Transcriptomic discovery and comparative analysis of neuropeptide precursors in sea cucumbers (Holothuroidea)

Saowaros Suwansa-ard[1,2,¶], Arada Chaiyamoon[1,¶], Alzbeta Talarovicova[3], Ruchanok Tinikul[4], Yotsawan Tinikul[1,7], Tanes Poomtong[6], Maurice R. Elphick[3], Scott F. Cummins[2], Prasert Sobhon[1,5,*]

[1]Department of Anatomy, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

[2]Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore, Queensland 4558, Australia

[3]School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom

[4]Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

[5]Faculty of Allied Health Sciences, Burapha University, Chonburi 20131, Thailand

[6]Coastal Fisheries Research and Development Center, Klongwan, Prachuab Khiri Khan 77000, Thailand

[7]Mahidol University, Nakhonsawan Campus, Nakhonsawan 60130, Thailand

[¶]These authors contributed equally to this work

Author for correspondence:

Prasert Sobhon
Department of Anatomy, Faculty of Science, Mahidol University, Bangkok 10400, Thailand
Fax: +66 0 2354 7168
Tel: +66 0 2201 5406
E-mail: prasert.sob@mahidol.ac.th

Running title:

Neuropeptide precursors in the sea cucumbers *Holothuria*

Conflicts of interest: none

Keywords: Neuropeptides, transcriptomes, nervous tissue, gonads, echinoderm, *Holothuria scabra*

Data accessibility: Raw sequence data of *Holothuria scabra* transcriptome has been deposited in the Sequence Read Archive (SRA) under accession number SRR5755244. Additional data has been provided in the supplementary materials of this article.

Authors' contributions: SS and AC performed the sample preparation, bioinformatic analysis of neuropeptides in *H. scabra*, investigation of gene expression, and prepared the first draft of the manuscript. AT performed the bioinformatic analysis of neuropeptides in *H. glaberrima* and prepared the result data. RT and YT performed the sample preparation for RNA sequencing. PS and TP provided the chemicals and reagents. MRE, SFC and PS conceptualized the work and supervised manuscript preparation and revision. All authors read and approved the final manuscript.

Funding: This research was supported by Agricultural Research Development Agency (Public Organization), Thailand.

Abstract

Neuropeptides synthesized and released by neuronal cells play important roles in the regulation of many processes, e.g. growth, feeding, reproduction, and behavior. In the past decade, next-generation sequencing technologies have helped to facilitate the identification of multiple neuropeptide genes in a variety of taxa, including arthropods, molluscs and echinoderms. In this study, we extend these studies to *Holothuria scabra,* a sea cucumber species that is widely cultured for human consumption. *In silico* analysis of *H. scabra* neural and gonadal transcriptomes enabled the identification of 28 transcripts that encode a total of 26 bilaterian and echinoderm-specific neuropeptide precursors. Furthermore, publicly available sequence data from another sea cucumber, *Holothuria glaberrima*, allowed a more in-depth comparative investigation. Interestingly, two isoforms of a calcitonin-type peptide precursor (CTPP) were deduced from the *H. scabra* transcriptome - *HscCTPP-long* and *HscCTPP-short,* likely the result of alternative splicing. We also identified a sea cucumber relaxin-type type peptide precursor, which is of interest because relaxin-type peptides have been shown to act as gonadotropic hormones in starfish. Two neuropeptides that appear to be holothurian-specific are GLRFA, and GN-19. In *H. scabra*, the expression of *GLRFA* was restricted to neural tissues, while *GN-19* expression was additionally found in the longitudinal muscle and intestinal tissues. In conclusion, we have obtained new insights into the neuropeptide signaling systems of holothurians,

which will facilitate physiological studies that may enable advances in the aquaculture of sea cucumbers.

1. Introduction

Neuropeptides (NPs) play important roles in cell communication that are essential for maintaining homeostasis and controlling various physiological activities in eumetazoans. NPs are synthesized by neurons and are derived from NP precursors (NPPs) [1], which are routed into the secretory pathway where they are processed and modified by a variety of enzymes, resulting in production of shorter bioactive NPs [2, 3]. The principal role of NPs is to mediate neuronal chemical communication with other cells, but the actions of NPs can be diverse depending on the spatial scale over which they act. For instance, NPs can act as neurotransmitters to mediate rapid communication between adjacent neurons [4]. As neuromodulators, NPs are released from one neuron and act on other neurons or other cell types by binding to cognate receptors and modifying the activity of target cells [2]. Functioning as neurohormones, NPs are secreted into the circulatory system and thereby can reach distant target cells and exert effects systemically.

There are currently more than 80 NP families described, in which distinct and well-conserved family-specific domains may be observed [5]. The high-throughput identification of NPs within a single species is now rendered relatively easy through genome/transcriptome sequencing and mass spectrometry-based proteomics. This has been exemplified in several investigations focusing on the neural tissue of non-model aquatic invertebrates, including crustaceans [6, 7], molluscs [8, 9], and echinoderms [10-12]. Ultimately, the discovery of NPs in various species will facilitate research investigating the evolution and comparative physiology of NPs. Furthermore, NP signaling systems represent potentially attractive targets to manipulate the physiology of economically important species, including aquaculture-reared species; for example, as agents to control induction of gonadal maturation and spawning [13].

Aquaculture of sea cucumbers (class Holothuroidea, Phylum Echinodermata) has dramatically increased since the end of the 20th century to keep up with consumer demand. *Holothuria scabra* (sandfish) is a tropical sea cucumber species considered to be one of the most commercially important because of its wide distribution, as well as high market demand and value, especially in Asian countries [14]. The consumption of sea cucumbers has increased with recognition that their nutritional composition has numerous health benefits, including the prevention or treatment of various diseases (for reviews, see [15-17]). Therefore, there is interest in understanding the biology of sea cucumbers in

order to improve their aquaculture [18-20]; for example, by identifying NPs that can enhance growth and reproduction. Pioneering studies on NPs in the sea cucumber *Holothuria glaberrima* identified two NPs belonging to the echinoderm SALMFamide family [21]. Functional characterization of these SALMFamides using immunohistochemical and pharmacological methods revealed that they are expressed in neuromuscular organs and act as muscle relaxants [22, 23]. Myoactive NPs have also been identified in the Japanese sea cucumber *Apostichopus japonicus* [24, 25]; for example, the myoactive NP NGIWYamide, which also triggers spawning in *A. japonicus* [13, 26, 27].

Recently, investigation of the occurrence of NPs in sea cucumbers and other echinoderms has been accelerated by transcriptomic analyses. Thus, many NPPs have been identified in sea urchins (Class Echinoidea), starfish (Class Asteroidea) and brittle stars (Class Ophiuroidea) [10, 12, 28, 29]. However, to date, the sea cucumber neuropeptidome has only been explored in *A. japonicus*, where several NPP transcripts have been identified by transcriptome analysis [11, 29, 30]. In this study, we performed an *in silico* analysis of *H. scabra* transcriptomes derived from the RNA of nerve ring, radial nerve, ovary, and testis tissue. NP transcripts were identified and annotated using well-established bioinformatics workflows. We also compared our findings from *H. scabra* with data obtained from a closely related species – *H. glaberrima*, utilizing publically available transcript sequences [31]. The NPPs identified in *H. scabra* and *H. glaberrima* were compared with data from other echinoderms and other bilaterians, enabling identification of NPs that have a widespread phylogenetic distribution, as well as NPs that may be unique to echinoderms or unique to sea cucumbers.

