Similarities between decapod and insect neuropeptidomes

Jan A Veenstra

Background. Neuropeptides are important regulators of physiological processes and behavior. Although they tend to be generally well conserved, recent results using trancriptome sequencing on decapod crustaceans give the impression of significant differences between species, raising the question whether such differences are real or artefacts.

Methods. The BLAST+ program was used to find short reads coding neuropeptides and neurohormons in publicly available short read archives. Such reads were then used to find similar reads in the same archives and the DNA assembly program Trinity was employed to construct contigs encoding the neuropeptide precursors as completely as possible.

Results. The seven decapod species analyzed in this fashion, the crabs *Eriocheir sinensis, Carcinus maenas* and *Scylla paramamosain,* the shrimp *Litopenaeus vannamei,* the lobster *Homarus americanus,* the fresh water prawn *Macrobrachium rosenbergii* and the crayfish *Procambarus clarkii* had remarkably similar neuropeptidomes. Although some neuropeptide precursors could not be assembled, in many cases individual reads pertaining to the missing precursors show unambiguously that these neuropeptides are present in these species. In other cases the tissues that express those neuropeptides were not used in the construction of the cDNA libraries. One novel neuropeptide was identified, elongated PDH (pigment dispersing hormone), a variation on PDH that has a two amino acid insertion in its core sequence. Hyrg is another peptide that is ubiquitously present in decapods and is likely a novel neuropeptide precursor.

Discussion. Many insect species have lost one or more neuropeptide genes, but apart from elongated PDH and hyrg all other decapod neuropeptides are present in at least some insect species and allatotropin is the only insect neuropeptide missing from decapods. This strong similarity between insect and decapod neuropeptidomes makes it possible to predict the receptors for decapod neuropeptides that have been deorphanized in insects. This includes the androgenic insulin like peptide that seems to be homologous to drosophila insulin-like peptide 8.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

Similarities between decapod and insect neuropeptidomes

Jan A. Veenstra[1]
[1] INCIA, UMR 5287 CNRS, Université de Bordeaux, Pessac, France

Corresponding Author:
Jan Veenstra, INCIA UMR 5287 CNRS, Université de Bordeaux, allée Geoffroy St Hillaire, CS 50023, 33 615 Pessac Cedex, France
Email address: jan.veenstra@u-bordeaux.fr

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

Abstract

Background. Neuropeptides are important regulators of physiological processes and behavior. Although they tend to be generally well conserved, recent results using trancriptome sequencing on decapod crustaceans give the impression of significant differences between species, raising the question whether such differences are real or artefacts.

Methods. The BLAST+ programm was used to find short reads coding neuropeptides and neurohormons in publicly available short read archives. Such reads were then used to find similar reads in the same archives and the DNA assembly program Trinity was employed to construct contigs encoding the neuropeptide precursors as completely as possible.

Results. The seven decapod species analyzed in this fashion, the crabs *Eriocheir sinensis, Carcinus maenas* and *Scylla paramamosain,* the shrimp *Litopenaeus vannamei,* the lobster *Homarus americanus,* the fresh water prawn *Macrobrachium rosenbergii* and the crayfish *Procambarus clarkii* had remarkably similar neuropeptidomes. Although some neuropeptide precursors could not be assembled, in many cases individual reads pertaining to the missing precursors show unambiguously that these neuropeptides are present in these species. In other cases the tissues that express those neuropeptides were not used in the construction of the cDNA libraries. One novel neuropeptide was identified, elongated PDH (pigment dispersing hormone), a variation on PDH that has a two amino acid insertion in its core sequence. Hyrg is another peptide that is ubiquitously present in decapods and is likely a novel neuropeptide precursor.

Discussion. Many insect species have lost one or more neuropeptide genes, but apart from elongated PDH and hyrg all other decapod neuropeptides are present in at least some insect species and allatotropin is the only insect neuropeptide missing from decapods. This strong similarity between insect and decapod neuropeptidomes makes it possible to predict the receptors for decapod neuropeptides that have been deorphanized in insects. This includes the androgenic insulin like peptide that seems to be homologous to drosophila insulin-like peptide 8.

Introduction

Lobsters, crayfish, prawns, crabs and shrimps are all crustaceans belonging to the order of the decapods. Many of these species are part of the human diet, are sometimes a major source of protein and are often considered a delicacy. While some species are caught in the wild, others, such as *e.g.* the freshwater prawn *Macrobrachium rosenbergii,* are mainly obtained from commercially important cultures. Many of these species are also sufficiently large to allow physiological experiments that are more difficult to perform on smaller arthropods. For these reasons decapods probably constitute the second best studied group of arthropods after insects. Neuropeptides have also been extensively researched in decapods and several neuropeptides were initially identified in these crustaceans before they were found in other arthropods such as insects (*e.g.* Kegel et al., 1989; Stangier et al., 1987, 1992).

The sizes of their genomes tend to be large (*e.g.* Yu et al., 2015; Song et al., 2016) and so far no complete decapod genome is available. Initially sequences of crustacean neuropeptides were determined by classical peptide isolation and Edman degradation (*e.g.* Kegel et al., 1989; Stangier et al., 1987, 1992; Bungart et al., 1995; Duve et al., 1997), but in the last decade numerous decapod peptides have been sequenced by mass spetrometry (*e.g.* Dickinson et al., 2008, 2009a,b; Ma et al., 2008, 2009, 2010; Stemmler 2007a,b, 2010). In the last two years identification of the decapod neuropeptidomes has further accelerated using next-generation sequencing methodology. As a consequence we now have fairly long lists of neuropeptides for several decapods. These include *Sagmariasus verrauxi* (Ventura et al., 2014), *Macrobrachium rosenbergii* (Suwansa-Ard et al., 2015), *Procambarus clarkii* (Veenstra, 2015), *Scylla paramamosain* (Bao et al., 2015) and *Homarus americanus* (Christie et al, 2015), while for other decapods significant amounts of data are available to analyze their neuropeptidomes. This is for the example the case for *Carcinus maenas, Litopenaeus vannamei* and *Eriocheir sinensis* (Li et al., 2012; Ghaffari et al., 2014; Verbruggen et al., 2015; Xu et al., 2015) . Some of the ESTs (expressed sequence tags) present in the publicly available databases have been summarized by Christie and his collaborators (Ma et al., 2009, 2010; Christie, 2014; Christie & Chi, 2015).

I have previously used the published short read archives for *Procambarus* to look for neuropeptide transcripts and could deduce complete or partial sequences for a surprisingly large number of neuropeptide precursors (Veenstra, 2015). When comparing the results obtained in

this species, with the lists of neuropeptide transcripts from other decapods, several differences appear. While several neuropeptides are consistently found in all species, others are only identified in some. The question is whether these differences are real or represent artefacts. For example, some peptides may not have been searched for in the assembled reads, or there were simply too few reads to allow assembly of a contig, while in other cases the tissue where the particular gene is predominantly expressed was perhaps not included in the analysis. I here try to answer these questions by reanalyzing published short sequence read archives for a number of decapods.

Materials & Methods

DNA sequences

The following short read archives (SRAs) were downloaded from NCBI using the prefetch command from the SRA Toolkit (http://www.ncbi.nlm.nih.gov/books/NBK158900/) : for *Carcinus maenas*: SRR1564428, SRR1572181, SRR1586326, SRR1589617, SRR1612556, SRR1632279, SRR1632285, SRR1632289, SRR1632290, SRR1632291, SRR1632292 and SRR1632293 (Verbruggen et al., 2015); for *Procambarus clarkii*: SRR1144630, SRR1144631, SRR1265966, SRR1509456, SRR1509457, SRR1509458 and SRR870673 (Jiang et al., 2014; Tom et al., 2014; Shen et al., 2014; Manfrin et al., 2015); for *Macrobrachium rosenbergii*: DRR023219, SRR1559288, SRR345608, SRR572725, DRR023253, SRR1653452, SRR345609, SRR896637, SRR1138560, SRR1653453, SRR345610, SRR896638, SRR1138561, SRR1653454, SRR345611, SRR896645, SRR1138562, SRR567391, SRR896646, SRR1138563, SRR572719, SRR896647, SRR1138564, SRR572720, SRR896649, SRR1138565, SRR572721, SRR896650, SRR1138572, SRR2082768, SRR572722, SRR896651, SRR1138573, SRR2082769, SRR572723, SRR1559287, SRR2082770, SRR572724 (Jung et al., 2011; Ventura et al., 2013; Suwansa-Ard et al., 2015); for *Scylla paramamosain*: SRR1310332, SRR1310333, SRR1205999, SRR3086589, SRR834579, SRR1206015, SRR3086590, SRR834580, SRR1310331 and SRR3086592 (Gao et al., 2014; Ma et al., 2014; Bao et al., 2015); for *Litopenaeus vanamei*: SRR1037362, SRR1407789, SRR1460505, SRR1952625, SRR2103853, SRR2103860, SRR2895158, SRR1037363, SRR1104812, SRR1407790, SRR1609917,

SRR2060962, SRR2103854, SRR2103861, SRR346404, SRR1037364, SRR1105791, SRR1407791, SRR1618514, SRR2060963, SRR2103855, SRR2103862, SRR554363, SRR1037365, SRR114084, SRR114085, SRR1460493, SRR1951370, SRR2060964, SRR2103856, SRR2103863, SRR554364, SRR1037366, SRR1184416, SRR1460494, SRR1951371, SRR2060965, SRR2103857, SRR2103864, SRR554365, SRR1039534, SRR1407787, SRR1460495, SRR1951372, SRR2103851, SRR2103858, SRR2103865, SRR556131, SRR1104083, SRR1104080, SRR1104086, SRR1104087, SRR1407788, SRR1460504, SRR1951373, SRR2103852, SRR2103859 and SRR2103866 (Li et al., 2012; Chen et al., 2013; Wei et al., 2014; Gao et al., 2015; Peng et al., 2015); for *Eriocheir sinensis*: ERR336998, SRR1555734, SRR2170964, SRR579530, SRR1199039, SRR1576649, SRR2170970, SRR579531, SRR1199053, SRR1735503, SRR2180019, SRR579532, SRR1199058, SRR1735536, SRR2180020, SRR769751, SRR1199228, SRR1735537, SRR546086, SRR770582, SRR1205971, SRR2073826 and SRR579529 (He et al., 2012; Hui et al., 2014; Li et al., 2014; Sun et al., 2014; Liu et al., 2015; Xu et al., 2015; Cui et al., 2015; Song et al., 2015; Wang et al., 2016); and for *Homarus americanus*: SRR2889572 and SRR2891007 (Christie et al., 2015). From *Euphausia crystallorophias* I analyzed ERR264582 (Toullec et al., 2013) for the presence of a novel putative neuropeptide that was found in the decapod transcriptomes.

The *Eriocheir sinensis* genome was downloaded from http://gigadb.org/dataset/100186, made into a BLAST database and searched for neuropeptide genes as described previously (Veenstra, 2014).

Data analysis

The fasta files were extracted from the SRAs using the fastq command from the SRA Toolkit from NCBI and then made into BLAST databases using BLAST+ (Camacho et al., 2009). Using the *Procambarus* predicted neuropeptide precursors as well as a few other peptides as queries those databases were then searched using the tblastn command. A few neuropeptide receptors were also analyzed. Identified reads that appeared to belong to the orthologous gene were extracted from the database and then used to identify similar reads using the blastn command. The latter were used as input for the Trinity program (Grabherr et al., 2011) and resulting transcripts were recursively used as input until either the transcript stopped increasing

in length or it was judged to be complete based on the location of in-frame stop codons and/or a signal peptide at the N-terminal of the protein predicted from the transcript. Calculations were run on a desktop computer with a AMD FX(tm)-6100 six-core processor and 15.4 Gb of memory under Ubuntu Linux.

This method is very efficient for the extraction of transcripts from single copy genes. However, when there are several paralog genes that have not evolved a lot since their separation, some paralogs may be missed, particularly when their expression levels are low. In those cases, a selection of the particular neuropeptide precursors from which the non-conserved regions (such as the signal peptides) had been removed was used as a query in a tblastn command and all the obtained reads were then fed as input to the Trinity program. It can not be excluded that some less well expressed paralogs of those genes that exist in multiple copies (neuroparsin, CHH (crustacean hyperglycemic hormone), PDH (pigment dispersing hormone) and possibly CFSH (crustacean female sex hormone) have been missed.

Clustal Omega (Sievers et al., 2011) was used for sequence alignments and those were inspected and when needed manually corrected using Seaview (Gouy, Guindon & Gascuel, 2010), which was also used to extract the regions for making phylogenetic trees with FastTree (Price, Dehal & Arkin, 2010).

Results

Trinity is a fantastic tool to reconstruct large DNA sequences from very short reads. However, not every sequence corresponds necessarily to a correct cDNA sequence or is biologically interesting. One regularly finds more than one sequence derived from the same gene. In the absence of a genomic sequence, as is the case here, it is not always possible to determine which is the correct one. There are several common causes for the failure to produce a single complete cDNA sequence. First, there may simply be insufficient reads available to produce a complete contig. Secondly, there may be allelic variation that causes the elongation to stop. Thirdly, alternative splicing, as is the case for genes encoding the agatoxin-like peptide, Neuropeptide F 1, CNMamide, calcitonin and CHH, may have the same effect. Fourth, recombining short sequences into a long one becomes very difficult in the case of repetitive sequences. One or more reads containing a sequencing error can aggravate some of the other problems, *i.e.* lack of sufficient reads, alternative splicing or allelic variants.

Most of the data analyzed here come from natural or almost natural populations that show much larger genetic variation than that found in the typical laboratory animals like mice or rats. Furthermore, many neuropeptide genes code for a number of highly similar neuropeptide paracopies and this makes it no doubt difficult to reconstruct the complete cDNA encoding such precursors and when the various copies are only separated by convertase cleavage sites, the problem may become acute. In one attempt to produce the *Eriocheir* FMRFamide precursor mRNA Trinity produced a partial transcript that had a perfect internal repeat of 164 nucleotides (Fig. S1), that must have been an artefact; a similar phenomenon is also present in the second predicted orcokinin precursor from *Scylla* (Bao et al., 2015; Fig. S1). Furthermore, I have previously shown that some neuropeptide genes have alleles that differ in the number of neuropeptide paracopies that they encode (Veenstra, 2010a; Veenstra, 2015). It is therefore not surprising that a relatively large number of transcripts for neuropeptide precursors containing multiple paracopies, such as FMRFamide, tachykinin, leucokinin, EFLamide etc, are incomplete even though significant numbers of individual reads are found in the various SRAs. Predictions by Trinity of neuropeptide precursors containing various paracopies may, for the same reasons, contain errors. For example, the allatostatin A precursor from *Carcinus* does not code for some of the previously identified peptides from this species (Duve et al., 1997), while the Trinity transcripts of several other neuropeptide precursor sequences from the same species that have been obtained by screening of classical cDNA libraries are identical (Klein et al., 1992, 1993; Linck et al., 1993; Chung et al., 2006; Wilcockson and Webster 2008). Other transcripts that are incomplete are often due to low expression levels.

While this work was in progress a draft genome for *Eriocheir sinensis* was published (Song et al., 2016). This sequence was prepared using short sequence reads and therefore suffers from the problems associated with this methodology (Richards and Murali, 2015). It is estimated that about 67% of total sequence is present in the current draft. Several of the transcripts identified here are not at all or only partially present in this genome and different exons of the same transcripts are regularly found on different contigs. Its usefulness was, therefore, limited.

The decapod neuropeptide genes that were found are indicated in Fig. 1, where for comparison the presence of neuropeptide genes of *Daphnia pulex*, a crustacean, and two insects, the termite *Zootermopis nevadensis* and the fruit fly *Drosophila melanogaster*, is also shown. Many of the neuropeptide precursor transcripts seem complete, at least as far at the coding region

is concerned, while for others very significant parts were found. Since one of the questions raised here is the presence of a particular neuropeptide gene, I have also added neuropeptide genes for which individual reads from an SRA provide evidence for its existence in the particular species, even though Trinity produced no contigs for transcripts from these genes. All the sequences, both DNA and deduced amino acids, are listed in Tables S1-S8 in the supplementary excel file.

Distribution.

Having all the SRAs it seemed interesting to look at where the various genes might be expressed. Although it is possible to do this for all species involved, some are not very interesting as there is a very limited number of tissues sampled, while in other species the different tissues were sampled on different occasions and analyzed in different fashions, making direct comparisons difficult. However, in the case of *Carcinus* a single publication reports SRAs for a large variety of tissues (Verbruggen et al., 2015). Therefore, I used this species to look at the expression of the various neuropeptide genes in different tissues. Those neuropeptide receptors for which a contig of a significant size could be obtained and for which a likely ligand could be deduced based on homology to a deorphanized protostomian GPCR (see Veenstra, 2016) were also include. Even though the actual number of individual reads is often small and quantification of RNAseq reads is tricky due to the PCR amplification protocol used to create these libraries, some interesting data are apparent (Table 1). Both the neuroparsins and the CHHs are expressed in virtually every tissue. In the case of the neuroparsin it is the neuroparsin 1 gene that is most abundantly expressed in all tissues, with the other two neuroparsin transcripts present at much lower levels. However, the two identified CHH transcripts are differentially expressed, one hormone is most abundant in the central nervous system and the eye (eyestalk), and the other in the intestine. Other neuropeptides found in the intestine are tachykinin, allatostatin C, the B transcript of the calcitonin gene, elevenin, orcokinin, the agatoxin-like peptide and hyrg. The expression of CCHamide 2 and trissin in the hemolymph, presumably in hemocytes, is interesting to note as is the relatively large number of neuropeptides found in the SRA derived from the epidermis.

Paralogs

There are several neuropeptide genes that have one or more paralog genes. These are

allatostatin C, CHH, moult inhibiting hormone (MIH), CCHamide, eclosion hormone, neuroparsin, PDH, insulin and CFSH. In some cases these are sufficiently different within the same species and sufficient similar between different species, that they clearly derive from different genes. This is the case for allatostatin C, CCHamide, insulin, neuropeptide F and eclosion hormone.

