 PMID: 18652396

27. Yum S, Takahashi T, Koizumi O, Ariura Y, Kobayakawa Y, Mohri S, et al. A novel neuropeptide, Hym-176, induces contraction of the ectodermal muscle in *Hydra*. Biochem Biophys Res Commun. 1998; 248: 584–90. https://doi.org/10.1006/bbrc.1998.8831 PMID: 9703970

28. Watanabe H, Fujisawa T, Holstein TW. Cnidarians and the evolutionary origin of the nervous system. Dev Growth Differ. 2009; 51: 167–83. https://doi.org/10.1111/j.1440-169X.2009.01103.x PMID: 19379274

29. Anctil M. Chemical transmission in the sea anemone *Nematostella vectensis*: A genomic perspective. Comp Biochem Physiol Part D Genomics Proteomics. Elsevier Inc.; 2009; 4: 268–89. https://doi.org/10.1016/j.cbd.2009.07.001 PMID: 20403752

30. Hand C, Uhlinger KR. The culture, sexual and asexual reproduction, and growth of the sea anemone *Nematostella vectensis*. Biol Bull. 1992; 182: 169–176. https://doi.org/10.2307/1542110 PMID: 29303672

31. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science. 2007; 317: 86–94. https://doi.org/10.1126/science.1139158 PMID: 17615350

32. Grimmelikhuijzen CJ, Rinehart KL, Spencer AN. Isolation of the neuropeptide less than Glu-Trp-Leu-Lys-Gly-Arg-Phe-NH2 (Pol-RFamide II) from the hydromedusa *Polyorchis penicillatus*. Biochem Biophys Res Commun. 1992; 183: 375–82. https://doi.org/10.1016/0006-291x(92)90491-3 PMID: 1550547

33. Darmer D, Schmutzler C, Diekhoff D, Grimmelikhuijzen CJ. Primary structure of the precursor for the sea anemone neuropeptide Antho-RFamide (less than Glu-Gly-Arg-Phe-NH2). Proc Natl Acad Sci U S A. 1991; 88: 2555–9. https://doi.org/10.1073/pnas.88.6.2555 PMID: 1706527

34. Leviev I, Grimmelikhuijzen CJ. Molecular cloning of a preprohormone from sea anemones containing numerous copies of a metamorphosis-inducing neuropeptide: a likely role for dipeptidyl aminopeptidase in neuropeptide precursor processing. Proc Natl Acad Sci U S A. 1995; 92: 11647–51. https://doi.org/10.1073/pnas.92.25.11647 PMID: 8524821

35. Moosler A, Rinehart KL, Grimmelikhuijzen CJ. Isolation of four novel neuropeptides, the hydra-RFamides I-IV, from *Hydra magnipapillata*. Biochem Biophys Res Commun. 1996; 229: 596–602. https://doi.org/10.1006/bbrc.1996.1849 PMID: 8954943

36. Darmer D, Hauser F, Nothacker HP, Bosch TC, Williamson M, Grimmelikhuijzen CJ. Three different prohormones yield a variety of Hydra-RFamide (Arg-Phe-NH2) neuropeptides in *Hydra magnipapillata*. Biochem J. 1998; 332 (Pt 2: 403–12. https://doi.org/10.1042/bj3320403 PMID: 9601069

37. Hayakawa E, Takahashi T, Nishimiya-Fujisawa C, Fujisawa T. A novel neuropeptide (FRamide) family identified by a peptidomic approach in *Hydra magnipapillata*. FEBS J. 2007; 274: 5438–48. https://doi.org/10.1111/j.1742-4658.2007.06071.x PMID: 17894820

38. Gajewski M, Leitz T, Schloßherr J, Plickert G. LWamides from Cnidaria constitute a novel family of neuropeptides with morphogenetic activity. Roux's Arch Dev Biol. 1996; 205: 232–242. https://doi.org/10.1007/BF00365801 PMID: 28306026