2. Materials and Methods

2.1. Animals, tissue collections, and RNA preparation

Specimens of *H. scabra* were obtained from Prachuap Khiri Khan Coastal Fisheries Research and Development Center, Department of Fisheries, Khlongwan, Prachuap Khiri Khan Province, Thailand. Animals (mean weight ~ 250 g) were anesthetized by immersion in ice-cold sea water for 30 min and then their tissues were collected. For neural tissue, the circumoral nerve rings (CNR) and radial nerve cords (RNC) were collected from 40 individuals (20 males and 20 females). The testes and ovaries were collected from male and female *H. scabra* (3 individuals each). All tissues were immediately frozen in liquid nitrogen and stored at -80^{o}C until used. RNA extraction was performed using TriPure isolation reagent (Roche, IN, USA) following the manufacturer's protocol. RNA quality and concentration were checked by gel electrophoresis and spectrophotometry (NanoDrop 1000; Thermo Fisher Scientific, DE, USA).

2.2. Transcriptome preparation and sequence assembly

RNA from different tissues were pooled and subsequently dried before being sent to Beijing Genomics Institute (BGI, Hong Kong) for library construction using their standard workflow for *de novo* RNA-seq transcriptomes (http://bgiamericas.com/). Briefly, total RNA was subjected to oligo-dT selection for mRNA purification. Enriched mRNA was used for complementary DNA (cDNA) library construction with a normalization method. The cDNA library was then sequenced using an Illumina HiSeq2000 instrument (Illumina Inc.). Trimmed reads from Illumina sequencing were used for *de novo* assembly using CLC Genomic workbench software (CLC Bio-Qaigen, AsiaPac, Taiwan) with parameters set as follows: seqType, fq; minimum kmer coverage = 4; minimum contig length of 100 bp; group pair distance = 250. Estimation of transcript expression was performed using the RNA-Seq analysis tool on the CLC Genomic workbench software and transcripts per kilobase million (TPM) was calculated. The raw sequence dataset was deposited in the NCBI Sequence Read Archive (SRA) database under the accession number SRR5755244.

2.3. NPP transcript mining, comparative sequence analysis, and phylogenetic analysis

Transcripts encoding putative NPPs were identified by tBLASTn searches using the CLC Main Workbench Version 7.7 (CLC Bio-Qaigen, AsiaPac, Taiwan). Briefly, the amino acid (aa) sequences of 57 known putative NPPs from other echinoderm species (*Strongylocentrotus purpuratus* [10], *Apostichopus japonicus* [11], *Asterias rubens* [12]), and *Acanthaster planci* [28] were used for tBLASTn searches with parameters set as follows: matrix, BLOSUM62; e-value, 100. BLAST hits were collected, translated, and then manually analyzed based on their similarities to the orthologous proteins from other species and the presence of conserved motifs. Signal peptide sequences were predicted using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) [32]. Sequence alignment and similarity were analysed by using MEGA6 software [33]. Enzymatic cleavage sites, putative bioactive peptide(s) and post-translational modifications were predicted based on previously known consensus cleavage motifs and modifications, and/or by using the NeuroPred program [34]. In addition to *H. scabra* transcriptomes, the search for NPPs was also carried out by analysis of neural transcriptome sequence data from *H. glaberrima*, which has been generated previously [31]. BLAST searches on the *H. glaberrima* data were performed by using SequenceServer software (http://www.sequenceserver.com) [35], with known NPP sequences from other echinoderms submitted as queries. Hit transcripts were collected and used for further sequence analyses as described earlier. The representation of NPPs in different species of echinoderms was visualized using Cytoscape software [36]. Finally, all deduced NPPs from *H. scabra* and *H. glaberrima* were subjected to additional BLASTp searches for further confirmation of sequences and relationships with known NPPs. Phylogenetic tree based analysis of

NPPs was performed using the maximum likelihood method, using the Phylogeny.fr webserver [37]. The parameters were set as follows: approximate likelihood-ratio test, SH-like; substitution model, WAG; substitution rate categories, 4; gamma distribution, 1. For sequence alignment and phylogenetic tree analysis, an additional search of target NPPs in public databases was carried out using the NCBI tBLASTn program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with parameter settings as follows: expect threshold, 100; word size, 3; matrix, BLOSUM62; Gap costs, existence=11, extension=1; database was limited to echinoderms (taxid:7586). Furthermore, to facilitate comparison of holothurian NPPs and NPs with related proteins/peptides in other echinoderms, publicly available transcriptome data from two brittle star (Class Ophiuroidea) species, *Amphiura filiformis* and *Ophiopsila aranea* [38], were also analyzed to identify NPPs as described above using CLC Main Workbench software.

2.4. Analysis of tissue expression of target genes by reverse transcription polymerase chain reaction (RT-PCR)

Tissues from the sea cucumber *H. scabra*, including CNR, RNC, ovary, testis, longitudinal muscle (LM), body wall, intestine and respiratory tree, were collected for total RNA isolation (N=5). Non-gonadal tissues were pooled from males and females, while gonadal tissues, testis and ovary, were collected separately. Total RNA was then used for complementary DNA synthesis (cDNA) (RevertAid reverse transcriptase; Thermo Scientific, USA). Gene-specific primers for target genes were designed using the Primer-BLAST program (https://www.ncbi.nlm.nih.gov/tools/primer-blast) (**Table S1a**). PCR conditions followed a routine protocol (a denaturing step of 3 minutes at 95 °C, followed by 35 cycles of 30 seconds at 95 °C, 30 seconds at 55 °C, and 30 seconds at 72 °C, followed by an extension at 72 °C for 5 minutes). PCR products were analyzed by agarose gel electrophoresis and subsequently photographed. To confirm the cDNA sequences, amplicons were purified (QIAquick gel extraction kit, QIAGEN, Germany) and then sent for DNA sequencing (Macrogen, Korea). Amplification of the *H. scabra* 16S RNA gene was used as the internal control, while the negative control was performed by using non-RT cDNA, which was prepared from pooled tissue RNA, as the cDNA template.

3. Results

3.1. Transcriptome production, sequence assembly, and identification of sea cucumber NPP transcripts

Details of *H. scabra* transcriptome sequencing, assembly and quantification are summarized in **Table 1**. Illumina sequencing of *H. scabra* neural and gonad tissue generated 35,377,444 clean reads with an average read length of 90 base pairs (bp). *De novo* assembly resulted in a total of 68,064 contigs, with an N50 of 868 bp, which includes 31,085,702 matched reads and 4,291,742 singletons.

Our *in silico* BLASTp analysis identified 28 *H. scabra* transcripts that encode 26 mostly full-length NPPs, as determined by the presence of a predicted N-terminal signal peptide and a stop codon within the coding sequence (**Table 2** and **File S1**). Similarly, from the *H. glaberrima* transcriptome, 27 NPPs were deduced from 29 transcripts (**Table 2** and **File S1**). A comparison of NPs identified from six echinoderm species, representing echinoids (*S. purpuratus*), asteroids (*A. rubens* and *A. planci*), and holothurians (*A. japonicus*, *H. scabra*, and *H. glaberrima*), showed that 23 NPPs are common to all, while some are class-specific (**Fig. 1**). Sea cucumber precursors of NPs belonging to bilaterian NP families include precursors of bursicons, calcitonin-type peptide (CTP), cholecystokinin (CCK), corticotropin-releasing hormone type peptides (CRHs), gonadotropin-releasing hormone (GnRH) family peptides (including GnRH and corazonin), glycoprotein hormones, kisspeptin, luqin-type peptide, melanin-concentrating hormone (MCH), NG peptide (NPS/CCAP-type), orexins, oxytocin/vasopressin (OT/VT), pedal peptide-like neuropeptides (PPLNPs), pigment-dispersing factor (PDF), thyrotropin-releasing hormone (TRH), relaxin/gonad-stimulating substance (GSS), somatostatin (SS) and tachykinin-like peptides (TKs). Thirteen of these were identified in *H. scabra* and *H. glaberrima* (**Fig. 2 and File S1**).