In the case of PDH, it is a bit more complicated. Variable numbers of PDH precursors were found in the seven decapod species. One group of precursors encoding PDH-like peptides distinguishes itself by an insertion of two amino acid residues in the predicted mature PDH. Such a predicted peptide was first found in *Procambarus* (Veenstra, 2015), but since it was based on a single read in one species, it seemed premature to give a distinct name. Now that complete precursor sequences are available and this peptide appears to be ubiquitously present in decapods, I propose to call it elongated PDH, or ePDH, to distinguish it from the more classical forms of these peptides (Fig. 2). The ePDH gene is one of the few genes that is present on a single contig of the draft genome from *Eriocheir*. It consists of three exons of which the first one is non-coding (Fig. 3). Partial sequence for one of the classical PDH genes show the intron between the two coding exons to be conserved.

In the case of neuroparsin, PDH, CHHs and its homolog MIH it is not always as clear that they represent different genes with unambiguous orthologs in different species. In some cases the observed differences could reflect allelic variations of a single gene or recent local gene duplications. Although no decapod genomes have been completely sequenced and the *Eriocheir* CHH genes are mostly very fragmentary, such local gene duplications are well known for CHH in decapods (Gu & Chan, 1998; Gu, Yu & Chan, 2000; Dircksen et al., 2001; Webster et al., 2012) as well as Chelicerates (Veenstra, 2016) and particularly in decapods the number of CHH genes can be quite large (Webster, Keller & Dircksen, 2012).

CCH/MIH

The CCH/MIH neuropeptide family is characterized by CHH, MIH, mandibular organ-inhibiting hormone (MOIH), vitellogenesis-inhibiting hormone (VIH) and gonad-inhibiting hormone (GIH). These hormones have been identifed by different physiological assays, but are in many cases pleiotropic. These peptides can be subdivided in two subfamilies, the CCHs proper and the other peptides. The precursors from the two groups differ in that the CCHs are

produced together with a CCH-precursor related peptide, while the prepropeptides from the other homornes consist exclusively of a signal peptide and the sequence of the mature hormone (Webster, Keller & Dircksen 2012). Three of the CHH/MIH transcripts identified here defy those rules as they do not have the CCH-precursor related peptide, yet on phylogenetic trees they form a separate branch that is closer to the the CHH than to the MIH cluster (Fig. 4). Adding more sequences to the tree does not change this (data not shown). In the *Eriocheir* draft genome many sequences corresponding to these hormones are located on small scaffolds making it impossible to ascertain whether or not these genes are clustered.

CFSH

The CFSH is a recent discovery (Zmora & Chung, 2014) and consequently we still know very little of this very interesting hormone. In *Procambarus* three related proteins were previously identified (Veenstra, 2015). In five of the other six decapod species two to four such hormones were found, but not in *Homarus*, where there are no ovary transcriptomes. The primary sequence of these different putative hormones is not very well conserved, but the cysteine residues are (Fig. 5). The phylogenetic tree of these hormones suggests an initial gene duplication giving rise to two types of CFSH, that I have arbitrarily called CFSH 1 and 2 (Fig. 6). In most species both CFSH 1 and 2 were found, but in in *Litopenaeus* only three CFSH 1 paralogs were found and no CFSH 2. In the draft genome of *Eriocheir* CFSH 1 and 2b genes contain a single coding exon. The CFSH gene 2a transcript is incomplete and it is not clear from the genomic sequence what it is. This hormone was initially isolated from the eyestalk of the crab *Callinectes* (Zmora & Chung, 2014), while it in the crayfish *Procambarus* it seemed to be strongly expressed in the ovary. It seemed therefore of interest to see whether these hormones might be expressed in the ovary in other decapods also. No significant expression was found in the ovaries of *Macrobrachium* and *Litopenaeus* [1 to 2 reads maximum for each hormone in an SRA], but 9 reads corresponding to *Litopenaeus* CFSH 1c (as well as 1 each for 1a and 1b) are present in SRR2060962 from the *Litopenaeus* testis. In *Eriocheir* expression is similar to that in *Carcinus* (Table 1), high expression levels in the eyestalk and a few reads only in the ovary. For *Scylla* and *Homarus* there are insufficient data to answer this question.

Neuroparsins and receptors

Three to four neuroparsin transcripts were identified in each of the seven decapod species. Three of the *Eriocheir* genes were found in the draft genome, two of these (neuroparsins 3 and 4) are on the same scaffold in a tail to tail configuration, where the start and stop codons of the two genes are separated by 11 960 and 9 045 nucleotides respectively (Fig. 7). These are the two *Eriocheir* genes that are most similar to one another, suggesting that they may reflect the most recent neuroparsin gene duplication in this species. As both these genes seem to have direct orthologs in *Scylla* and *Caracinus*, that particular gene duplication possibly occurred in a common ancestor of the three crab species (Fig. 8). The neuroparsin receptor was recently identified as a venus kinase receptor (Vogel et al., 2015); two such receptors are found in all seven decapod species (Table S8). The phylogenetic tree made of the various venus kinase receptors suggests that the other arthropods venus kinase receptors are equally similar to both decapod receptors (Fig. 9).

Insulin-like peptides and receptors

Three different insulin-related peptides were identified. These are the well known androgenic insulin-related peptide (Fig. 10), an insulin-like peptide (Fig. 11) that seems most similar to the *Drosophila* insulin-like peptides 1-6 (Nässel & vanden Broeck, 2015), and a peptide that is orthologous to *Drosophila* insulin-like peptide 7 and that has been called relaxin (Fig. 12). The latter was previously identified in *Sagmariasius* and *Procambarus* (Chandler et al., 2014; Veenstra, 2015). As can be seen from the figures, the androgenic insulin-like peptide is the least conserved of those three (Figs. 10-12). Insulin-related peptides use two different types of receptors, the typical tyrosine kinase receptor and GPCRs. Insects typical have one gene coding an insulin tyrosine kinase receptor and have one or two GPCRs that are related to the vertebrate relaxin receptors RXFP1 and RXFP2. Given the interest in the androgenic insulin-like peptide both for its intriguing physiology as a peptidergic sex hormone and for its commercial potential (Ventura & Sagi, 2012), I also analyzed the likely insulin receptors.

The typical insulin tyrosine kinase receptor, similar to the one recently described from *Macrobrachium* (Sharabi et al., 2016), was also found in the other six decapods (Table S8). Two receptors similar to the vertebrate relaxin receptors RXFP1 and RXPF2 were also identified. Those two GPCRs are most similar to the *Drosophila* receptors CG31096 and CG34411, also known as leucine-rich repeat containing GPCR- 3 and 4 (LGR3 and LGR4) respectively.

However, the much weaker expression of those receptors made it impossible to deduce their complete cDNA sequences and in some cases no contigs could be obtained. Interestingly the SRA from the *Eriocheir* accesory gland (SRR2170964) not only shows very large number of reads for the androgenic Insulin-like peptide, but also very significant expression of the insulin tyrosine kinase receptor and a somewhat lower expression of the ortholog of *Drosophila* LGR3.

Splice variants

There were a number of neuropeptide encoding cDNAs that revealed splice variants. Those that concerned the untranslated regions were ignored, but there are five neuropeptide genes that have alternative transcripts producing different precursors: the CHHs, CNMamide, neuropeptide F 1, calcitonin and the agatoxin-like peptide. In the case of neuropeptide F 1, there is an extra exon sliced into the sequence of the peptide, as described previously from insects (Roller et al., 2008; Nuss et al., 2010; Dircksen et al., 2011). The CNMamide gene in the termite *Zootermopsis* contains five coding exons, the last two of which are alternatively added to the first three and then produces a different CNMamide-like peptide. In four of the seven decapods similar alternative splice products were found for the CNMamide precursor. However, while the mature peptide derived from the major splice form is well conserved, the second is much less so (Fig. 13). Two to four splice variants (Fig. S2) were found for the recently discovered μ-agatoxin-like peptide (Sturm et al., 2015). As in some insects (Veenstra, 2014), the calcitonin gene produces two different transcripts, encoding different types of calcitonin, that are similar to the insect calcitonins (Fig. S3). In *Litopenaeus, Macrobrachium, Homarus* and *Procambarus* the second transcripts are predicted to produce a calcitonin-like peptide that does not have one but two cysteine bridges at is N-terminus (Fig. 14). The calcitonin gene is absent from the *Eriocheir* draft genome, and hence it is impossible to compare the insect and decapod calcitonin gene structures.

Other peptides.

In several cases novel neuropeptides have been detected by mass spectrometry. These are often structural variants of well known neuropeptides such as the RFamides, tachykinins and allatostatins A or B (*e.g.* Ma et al., 2008, 2009, 2010). However, not all peptide sequences identified this way belong to known neuropeptides. From *Homarus*, *Carcinus* and *Litopenaeus* other peptide sequences have been reported. The ones from *Carcinus* have previously been

suggested to represent fragments of cryptocyanin (Ma et al., 2009), and this was confirmed (Fig. S4). Several of the peptides from *Homarus* are shown here to represent fragments of thymosin, actin or histone 2A, however the origins of others remain unclear (Fig. S4). The one peptide reported from *Litopenaeus*, L/IPEPEDPMAEAGHEL/I (Ma et al., 2010), is more interesting, as it could potentially be part of a novel neuropeptide (precursor). This sequence is part of a small protein that has a signal peptide followed by a peptide containing a small piece that is very well conserved (Fig. 15). However, it lacks the classical convertase cleavage sites that one usually finds in neuropeptide precursors and hence its status as a neuropeptide is unclear. Such proteins are also found in the other decapods. Although it was not possible for Trinity to produce a complete contig for *Scylla paramosain*, a similar sequence is present in the databanks for *S. olivacea.* I was unable to find similar proteins in insects, but an orthologous protein was detected in the SRA from *Euphausia crystallorophias*. The latter sequence shows that the only conserved part is the same as in the decapods (Fig. 15). This peptide was called hyrg (pronounced hirg), for four of the conserved amino acids. Interestingly, the eyestalk seems to be the tissue where expression of hyrg is the highest (Table 1), thus suggesting that it is likely a neurohormone.

Discussion

Insects and decapods are estimated to have had their last common ancestor about 596 Mya, while similar estimates for the common ancestor of crabs and lobsters on the one hand and that for termites and flies on the other are 322 and 348 Mya respectively (Hedges et al., 2015). The gross morphology of decapods has changed a lot less than that of insects and when one compares the respective neuropeptidomes of those two groups, it is clear that those are similarly much better conserved in decapods than in insects (Fig. 1). Most of the changes in insects are losses of neuropeptides that are particularly pronounced in flies, and perhaps even more so in *D. melanogaster*.

Whenever in this study a particular gene has not been identified from a decapod species, either one of the following is true: (1), the gene is not expressed at high levels and there are relatively small amounts of RNAseq reads, (2) the gene is expressed predominantly in tissues that have not been sampled in the species in question or (3) the gene has several paralog genes (PDH, CHH, neuroparsin) and it may not have the same number of paralogs in all species and/or some of those paralogs may be expressed in tissues that were not sampled. A combination of (1)

and (2) likely explains the absence of some of the neuropeptide genes in *Scylla.* From that species the eyestalk was not analyzed, even though this tissue is by far the richest source of neuropeptides. Nevertheless, the existence of several *Scylla* neuropeptide genes could be inferred from individual RNAseq reads, while the few genes that are completely lacking are only weakly expressed in the other species. The androgenic peptide was found neither in *Carcinus* nor *Homarus*. As in *Homarus* only the nervous system was included in the analysis, this is to be expected. In the case of *Carcinus*, it is plausible that the testis samples did not include the accessory gland, or that the sample was taken at a moment in the life cycle of the animal that expression of this peptide is low or non-existent. With one exception, in all other instances where a transcript seems to be lacking it is either from a gene for which an alternative transcript was found (*e.g.* in the case of the CNMamide and Neuropeptide F 1 genes), or the number of paralogs may differ between the various species (neuroparsin, CHH, MIH, PDF). The only exception is MIH in *Homarus*. Although this peptide has been reported by mass spectrometry from the stomatogastric ganglion of *H. americanus* (Ma et al., 2008) and in spite of using this and the MIH sequence of the closely related species *H. gammarus* (Ollivaux, 2006) as queries in the BLAST program, no MIH transcript was found in the *H. americanus* SRAs.

Neuropeptide evolution

It thus appears that the neuropeptidome of decapods has been remarkably well conserved during evolution. Differences that are found between the insect and decapod neuropeptidomes are the loss or the gain of a neuropeptide. Although there possibly still remain arthropod neuropeptides to be discovered, it appears that the loss of neuropeptides in decapods is limited to a single neuropeptide, *i.e.* allatotropin. Allatotropin is present in mollusks, annelids as well as chelicerates (Veenstra 2010a, 2011, 2016) and hence, it must have been present in the arthropod ancestor. Small peptides are sometimes hard to find using the BLAST program and allatotropin is no exception to this rule (Veenstra, Rodriguez & Weaver, 2012). Nevertheless, as I was neither able to find even a single read correponding to its receptor, including in the very abundant number of transcriptome reads from *Homarus*, I conclude that this peptide was most likely lost. In the termite and the fruit fly on the other hand, more neuropeptides are missing, particularly in *Drosophila*. At first sight insects, as a group, lack EFLamide, the androgenic insulin-like hormone, CFSH and ePDH. However, the recent identification of an EFLamide

receptor in *Platynereis dumerlii* as a TRH GPCR ortholog (Bauknecht & Jékely, 2015) and the presence of such a GPCR in *Nilaparvata lugens* (Tanaka et al., 2014) suggests that some insects may have such a gene. As described below, it is plausible that the androgenic peptide has an insect ortholog. What seems really different is that many insect species, in particular holometabolous species, have lost several neuropeptides (Derst et al., 2016). Thus *Drosophila* no longer has genes for elevenin, vasopressin, allatotropin, allatostatin CCC, EFLamide, neuroparsin, calcitonin, ACP, eclosion hormone 2, neuropeptide F 2 and it also lost the possibility to produce alternative transcripts from the CNMamide and neuropeptide F1 genes. The beetle *Tribolium castaneum* on the other hand still has most of those neuropeptides, but lost allatostatin A, corazonin and leucokinin.

New neuropeptides

Since the last common ancestor of decapods and insects - estimated to have lived 596 Mya (Hedges et al., 2015) - very few neuropeptides seem to have been added to either of the two lineages. Novel neuropeptide genes that have appeared seem all to have originated by duplication from existing ones and are easily recognized as the paralogs of the parent genes. Examples of such genes are the various paralogs of CHH and MIH, PDH and neuroparsin in crustaceans and in insects the typtopyrokinin and SIFamide paralogs as well as the great variety of adipokinetic hormones (all orthologs of crustacean RPCH). The only exception may be hyrg, the precursor for the peptide initially identified from *Litopenaeus* (Ma et al., 2010). This peptide, that is well expressed in the eyestalk and the midgut, has a distribution typical of a neuroendocrine peptide. As I was unable to find it outside of crustaceans, it could be a novel invention of this group. The structure of this putative neuropeptide precursor is somewhat reminiscent of limostatin, a small neuroendocrine protein discovered in *Drosophila* that intereacts with a GPCR (Alfa et al., 2015) previously identified as the receptor for neuropeptide pyrokinin 1 (Cazzamali et al., 2005). The similarity between limostatin and hyrg resides in the apparent absence of conventional convertase sites in these putative neuropeptide precursors [those postulated to function in the limostatin precursor (Alfa et al., 2015) seem highly unusual (Veenstra, 2000)]. In the same context the *Drosophila* sex peptide comes to mind, as it also acts on a neuropeptide receptor without having neither a well conserved structure nor the typical neuropeptide convertase cleavage sites (Kim et al., 2010). Perhaps one or more of these proteins represent newly evolved

ligands for existing neuropeptide receptors that could potentially become novel neuropeptides.

Missing neuropeptides

Many decapod neuropeptides have been identified by mass spectrometry over the years (*e.g.* Stemmler et al., 2007a,b; 2010; Christie et al., 2008; Ma et al., 2008, 2009, 2010; Dickinson et al., 2008, 2009a,b). Most of those were identified in the various SRAs, although not always in exactly the same molecular form. In particular I was unable to find some of the analogs of SIFamide that have been reported (*e.g.* Hui et al., 2012). I could neither find [Val^1]-SIFamide in any species, however this peptide seems to be present in the stomatogastric nervous system (Christie et al., 2006) and this might explain its absence from the various SRAs. Several of the peptides previously described from these data that did not appear to be neuropeptides could be identified as being part of well known proteins and it also allowed me to identify the hyrg trancript. However, there are three neuropeptides that either have been reported or suggested to be present in decapods that were not found in any of the SRAs from the seven decapod species studied here. These are a pituitary adenylate cyclase activating polypeptide (PACAP) from *Litopenaeus vannamei* (Lugo et al., 2013), a GnRH-like peptide from *Procambarus clarkii* (Guan et al., 2014) and two kisspeptins from *Macrobrachium rosenbergii* (Thongbuakaew et al., 2016). None of these peptides could be found in any of the SRAs, neither those from the species from which they were reported, nor from any of the other species. In two cases (PACAP and GnRH), the amino acid sequences of the peptides have been published from the same species used here, so my inability to find these peptides is not due to significant sequence differences between the species used for bioinformatic analysis and those from which the peptides were identified. I was neither able to find evidence for the receptors for such peptides in any of decapods. The GnRH receptor identified from the ovary of the oriental river prawn *Macrobrachium nipponense* is the corazonin receptor (Du, Ma & Qiu, 2015), clearly suggesting that corazonin is the decapod GnRH. Given the strong conservation of the decapod neuropeptidome described here, I conclude that is highly unlikely that any of those three peptides is present in decapods.