39. Takahashi T, Koizumi O, Ariura Y, Romanovitch A, Bosch TC, Kobayakawa Y, et al. A novel neuropeptide, Hym-355, positively regulates neuron differentiation in *Hydra*. Development. 2000; 127: 997–1005. https://doi.org/10.1016/S1095-6433(99)90367-7 PMID: 10662639

40. Zhou A, Webb G, Zhu X, Steiner DF. Proteolytic processing in the secretory pathway. J Biol Chem. 1999; 274: 20745–8. https://doi.org/10.1074/jbc.274.30.20745 PMID: 10409610

41. Fälth M, Sköld K, Svensson M, Nilsson A, Fenyö D, Andren PE. Neuropeptidomics strategies for specific and sensitive identification of endogenous peptides. Mol Cell Proteomics. 2007; 6: 1188–97. https://doi.org/10.1074/mcp.M700016-MCP200 PMID: 17401030

42. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011; 8: 785–6. https://doi.org/10.1038/nmeth.1701 PMID: 21959131

43. Lengfeld T, Watanabe H, Simakov O, Lindgens D, Gee L, Law L, et al. Multiple Wnts are involved in *Hydra* organizer formation and regeneration. Dev Biol. 2009; 330: 186–99. https://doi.org/10.1016/j.ydbio.2009.02.004 PMID: 19217898

44. Kusserow A, Pang K, Sturm C, Hrouda M, Lentfer J, Schmidt HA, et al. Unexpected complexity of the Wnt gene family in a sea anemone. Nature. 2005; 433: 156–60. https://doi.org/10.1038/nature03158 PMID: 15650739

45. Yum S, Takahashi T, Hatta M, Fujisawa T. The structure and expression of a preprohormone of a neu-ropeptide, Hym-176 in *Hydra magnipapillata*. FEBS Lett. 1998; 439: 31–4. https://doi.org/10.1016/s0014-5793(98)01314-3 PMID: 9849871

46. Schmutzler C, Darmer D, Diekhoff D, Grimmelikhuijzen CJ. Identification of a novel type of processing sites in the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH2) from *Anthopleura elegantissima*. J Biol Chem. 1992; 267: 22534–41. Available: http://www.ncbi.nlm.nih.gov/pubmed/1429603 PMID: 1429603

47. Reinscheid RK, Grimmelikhuijzen CJ. Primary structure of the precursor for the anthozoan neuropep-tide antho-RFamide from *Renilla köllikeri*: evidence for unusual processing enzymes. J Neurochem. 1994; 62: 1214–22. https://doi.org/10.1046/j.1471-4159.1994.62031214.x PMID: 7906718

48. Lundquist CT, Clottens FL, Holman GM, Nichols R, Nachman RJ, Nässel DR. Callitachykinin I and II, two novel myotropic peptides isolated from the blowfly, *Calliphora vomitoria*, that have resemblances to tachykinins. Peptides. 1994; 15: 761–8. https://doi.org/10.1016/0196-9781(94)90027-2 PMID: 7984492

49. Neupert S, Russell WK, Russell DH, López JD, Predel R, Nachman RJ. Neuropeptides in Heteroptera: identification of allatotropin-related peptide and tachykinin-related peptides using MALDI-TOF mass spectrometry. Peptides. 2009; 30: 483–8. https://doi.org/10.1016/j.peptides.2008.11.009 PMID: 19084564

50. Anastasi A, Erspamer V. The isolation and amino acid sequence of eledoisin, the active endecapeptide of the posterior salivary glands of *Eledone*. Arch Biochem Biophys. 1963; 101: 56–65. https://doi.org/10.1016/0003-9861(63)90533-2 PMID: 14012712

51. Chiwakata C, Brackmann B, Hunt N, Davidoff M, Schulze W, Ivell R. Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse testis. Endocrinology. 1991; 128: 2441–8. https://doi.org/10.1210/endo-128-5-2441 PMID: 1708336

52. Veenstra JA. Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. Gen Comp Endocrinol. 2011; 171: 160–75. https://doi.org/10.1016/j.ygcen.2011.01.005 PMID: 21241702