Table 1. Summary of transcriptome sequencing and sequence assembly

Number of reads (reads)	35,377,444
Matched reads (reads)	31,085,702
Singletons (reads)	4,291,742
Average read length (bp)	90
Number of contigs (contigs)	68,064
N50 (bp)	868
Minimum contig length (bp)	153
Maximum contig length (bp)	25,282
Average contig length (bp)	653

Table 2. Summary of NPPs present in various echinoderm species. The presence of NPPs in *S. purpuratus*, *A. rubens*, *A. planci*, *A. japonicus*, *H. scabra*, *H. glaberrima* are shown.

Neuropeptides	***Strongylocentrotus purpuratus***	***Asterias rubens***	***Acanthaster planci***	***Apostichopus japonicus***	***Holothuria scabra***	***Holothuria glaberrima***

Bursicons alpha (Bur-α)	+[j]	+[l]	+[q]	+[b]	-	-
Bursicons beta (Bur-β)	+[j]	+[l]	+[q]	+[b]	-	-
Cholecystokinin (CCK)	+[n]	+[l]	+[q]	+[r]	-	-
Corticotropin-releasing hormone type precursor (CRH)	-**	+[l]	-[q]	-**	-	-
NG peptides (NGIWYamide/NGIWFamide/NGFFFamide/NGFFYamide)	+[fh]	+[lm]	+[q]	+[ace]	+	+
GLRFA	-[b]	-*	-[q]	+[ace]	+	+
Glycoprotein hormone alpha-2 (GPA-2)	+[j]	+[l]	+[q]	+[b]	+	+
Glycoprotein hormone beta-5 (GPB-5)	+[j]	+[l]	+[q]	-[b]	+	+
GN-19	-[b]	-*	-[q]	+[ace]	+	+
Insulin-like growth factors (IGFs)	+[j]	+[l]	+[q]	-*	-	+
Kisspeptin (Kiss)	+[l]	+[l]	+[q]	+*	+	+
Luqin	+[l]	+[l]	+[q]	+[b]	+	+
Np1F - F-type SALMFamide (SALMFa-F)	+[f,g]	+[l]	+[q]	+[ac]	+	+
Np1L - L-type SALMFamide (SALMFa-L)	+[fi]	+[l]	+[q]	+[d]	+	+
Np2 - GnRH precursor	+[f]	+[k,l]	+[q]	-[b]	-	-
Np3 - Thyrotropin-releasing hormone (TRH) precursor	+[f]	+[l]	+[q]	+[b]	+	+
Np4 - Calcitonin-type peptide (CTP) precursor	+[f]	+[l]	+[q]	+[b]	+	+
Np5 - AN peptide precursor	+[f]	+[l]	+[q]	+[b]	+	+
Np6 - pedal peptide-type neuropeptide-1 (PPLNP-1; homologs of molluscan pedal peptides and arthropod orcokinins)	+[f]	+[l]	+[q]	-[b]	-	-
Np7 - pedal peptide-type neuropeptide-2 (PPLNP-2; homologs of molluscan pedal peptides and arthropod orcokinins)	+[f]	-[l]	-[q]	+[b]	+	+
Np8	+[f]	-[l]	-[q]	+*	+	+
Np9	+[f]	-[l]	-[q]	+[b]	-	-
Np10	+[f]	-[l]	-[q]	-[b]	-	-
Np11	+[f]	+[l]	+[q]	+*	+	-
Np12 – Corazonin (Crz)	+[f]	+[k,l]	+[q]	+*	+	+
Np13	+[f]	-[l]	-[q]	-[b]	-	-
Np14 - Melanin-concentrating hormone (MCH)	+[f]	+[l]	+[q]	-[b]	-	-
Np15	+[f]	+[l]	+[q]	+*	+	+
Np16	+[f]	-[l]	-[q]	+[b]	+	+
Np17	+[f]	-[l]	-[q]	+*	+	+
Np18	+[f]	+[l]	+[q]	+*	+	+
Np19 - Somatostatin (SS)	+[f]	+[l]	+[q]	+[r]	+	+
Np20	+[f]	-[l]	-[q]	-[b]	-	-
Np21	-*	+[l]	-[q]	-*	-	-
Np22	-*	+[l]	+[q]	-*	-	-
Np23	+*	+[l]	+[q]	+*	+	+
Np24	-*	+[l]	-[q]	-*	-	-
Np25	-*	+[l]	+[q]	-*	-	-
Np26	-*	+*	+[q]	-*	-	-
Np27	-*	+*	+[q]	-*	-	-
Np28	-*	-*	+[q]	-*	-	-
Np29	-*	+*	+[q]	-*	-	-
Np30	-*	-*	+[q]	-*	-	-
Np31	-*	+*	+[q]	-*	-	-

Np32	-*	-*	+[q]	-*	-	-
Np33	-*	+*	+[q]	-*	-	-
Np34	-*	+*	+[q]	-*	-	-
Np35	-*	+*	+[q]	-*	-	-
Orexin-1 (OX-1)	+[no]	+[l]	+[q]	+*	+	+
Orexin-2 (OX-2)	+[no]	+[l]	+[q]	+*	+	+
Oxytocin/Vasopressin (OT/VP; Echinotocin/Asterotocin/Holotocin)	+[j]	+[l]	+[q]	+*	+	+
Pigment-dispersing factor (PDF)	+[o]	+[l]	-[q]	+[b]	+	+
Relaxin-like peptide-1 (gonad-stimulating substance type)	-*	+[l]	+[q]	+*	+	+
Relaxin-like peptide-2	-*	+[l]	-[q]	-*	-	-
Starfish myorelaxant peptide (SMP)	-*	+[p]	+[q]	-*	-	-
Stichopin	-[b]	-*	-[q]	+[ace]	-	-
Tachykinin (TK)	-*	+[l]	+[q]	-*	-	-

Note: +, presence; -, absence or transcript was not found; *, BLAST search against the available genome/transcriptome database of that given species was performed by the current authors; **, previously identified CRHs in *S. purpuratus* (Jekely, 2013) and *A. japonicus* (Rowe *et al.*, 2014) are identified as PDF precursors following Semmens *et al*., 2016. Lower case letters indicate the references where the genes have been reported: a) Elphick, 2012; b) Rowe *et al*., 2014; c) Ohtani *et al*., 1999; d) Elphick *et al*., 2013; e) Iwakoshi *et al*., 1995; f) Rowe, Elphick, 2012; g) Elphick, Thorndyke, 2005; h) Elphick, Rowe, 2009; i) Rowe, Elphick, 2010; j) Burke *et al*., 2006; k) Tian *et al*., 2016; l) Semmens *et al*., 2016; m) Semmens *et al*., 2013; n) Mirabeau, Joly, 2013; o) Jekely, 2013; p) Kim *et al*., 2016; q) Smith et al., 2017; r) Zandawala et al., 2017.