Functional aspects

Conservation of structure does not necessarily imply conservation of function. The function

of crustacean RPCH and its insect ortholog AKH are distinctly different. A neuropeptide sequence does not reveal its function, but the distribution of its receptor give some clues. FMRFamide is known to effect muscle contraction in decapods (Worden, Kravitz & Goy, 1995), the expression of its putative receptor in muscle, heart and the epidermis (that contains muscle as well) suggests that it has similar effects. The simultaneous expression of elevenin and a putative elevenin receptor in the midgut suggests that is has a digestive function. The hormone GPA2/GPB5 was suggested to be an antidiuretic hormone in *Drosophila* (Sellami, Agricola & Veenstra, 2011) and was subsequently shown to stimulate sodium reabsorption in the mosquito hindgut (Paluzzi, Vanderveken & O'Donnell, 2014). The very abundant expression of its receptor in the gill suggests that its function *Carcinus* may well be similar. An interesting difference between insects and decapods is the presence of ecydysis triggering hormone in the decapod nervous system and eye(stalk); in insects this peptide seems to be exclusively present in cells associated with the tracheal system and absent from the central nervous system (Roller et al., 2010). It will be interesting to know whether the function of ecdysis triggering hormone within the decapod nervous system is related to ecdysis behavior.

Intestine

Neuropeptides in the intestine are typically produced by enteroendocrine cells. CHH (Chung, Dircksen & Webster, 1999), SIFamide and tachykinin immunoreactive enteroendocrine cells (Christie et al., 2007) have been previously described from decapods. No SIFamide reads were found in the *Carcinus* intestine SRA, but allatostatin C, calcitonin-B, elevenin, orcokinin and hyrg were all present in seemingly significant numbers of reads (Table 1). This ensemble of gut neuropeptides differs significantly from what is known from the *Drosophila* midgut (Veenstra & Ida, 2014), even though tachykinin, allatostatin C and orcokinins are present in both, while the calcitonin B transcript is abundant in phasmid midgut SRAs (Veenstra, 2014).

CHH and MIH

The neuropeptides related to CHH are amongst the best known crustacean hormones (excellent review by Webster, Keller & Dircksen, 2012). As was expected based on the literature, several molecular forms were found. There are reasons to think there may be more of these hormones than reported here. First of all, the few decapod CHH genes that have been

identified are typically present in clusters and in *Metapaeneus ensis* 16 such genes have been found (Gu & Chan, 1998). Secondly, as shown here and elsewhere (*e.g.* Hsu et al., 2006; Li et al., 2010; Ventura-López et al., 2016) some of these genes are differentially expressed. Thus, if a gene is predominantly expressed in a tissue not included in the analysis, it may not be found. Finally, since these hormones are similar in structure, it is possible that Trinity would have problems producing all contigs. Indeed the number of different CHH cDNAs reported from *Carcinus maenas* (Dircksen et al., 2001) is larger than found here.

The biological activities of these hormones vary widely and the hormones with very similar sequences may have quite different physiological effects (*e.g.* Webster, Keller & Dircksen, 2012; Luo et al., 2015). It is for this reason that is impossible to interpret the meaning of the three predicted hormones that defy classification as either a CHH-like or MIH-like hormone (Figs. 5).

PDH

There are generally within the same species several precursors coding the shorter, more classical, PDHs, those different precursors code sometimes for the same mature peptide. It seems plausible that some of these differences reflect either allelic variations of a single gene or recent local gene duplications. Most of the species have two or more different predicted mature PDH peptides. It has previously been shown that the two PDHs from the crab *Cancer productus* have different functions, one is released as a hormone into the hemolymph, while others is used within the central nervous system (Hsu et al., 2010). As the tissue used for the *Scylla* transcriptome did not include the eyestalk it is thus not surprising that the hormonal PDH is lacking from the deduced trancriptome in this species. ePDH is not expressed in the eyestalk and one might therefore be tempted to think it is not released into the hemolymph. However it is present in the *Litopenaeus* transcriptome that was produced from abdominal muscle, hepatopancreas, gills and pleopods (Ghaffari et al., 2014) and thus is likely produced somewhere in the periphery (this transcriptome contains relatively few neuropeptides as it includes neither the central nervous system nor the intestine).

CFSH

CFSH was discovered very recently in the crab *Callinectes sapidus* (Zmora & Chung, 2014) and consequently we know still very little of this extraordinarily interesting hormone. I

previously reported the presence of both CFSH and two homologous proteins in *Procambarus* (Veenstra, 2015). Now that there are a few more sequences available, it appears that this gene commonly has several paralogs. Some of these seem to have a relatively recent origin, as the most closely related sequence comes from the same species (Fig. 6). The independent gene duplications of these proteins as well the great sequence variability between and within species may indicate that all these hormones act on the same receptor. Given the relatively large size of these hormones one might expect a leucine rich repeat G-protein coupled receptor or a dimeric proteine kinase, perhaps one of the two venus kinase receptors, but this remains speculation. The primary structure of CFSH is not very well conserved and its receptor is unknown. Hence, we don't know whether an orthologous hormonal regulatory system might be present in other arthropods, like *e.g.* insects (given the great similarity in their neuropeptidomes this seems a distinct possibility, at least in the more primitive insects). It seems that the expression of this hormone in the ovary of *Procambarus* (Veenstra, 2015) is unusual, as it was not found in any of the other decapods for which an ovary SRA is available.

Insulin and neuroparsin

Other intriguing neuropeptides are the neuroparsins and the insulin-related hormones. There are three different insulin-like hormones. There are also three different insulin receptors, the classical tyrosine kinase and two G-protein coupled receptors. What I have called insulin is the hormone most similar to the *Drosophila* insulin-like peptides 1-6, which function as growth hormones and are also important for reproduction and that signal through the classical tyrosine kinase receptor (Nässel & vanden Broeck, 2015). The same receptor is also present in decapods as shown here and elsewhere (Veenstra, 2015); it has recently been characterized in two decapods (Aizen et al., 2016; Sharabi et al., 2016). Both insulin and neuroparsins activate tyrosine kinase receptors. However, whereas the actions of insulin in insects are relatively well understood due to very extensive research on these peptides in *Drosophila* (Nässel & vanden Broeck, 2016), the function of neuroparsin is less clear, as it is absent from *Drosophila melanogaster* (Veenstra, 2010b). It is interesting to note that some species have several insulin genes and few if any neuroparsin genes (*Drosophila, Acyrthosiphon, Zootermopsis*), while decapods and *Locusta* have several neuroparsin transcripts and only a single insulin gene, suggesting some complementation between these two hormones. Indeed, in some cases, such as

568 vitellogenesis in the mosquito both hormones have synergistic effects (Brown et al., 1998; Dhara
569 et al., 2013), however in the migratory locust they act antagonistically (Badisco et al., 2011).
570 Initially isolated from the migratory locust *L. migratoria* (Girardie et al., 1989) neuroparsin was
571 shown to have strong anti-juvenile hormone effects, effecting both reproduction and
572 metamorphosis (Girardie et al., 1987). It has been shown that neuroparsin RNAi also inhibits
573 vitellogenesis, and hence reproduction, in the decapod *Metapenaeus ensis* (Yang et al., 2014).
574 The receptor for this hormone was recently identified in mosquitoes as a venus kinase receptor
575 (Vogel et al., 2015), a type of receptor that was lost in chordates during evolution (Dissous,
576 2015). Although orthologous venus kinase receptors are present in other arthropod genomes
577 (notably *Limulus*, *Strigamia* and *Stegodyphus*, Table S8) as well as mollusks (Vanderstraete et
578 al., 2013), no neuroparsin orthologs could be found in those species. The evolutionary origin of
579 neuroparsin is therefore unclear and it is not known whether species that seem to lack
580 neuroparsin need a hormone ligand to activate the venus kinase receptor (Dissous, 2015). The
581 presence of two such receptors in decapods is intriguing, but has also been found in Lepidoptera
582 and trematodes (Dissous, 2015).
583
584 The other insulin-like peptides
585 Insects and decapods share many neuropeptides and it is not surprising that the various
586 decapod insulin-related hormones also have insect orthologs. The insulin-like hormone I have
587 called relaxin is an ortholog of *Drosophila* insulin-like peptide 7. This hormone is not only
588 preseant in insects, but also in ticks, spider mites, mollusks and even acorn worms (Veenstra,
589 2010a; Veenstra, Rombouts & Grbić , 2012). As previously pointed out, this hormone is only
590 present in the genomes of those species that also have an ortholog of *Drosophila* gene CG34411,
591 that encodes LGR4 that is homologous to vertebrate relaxin GPCRs (Veenstra, Rombouts &
592 Grbić , 2012 ; Veenstra, 2014). This suggests that this GPCR functions as a receptor for the
593 arthropod relaxins. It must be noted that this does not exclude the possibility that arthropod
594 relaxins may also signal through the classical insulin tyrosine kinase receptor. In fact, there is
595 evidence from *Drosophila* that this is so (Linneweber et al., 2014).
596 *Drosophila* has an eighth insulin-like hormone that was initially discovered because it is
597 secreted by the imaginal discs (Colombani et al., 2015; Garelli et al., 2015). However, data from
598 fly atlas (Chintapalli, Wang & Dow, 2007) show that it is also expressed by the ovary. This

hormone was suggested to be acting through the GPCR encoded by *Drosophila* gene CG31096 encoding LGR3 (Veenstra, 2014) and this has now been confirmed (Vallejo et al., 2015; Garelli et al., 2015). LGR3 is also related to vertebrate GPCRs binding relaxin. As reported previously it has a *Procambarus* ortholog (Veenstra, 2016), and as shown here is generally present in decapods. Combined these data suggest that LGR3 is the receptor for the androgenic insulin-like peptide from the accessory gland. The absence of clear sequence homology between the *Drosophila* and decapod peptides is not surprising, as the primary sequence of this hormone is poorly conserved in both decapods (Fig. 10) and insects [other insects almost certainly have such a peptide, since they have the receptor, but only within flies is it possible to find orthologs using the BLAST program]. Interestingly, both these hormones are produced by gonads or associated accessory glands. At first sight it seems that in crustaceans it is predominantly the male that produces it, while in adult flies it is the female. However, work on the expression of LGR3 in *Drosophila* shows it to be important for the development of both male and female specific sexual characters (Meissner et al., 2016) and it is perhaps better considered a (sexual ?) maturation hormone for both sexes. This would also make it easier to understand how during evolution it was coopted by the imaginal discs. In decapods the male has two Z chromosomes and is the default sex (Cui et al., 2015). Therefore, one would expect females to have a mechanism (not necessarily hormonal) to escape becoming a male and might thus expect a gynogenic rather than an androgenic hormone (this is one of the reasons why CFSH is so interesting). Even in decapods there is now evidence that the androgenic insulin-like peptide is not specific for males. Thus, in *Scylla paramosain* it is also expressed by the ovary and at higher levels at the end of vitellogenesis (Huang et al, 2014). While the relative levels of expression may seem low as compared to those of actin (Huang et al, 2014), the actual quantities of peptide produced may well rival those made by the accessory gland, considering that the ovary is so much larger [could the accessory gland be the remnants of an embryonic ovary anlage ?]. Given the effectiveness of RNAi in crustaceans and the strong phenotypes obtained in the absence of androgenic peptide (Ventura et al., 2009), the hypothesis that LGR3 is important in the transduction of the androgenic peptide signal can be tested. As with relaxin, a GPCR specifically activated by the androgenic insulin-like peptide does not exclude the possibility that it may also act on the classical insulin tyrosine kinase receptor, as suggested by recent experiments in the decapod *Sagmariasus* (Aizen et al., 2016). Possible relations between the decapod insulin-related peptides

and their receptors are illustrated in Figure 16.

It is of interest to note that the mammalian GPCR most similar to LGR3 is RXFP2, the receptor for insulin-related peptide 3. The latter hormone was initially discovered from the testis and is important not only to insure testicular descent (Adhama, Emmen & Engel, 2000) but also in the female reproductive system (Satchell et al., 2013). Thus the data suggest that not only the structures of the receptor and its ligand are recognizably similar, but so might be their function. This is rather interesting, as most neuropeptides with orthologs in both proto- and deuterostomia have quite different functions in these two groups.

Conclusions

Decapod neuropeptidomes are highly conserved and share many neuropeptides with insects. Although a shared neuropeptide structure does not necessarily translate into a shared function, it should allow for the rapid identification receptors in decapods in those cases where the orthologous insect receptor is known.

Acknowledgements

References

Adhama IM, Emmen JMA, Engel W. 2000. The role of the testicular factor INSL3 in establishing the gonadal position. *Molecular and Cellular Endocrinology* 160:11-16. doi:10.1016/S0303-7207(99)00188-4.

Aizen J, Chandler JC, Fitzgibbon QP, Sagi A, Battaglene SC, Elizur A, Ventura T. 2016. Production of recombinant insulin-like androgenic gland hormones from three decapod species: in vitro testicular phosphorylation and activation of a newly identified tyrosine kinase receptor from the Eastern spiny lobster, *Sagmariasus verreauxi. General and Comparative Endocrinology* 229:8-18. doi: 10.1016/j.ygcen.2016.02.013.

Alfa RW, Park S, Skelly KR, Poffenberger G, Jain N, Gu X, Kockel L, Wang J, Liu Y, Powers AC, Kim SK. 2015. Suppression of insulin production and secretion by a decretin hormone. *Cell Metabolism* 21:323-33. doi: 10.1016/j.cmet.2015.01.006.

Badisco L, Marchal E, Van Wielendaele P, Verlinden H, Vleugels R, Vanden Broeck J 2011. RNA interference of insulin-related peptide and neuroparsins affects vitellogenesis in the desert locust *Schistocerca gregaria. Peptides* 32:573–580. doi:10.1016/j.peptides.2010.11.008.

Bao C, Yang Y, Huang H, Ye H. 2015. Neuropeptides in the cerebral ganglia of the mud crab,

Scylla paramamosain: transcriptomic analysis and expression profiles during vitellogenesis. *Scientific Reports* 5:17055. doi: 10.1038/srep17055.

Bauknecht P, Jékely G. 2015. Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs. *Cell Reports* 12:684-693. doi: 10.1016/j.celrep.2015.06.052.

Brown MR, Graf R, Swiderek KM, Fendley D, Stracker TH, Champagne DE, Lea AO. 1998. Identification of a steroidogenic neurohormone in female mosquitoes. *Journal of Biological Chemistry* 273:3967-3971.

Bungart D, Hilbich C, Dircksen H, Keller R. 1995. Occurrence of analogues of the myotropic neuropeptide orcokinin in the shore crab, *Carcinus maenas*: evidence for a novel neuropeptide family . *Peptides* 16:67-72. doi:10.1016/0196-9781(94)00145-V.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421.

Cazzamali G, Torp M, Hauser F, Williamson M, Grimmelikhuijzen CJP. 2005. The *Drosophila* gene CG9918 codes for a pyrokinin-1 receptor. *Biochemical and Biophysical Research Commununications* 335(1):14-9. doi:10.1016/j.bbrc.2005.07.038.

Chandler JC, Aizen J, Elizur A, Hollander-Cohen L, Battaglene SC, Ventura T. 2015. Discovery of a novel insulin-like peptide and insulin binding proteins in the Eastern rock lobster *Sagmariasus verreauxi*. *General and comparative Endocrinology* 215:76-87. doi: 10.1016/j.ygcen.2014.08.018.

Chen X, Zeng D, Chen X, Xie D, Zhao Y, Yang C, Li Y, Ma N, Li M, Yang Q, Liao Z, Wang H. 2013. Transcriptome analysis of *Litopenaeus vannamei* in response to white spot syndrome virus infection. *PLoS One* 8:e73218. doi: 10.1371/journal.pone.0073218

Chintapalli VR, Wang J, Dow JAT. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nature Genetics* 39:715-720. doi:10.1038/ng2049.

Christie AE. 2014. Expansion of the *Litopenaeus vannamei* and *Penaeus monodon* peptidomes using transcriptome shotgun assembly sequence data. *General and Comparative Endocrinology* 206:235-254. doi: 10.1016/j.ygcen.2014.04.015.

Christie AE, Chi M. 2015. Prediction of the neuropeptidomes of members of the Astacidea (Crustacea, Decapoda) using publicly accessible transcriptome shotgun assembly (TSA) sequence data. *General and comparative Endocrinology* 224:38-60. doi: 10.1016/j.ygcen.2015.06.001.

Christie AE, Stemmler EA, Peguero B, Messinger DI, Provencher HL, Scheerlinck P, Hsu YW, Guiney ME, de la Iglesia HO, Dickinson PS. 2006. Identification, physiological actions, and distribution of VYRKPPFNGSIFamide (Val1)-SIFamide) in the stomatogastric nervous system

of the American lobster *Homarus americanus*. *Journal of comparative Neurology* 496:406-421.

Christie AE, Cashman CR, Brennan HR, Ma M, Sousa GL, Li L, Stemmler EA, Dickinson PS. 2008. Identification of putative crustacean neuropeptides using in silico analyses of publicly accessible expressed sequence tags. *General and Comparative Endocrinology* 156:246-64. doi: 10.1016/j.ygcen.2008.01.018.

Christie AE, Kutz-Naber KK, Stemmler EA, Klein A, Messinger DI, Goiney CC, Conterato AJ, Bruns EA, Hsu YW, Li L, Dickinson PS. 2007. Midgut epithelial endocrine cells are a rich source of the neuropeptides APSGFLGMRamide (*Cancer borealis* tachykinin-related peptide Ia) and GYRKPPFNGSIFamide (Gly1-SIFamide) in the crabs *Cancer borealis, Cancer magister* and *Cancer productus*. *Journal of experimental Biology* 210):699-714. doi: 10.1242/jeb.02696.

Christie AE, Chi M, Lameyer TJ, Pascual MG, Shea DN, Stanhope ME, Schulz DJ, Dickinson PS. 2015. Neuropeptidergic signaling in the American lobster *Homarus americanus*: New insights from high-throughput nucleotide sequencing. *PLoS One* 10:e0145964. doi: 10.1371/journal.pone.0145964.

Chung JS, Dircksen H, Webster SG. 1999. A remarkable, precisely timed release of hyperglycemic hormone from endocrine cells in the gut is associated with ecdysis in the crab *Carcinus maenas*. *Proceedings of the National Academy of Sciences of the United States of America* 96:13103-13107. doi: 10.1073/pnas.96.23.13103.