53. Sato Y, Oguchi M, Menjo N, Imai K, Saito H, Ikeda M, et al. Precursor polyprotein for multiple neuropep-tides secreted from the suboesophageal ganglion of the silkworm *Bombyx mori*: characterization of the cDNA encoding the diapause hormone precursor and identification of additional peptides. Proc Natl Acad Sci U S A. 1993; 90: 3251–5. https://doi.org/10.1073/pnas.90.8.3251 PMID: 8475067

54. Duportets L, Gadenne C, Couillaud F. A cDNA, from *Agrotis ipsilon* that encodes the pheromone bio-synthesis activating neuropeptide (PBAN) and other FXPRL peptides. Peptides. 1999; 20: 899–905. https://doi.org/10.1016/s0196-9781(99)00079-0 PMID: 10503766

55. Hauser F, Williamson M, Cazzamali G, Grimmelikhuijzen CJP. Identifying neuropeptide and protein hor-mone receptors in *Drosophila melanogaster* by exploiting genomic data. Brief Funct Genomic Proteo-mic. 2006; 4: 321–30. https://doi.org/10.1093/bfgp/eli003 PMID: 17202123

56. Mori K, Miyazato M, Ida T, Murakami N, Serino R, Ueta Y, et al. Identification of neuromedin S and its possible role in the mammalian circadian oscillator system. EMBO J. 2005; 24: 325–35. https://doi.org/10.1038/sj.emboj.7600526 PMID: 15635449

57. Brighton PJ, Szekeres PG, Willars GB. Neuromedin U and its receptors: structure, function, and physio-logical roles. Pharmacol Rev. 2004; 56: 231–48. https://doi.org/10.1124/pr.56.2.3 PMID: 15169928

58. Williamson M, Lenz C, Winther AM, Nässel DR, Grimmelikhuijzen CJ, Winther ME. Molecular cloning, genomic organization, and expression of a B-type (cricket-type) allatostatin preprohormone from *Dro-sophila melanogaster*. Biochem Biophys Res Commun. 2001; 281: 544–50. https://doi.org/10.1006/bbrc.2001.4402 PMID: 11181081

59. Yamanaka N, Hua Y-J, Roller L, Spalovská-Valachová I, Mizoguchi A, Kataoka H, et al. Bombyx prothoracicostatic peptides activate the sex peptide receptor to regulate ecdysteroid biosynthesis. Proc Natl Acad Sci U S A. 2010; 107: 2060–5. https://doi.org/10.1073/pnas.0907471107 PMID: 20133850

60. Schoofs L, Holman GM, Hayes TK, Nachman RJ, De Loof A. Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MIP), a novel biologically active neuropeptide from *Locusta migra-toria*. Regul Pept. 1991; 36: 111–9. Available: https://doi.org/10.1016/0167-0115(91)90199-q PMID: 1796179

61. Fu Q, Tang LS, Marder E, Li L. Mass spectrometric characterization and physiological actions of VPNDWAHFRGSWamide, a novel B type allatostatin in the crab, *Cancer borealis*. J Neurochem. 2007; 101: 1099–107. https://doi.org/10.1111/j.1471-4159.2007.04482.x PMID: 17394556

62. Moroz LL, Edwards JR, Puthanveettil S V., Kohn AB, Ha T, Heyland A, et al. Neuronal transcriptome of *Aplysia*: Neuronal compartments and circuitry. Cell. 2006; 127: 1453–1467. https://doi.org/10.1016/j.cell.2006.09.052 PMID: 17190607

63. Ahn S-J, Martin R, Rao S, Choi M-Y. Neuropeptides predicted from the transcriptome analysis of the gray garden slug D*eroceras reticulatum*. Peptides. 2017; 93: 51–65. https://doi.org/10.1016/j.peptides.2017.05.005 PMID: 28502716

64. Conzelmann M, Williams EA, Tunaru S, Randel N, Shahidi R, Asadulina A, et al. Conserved MIP receptor-ligand pair regulates *Platynereis* larval settlement. Proc Natl Acad Sci U S A. 2013; 110: 8224–9. https://doi.org/10.1073/pnas.1220285110 PMID: 23569279