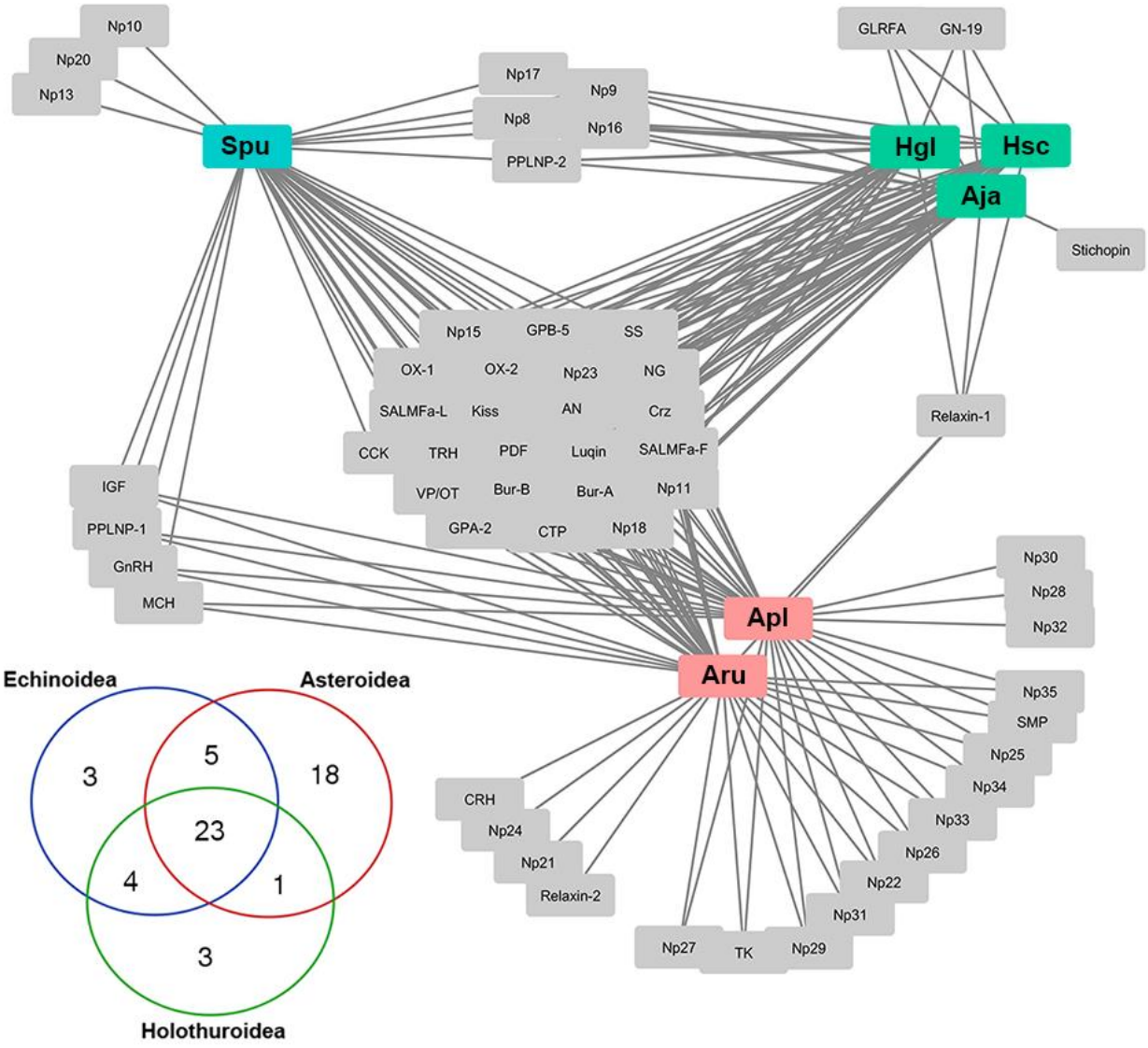

Figure 1. Clustering of neuropeptide precursors across different echinoderm classes. Included are the sea urchin *S. purpuratus* (Spu) as a representative for the class Echinoidea, the starfish *A. rubens* (Aru) and *A. planci* (Apl) as representatives of the class Asteroidea, and the sea cucumbers *A. japonicus* (Aja), *H. scabra* (Hsc) and *H. glaberrima* (Hgl) as representatives of the class Holothuroidea. The Venn diagram shows the number of neuropeptide precursors that are unique/common in the different echinoderm classes. The total number of neuropeptide precursors is 57 (for a list of neuropeptide precursors and abbreviations, refer to **Table 2**).

Figure 2. Amino acid sequences of neuropeptide precursors in *H. scabra* that belong to protein families that are conserved throughout the Bilateria. Gray, signal peptide; red, cleavage sites; green, glycine residue responsible for C-terminal amidation; blue, cysteine residue; +, missing amino acid residues as a result of partial transcript.

Sea cucumber NPPs that may be unique to echinoderms (homologs not yet discovered in other phyla), include the precursors of AN peptides, GN-19, SALMFamides, GLRFA, neuropeptide (Np) 8 to Np11, Np13, Np15 to Np18, and Np20 to Np35. Eleven of these were found in *H. scabra* and *H. glaberrima* (**Fig. 3** and **File S1**). Multiple sequence alignment of their precursor/mature peptides indicated high conservation among different echinoderm species (**File S2**); however, the GLRFA and GN-19 precursors appear to be restricted to the holothurians, as discussed in more detail below.

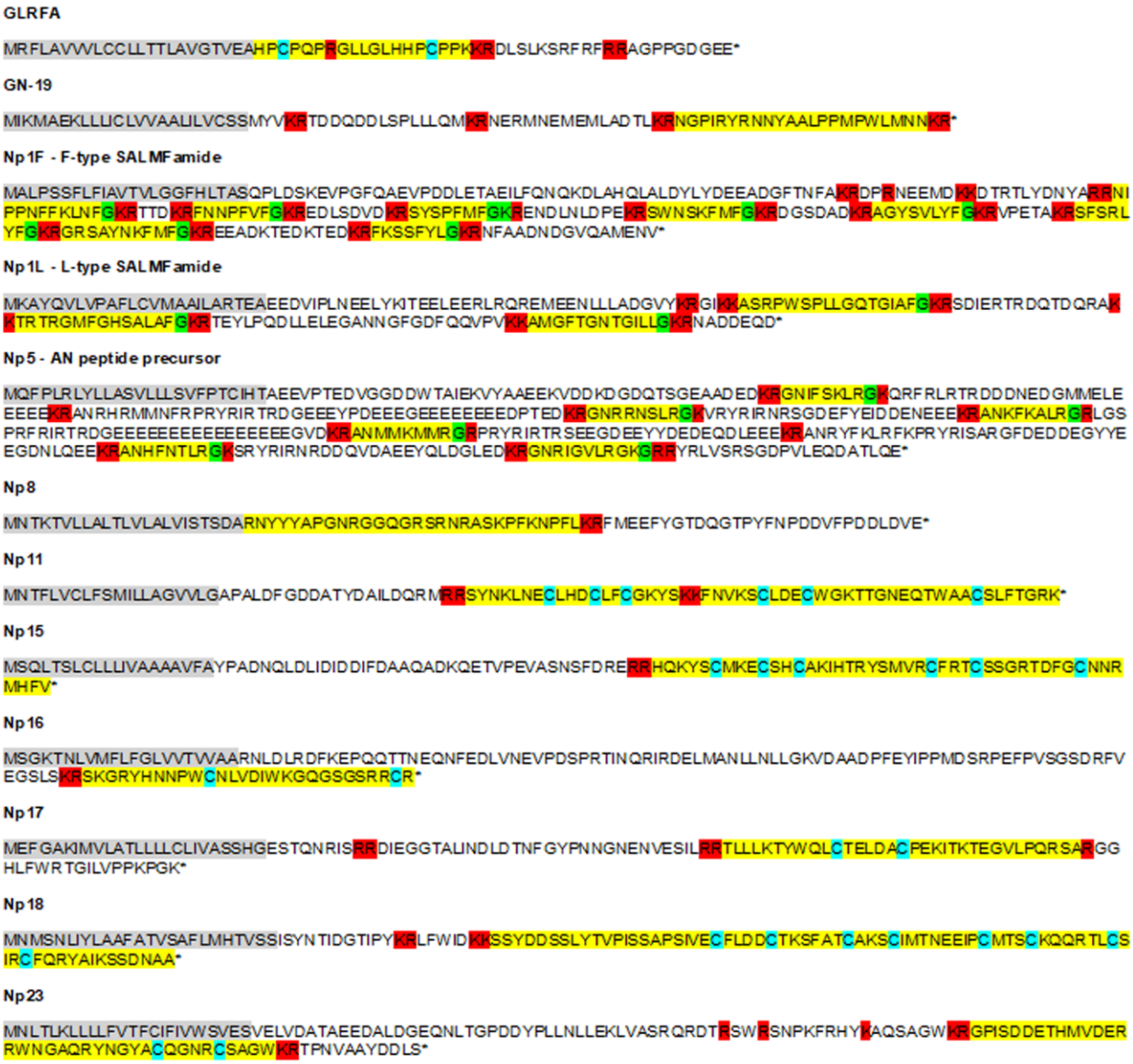

Figure 3. Amino acid sequences of neuropeptide precursors in *H. scabra* that belong to protein families that may be unique to echinoderms. Gray, signal peptide; red, cleavage sites; green, glycine residue responsible for C-terminal amidation; blue, cysteine residue; +, missing amino acid residues as a result of partial transcript.