Chung JS, Wilcockson DC, Zmora N, Zohar Y, Dircksen H, Webster SG. 2006. Identification and developmental expression of mRNAs encoding crustacean cardioactive peptide (CCAP) in decapod crustaceans . *Journal of experimental Biology* 209:3862-3872. doi: 10.1242/jeb.02425.

Colombani J, Andersen DS, Léopold P. 2012. Secreted peptide Dilp8 coordinates *Drosophila* tissue growth with developmental timing. *Science* 336:582–585. doi: 10.1126/science.1216689.

Cui Z, Hui M, Liu Y, Song C, Li X, Li Y, Liu L, Shi G, Wang S, Li F, Zhang X, Liu C, Xiang J, Chu KH. 2015. High-density linkage mapping aided by transcriptomics documents ZW sex determination system in the Chinese mitten crab *Eriocheir sinensis*. *Heredity (Edinb)* 115:206-215. doi: 10.1038/hdy.2015.26.

Derst C, Dircksen H, Meusemann K, Zhou X, Liu S, Predel R. 2016. Evolution of neuropeptides in non-pterygote hexapods. *BMC Evolutionary Biology* 16:51. doi: 10.1186/s12862-016-0621-4.

Dhara A, Eum JH, Robertson A, Gulia-Nuss M, Vogel KJ, Clark KD, Graf R, Brown MR, Strand MR. 2013. Ovary ecdysteroidogenic hormone functions independently of the insulin receptor in the yellow fever mosquito, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* 43:1100-1108. doi: 10.1016/j.ibmb.2013.09.004.

Dickinson PS, Stemmler EA, Cashman CR, Brennan HR, Dennison B, Huber KE, Peguero B, Rabacal W, Goiney CC, Smith CM, Towle DW, Christie AE. 2008. SIFamide peptides in clawed lobsters and freshwater crayfish (Crustacea, Decapoda, Astacidea): a combined molecular, mass

spectrometric and electrophysiological investigation. *General and Comparative Endocrinology* 156:347-60. doi: 10.1016/j.ygcen.2008.01.011.

Dickinson PS, Stemmler EA, Barton EE, Cashman CR, Gardner NP, Rus S, Brennan HR, McClintock TS, Christie AE. 2009a. Molecular, mass spectral, and physiological analyses of orcokinins and orcokinin precursor-related peptides in the lobster *Homarus americanus* and the crayfish *Procambarus clarkii. Peptides* 30:297-317. doi: 10.1016/j.peptides.2008.10.009.

Dickinson PS, Wiwatpanit T, Gabranski ER, Ackerman RJ, Stevens JS, Cashman CR, Stemmler EA, Christie AE. 2009b. Identification of SYWKQCAFNAVSCFamide: a broadly conserved crustacean C-type allatostatin-like peptide with both neuromodulatory and cardioactive properties. *Journal of experimental Biology* 212:1140-52. doi: 10.1242/jeb.028621.

Dircksen H, Böcking D, Heyn U, Mandel C, Chung JS, Baggerman G, Verhaert P, Daufeldt S, Plosch T, Jaros PP, Waelkens E, Keller R, Webster SG. 2001. Crustacean hyperglycaemic hormone (CHH)-like peptides and CHH-precursor-related peptides from pericardial organ neurosecretory cells in the shore crab, *Carcinus maenas*, are putatively spliced and modified products of multiple genes . *Biochemical Journal* 356:159-170. doi: 10.1042/bj3560159.

Dircksen H, Neupert S, Predel R, Verleyen P, Huybrechts J, Strauss J, Hauser F, Stafflinger E, Schneider M, Pauwels K, Schoofs L, Grimmelikhuijzen CJP. 2011. Genomics, transcriptomics, and peptidomics of *Daphnia pulex* neuropeptides and protein hormones. *Journal of Proteome Research* 10:4478-1504. doi: 10.1021/pr200284e.

Dissous C. 2015. Venus kinase receptors at the crossroads of insulin signaling: Their role in reproduction for helminths and insects. *Frontiers in Endocrinology (Lausanne)* 6:118. doi: 10.3389/fendo.2015.00118.

Du YX, Ma KY, Qiu GF. 2015. Discovery of the genes in putative GnRH signaling pathway with focus on characterization of GnRH-like receptor transcripts in the brain and ovary of the oriental river prawn *Macrobrachium nipponense. Aquaculture* 442:1-11. doi:10.1016/j.aquaculture.2015.02.016.

Duve H, Johnsen AH, Maestro JL, Scott AG, Jaros PP, Thorpe A. 1997. Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab *Carcinus maenas. European Journal of Biochemistry* 250:727-734. doi: 10.1111/j.1432-1033.1997.00727.x

Gao J, Wang X, Zou Z, Jia X, Wang Y, Zhang Z. 2014. Transcriptome analysis of the differences in gene expression between testis and ovary in green mud crab (*Scylla paramamosain*). *BMC Genomics* 15:585. doi: 10.1186/1471-2164-15-585.

Gao Y, Zhang X, Wei J, Sun X, Yuan J, Li F, Xiang J. 2015. Whole transcriptome analysis provides insights into molecular mechanisms for molting in *Litopenaeus vannamei. PLoS One* 10:e0144350. doi: 10.1371/journal.pone.

Garelli A, Gontijo AM, Miguela V, Caparros E, Dominguez M. 2012. Imaginal discs secrete insulin-like peptide 8 to mediate plasticity of growth and maturation. *Science* 336:579–582. doi: 10.1126/science.1216735.

Garelli A, Heredia F, Casimiro AP, Macedo A, Nunes C, Garcez M, Dias AR, Volonte YA, Uhlmann T, Caparros E, Koyama T, Gontijo AM. 2015. Dilp8 requires the neuronal relaxin receptor Lgr3 to couple growth to developmental timing. *Nature Commununications* 6:8732. doi: 10.1038/ncomms9732.

Ghaffari, Noushin; Sanchez-Flores, Alejandro; Ryan, Doan; Garcia-Orozco, Karina D; Chen, Patricia L.; Ochoa-Leyva, Adrian; Lopez-Zavala, Alonso A.; Carrasco, J. Salvador; Hong, Chris; Brieba, Luis G.; Rudino-Pinera, Enrique; Blood, Philip D.; Jason A., Sawyer; Johnson, Charles D.; Dindot, Scott V.; Sotelo-Mundo, Rogerio R.; Criscitiello, Michael F. 2014. Novel transcriptome assembly and improved annotation of the whiteleg shrimp (*Litopenaeus vannamei*), a dominant crustacean in global seafood mariculture. *Scientific Reports* 4:7081 doi:10.1038/srep07081.

Girardie J, Bourême D, Couillaud F, Tamarelle M, Girardie A. 1987. Anti-juvenile effect of neuroparsin A, a neuroprotein isolated from locust corpora cardiaca. *Insect Biochemistry* 17:977-983. doi:10.1016/0020-1790(87)90106-5.

Girardie J, Girardie A, Huet JC, Pernollet JC. 1989. Amino acid sequence of locust neuroparsins. *FEBS Letters* 245:4-8.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* 27:221-4. doi: 10.1093/molbev/msp259.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644-652. doi: 10.1038/nbt.1883.

Gu P-L, Chan SM. 1998. The shrimp hyperglycemic hormone-like neuropeptide is encoded by multiple copies of genes arranged in a cluster. *FEBS Letters* 441:397–403. doi: 10.1016/S0014-5793(98)01573-7

Gu P-L, Yu KL, Chan SM. 2000. Molecular characterization of an additional shrimp hyperglycemic hormone: cDNA cloning, gene organization, expression and biological assay of recombinant proteins. *FEBS Letters* 472:122–128. doi: 10.1016/S0014-5793(00)01420-4

Guan ZB, Shui Y, Liao XR, Xu ZH, Zhou X. 2014. Primary structure of a novel gonadotropin-releasing hormone (GnRH) in the ovary of red swamp crayfish *Procambarus clarkii.* *Aquaculture* (418–419):67-71. tdoi: 10.1016/j.aquaculture.2013.10.010.

He L, Wang Q, Jin X, Wang Y, Chen L, Liu L, Wang Y. 2012. Transcriptome profiling of testis during sexual maturation stages in *Eriocheir sinensis* using Illumina sequencing. *PLoS One* 7:e33735. doi: 10.1371/journal.pone.0033735.

Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015. Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* 32:835-845. doi: 10.1093/molbev/msv037.

Hsu YW, Messinger DI, Chung JS, Webster SG, de la Iglesia HO, Christie AE. 2006. Members of the crustacean hyperglycemic hormone (CHH) peptide family are differentially distributed both between and within the neuroendocrine organs of *Cancer* crabs: implications for differential release and pleiotropic function. *Journal of experimental Biology* 209:3241-3256. doi: 10.1242/jeb.02372.

Hsu YW, Stemmler EA, Messinger DI, Dickinson PS, Christie AE, de la Iglesia HO. 2010. Cloning and differential expression of two beta-pigment-dispersing hormone (beta-PDH) isoforms in the crab *Cancer productus*: evidence for authentic beta-PDH as a local neurotransmitter and beta-PDH II as a humoral factor. *Journal of comparative Neurology* 508:197-211. doi: 10.1002/cne.21659.

Huang X, Ye H, Huang H, Yang Y, Gong J. 2014. An insulin-like androgenic gland hormone gene in the mud crab, *Scylla paramamosain*, extensively expressed and involved in the processes of growth and female reproduction. *General and comparative Endocrinology* 204:229-38. doi: 10.1016/j.ygcen.2014.06.002.

Hui L, Xiang F, Zhang Y, Li L. 2012. Mass spectrometric elucidation of the neuropeptidome of a crustacean neuroendocrine organ. *Peptides* 36:230-9. doi: 10.1016/j.peptides.2012.05.007.

Hui M, Liu Y, Song C, Li Y, Shi G, Cui Z. 2014. Transcriptome changes in *Eriocheir sinensis* megalopae after desalination provide insights into osmoregulation and stress adaption in larvae. Transcriptome changes in *Eriocheir sinensis* megalopae after desalination provide insights into osmoregulation and stress adaption in larvae. *PLoS One* 9:e114187. doi: 10.1371/journal.pone.0114187.

Jiang H, Xing Z, Lu W, Qian Z, Yu H, Li J. 2014. Transcriptome analysis of red swamp crawfish *Procambarus clarkii* reveals genes involved in gonadal development. *PLoS One* 9:e105122. doi: 10.1371/journal.pone.0105122.

Jung H, Lyons RE, Dinh H, Hurwood DA, McWilliam S, Mather PB. 2011. Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): de novo assembly, annotation and marker discovery. *PLoS One* 6:e27938. doi: 10.1371/journal.pone.0027938.

Kegel G, Reichwein B, Weese S, Gaus G, Peter-Katalinić J, Keller R. 1989 .Amino acid sequence of the crustacean hyperglycemic hormone (CHH) from the shore crab, *Carcinus maenas*. *FEBS Letters* 255:10-14. doi: 10.1111/j.1471-4159.1987.tb02902.x

Kim YJ, Bartalska K, Audsley N, Yamanaka N, Yapici N, Lee JY, Kim YC, Markovic M, Isaac E, Tanaka Y, Dickson BJ. 2010. MIPs are ancestral ligands for the sex peptide receptor. *Proceedings of the National Academy of Sciences of the United States of America* 107:6520-6525. doi: 10.1073/pnas.0914764107.

Klein JM, de Kleijn DP, Keller R, Weidemann WM. 1992. Molecular cloning of crustacean pigment dispersing hormone precursor . *Biochemical and Biophysical Research Commununications* 189:1509-1514. doi:10.1016/0006-291X(92)90246-H.

Klein JM, Mangerich S, de Kleijn DP, Keller R, Weidemann,W.M. 1993. Molecular cloning of crustacean putative molt-inhibiting hormone (MIH) precursor . *FEBS Letters* 334:139-142. doi: 10.1016/0014-5793(93)81699-Z.

Li S, Li F, Wang B, Xie Y, Wen R, Xiang J. 2010. Cloning and expression profiles of two isoforms of a CHH-like gene specifically expressed in male Chinese shrimp, *Fenneropenaeus chinensis. General and comparative Endocrinology* 167:308-316. doi:10.1016/j.ygcen.2010.03.028.

Li C, Weng S, Chen Y, Yu X, Lü L, Zhang H, He J, Xu X. 2012. Analysis of *Litopenaeus vannamei* transcriptome using the next-generation DNA sequencing technique. *PLoS One* 7:e47442. doi: 10.1371/journal.pone.0047442.

Li Y, Hui M, Cui Z, Liu Y, Song C, Shi G. 2014. Comparative transcriptomic analysis provides insights into the molecular basis of the metamorphosis and nutrition metabolism change from zoeae to megalopae in *Eriocheir sinensis*. *Comparative Biochemistry and Physiology Part D Genomics and Proteomics* 13:1-9. doi: 10.1016/j.cbd.2014.10.002.

Linck B, Klein JM, Mangerich S, Keller R, Weidemann WM . 1993. Molecular cloning of crustacean red pigment concentrating hormone precursor. *Biochemical and Biophysical Research Commununications* 195:807-813. doi:10.1006/bbrc.1993.2117.

Linneweber GA, Jacobson J, Busch KE, Hudry B, Christov CP, Dormann D, Yuan M, Otani T, Knust E, de Bono M, Miguel-Aliaga I. 2014. Neuronal control of metabolism through nutrient-dependent modulation of tracheal branching. *Cell* 156:69-83. doi: 10.1016/j.cell.2013.12.008.

Liu Y, Hui M, Cui Z, Luo D, Song C, Li Y, Liu L. 2015. Comparative transcriptome analysis reveals sex-biased gene expression in juvenile Chinese mitten crab *Eriocheir sinensis*. *PLoS One* 10:e0133068. doi: 10.1371/journal.pone.0133068.

Lugo JM, Carpio Y, Morales R, Rodríguez-Ramos T, Ramos L, Estrada MP. 2013. First report of the pituitary adenylate cyclase activating polypeptide (PACAP) in crustaceans: conservation of its functions as growth promoting factor and immunomodulator in the white shrimp *Litopenaeus vannamei. Fish & Shellfish Immunology* 35:1788-96. doi: 10.1016/j.fsi.2013.08.028.

Luo X, Chen T, Zhong M, Jiang X, Zhang L, Ren C, Hu C. 2015. Differential regulation of hepatopancreatic vitellogenin (VTG) gene expression by two putative molt-inhibiting hormones

(MIH1/2) in Pacific white shrimp (*Litopenaeus vannamei*). *Peptides* 68:58-63. doi: 10.1016/j.peptides.2014.11.002.

Ma M, Chen R, Sousa GL, Bors EK, Kwiatkowski MA, Goiney CC, Goy MF, Christie AE, Li L. 2008. Mass spectral characterization of peptide transmitters/hormones in the nervous system and neuroendocrine organs of the American lobster *Homarus americanus. General and Comparative Endocrinology* 156:395-409. doi: 10.1016/j.ygcen.2008.01.009.

Ma M, Bors EK, Dickinson ES, Kwiatkowski MA, Sousa GL, Henry RP, Smith CM, Towle DW, Christie AE, Li L. 2009. Characterization of the *Carcinus maenas* neuropeptidome by mass spectrometry and functional genomics. *General and Comparative Endocrinology* 161:320-334. doi: 10.1016/j.ygcen.2009.01.015.

Ma M, Gard AL, Xiang F, Wang J, Davoodian N, Lenz PH, Malecha SR, Christie AE, Li L. 2010. Combining in silico transcriptome mining and biological mass spectrometry for neuropeptide discovery in the Pacific white shrimp *Litopenaeus vannamei. Peptides* 31:27-43. doi: 10.1016/j.peptides.2009.10.007.

Ma H, Ma C, Li S, Jiang W, Li X, Liu Y, Ma L. 2014. Transcriptome analysis of the mud crab (*Scylla paramamosain*) by 454 deep sequencing: assembly, annotation, and marker discovery. *PLoS One* 9:e102668. doi: 10.1371/journal.pone.0102668.

Manfrin C, Tom M, De Moro G, Gerdol M, Giulianini PG, Pallavicini A. 2015. The eyestalk transcriptome of red swamp crayfish *Procambarus clarkii. Gene* 557:28-34. doi: 10.1016/j.gene.2014.12.001.

Meissner GW, Luo SD, Dias BG, Texada MJ, Baker BS. 2016. Sex-specific regulation of Lgr3 in *Drosophila* neurons. *Proceedings of the National Academy of Sciences of the United States of America* 113:E1256-65. doi: 10.1073/pnas.1600241113.

Nässel DR, vanden Broeck J. 2015. Insulin/IGF signaling in *Drosophila* and other insects: factors that regulate production, release and post-release action of the insulin-like peptides. *Cellular and Molecular Life Sciences* 216:271-290. doi: 10.1007/s00018-015-2063-3.

Nuss AB, Forschler BT, Crim JW, TeBrugge V, Pohl J, Brown MR. 2010. Molecular characterization of neuropeptide F from the eastern subterranean termite *Reticulitermes flavipes* (Kollar) (Isoptera: Rhinotermitidae). *Peptides* 31:419-428. doi: 10.1016/j.peptides.2009.09.001.

Ohler U. 2006. Identification of core promoter modules in *Drosophila* and their application in accurate transcription start site prediction. *Nucleic Acids Research* 34:5943-5950. doi: 10.1093/nar/gkl608.

Ollivaux C, Vinh J, Soyez D, Toullec JY. 2006. Crustacean hyperglycemic and vitellogenesis-inhibiting hormones in the lobster *Homarus gammarus. FEBS Journal* 273,2151-2160. doi: 10.1111/j.1742-4658.2006.05228.x.