65. Predel R, Nachman RJ, Gäde G. Myostimulatory neuropeptides in cockroaches: structures, distribution, pharmacological activities, and mimetic analogs. J Insect Physiol. 2001; 47: 311–24. https://doi.org/10.1016/s0022-1910(00)00129-3 PMID: 11166295

66. Clynen E, Schoofs L. Peptidomic survey of the locust neuroendocrine system. Insect Biochem Mol Biol. 2009; 39: 491–507. https://doi.org/10.1016/j.ibmb.2009.06.001 PMID: 19524670

67. Lorenz MW, Kellner R, Hoffmann KH. A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket, *Gryllus bimaculatus*. J Biol Chem. 1995; 270: 21103–8. https://doi.org/10.1074/jbc.270.36.21103 PMID: 7673141

68. Jang Y-H, Chae H-S, Kim Y-J. Female-specific myoinhibitory peptide neurons regulate mating receptivity in *Drosophila melanogaster*. Nat Commun. 2017; 8: 1630. https://doi.org/10.1038/s41467-017-01794-9 PMID: 29158481

69. Oh Y, Yoon S-E, Zhang Q, Chae H-S, Daubnerová I, Shafer OT, et al. A homeostatic sleep-stabilizing pathway in *Drosophila* composed of the sex peptide receptor and its ligand, the myoinhibitory peptide. PLoS Biol. 2014; 12: e1001974. https://doi.org/10.1371/journal.pbio.1001974 PMID: 25333796

70. Peymen K, Watteyne J, Borghgraef C, Van Sinay E, Beets I, Schoofs L. Myoinhibitory peptide signaling modulates aversive gustatory learning in *Caenorhabditis elegans*. PLoS Genet. 2019; 15: e1007945. https://doi.org/10.1371/journal.pgen.1007945 PMID: 30779740

71. Marlow HQ, Srivastava M, Matus DQ, Rokhsar D, Martindale MQ. Anatomy and development of the nervous system of *Nematostella vectensis*, an anthozoan cnidarian. Dev Neurobiol. 2009; 69: 235–54. https://doi.org/10.1002/dneu.20698 PMID: 19170043

72. Koizumi O. Nerve ring of the hypostome in *Hydra*: is it an origin of the central nervous system of bilaterian animals? Brain Behav Evol. 2007; 69: 151–9. https://doi.org/10.1159/000095204 PMID: 17230023

73. Watanabe H, Kuhn A, Fushiki M, Agata K, Özbek S, Fujisawa T, et al. Sequential actions of β-catenin and Bmp pattern the oral nerve net in *Nematostella vectensis*. Nat Commun. 2014; 5: 5536. https://doi.org/10.1038/ncomms6536 PMID: 25534229

74. Petruzziello F, Fouillen L, Wadensten H, Kretz R, Andren PE, Rainer G, et al. Extensive characterization of Tupaia belangeri neuropeptidome using an integrated mass spectrometric approach. J Proteome Res. 2012; 11: 886–96. https://doi.org/10.1021/pr200709j PMID: 22070463

75. Zhang X, Petruzziello F, Zani F, Fouillen L, Andren PE, Solinas G, et al. High identification rates of endogenous neuropeptides from mouse brain. J Proteome Res. 2012; 11: 2819–27. https://doi.org/10.1021/pr3001699 PMID: 22424378

76. Sasaki K, Osaki T, Minamino N. Large-scale identification of endogenous secretory peptides using electron transfer dissociation mass spectrometry. Mol Cell Proteomics. 2013; 12: 700–9. https://doi.org/10.1074/mcp.M112.017400 PMID: 23250050

77. Southey BR, Amare A, Zimmerman T a, Rodriguez-Zas SL, Sweedler J V. NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. Nucleic Acids Res. 2006; 34: W267–72. https://doi.org/10.1093/nar/gkl161 PMID: 16845008