Two isoforms of a calcitonin-type peptide precursor (CTPP) were identified in *H. scabra* transcriptome. The HscCTPP-long has 149 residues and consists of a signal peptide followed by a 126 aa CTP propeptide that contains a 35 aa CTPP-related peptide and two amidated CTPs - HscCTP-1 and -2 (35 and 36 amino acid residues in length, respectively). The HscCTPP-short has 99 residue and contains a signal peptide followed by a 35 aa CTPP-related peptide and a mature CTP peptide that is identical to the HscCTP-2 of HscCTPP-long (therefore named HscCTP-2; **Fig. 4A**). Sequence alignment of all echinoderm CTPs revealed two highly conserved motifs; a 'G/S-C-$X_{(2-3)}$-F-X-G-C' motif at the N-terminus, and a 'F-G-$X_{(2)}$-G/S-P-NH_2' motif at the C-terminus (**Fig. 4B**). Phylogenetic analysis separates the holothurian CTPPs from the CTPPs of the asteroids, ophiuroids, and echinoids, regardless of CTP number (**Figure 4C**). Our tissue-specific RT-PCR analysis demonstrates that in *H. scabra* both *CTPs (long and short)* are co-expressed in the neural tissues (CNR and RNC), as well as the LM and intestine (**Fig. 5**).

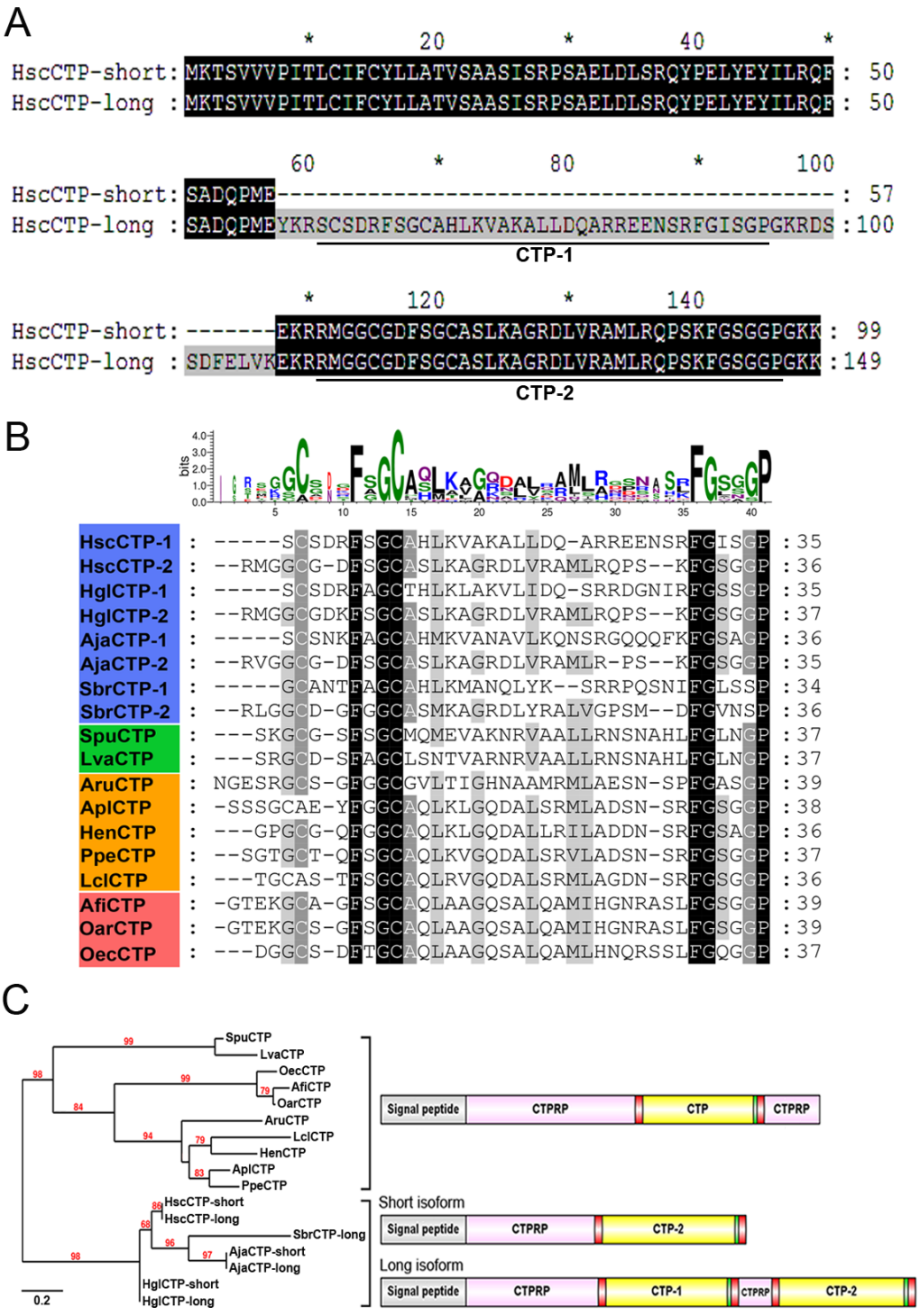

Figure 4. Calcitonin-type peptides (CTP) in *H. scabra*. (A) Alignment of short and long isoforms of *H. scabra* CTP precursors (HscCTP-short and -long, respectively). Underline indicates the predicted bioactive CTPs. (B) Alignment of CTPs from various echinoderm species. Black shading indicates identical amino acid residues whereas gray shading indicates similar amino acid residues. The sequence logo above the alignment shows the conservation of amino acid residues. The classes of representative species are indicated by different colors: green, Echinoidea; orange, Asteroidea; red, Ophiuroidea; blue, Holothuroidea. (C) Phylogenetic analysis of echinoderm CTP precursors based on

maximum likelihood estimation. The scale bar indicates the estimated amino acid substitutions per site. Branch support values are shown at nodes (≥50%). Protein illustration shows the general structure of CTP precursors in different echinoderm classes. CTPRP, calcitonin precursor-related peptide. For species abbreviations and sequences used in the amino acid alignment and phylogenetic tree, see **Table S1c**.