990 Paluzzi JP, Vanderveken M, O'Donnell MJ. 2014. The heterodimeric glycoprotein hormone,
991 GPA2/GPB5, regulates ion transport across the hindgut of the adult mosquito, *Aedes aegypti.*
992 *PLoS One* 9:e86386. doi: 10.1371/journal.pone.0086386.
993
994 Peng J, Wei P, Zhang B, Zhao Y, Zeng D, Chen X, Li M, Chen X. 2015. Gonadal transcriptomic
995 analysis and differentially expressed genes in the testis and ovary of the Pacific white shrimp
996 *(Litopenaeus vannamei*). *BMC Genomics* 16:1006. doi: 10.1186/s12864-015-2219-4.
997
998 Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees for
999 large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490.
1000
1001 Richards S, Murali SC. 2015. Best practices in insect genome sequencing: what
1002 works and what doesn't. *Current Opinion in Insect Science* 7:1–7. doi:
1003 10.1016/j.cois.2015.02.013.
1004
1005 Roller L, Yamanaka N, Watanabe K, Daubnerová I, Zitnan D, Kataoka H, Tanaka Y. 2008. The
1006 unique evolution of neuropeptide genes in the silkworm *Bombyx mori. Insect Biochemistry and*
1007 *Molecular Biology* 38:1147-1157. doi:10.1016/j.ibmb.2008.04.009.
1008
1009 Roller L, Zitnanová I, Dai L, Simo L, Park Y, Satake H, Tanaka Y, Adams ME, Zitnan D. 2010.
1010 Ecdysis triggering hormone signaling in arthropods. *Peptides* 31:429-41. doi:
1011 10.1016/j.peptides.2009.11.022.
1012
1013 Satchell L, Glister C, Bleach EC, Glencross RG, Bicknell AB, Dai Y, Anand-Ivell R, Ivell R,
1014 Knight PG. 2013. Ovarian expression of insulin-like peptide 3 (INSL3) and its receptor (RXFP2)
1015 during development of bovine antral follicles and corpora lutea and measurement of circulating
1016 INSL3 levels during synchronized estrous cycles. *Endocrinology* 154:1897-906. doi:
1017 10.1210/en.2012-2232.
1018
1019 Sellami A, Agricola HJ, Veenstra JA. 2011. Neuroendocrine cells in *Drosophila melanogaster*
1020 producing GPA2/GPB5, a hormone with homology to LH, FSH and TSH. *General and*
1021 *comparative Endocrinology* 170:582-588. doi: 10.1016/j.ygcen.2010.11.015.
1022
1023 Sharabi O, Manor R, Weil S, Aflalo ED, Lezer Y, Levy T, Aizen J, Ventura T, Mather PB,
1024 Khalaila I, Sagi A. 2016. Identification and characterization of an insulin-like receptor involved
1025 in crustacean reproduction. *Endocrinology* 157:928-941. doi: 10.1210/en.2015-1391.
1026
1027 Shen H, Hu Y, Ma Y, Zhou X, Xu Z, Shui Y, Li C, Xu P, Sun X. 2014. In-depth transcriptome
1028 analysis of the red swamp crayfish *Procambarus clarkii. PLoS One* 9:e110548. doi:
1029 10.1371/journal.pone.0110548.
1030
1031 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert
1032 M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein
1033 multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7:539. doi:
1034 10.1038/msb.2011.75.
1035

Song C, Cui Z, Hui M, Liu Y, Li Y, Li X. 2015. Comparative transcriptomic analysis provides insights into the molecular basis of brachyurization and adaptation to benthic lifestyle in *Eriocheir sinensis*. *Gene* 558:88-98. doi: 10.1016/j.gene.2014.12.048.

Song L, Bian C, Luo Y, Wang L, You X, Li J, Qiu Y, Ma X, Zhu Z, Ma L, Wang Z, Lei Y, Qiang J, Li H, Yu J, Wong A, Xu J, Shi Q, Xu P. 2016. Draft genome of the Chinese mitten crab, *Eriocheir sinensis*. *Gigascience* 5:5. doi: 10.1186/s13742-016-0112-y.

Stangier J, Hilbich C, Beyreuther K, Keller R. 1987. Unusual cardioactive peptide (CCAP) from pericardial organs of the shore crab Carcinus maenas. *Proceedings of the National Academy of Sciences of the United States of America* 84:575-579.

Stangier J, Hilbich C, Burdzik S, Keller R. 1992. Orcokinin: a novel myotropic peptide from the nervous system of the crayfish, *Orconectes limosus*. *Peptides*13:859-864. doi:10.1016/0196-9781(92)90041-Z.

Stemmler EA, Bruns EA, Gardner NP, Dickinson PS, Christie AE. 2007a. Mass spectrometric identification of pEGFYSQRYamide: a crustacean peptide hormone possessing a vertebrate neuropeptide Y (NPY)-like carboxy-terminus. *General and Comparative Endocrinology* 152:1-7.

Stemmler EA, Cashman CR, Messinger DI, Gardner NP, Dickinson PS, Christie AE. 2007b. High-mass-resolution direct-tissue MALDI-FTMS reveals broad conservation of three neuropeptides (APSGFLGMRamide, GYRKPPFNGSIFamide and pQDLDHVFLRFamide) across members of seven decapod crustaean infraorders. *Peptides* 28:2104-15.

Stemmler EA, Bruns EA, Cashman CR, Dickinson PS, Christie AE. 2010. Molecular and mass spectral identification of the broadly conserved decapod crustacean neuropeptide pQIRYHQCYFNPISCF: the first PISCF-allatostatin (*Manduca sexta*- or C-type allatostatin) from a non-insect. *General and Comparative Endocrinology* 165:1-10. doi: 10.1016/j.ygcen.2009.05.010.

Sturm S, Ramesh D, Brockmann A, Neupert S, Predel R. 2016. Agatoxin-like peptides in the neuroendocrine system of the honey bee and other insects. *Journal of Proteomics* 132:77-84. doi: 10.1016/j.jprot.2015.11.021.

Sun Y, Zhang Y, Liu Y, Xue S, Geng X, Hao T, Sun J. 2014. Changes in the organics metabolism in the hepatopancreas induced by eyestalk ablation of the Chinese mitten crab *Eriocheir sinensis* determined via transcriptome and DGE analysis. *PLoS One* 9:e95827. doi: 10.1371/journal.pone.0095827.

Suwansa-Ard S, Thongbuakaew T, Wang T, Zhao M, Elizur A, Hanna PJ, Sretarugsa P, Cummins SF, Sobhon P. 2015. *In silico* neuropeptidome of female *Macrobrachium rosenbergii* based on transcriptome and peptide mining of eyestalk, central nervous system and ovary. *PLoS One* 10:e0123848. doi: 10.1371/journal.pone.0123848.

Tanaka Y, Suetsugu Y, Yamamoto K, Noda H, Shinoda T. 2014. Transcriptome analysis of neuropeptides and G-protein coupled receptors (GPCRs) for neuropeptides in the brown planthopper *Nilaparvata lugens*. *Peptides* 53:125-33. doi: 10.1016/j.peptides.2013.07.027.

Thongbuakaew T, Saetan J, Suwansa-ard A, Kankoun W, Sumpownon C, Parhar I, Meeratana P, Sobhon P, Sretarugsa P. 2016. The existence of kisspeptin-like peptides and effects on ovarian development and maturation in the giant freshwater prawn *Macrobrachium rosenbergii*. *Aquaculture* 455:50-62. doi: 10.1016/j.aquaculture.2016.01.006.

Tom M, Manfrin C, Chung SJ, Sagi A, Gerdol M, De Moro G, Pallavicini A, Giulianini PG. 2014. Expression of cytoskeletal and molt-related genes is temporally scheduled in the hypodermis of the crayfish *Procambarus clarkii* during premolt. *Journal of experimental Biology* 217:4193-202. doi: 10.1242/jeb.109009.

Toullec JY, Corre E, Bernay B, Thorne MA, Cascella K, Ollivaux C, Henry J, Clark MS. 2013. Transcriptome and peptidome characterisation of the main neuropeptides and peptidic hormones of a euphausiid: the Ice Krill, *Euphausia crystallorophias*. *PLoS One* 8:e71609. doi: 10.1371/journal.pone.0071609.

Vallejo DM, Juarez-Carreño S, Bolivar J, Morante J, Dominguez M. 2015. A brain circuit that synchronizes growth and maturation revealed through Dilp8 binding to Lgr3. *Science* 350:aac6767. doi: 10.1126/science.aac6767.

Vanderstraete M, Gouignard N, Ahier A, Morel M, Vicogne J, Dissous C. 2013. The venus kinase receptor (VKR) family: structure and evolution. *BMC Genomics* 14:361. doi: 10.1186/1471-2164-14-361.

Veenstra JA. 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Archives of Insect Biochemistry and Physiology* 43:49-63. doi: 10.1002/(SICI)1520-6327(200002)43:2<49::AID-ARCH1>3.0.CO;2-M.

Veenstra JA. 2010a. Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *General and Comparative Endocrinology* 167:86-103. doi: 10.1016/j.ygcen.

Veenstra JA. 2010b. What the loss of the hormone neuroparsin in the *melanogaster* subgroup of *Drosophila* can tell us about its function. *Insect Biochemistry and Molecular Biology* 40,354-61. doi: 10.1016/j.ibmb.2010.03.001.

Veenstra JA. 2011. Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *General and comparative Endocrinology* 171:160-175. doi: 10.1016/j.ygcen.2011.01.005.

Veenstra JA. 2014.The contribution of the genomes of a termite and a locust to our understanding of insect neuropeptides and neurohormones. *Frontiers in Physiology* 5:454. doi: 10.3389/fphys.2014.00454.

Veenstra JA. 2015. The power of next-generation sequencing as illustrated by the neuropeptidome of the crayfish *Procambarus clarkii. General and comparative Endocrinology* 224:84-95. doi: 10.1016/j.ygcen.2015.06.013.

Veenstra JA. 2016. Neuropeptide evolution: Chelicerate neurohormone and neuropeptide genes may reflect one or more whole genome duplications. *General and comparative Endocrinology* 229:41-55. doi: 10.1016/j.ygcen.2015.11.019.

Veenstra JA, Ida T. 2014. More *Drosophila* enteroendocrine peptides: Orcokinin B and the CCHamides 1 and 2. *Cell and Tissue Research* 357:607-621. doi: 10.1007/s00441-014-1880-2.

Veenstra JA, Rodriguez L, Weaver RJ. 2012. Allatotropin, leucokinin and AKH in honey bees and other Hymenoptera. *Peptides* 35,122-130. doi: 10.1016/j.peptides.2012.02.019.

Veenstra JA, Rombauts S, Grbić M. 2012. *In silico* cloning of genes encoding neuropeptides, neurohormones and their putative G-protein coupled receptors in a spider mite. *Insect Biochemistry and Molecular Biology* 42:277-295. doi: 10.1016/j.ibmb.2011.12.009.

Ventura T, Sagi A. 2012. The insulin-like androgenic gland hormone in crustaceans: From a single gene silencing to a wide array of sexual manipulation-based biotechnologies. *Biotechnology Advances* 30:1543-50. doi: 10.1016/j.biotechadv.2012.04.008.

Ventura T, Manor R, Aflalo ED, Weil S, Raviv S, Glazer L, Sagi A. 2009. Temporal silencing of an androgenic gland-specific insulin-like gene affecting phenotypical gender differences and spermatogenesis. *Endocrinology* 150:1278-1286. doi: 10.1210/en.2008-0906.

Ventura T, Manor R, Aflalo ED, Chalifa-Caspi V, Weil S, Sharabi O, Sagi A. 2013. Post-embryonic transcriptomes of the prawn *Macrobrachium rosenbergii*: multigenic succession through metamorphosis. *PLoS One* 8 :e55322. doi: 10.1371/journal.pone.0055322.

Ventura T, Cummins SF, Fitzgibbon Q, Battaglene S, Elizur A. 2014. Analysis of the central nervous system transcriptome of the eastern rock lobster *Sagmariasus verreauxi* reveals its putative neuropeptidome. *PLoS One* 9:e97323. doi: 10.1371/journal.pone.0097323.

Ventura-López C, Gómez-Anduro G, Arcos FG, Llera-Herrera R, Racotta IS, Ibarra AM. 2016. A novel CHH gene from the Pacific white shrimp *Litopenaeus vannamei* was characterized and found highly expressed in gut and less in eyestalk and other extra-eyestalk tissues. *Gene in press*. doi: 10.1016/j.gene.2016.02.011.

Verbruggen B, Bickley LK, Santos EM, Tyler CR, Stentiford GD, Bateman KS, van Aerle R. 2015. De novo assembly of the *Carcinus maenas* transcriptome and characterization of innate immune system pathways. *BMC Genomics* 16:458. doi: 10.1186/s12864-015-1667-1.

Vogel KJ, Brown MR, Strand MR. 2015. Ovary ecdysteroidogenic hormone requires a receptor tyrosine kinase to activate egg formation in the mosquito Aedes aegypti. *Proceedings of the*

National Academy of Sciences of the United States of America 112:5057-62. doi: 10.1073/pnas.1501814112.

Wang B, Ning Q, Hao T, Yu A, Sun J. 2016. Reconstruction and analysis of a genome-scale metabolic model for *Eriocheir sinensis eyestalks*. *Molecular BioSystems* 12:246-252. doi: 10.1039/c5mb00571j.

Webster SG, Keller R, Dircksen H. 2012. The CHH-superfamily of multifunctional peptide hormones controlling crustacean metabolism, osmoregulation, moulting, and reproduction. *General and comparative Endocrinology* 175:217-233. doi: 10.1016/j.ygcen.2011.11.035.

Wei J, Zhang X, Yu Y, Li F, Xiang J. 2014. RNA-Seq reveals the dynamic and diverse features of digestive enzymes during early development of Pacific white shrimp *Litopenaeus vannamei. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 11:37-44. doi: 10.1016/j.cbd.2014.07.001.

Wilcockson DC, Webster SG. 2008. Identification and developmental expression of mRNAs encoding putative insect cuticle hardening hormone, bursicon in the green shore crab *Carcinus maenas. General and comparative Endocrinology* 156:113-125. doi:10.1016/j.ygcen.2007.12.003.

Worden MK, Kravitz EA, Goy MF. 1995. Peptide F1, an N-terminally extended analog of FMRFamide, enhances contractile activity in multiple target tissues in lobster. *Journal of experimental Biology* 198:97-108.

Xu Z, Zhao M, Li X, Lu Q, Li Y, Ge J, Pan J. 2015. Transcriptome profiling of the eyestalk of precocious juvenile Chinese mitten crab reveals putative neuropeptides and differentially expressed genes. *Gene* 569:280-286. doi: 10.1016/j.gene.2015.05.075.

Yang SP, He JG, Sun CB, Chan SF. 2014. Characterization of the shrimp neuroparsin (MeNPLP): RNAi silencing resulted in inhibition of vitellogenesis. *FEBS Open Biology* 4:976-86. doi: 10.1016/j.fob.2014.09.005.

Yu Y, Zhang X, Yuan J, Li F, Chen X, Zhao Y, Huang L, Zheng H, Xiang J. 2015. Genome survey and high-density genetic map construction provide genomic and genetic resources for the Pacific white shrimp *Litopenaeus vannamei. Scientific Reports* 5:15612. doi: 10.1038/srep15612.

Zmora N, Chung JS. 2014. A novel hormone is required for the development of reproductive phenotypes in adult female crabs. *Endocrinology* 155:230-239. doi: 10.1210/en.2013-1603.

Figure 1(on next page)

Overview of the presence neuropeptide genes in seven decapods, *Daphnia pulex* and two insect species.

Dark blue: neuropeptide precursors that have been published previously; light blue, neuropeptide precursors (or significant parts therefore) that can be deduced directly from publicly available TSAs; red: precursors assembled here; yellow: precursors that could not be assembled, but for which individual reads in TSAs demonstrate their existence in the particular species. Asterisks indicate the existence of more than one gene.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