78. Darmer D, Hauser F, Nothacker HP, Bosch TC, Williamson M, Grimmelikhuijzen CJ. Three different prohormones yield a variety of Hydra-RFamide (Arg-Phe-NH2) neuropeptides in *Hydra magnipapillata*. Biochem J. 1998; 332 (Pt 2: 403–12. https://doi.org/10.1042/bj3320403 PMID: 9601069

79. Zhang X, Pan H, Peng B, Steiner DF, Pintar JE, Fricker LD. Neuropeptidomic analysis establishes a major role for prohormone convertase-2 in neuropeptide biosynthesis. J Neurochem. 2010; 112: 1168–79. https://doi.org/10.1111/j.1471-4159.2009.06530.x PMID: 19968759

80. Lim J, Berezniuk I, Che F-Y, Parikh R, Biswas R, Pan H, et al. Altered neuropeptide processing in prefrontal cortex of Cpe (fat/fat) mice: implications for neuropeptide discovery. J Neurochem. 2006; 96: 1169–81. https://doi.org/10.1111/j.1471-4159.2005.03614.x PMID: 16417576

81. Hoshino A, Lindberg I. Peptide Biosynthesis: Prohormone Convertases 1/3 and 2. Colloq Ser Neuropeptides. 2012; 1: 1–112. https://doi.org/10.4199/C00050ED1V01Y201112NPE001

82. Boonen K, Baggerman G, D'Hertog W, Husson SJ, Overbergh L, Mathieu C, et al. Neuropeptides of the islets of Langerhans: a peptidomics study. Gen Comp Endocrinol. 152: 231–41. https://doi.org/10.1016/j.ygcen.2007.05.002 PMID: 17559849

83. Van Camp KA, Baggerman G, Blust R, Husson SJ. Neuropeptidomic analysis of zebrafish brain. Methods Mol Biol. 2018; 1719: 241–246. https://doi.org/10.1007/978-1-4939-7537-2_16 PMID: 29476516

84. Leung PS, Shaw C, Maule AG, Thim L, Johnston CF, Irvine GB. The primary structure of neuropeptide F (NPF) from the garden snail, *Helix aspersa*. Regul Pept. 1992; 41: 71–81. https://doi.org/10.1016/0167-0115(92)90515-v PMID: 1472263

85. Brown MR, Crim JW, Arata RC, Cai HN, Chun C, Shen P. Identification of a *Drosophila* brain-gut peptide related to the neuropeptide Y family. Peptides. 1999; 20: 1035–42. https://doi.org/10.1016/s0196-9781(99)00097-2 PMID: 10499420

86. Cupo A, Proux J. Biochemical characterization of a vasopressin-like neuropeptide in *Locusta migratoria*. Evidence of high molecular weight protein encoding vasopressin sequence. Neuropeptides. 1983; 3: 309–18. https://doi.org/10.1016/0143-4179(83)90048-3 PMID: 6866218

87. Cruz LJ, de Santos V, Zafaralla GC, Ramilo CA, Zeikus R, Gray WR, et al. Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from *Conus geographus* and *Conus straitus* venoms. J Biol Chem. 1987; 262: 15821–4. PMID: 3680228

88. Schoofs L, Danger JM, Jegou S, Pelletier G, Huybrechts R, Vaudry H, et al. NPY-like peptides occur in the nervous system and midgut of the migratory locust, *Locusta migratoria* and in the brain of the grey fleshfly, *Sarcophaga bullata*. Peptides. 9: 1027–36. https://doi.org/10.1016/0196-9781(88)90084-8 PMID: 3244556

89. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science. 2007; 317: 86–94. https://doi.org/10.1126/science.1139158 PMID: 17615350

90. Jékely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. Proc Natl Acad Sci U S A. 2013; 110: 8702–7. https://doi.org/10.1073/pnas.1221833110 PMID: 23637342

91. Mirabeau O, Joly J-S. Molecular evolution of peptidergic signaling systems in bilaterians. Proc Natl Acad Sci U S A. 2013; 110: E2028–37. https://doi.org/10.1073/pnas.1219956110 PMID: 23671109