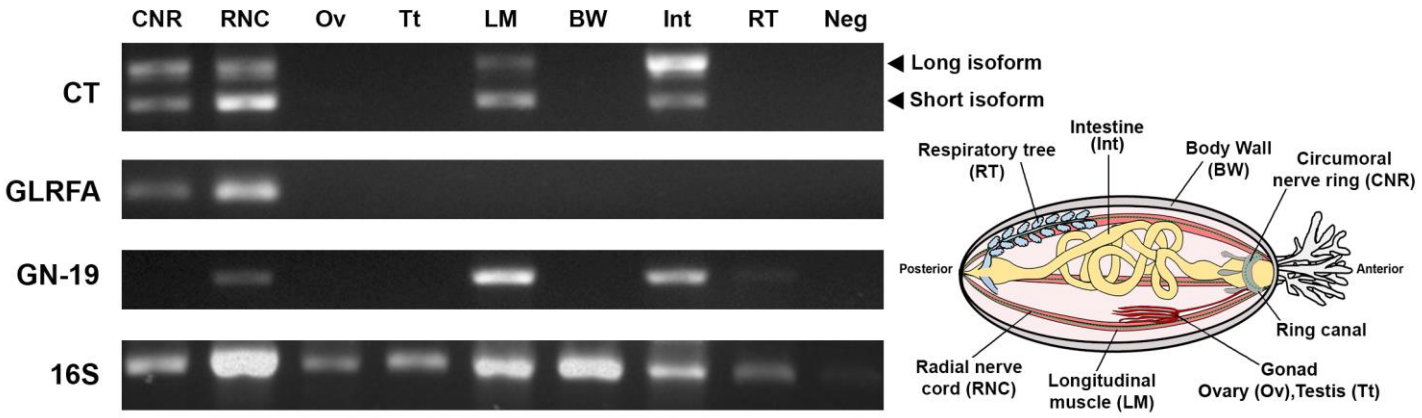

Figure 5. Expression of *calcitonin-type peptide* (*CT*), *GLRFA*, and *GN-19* genes in *H. sabra*. Spatial expression of CT, GLRFA, and GN-19 in various tissues of *H. scabra* are shown. The expression of 16S RNA transcripts in various tissues was used as the internal control. A negative control (Neg) was performed by using the non-RT cDNA as the template for RT-PCR.

3.2. Holothurian-specific neuropeptides: GLRFA and GN-19

GLRFA peptide: A *HscGLRFA* transcript encoding a 65 aa NPP (**Table 2** and **Fig. 3**) was identified in *H. scabra* that shows similarity (36-69%) to the GLRFA NPPs of *A. japonicus, H. glaberrima* and giant California sea cucumber (*Parastichopus californicus*). The possible enzymatic cleavage sites within the GLRFA propeptide were predicted based on a previous study in *A. japonicus* (*yellow boxes*, **Fig. 6**) as well as alternative cleavage sites which give rise two potential mature peptides (for example, 'HPCPQPRGLLGLHHPCPPK' and 'DLSLKSRFRF' for HscGLRFA) (*red boxes*, **Fig. 6**). In addition, an 'H-P-C-X_3' motif and two cysteine residues are conserved within the propeptide region of all identified sea cucumber GLRFAs. *HscGLRFA* is found in the CNR and RNC (**Fig. 5**).

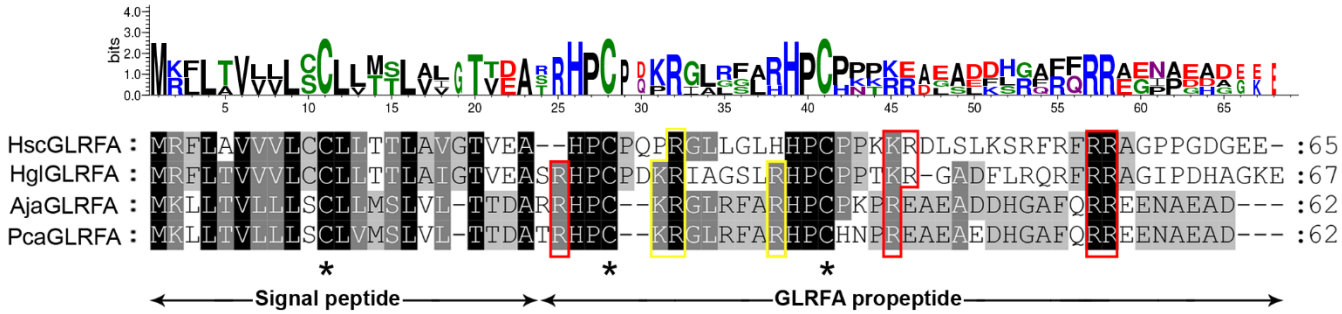

Figure 6. Comparative alignment of GLRFA precursor proteins. Black shading indicates identical amino acid residues whereas gray shading indicates similar amino acid residues. The yellow boxes show the previously known putative monobasic/dibasic cleavage sites, while the red boxes show possible alternative monobasic/dibasic cleavage sites of the GLRFA propeptide. The regions of the signal peptide and propeptide are indicated. The sequence logo above the alignment shows the conservation of amino acid residues. Asterisks show the conserved cysteine residues. For species abbreviations and sequences used in the amino acid alignment, see **Table S1c**.

GN-19: A full transcript encoding GN-19 was predicted from the *H. scabra* transcriptomes (HscGN-19; **Table 2**). The encoded NPP consists of 85 aa residues, which includes a 24 aa signal peptide and a 61 aa GN-19 propeptide (**Fig. 3**). A similar precursor characteristic was observed in the *H. glaberrima* GN-19 (**File S1**). The predicted GN-19 mature peptides in *H. scabra* and *H. glaberrima* are 'NGPIRYRNNYAALPPMPWLMNN' and 'NGGPRFRNNYAALPPMPWLMNN', respectively. The HscGN-19 NPP exhibits a high level of sequence similarity with GN-19 precursors of *A. japonicus* (46% identity) and *H. glaberrima* (94%) (**Fig. 7**). From RT-PCR, *HscGN-19* expression is detected in the RNC, LM, and intestine (**Fig. 5**).

Figure 7. Comparative alignments of GN-19 peptides in holothurian species. Black shading indicates identical amino acid residues whereas gray shading indicates similar amino acid residues. The

sequence logo above the alignment shows the conservation of amino acid residues. For species abbreviations and sequences used in the amino acid alignment, see **Table S1c**.

4. Discussion

Accumulating sequence data from echinoderms has led to an increase in our understanding of NPs that regulate key physiological activities. These include NPs that may control growth, feeding and reproduction. For the Holothuroidea and Asteroidea, an *in silico* analysis of the sea cucumber *A. japonicus* and starfish *A. rubens* neural transcriptomes had predicted 17 and 40 NPPs, respectively [11, 12]. Most recently, the identification of NPPs in the *A. planci* (crown-of-thorn starfish) neural transcriptome revealed an additional 10 potential novel NPPs [28], which expanded the number found in the phylum Echinodermata. For the Echinoidea, *in silico* identification of NPP transcripts from the sea urchin *S. purpuratus* radial nerve cDNA library revealed 20 different NPPs, half of which appeared to be echinoderm-specific [10]. In this study, we add two holothurians for NPP comparison, showing that at least 23 NPPs are common to echinoderms. While fewer echinoid- and holothurian-specific NPPs are observed, a number of NPPs in asteroids may be restricted to this class or the sub-phylum Asterozoa (Asteroidea + Ophiuroidea), including NPPs of CRH and TK, Np21, Np22, Np24 to Np35 (see **Fig. 1**). Sequence alignment of the common echinoderm NPPs (including those identified in brittle stars, class Ophiuroidea) with other bilaterian NPPs and mature NPs, shows high conservation (**File S2**). It is noteworthy that a larger number of class-specific NPPs were identified when more in-depth molecular analysis has been performed in multiple species, although the absence of some neuropeptides might be a result of the physiological status of the animals, e.g., reproductive stage and stress, in which certain NPs are not being expressed. Therefore, we expect that future studies on NPPs in various other echinoderm species will certainly expand the number of novel NPPs and help elucidate the NP signaling systems within this animal phylum.