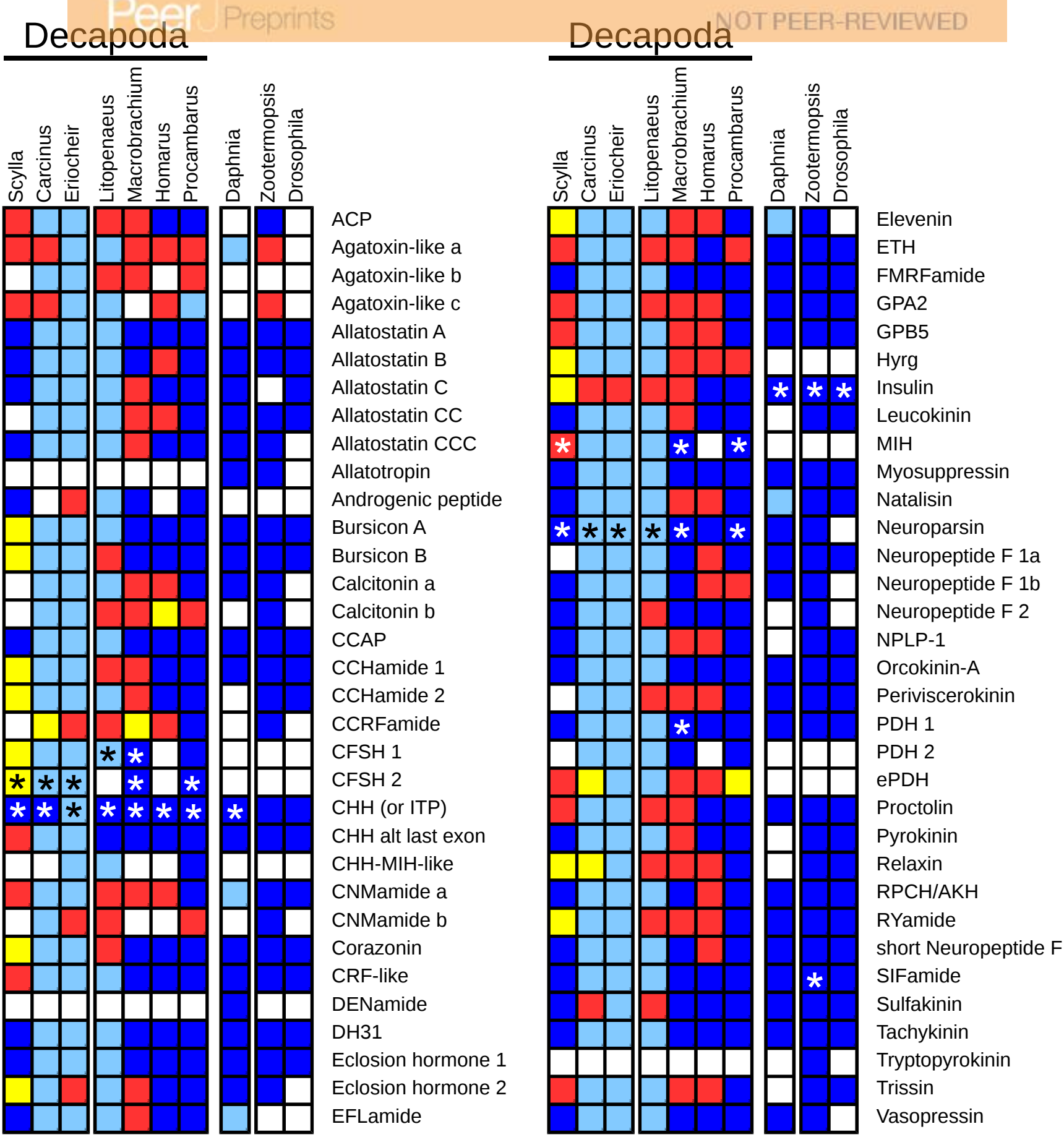
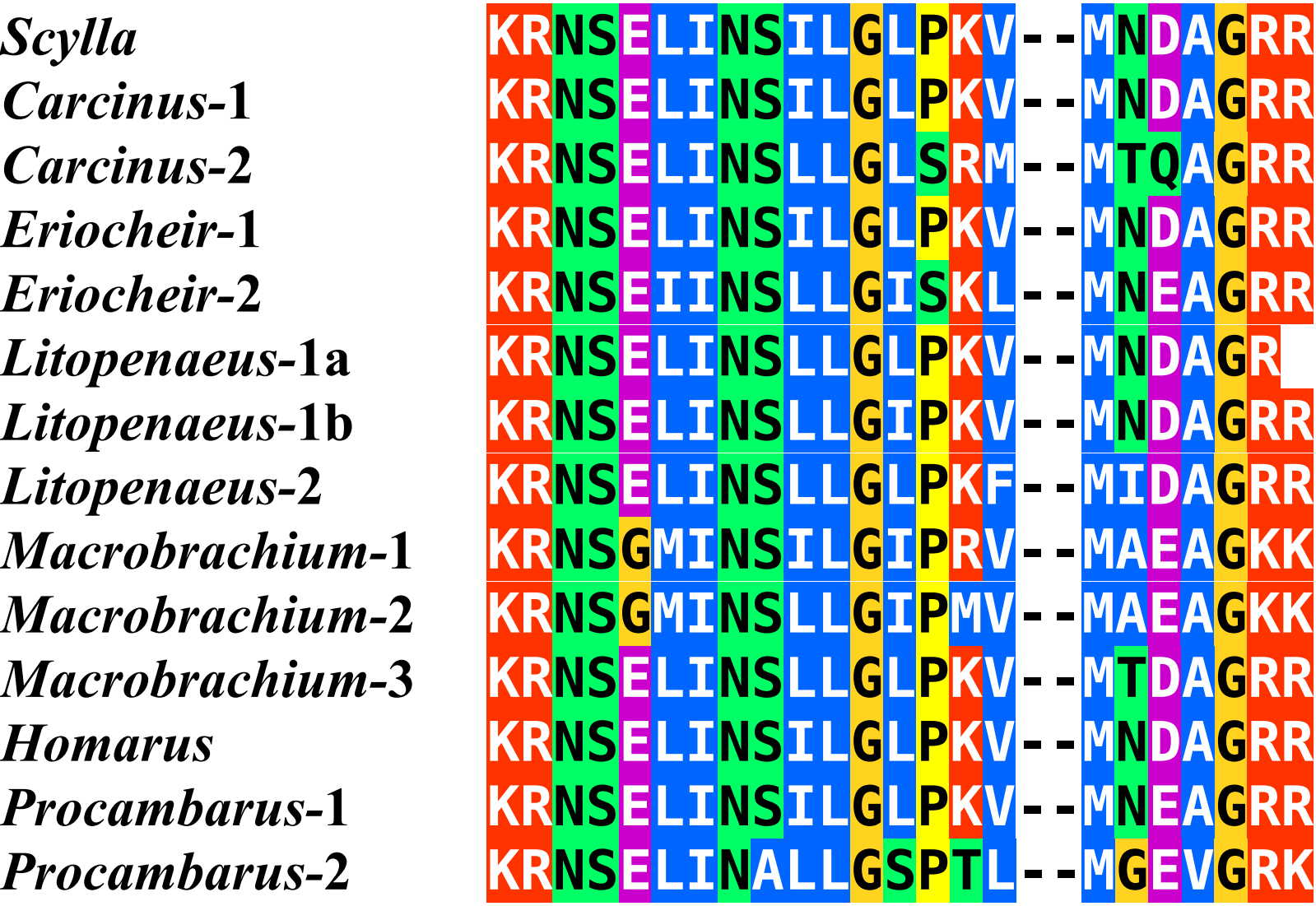


Figure 2(on next page)

Sequence allignment of PDH and ePDH.

Parts of the various PDH precursors including the convertase cleavage sites of the various decapod species. Note that the ubiquitous presence of ePDH that has a two aminon acid insertion.

PDH



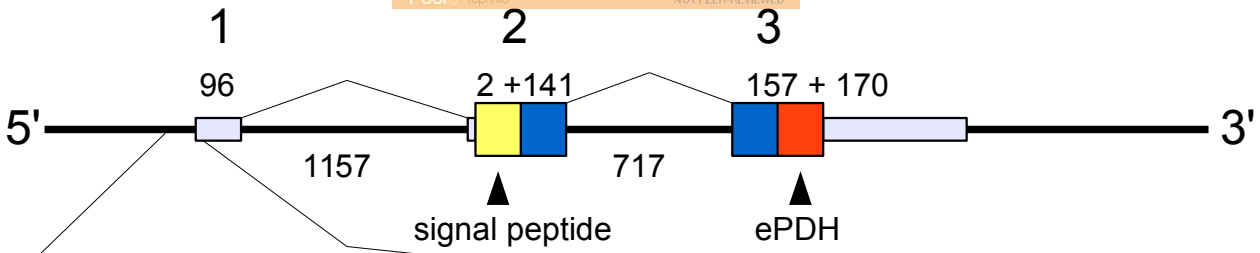
ePDH

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, pub: 20 Mar 2016

Figure 3(on next page)

Structure of the ePDH gene from Eriocheir sinensis.

The ePDH gene consists of three exons and two introns. DNA sequences coding the signal peptide in yellow, mature ePDH sequence in red and the remainder of the precursor in blue. Numbers indicate sizes of introns and exons in nucleotides. The DNA sequence containing the TATA box and a sequence that is recognizably similar to the *Drosophila* core promoter motif 1 (in blue, Ohler, 2006)and the start of the mRNA (in red) are also displayed; the red nucleotides at the end are part of the mRNA.

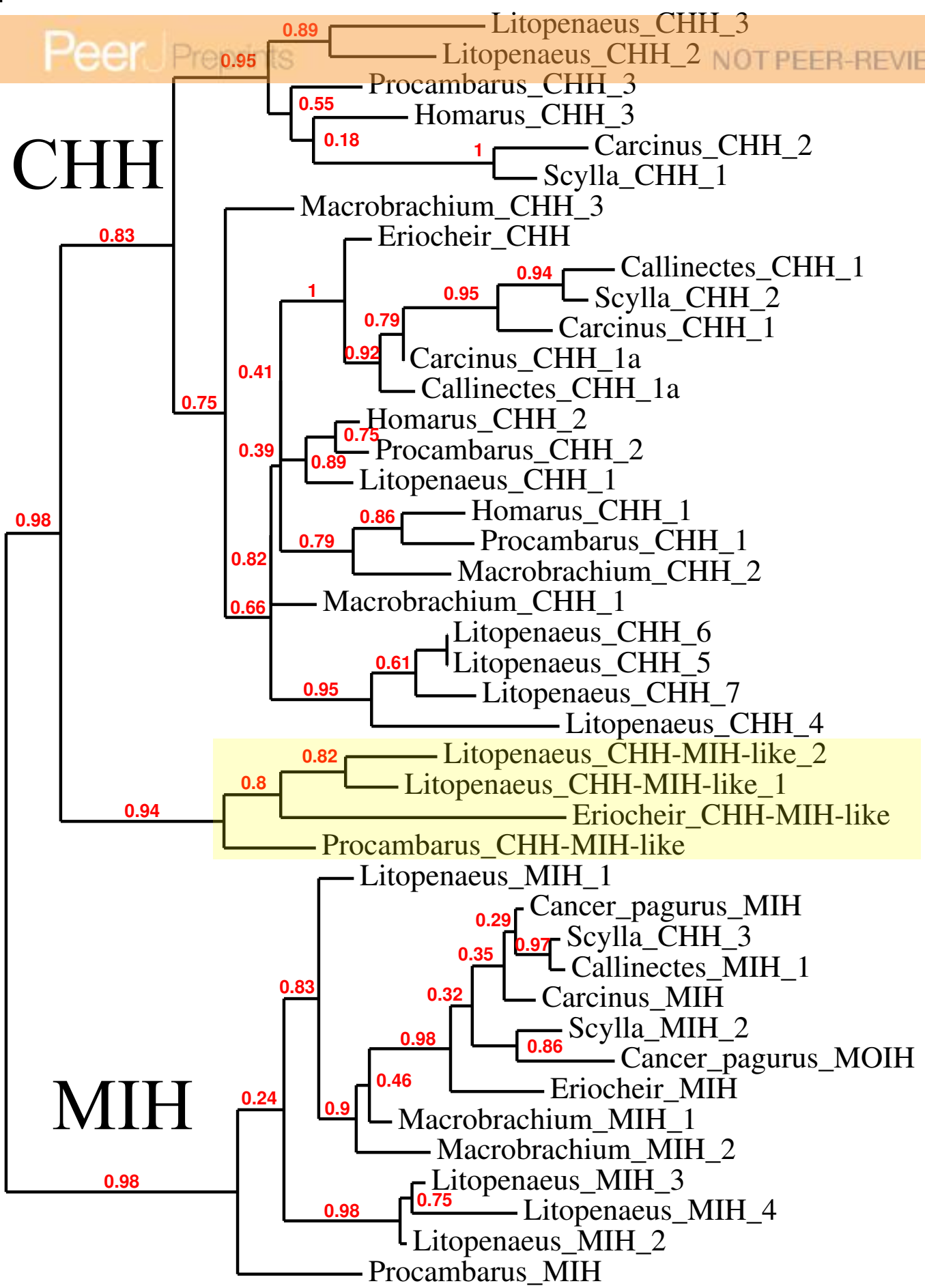


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

Figure 4(on next page)

Phylogenetic tree showing the evolutionary relationships between the CHH and MIH hormones.

Hormones are those identified from decapod SRAs as well as a few for which the biological activity has been described. Highlighted in yellow are the three sequences that on the tree are more similar to CHH, but lack the precursor-related peptide typically present in CHH



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

5

CFSH alignments.

The *S. paramosain* reads corresponding to the Scylla sequence in the tree are completely identical to *S. olivacea* sequence.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

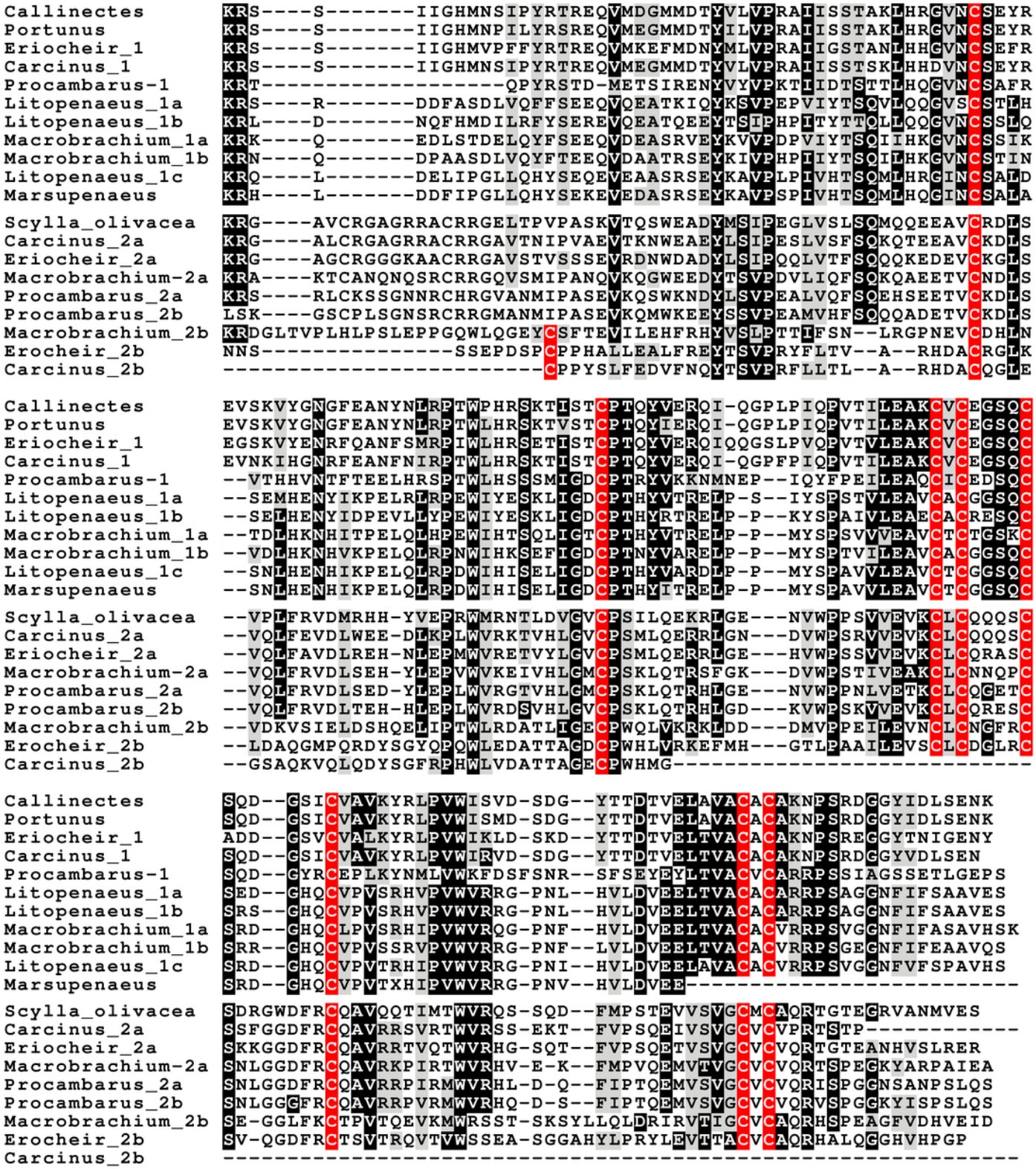
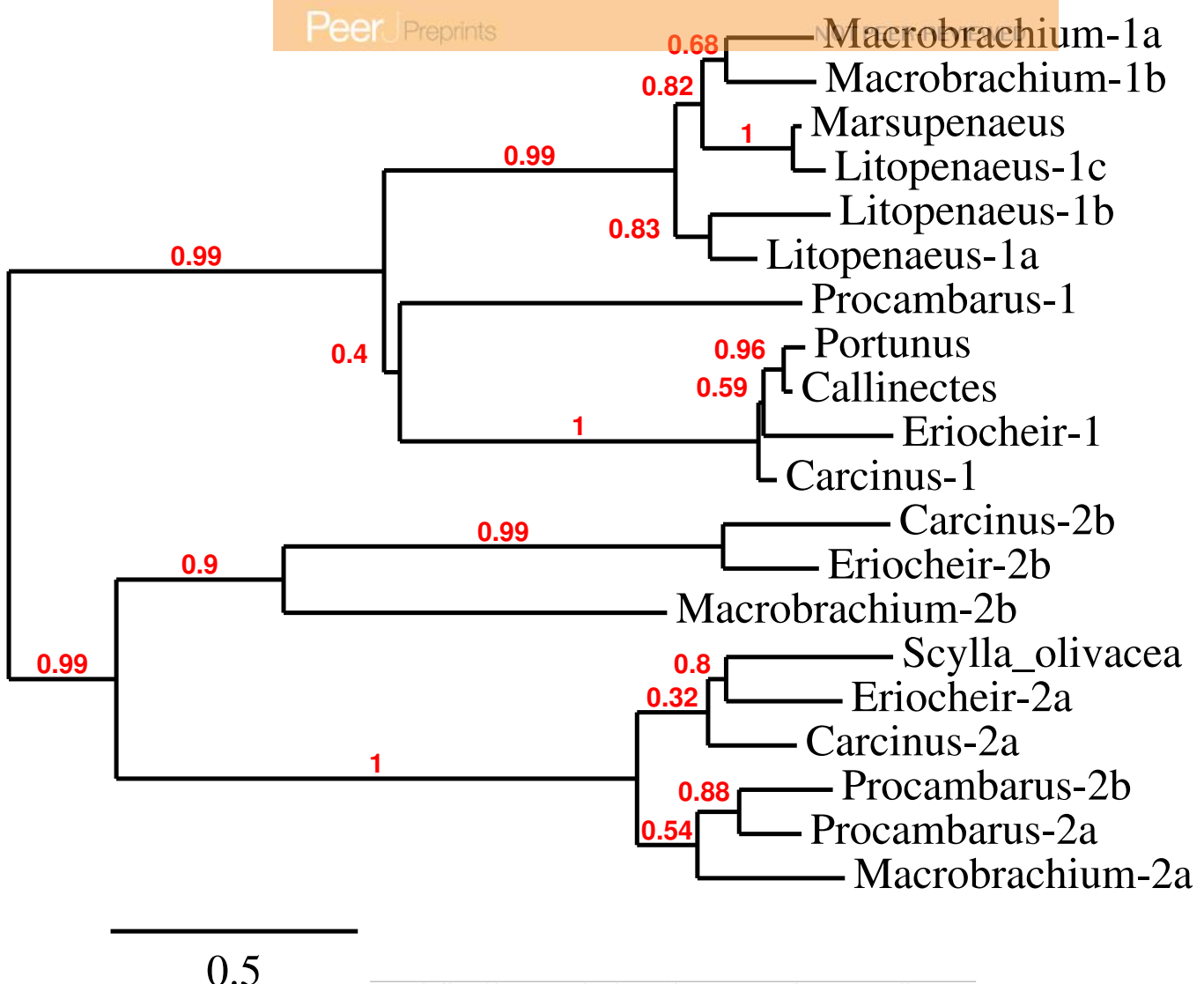


Figure 6(on next page)

CFSH phylogenetic tree.