We report for the first time a precursor of a relaxin-type peptide in sea cucumbers. In starfish, relaxin-type peptides have been well established for their role as a gonadotropic hormone [39]. Therefore, it will be of interest to investigate the actions of relaxin-type peptides as potential regulators of reproductive physiology in sea cucumbers. We also report the identification of two isoforms of a CTPP from the *H. scabra* transcriptome - *HscCTPP-long* and *HscCTPP-short*, which are likely the result of alternative splicing. The presence of two CTPP isoforms is a common feature of the holothurians since they are also found in *H. glaberrima*, *A. japonicus* and the ‘hairy’ sea cucumber, *Sclerodactyla briareus* (**File S1**). CTPPs are also found in the arthropods [known as the diuretic hormones (DH)-31],

which include two DH-31 precursors, one isoform encodes 1 mature CT, while the other encodes 3 mature CTs [40]. Their function has been implicated in water balance [41], ionic exchange [42], reproduction and ecdysteroidogenesis [43]. In mammals, CT is a 32 aa peptide that is primarily produced and secreted from the thyroid gland, where it regulates calcium metabolism [44]. However, CT is also produced from other mammalian tissues such as the brain, prostate and uterus, where it has other roles [45]. The mammalian neural tissue-derived CT-like peptide (known as calcitonin gene-related peptide; CGRP) is a product from CT gene alternative splicing, and its roles include stimulating vasodilation and neuromodulation [46, 47]. While the presence of CTP in the neural tissues of *H. scabra* confirms a neurohormone and/or neuromodulatory role, the non-neural CTP expression also implicates its role in coordinating muscular and gut physiology.

The GLRFA mature peptide of *A. japonicus* is 5 residues in length, following NPP processing at an N-terminal dibasic and C-terminal monobasic cleavage sites [30]. However, we find that the putative C-terminal monobasic cleavage site is absent in the *H. scabra* GLRFA NPP, and hence it is more likely that the mature peptide is larger. Meanwhile, there has been strong selective pressure for all holothurian GLRFA precursors to conserve an 'H-P-C-X_3' motif and two cysteine residues within the propeptide region. Furthermore, alternative enzymatic cleavage sites located in the C-terminal region were predicted (*see* **Fig. 6**). The two cysteine residues may enable intramolecular disulfide bridge formation. Clearly, further work is required to confirm the mature peptides that are derived from GLRFA NPPs in *Holothuria* and to investigate their bioactivities. In *A. japonicus*, GLRFA enhances electrically-induced contraction of the LM, and induces contraction of the intestine [48]. From our study, *HscGLRFA* expression in the CNR and RNC suggests a neuromodulatory role in *H. scabra*.

GN-19 was first identified in the *A. japonicus*, and was named according to its first and the last residues (glycine and asparagine, respectively) and its length (19 residues) [11]. We identified a transcript encoding a GN-19 NPP in *H. scabra*. Although a high conservation in amino acid composition could be observed throughout the length of peptide precursor of all sea cucumber GN-19s, the predicted GN-19 mature peptides in both *H. scabra* and *H. glaberrima* are 3 residues longer than the *A. japonicus* GN-19 peptide, and their N-terminal residue is an asparagine, not a glycine (see **Fig. 7**). Based on a previous study in *A. japonicus*, GN-19 regulates intestinal activity, inducing contraction or relaxation [48]. Expression of *HscGN-19* in the RNC, LM, and intestine supports a possible function for GN-19 in regulation of muscle activity in *H. scabra*. In *A. japonicus*, GN-19 has no effect on LM [48], yet the observed expression of the *HscGN-19* precursor in the LM suggests that its peptide products could possibly modulate LM activity.

5. Conclusions

In this study, transcriptome sequence data was analysed to identify 26 putative NPPs in the sea cucumbers *H. scabra* and *H. glaberrima.* Comparative sequence analysis with other echinoderms revealed that the majority of these NPPs had been reported previously. Two NPPs, GLRFA, and GN-19, appear to be restricted to the holothurians. The findings of this study provide deeper insights into neuropeptide signaling systems in holothurians and a basis for experimental investigation of their functions. Neuropeptides could be used in the future to help improve aquaculture for *H. scabra* and other economically important sea cucumber species.

Acknowledgements

This work was supported by Agricultural Research Development Agency (Public Organization), Bangkok, Thailand.

References

1. Nässel DR, Larhammar D: **Neuropeptides and Peptide Hormones**. In: *Neurosciences - From Molecule to Behavior: a university textbook.* Edited by Galizia GC, Lledo P-M. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013: 213-237.
2. Hokfelt T, Broberger C, Xu ZQ, Sergeyev V, Ubink R, Diez M: **Neuropeptides--an overview**. *Neuropharmacology* 2000, **39**(8):1337-1356.
3. Fricker LD: **Neuropeptide-processing enzymes: applications for drug discovery**. *The AAPS journal* 2005, **7**(2):E449-455.
4. Suudhof TC: **Neurotransmitter release**. *Handbook of experimental pharmacology* 2008(184):1-21.
5. Jekely G: **Global view of the evolution and diversity of metazoan neuropeptide signaling**. *Proc Natl Acad Sci U S A* 2013, **110**(21):8702-8707.
6. Suwansa-ard S, Thongbuakaew T, Wang T, Zhao M, Elizur A, Hanna PJ, Sretarugsa P, Cummins SF, Sobhon P: ***In silico* neuropeptidome of female *Macrobrachium rosenbergii* based on transcriptome and peptide mining of eyestalk, central nervous system and ovary**. *PLoS ONE* 2015, **10**(5):e0123848.
7. Ventura T, Cummins SF, Fitzgibbon Q, Battaglene S, Elizur A: **Analysis of the central nervous system transcriptome of the eastern rock lobster *Sagmariasus verreauxi* reveals its putative neuropeptidome**. *PLoS One* 2014, **9**(5):e97323.
8. Stewart MJ, Favrel P, Rotgans BA, Wang T, Zhao M, Sohail M, O'Connor WA, Elizur A, Henry J, Cummins SF: **Neuropeptides encoded by the genomes of the Akoya pearl oyster *Pinctata fucata* and Pacific oyster *Crassostrea gigas*: a bioinformatic and peptidomic survey**. *BMC Genomics* 2014, **15**:840.
9. Adamson KJ, Wang T, Zhao M, Bell F, Kuballa AV, Storey KB, Cummins SF: **Molecular insights into land snail neuropeptides through transcriptome and comparative gene analysis**. *BMC genomics* 2015, **16**:308.
10. Rowe ML, Elphick MR: **The neuropeptide transcriptome of a model echinoderm, the sea urchin *Strongylocentrotus purpuratus***. *Gen Comp Endocrinol* 2012, **179**(3):331-344.