Phylogenetic tree of the various CFSH orthologs identified here and elsewhere. The only *Scylla* sequence is from *S. olivacea* (GDRN01022056.1). *S. paramosain* has a very limited number of SRA reads that correspond to three orthologs found in *Carcinus* and *Eriocheir*. Note that *Macrobrachium*, *Litopenaeus* and *Procambarus* seem to have independently gone through relatively recent gene duplications.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

Table 1(on next page)

Configuration of *Eriocheir* neuroparsin genes 3 and 4.

The relative organization of the two neuroparsin genes relative to one another is indicated. The two genes are located on opposite strnads and each gene has four exons and three intron. Numbers indicated the lengths of the exons, introns and the intergenic distance in nucleotides.

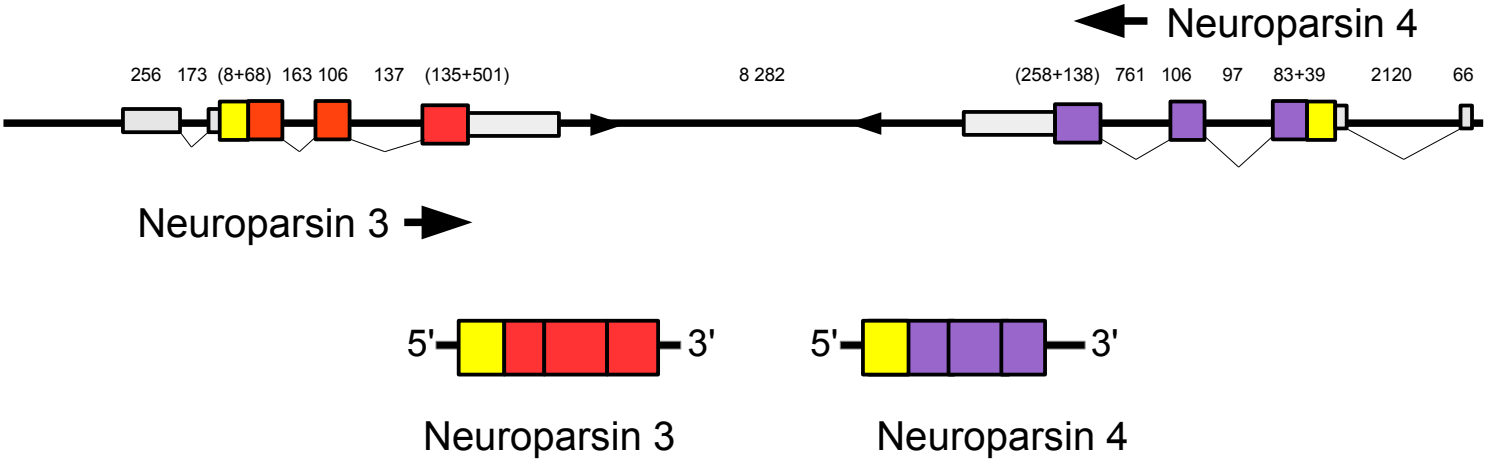
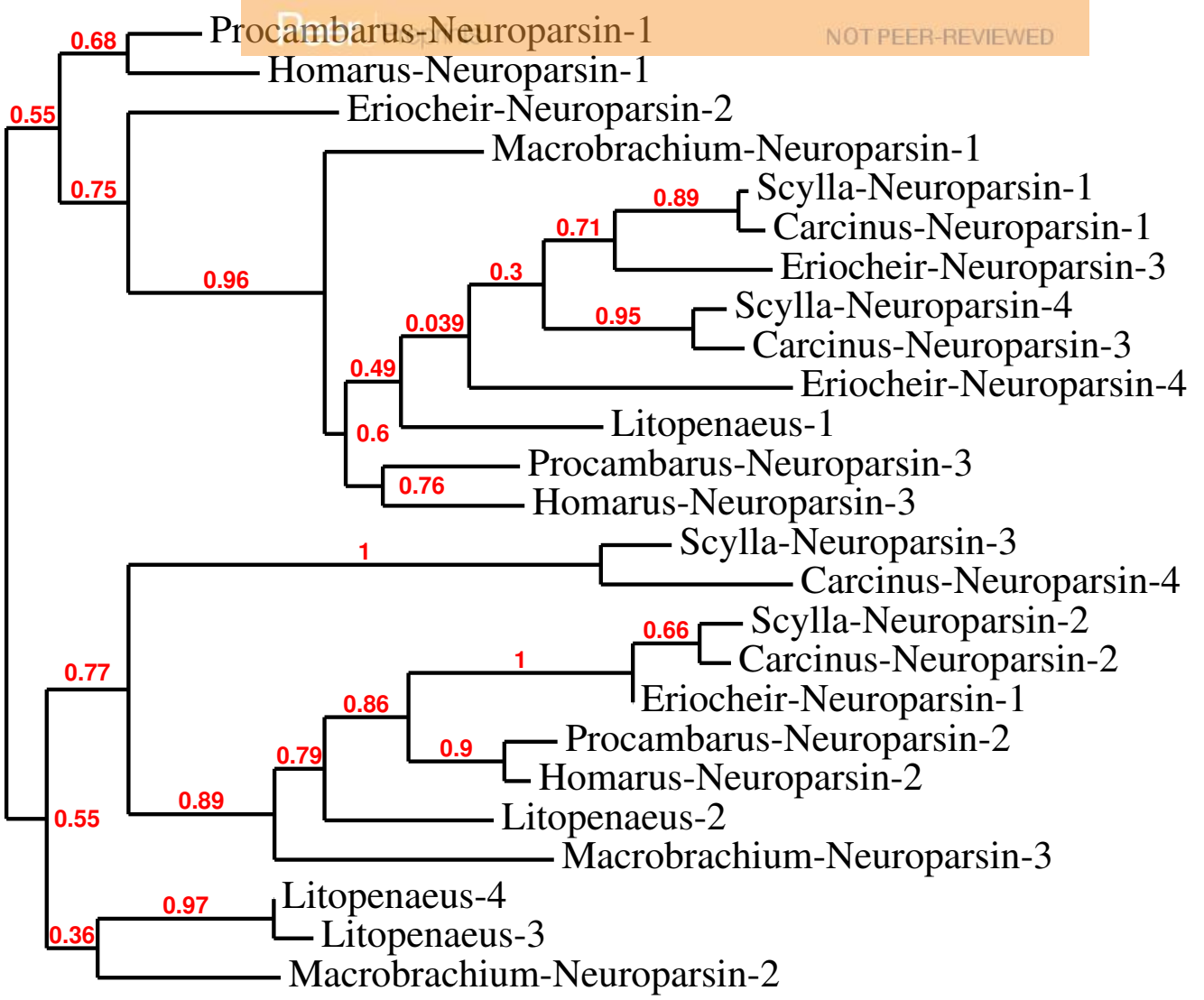


Figure 7 (on next page)

Neuroparsin phylogenetic tree.

The different decapod neuroparsin sequences found in the different species were used to make a phylogenetic tree. Note that the duplication of some neuroparsin likely occurred after the crabs separated from the other decapods.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

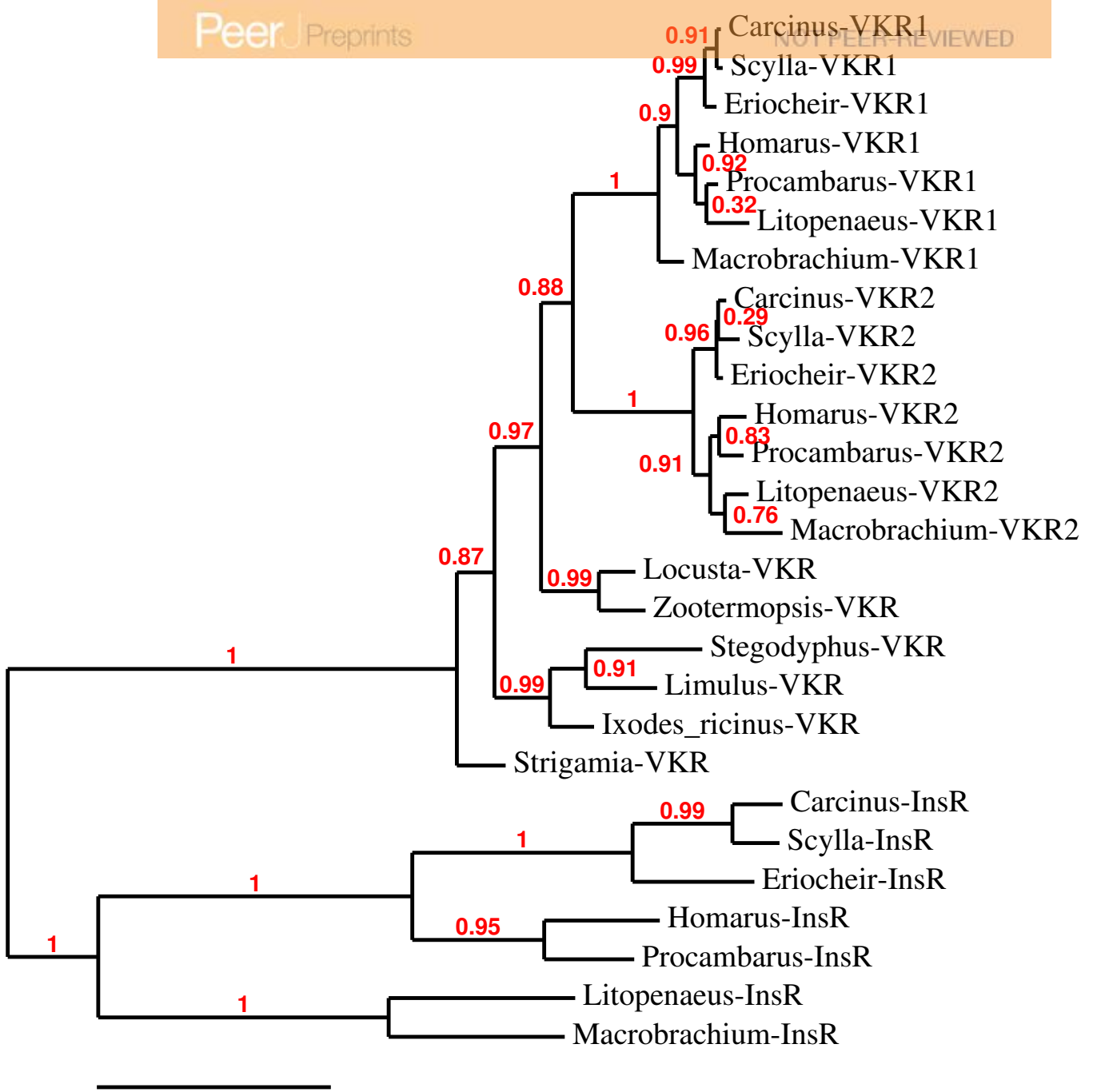


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

Figure 8(on next page)

Phylogenetic tree of the tyrosine kinase domains of the decapod insulin and venus kinase receptors.

Venus kinase receptors from the following species were added for increased resolution: *Limulus polyphemus, Stegodyphus mimosarum, Locusta migratoria, Ixodes ricinus* and *Zootermopsis nevadensis.* Note that the duplication of the venus kinase receptor gene is not generally present in arthropods and could thus be specific to crustaceans.



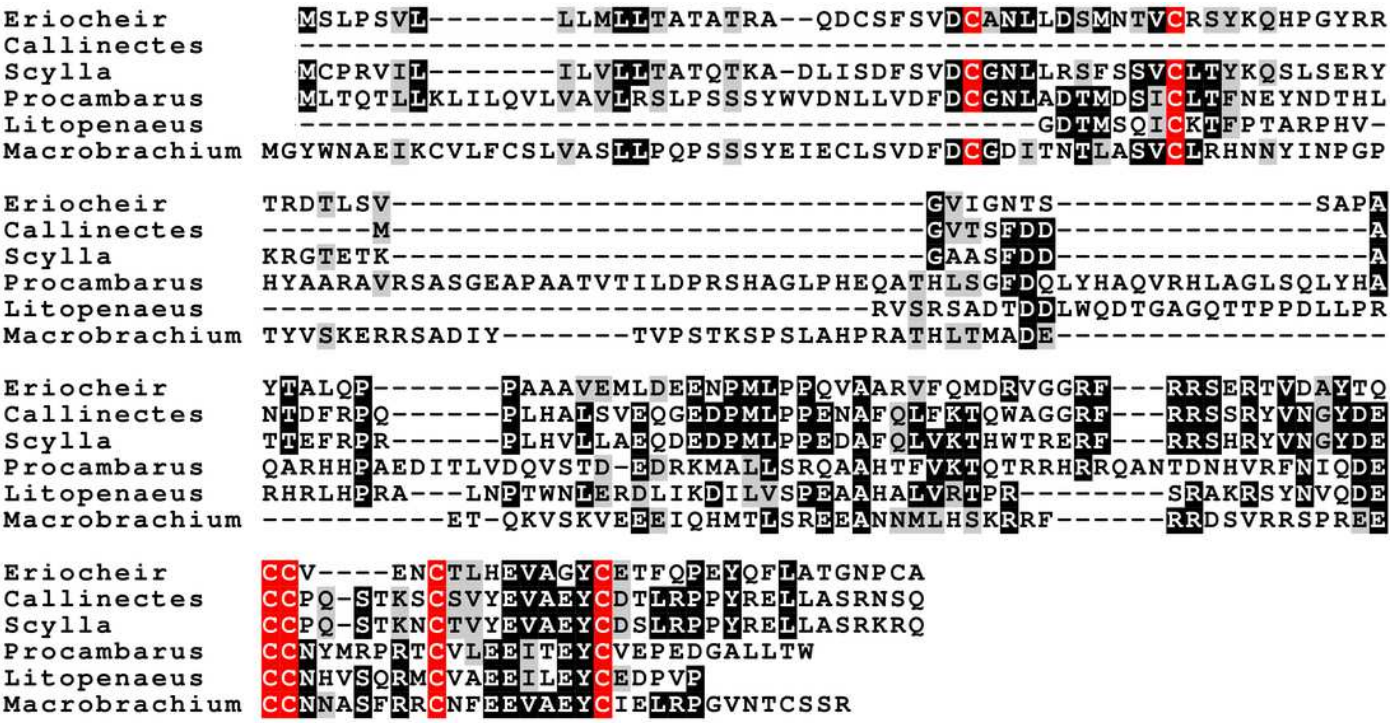
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016

9

Sequence alignment of the decapod adrogenic insulin-like peptides.

Note the relatively poor conservation of the primary sequences of these hormones.

Conserved residues indicated in black highlighting, and cysteine residues in red.

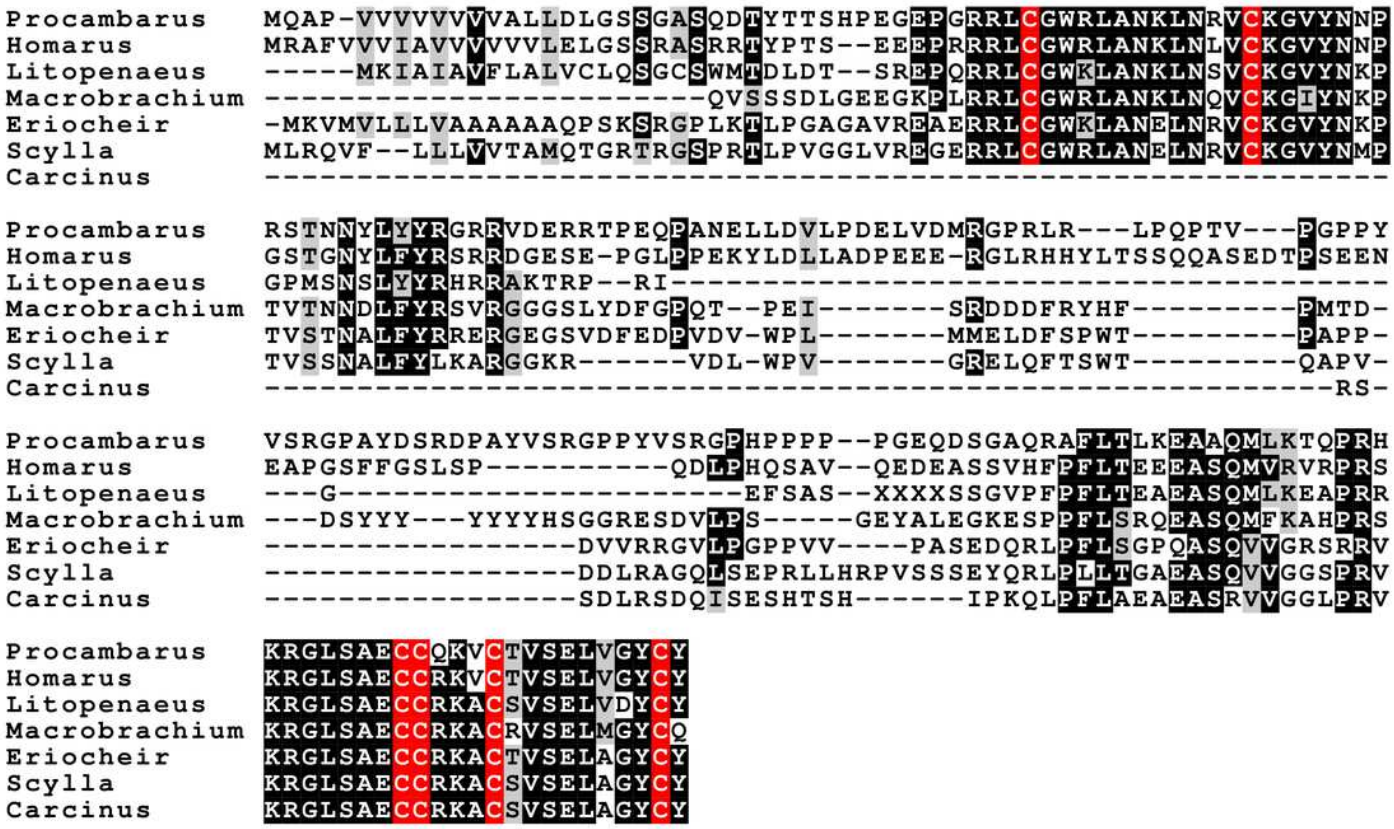


10

Sequence alignment of the decapod insulin-like peptides.