11. Rowe ML, Achhala S, Elphick MR: **Neuropeptides and polypeptide hormones in echinoderms: new insights from analysis of the transcriptome of the sea cucumber *Apostichopus japonicus***. *Gen Comp Endocrinol* 2014, **197**:43-55.
12. Semmens DC, Mirabeau O, Moghul I, Pancholi MR, Wurm Y, Elphick MR: **Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution**. *Open biology* 2016, **6**(2).
13. Kato S, Tsurumaru S, Taga M, Yamane T, Shibata Y, Ohno K, Fujiwara A, Yamano K, Yoshikuni M: **Neuronal peptides induce oocyte maturation and gamete spawning of sea cucumber, *Apostichopus japonicus***. *Dev Biol* 2009, **326**(1):169-176.
14. Conand C: **Present status of world sea cucumber resources and utilisation: an international overview**, vol. 463. Rome: Food and agriculture organization of the United nations; 2004.
15. Wen J, Hu C, Fan S: **Chemical composition and nutritional quality of sea cucumbers**. *Journal of the science of food and agriculture* 2010, **90**(14):2469-2474.
16. Bordbar S, Anwar F, Saari N: **High-value components and bioactives from sea cucumbers for functional foods--a review**. *Marine drugs* 2011, **9**(10):1761-1805.
17. Kiew PL, Don MM: **Jewel of the seabed: sea cucumbers as nutritional and drug candidates**. *International journal of food sciences and nutrition* 2012, **63**(5):616-636.
18. Conand C: **The Fishery Resources of Pacific Island Countries. Part 2. Holothurians**. Rome: Food and agriculture organization of the United nations; 1990.
19. Battaglene SC: **Culture of tropical sea cucumbers for stock restoration and enhancement**. *Naga* 1999, **22**(4):4-11.
20. Hamel J-F, Conand C, Pawson DL, Mercier A: **The sea cucumber *Holothuria scabra* (Holothuroidea: Echinodermata): Its biology and exploitation as Beche-de-mer**. In: *Advances in Marine Biology.* vol. Volume 41: Academic Press; 2001: 129-223.
21. Diaz-Miranda L, Price DA, Greenberg MJ, Lee TD, Doble KE, Garcia-Arraras JE: **Characterization of two novel neuropeptides from the sea cucumber *Holothuria glaberrima***. *The Biological Bulletin* 1992, **182**(2):241-247.
22. Diaz-Miranda L, Blanco RE, Garcia-Arraras JE: **Localization of the heptapeptide GFSKLYFamide in the sea cucumber *Holothuria glaberrima* (Echinodermata): a light and electron microscopic study**. *J Comp Neurol* 1995, **352**(4):626-640.
23. Diaz-Miranda L, Garcia-Arraras JE: **Pharmacological action of the heptapeptide GFSKLYFamide in the muscle of the sea cucumber *Holothuria glaberrima* (Echinodermata)**. *Comparative biochemistry and physiology Part C, Pharmacology, toxicology & endocrinology* 1995, **110**(2):171-176.
24. Iwakoshi E, Ohtani M, Takahashi T, Muneoka Y, Ikeda T, et al: **Comparative aspects of structure and action of bioactive peptides isolated from the sea cucumber Stichopus japonicus**. *Peptide Chemistry Osaka: Protein Research Foundation* 1995:261–264.
25. Ohtani M, Iwakoshi E, Muneoka Y, Minakata H, Nomoto K: **Isolation and characterization of bioactive peptides from the sea cucumber, *Stichopus japonicus***. In: *Peptide Science — Present and Future: Proceedings of the 1st International Peptide Symposium.* Edited by Shimonishi Y. Dordrecht: Springer Netherlands; 1999: 419-420.
26. Inoue M, Birenheide R, Koizumi O, Kobayakawa Y, Muneoka Y, Motokawa T: **Localization of the neuropeptide NGIWYamide in the holothurian nervous system and its effects on muscular contraction**. *Proceedings of the Royal Society B: Biological Sciences* 1999, **266**(1423):993-993.
27. Saha AK, Tamori M, Inoue M, Nakajima Y, Motokawa T: **NGIWYamide-induced contraction of tube feet and distribution of NGIWYamide-like immunoreactivity in nerves of the starfish *Asterina pectinifera***. *Zoological Science* 2006, **23**(7):627-632.
28. Smith MK, Wang T, Suwansa-ard S, Motti CA, Elizur A, Zhao M, Rowe ML, Hall MR, Elphick MR, Cummins SF: **The neuropeptidome of the Crown-of-Thorns Starfish, *Acanthaster planci***. *J Proteomics* 2017.
29. Zandawala M, Moghul I, Yañez Guerra LA, Delroisse J, Abylkassimova N, Hugall AF, O'Hara TD, Elphick MR: **Discovery of novel representatives of bilaterian neuropeptide families and reconstruction of neuropeptide precursor evolution in ophiuroid echinoderms**. *Open biology* 2017, **7**(9).
30. Elphick MR: **The protein precursors of peptides that affect the mechanics of connective tissue and/or muscle in the echinoderm *Apostichopus japonicus***. *PLOS ONE* 2012, **7**(8):e44492.
31. Mashanov VS, Zueva OR, Garcia-Arraras JE: **Transcriptomic changes during regeneration of the central nervous system in an echinoderm**. *BMC Genomics* 2014, **15**:357.

32. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nat Methods* 2011, **8**(10):785-786.
33. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S: **MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0**. *Mol Biol Evol* 2013, **30**(12):2725-2729.
34. Southey BR, Amare A, Zimmerman TA, Rodriguez-Zas SL, Sweedler JV: **NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides**. *Nucleic Acids Research* 2006, **34**(suppl 2):W267-W272.
35. Priyam A, Woodcroft BJ, Rai V, Munagala A, Moghul I, Ter F, Gibbins MA, Moon H, Leonard G, Rumpf W *et al*: **Sequenceserver: a modern graphical user interface for custom BLAST databases**. *bioRxiv* 2015.
36. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res* 2003, **13**(11):2498-2504.
37. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M *et al*: **Phylogeny.fr: robust phylogenetic analysis for the non-specialist**. *Nucleic Acids Res* 2008, **36**(Web Server issue):W465-469.
38. Delroisse J, Mallefet J, Flammang P: ***De novo* adult transcriptomes of two european brittle stars: Spotlight on opsin-based photoreception**. *PLoS One* 2016, **11**(4):e0152988.
39. Mita M, Yoshikuni M, Ohno K, Shibata Y, Paul-Prasanth B, Pitchayawasin S, Isobe M, Nagahama Y: **A relaxin-like peptide purified from radial nerves induces oocyte maturation and ovulation in the starfish, *Asterina pectinifera***. *Proc Natl Acad Sci U S A* 2009, **106**(23):9507-9512.
40. Veenstra JA: **The contribution of the genomes of a termite and a locust to our understanding of insect neuropeptides and neurohormones**. *Frontiers in Physiology* 2014, **5**:454.
41. Kwon H, Lu HL, Longnecker MT, Pietrantonio PV: **Role in diuresis of a calcitonin receptor (GPRCAL1) expressed in a distal-proximal gradient in renal organs of the mosquito *Aedes aegypti* (L.)**. *PLoS One* 2012, **7**(11):e50374.
42. Dubos MP, Badariotti F, Rodet F, Lelong C, Favrel P: **Molecular and physiological characterization of an invertebrate homologue of a calcitonin-related receptor**. *Biochem Biophys Res Commun* 2003, **310**(3):972-978.
43. Zandawala M, Li S, Hauser F, Grimmelikhuijzen CJP, Orchard I: **Isolation and functional characterization of calcitonin-like diuretic hormone receptors in *Rhodnius prolixus***. *PLoS ONE* 2013, **8**(11):e82466.
44. Copp DH, Cameron EC, Cheney BA, Davidson AG, Henze KG: **Evidence for calcitonin--a new hormone from the parathyroid that lowers blood calcium**. *Endocrinology* 1962, **70**:638-649.
45. Pondel M: **Calcitonin and calcitonin receptors: bone and beyond**. *International Journal of Experimental Pathology* 2000, **81**(6):405-422.
46. Goodman EC, Iversen LL: **Calcitonin gene-related peptide: novel neuropeptide**. *Life Sci* 1986, **38**(24):2169-2178.
47. Schou WS, Ashina S, Amin FM, Goadsby PJ, Ashina M: **Calcitonin gene-related peptide and pain: a systematic review**. *The journal of headache and pain* 2017, **18**(1):34.
48. Ohtani M, Iwakoshi E, Muneoka Y, Minakata H, Nomoto K: **Isolation and characterization of bioactive peptides from the sea cucumber, *Stichopus japonicus***. In: *Peptide Science — Present and Future: Proceedings of the 1st International Peptide Symposium.* Edited by Shimonishi Y. Dordrecht: Springer Netherlands; 2002: 419-420.