Note the much better conservation of the primary sequences of the A and B chains of these hormones. Conserved residues indicated in black highlighting, and cysteine residues in red. The *Carcinius* sequence, although incomplete, is clearly part of an insulin precursor.

Conserved residues indicated in black highlighting, and cysteine residues in red.



11

Sequence alignment of the decapod relaxins.

Note the relatively good sequences conservation between the different Decapod peptides and Dilp-7. Conserved residues indicated in black highlighting, and cysteine residues in red.

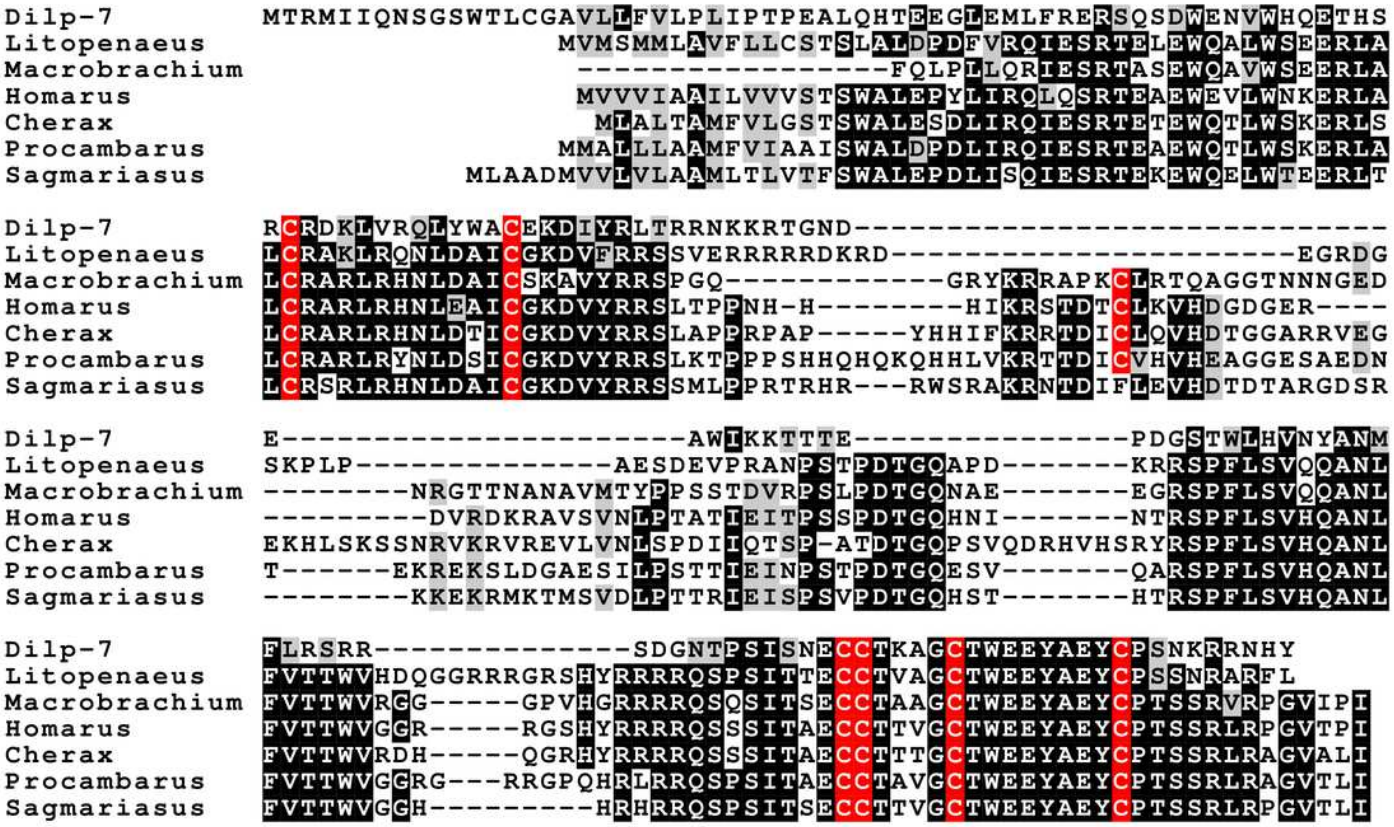
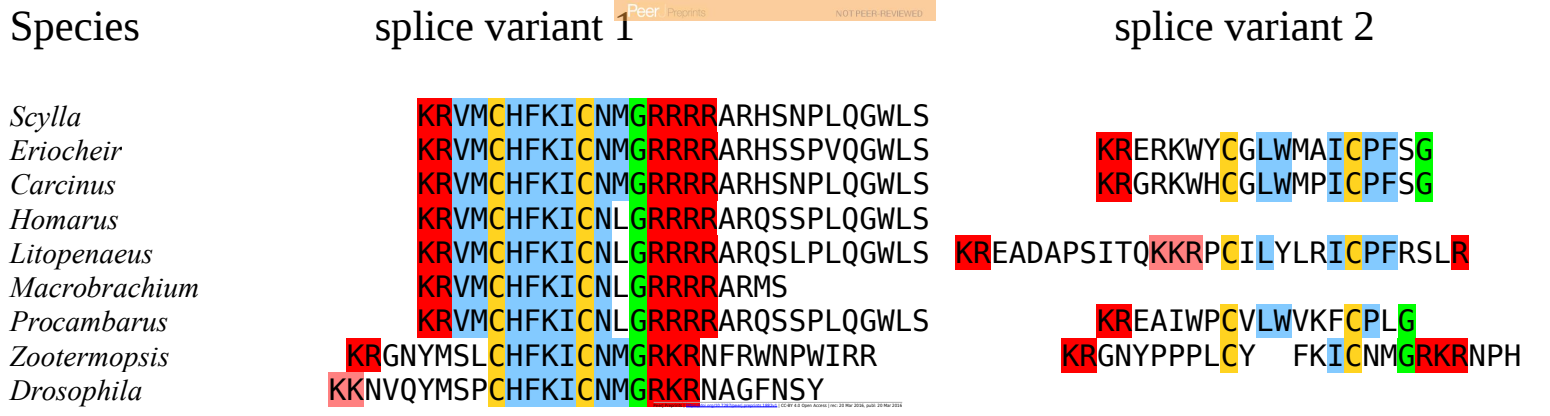


Table 2 (on next page)

Last parts of CNMamide precursors

Some arthropods produce alternatively spliced mRNA predicted to produce different CNMamides. Notice that the major splice variant produces a much better conserved neuropeptide than the alternative. Residues in red are predicted to be cleaved by convertase and removed by carboxypeptidase during processing; the green glycine residues will be transformed in C-terminal amides and the cysteine residues are orange. Residues conserved between the different species are in blue.



Sequence alignment of the decapod B-calcitonins.

Some of the decapod B-calcitonins are predicted to have two cysteine bridges in the N-terminal part of their sequence, rather than one.

13

Hyrg sequence alignment.

Note that only a small part of the sequence of this puative neuropeptide is conserved in both decapods as well as in *Euphausia crystallorophias.*

Table 3(on next page)

Ligand-receptor interactions of insulin-related peptides.

Figure indicates the postulated major interactions of the three decapod insulin-like peptides with three receptors. Secondary interactions are indicated by broken lines. Drosophila gene numbers for orthologous genes are indicated in red. LRR-GPCRs: Leucine-riche repeat GPCRs.

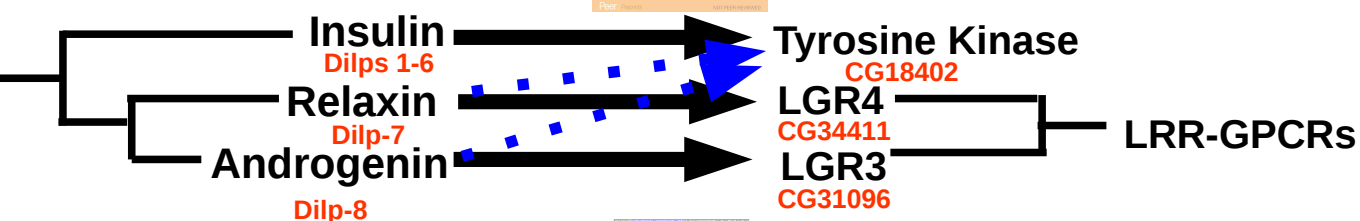

Insulin
Dilps 1-6
Relaxin
Dilp-7
Androgenin
Dilp-8
Tyrosine Kinase
CG18402
LGR4
CG34411
LGR3
CG31096
LRR-GPCRs

Table 4(on next page)

Tissue distribution of neuropeptides and neuropeptide GPCRs in various tissues. Part 1.

The number of individual reads found in different SRAs from eggs and eleven tissues or *Carcinus maenas.*

Eggs	Eye	Nerve	Intestine	Ovary	Testis	Epidermis	Muscle	Heart	Hepatopancreas	Gill	Haemolymph	
17	87	1	0	24	0	2	0	0	0	0	0	ACP
6	11	1	1	0	0	1	0	0	0	2	0	ACP-GPCR
119	823	1376	2	2	0	1	0	6	2	0	0	Agatoxin-like peptide
35	138	1093	1	1	3	4	1	0	2	1	83	Allatostatin A
14	45	115	12	11	1	42	143	25	2	1	0	AstA-GPCR
85	272	387	0	165	1	8	1	2	0	2	0	Allatostatin B (= mip)
8	5	37	8	27	3	27	3	0	29	7	4	AstB-GPCR
31	41	331	246	7	0	26	0	3	3	0	0	Allatostatin C
19	1	155	2	0	1	1	0	0	0	0	0	Allatostatin CC
7	46	76	0	53	0	3	0	0	0	1	1	Allatostatin CCC
19	17	10	2	105	4	86	1	0	0	0	0	AstC-GPCR
16	12	10	2	104	4	88	1	0	0	0	0	AstC-GPCR, splice variant
14	0	550	0	16	0	7	0	0	0	0	0	Bursicon-A
10	0	444	2	0	0	0	0	0	0	0	0	Bursicon-B
202	59	98	46	31	20	49	91	51	3	65	0	Bursicon-GPCR
24	12	11	178	0	0	0	0	0	0	0	0	Calcitonin
17	10	10	112	0	0	0	0	0	0	0	0	Calcitonin common exon
2	1	5	2	0	0	0	0	0	0	0	0	Calcitonin A-specific
18	10	1	121	0	0	0	0	0	0	0	0	Calcitonin B-specific
3	8	38	0	0	0	0	0	0	0	1	0	CCHamide 1
3	12	13	1	1	0	0	1	1	0	2	17	CCHamide 2
3	3	4	8	22	3	9	0	0	0	0	0	CCHamide-GPCR-1
7	2	9	0	25	126	7	0	0	0	0	0	CCHamide-GPCR-2
2	23	34	0	112	0	20	0	0	0	0	0	CNMamide
0	5	6	0	40	0	3	0	0	0	0	0	CNMa a specific
0	0	0	0	17	0	1	0	0	0	0	0	CNMa b specific
20	153	4	0	0	0	0	0	0	0	0	0	Corazonin
26	89	208	0	0	1	0	0	106	0	0	0	CRF-like diuretic hormone
10	6	9	0	0	0	0	0	1	0	0	0	CRF-like DH-GPCR
34	31	823	0	0	0	0	0	1	0	0	0	CCAP
22	3	23	6	25	1	11	14	9	2	1	0	CCAP-GPCRa
15	3	25	5	18	0	11	15	9	2	2	0	CCAP-GPCRb
1	353	0	0	0	0	2	0	0	0	0	0	CFSH 1
1	23	1	0	1	0	0	0	0	0	0	0	CSFH 2a
0	1	0	0	2	0	1	0	0	0	0	0	CFSH 2b
83	3418	476	101	78	25	46	26	28	14	61	4	CHH 1
111	4969	663	141	125	35	59	47	33	18	80	7	CHH 1 alternative splice product
24	366	330	2757	44	49	41	4	20	45	18	2	CHH 2
43	173	735	1	2	0	0	0	4	0	0	0	DH31
11	161	28	0	0	6	1	0	0	0	0	0	Eclosion hormone 1
10	0	0	0	0	0	1	0	0	0	0	0	Eclosion hormone 2
103	48	120	1	3	5	50	0	1	0	0	0	ETH
11	30	45	0	0	0	0	0	0	0	0	0	EFLamide
15	7	10	0	0	1	10	0	0	0	0	0	EFLamide-GPCR
5	12	32	156	0	4	3	0	0	2	0	1	Elevenin
5	8	6	81	1	0	4	0	1	3	4	4	Elevenin-GPCR-1
14	75	84	36	78	57	92	67	43	28	24	25	Elevenin-GPCR-2
33	233	191	0	1	1	34	0	1	0	0	0	FMRFamide
3	19	51	2	4	2	81	41	49	3	25	0	FMRFa-GPCR

Table 5(on next page)

Tissue distribution of neuropeptides and neuropeptide GPCRs in various tissues. Part 2.

The number of individual reads found in different SRAs from eggs and eleven tissues or *Carcinus maenas.*

Eggs	Eye	Nerve	Intestine	Ovary	Testis	Epidermis	Muscle	Heart	Hepatopancreas	Gill	Haemolymph	
12	46	43	0	0	0	1	0	0	0	0	0	GPA2
20	48	75	2	5	0	2	0	1	1	0	0	GPB5
80	216	788	45	41	7	64	36	80	0	1065	2	GPA2/GPB5-GPCR
1	571	3	30	0	12	1	1	0	2	0	1	Hyrg
3	1	0	0	0	0	2	0	0	0	0	0	Insulin
26	29	22	16	150	13	57	4	3	2	3	0	Insulin tyrosine kinase receptor
11	71	141	0	2	0	0	0	0	1	0	0	Leucokinin-a
11	120	228	0	0	0	0	0	0	1	0	0	Leucokinin-b
2	8	16	1	9	0	3	0	0	0	0	0	Leucokinin-GPCR
4	189	0	0	0	0	3	0	0	0	0	0	MIH
26	149	298	1	1	1	1	0	0	0	0	0	Myosuppressin
39	13	45	133	18	0	7	6	2	0	4	0	Myosuppressin-GPCR ?
28	87	81	0	0	0	5	0	0	0	0	0	Natalisin
130	952	4536	635	324	491	365	331	1444	161	1935	0	Neuroparsin 1
73	182	478	13	22	0	6	4	8	0	6	0	Neuroparsin 2
2	83	160	94	52	6	5	22	12	0	81	0	Neuroparsin 3
308	94	877	108	402	124	665	2910	986	928	727	5	Venus kinase receptor 1
357	258	1211	233	233	115	257	1119	519	131	199	17	Venus kinase receptor 2
5	28	15	0	0	0	2	1	2	0	1	1	Neuropeptide F 1a
3	22	12	0	0	0	2	1	2	0	1	1	Neuropeptide F 1b
1	1	5	0	0	0	0	0	0	0	0	0	NPF 1b specific
20	62	45	0	1	3	0	0	0	0	3	0	Neuropeptide F 2
50	603	449	8	0	0	10	0	0	0	0	1	Neuropeptide-like precursor 1
40	223	661	26	0	0	2	0	0	1	0	0	Orcokinin-A
8	5	20	1	1	2	3	0	0	0	1	0	Periviscerokinin
112	446	2	1	2	4	5	3	3	0	4	2	PDH 1
25	101	0	0	1	1	0	0	0	0	0	0	PDH 2
26	7	3	1	0	1	1	0	0	0	0	0	PDH-GPCR-1
16	18	3	1	0	0	1	0	0	0	0	0	PDH-GPCR-2
17	91	411	0	0	1	1	0	0	0	1	0	Proctolin
6	20	107	6	13	2	19	94	43	3	15	0	Proctolin-GPCR-1
12	18	119	0	7	1	10	0	3	0	0	0	Proctolin-GPCR-2
25	70	94	0	12	0	0	0	0	0	0	0	Pyrokinin
2	2	6	0	2	0	28	0	0	0	0	0	Pyrokinin-1-GPCR-2
42	182	65	0	0	0	2	0	0	0	0	0	RPCH
93	2	103	0	0	0	0	0	0	0	0	0	RYamide
4	5	5	8	2	0	2	0	0	8	0	0	Ryamide-GPCR-1
5	8	7	9	3	0	35	0	0	146	0	0	Ryamide-GPCR-2
3	26	19	2	10	1	6	0	0	0	0	0	sNPF
5	1	20	1	0	3	79	0	0	2	0	0	sNPF-GPCR-1
1	0	7	0	0	0	6	0	1	1	0	0	sNPF-GPCR-2
44	410	395	0	0	0	1	0	0	0	1	1	SIFamide
3	6	5	4	0	6	1	0	0	0	0	0	Sulfakinin
30	296	124	246	3	0	0	2	2	0	0	0	Tachykinin
8	12	31	9	18	1	45	4	4	5	8	169	Trissin
10	36	42	0	4	1	5	0	0	0	0	2	Vasopressin
4	3	2	0	0	0	11	0	0	0	1	0	Vasopressin-GPCR
9	10	7	3	52	0	1	0	0	0	0	1	CG31096 ortholog
1	2	0	0	3	0	0	1	0	0	0	0	CG34411 ortholog

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1882v1 | CC-BY 4.0 Open Access | rec: 20 Mar 2016, publ: 20 Mar 2016