

SCIENTIFIC REPORTS



OPEN

Transcriptomic characterization and curation of candidate neuropeptides regulating reproduction in the eyestalk ganglia of the Australian crayfish, *Cherax quadricarinatus*

Tuan Viet Nguyen, Scott F. Cummins, Abigail Elizur & Tomer Ventura

Received: 05 July 2016
Accepted: 11 November 2016
Published: 07 December 2016

The Australian redclaw crayfish (*Cherax quadricarinatus*) has recently received attention as an emerging candidate for sustainable aquaculture production in Australia and worldwide. More importantly, *C. quadricarinatus* serves as a good model organism for the commercially important group of decapod crustaceans as it is distributed worldwide, easy to maintain in the laboratory and its reproductive cycle has been well documented. In order to better understand the key reproduction and development regulating mechanisms in decapod crustaceans, the molecular toolkit available for model organisms such as *C. quadricarinatus* must be expanded. However, there has been no study undertaken to establish the *C. quadricarinatus* neuropeptidome. Here we report a comprehensive study of the neuropeptide genes expressed in the eyestalk in the Australian crayfish *C. quadricarinatus*. We characterised 53 putative neuropeptide-encoding transcripts based on key features of neuropeptides as characterised in other species. Of those, 14 neuropeptides implicated in reproduction regulation were chosen for assessment of their tissue distribution using RT-PCR. Further insights are discussed in relation to current knowledge of neuropeptides in other species and potential follow up studies. Overall, the resulting data lays the foundation for future gene-based neuroendocrinology studies in *C. quadricarinatus*.

Neuropeptides play an important role in the regulation of both vertebrate and invertebrate reproduction^{1–3}. By binding to receptors in target tissues, neuropeptides elicit a number of downstream cascades including changes in secondary messengers, phosphorylation and eventually altered transcription activity which leads to a biological response⁴. *In silico* neuropeptidome research aimed at understanding different physiological aspects of crustacean biology has been extensively studied in several crustacean species, including *Scylla paramamosain*⁵, *Macrobrachium rosenbergii*⁶, *Procambarus clarkii*⁷, *Homarus americanus* and *Sagmariasus verreauxi*^{8,9}.

The crustacean eyestalk (together with the supra-oesophageal and thoracic ganglia), is known to be the primary source of expression and secretion of key factors which regulate many aspects of crustacean behaviour and physiology, including growth and reproduction^{10–12}. These biological processes operate under the multifaceted control of biogenic amines, neuropeptides, hormones and lipid metabolites such as methyl farnesoate and ecdysteroids¹³. Crustacean neuropeptide research has focused on a number of key factors that are related to reproduction and sexual development. For instance, the immunoreactivity of antibodies raised against gonadotropin releasing hormone (GnRH) in the central nervous system of the giant freshwater prawn *M. rosenbergii* suggests a vertebrate-like reproductive system in crustaceans¹⁴. Similarly, a GnRH-like peptide was found in the white leg shrimp *Litopenaeus vannamei*¹⁵, and the primary structure of GnRH was then reported by mass spectral analysis

GeneCology Research Centre, Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Sunshine Coast, Queensland, Australia. Correspondence and requests for materials should be addressed to T.V. (email: tventura@usc.edu.au)

in the red swamp crayfish *P. clarkii*¹⁶. In addition, a recently discovered crustacean female sex hormone (CFSH) in the blue crab *Callinectes sapidus*¹⁷ has been implicated in female sexual maturation, leading to identification of orthologs in the crayfish *P. clarkii*⁷ and the Eastern spiny lobster *S. verreauxi*⁹. In addition, research has been focused into the crustacean androgenic gland specific insulin-like hormone^{18–22}, work which was recently supplemented with the characterization of its receptor^{23,24}.

In parallel, a well-known candidate for reproduction studies is the Crustacean Hyperglycemic Hormone (CHH) superfamily, which includes CHH itself, as well as Molt Inhibiting Hormone (MIH), Gonad Inhibiting Hormone (GIH, also referred to as Vitellogenesis Inhibiting Hormone (VIH)), and Ion transport protein (ITP) [reviewed by Webster, *et al.*²⁵, Chan, *et al.*¹³, and Subramoniam²⁶]. Still, it is important to note that in spite of the known function of some members in the CHH superfamily, their identity and functionality is vastly uncharacterised¹³. Lastly, there are numerous neuropeptides that have previously been demonstrated to have other functions (for example, cardio-activity, osmoregulation, pigmentation, stress response) but then later were assigned novel roles in reproduction and/or sexual development^{27–30}. These findings highlight that while our understanding of the molecular mechanisms of neuropeptides involved in crustacean reproduction processes is quite advanced, it is yet still far from being complete.

In the past, classical peptidomics studies used liquid chromatography and/or mass spectrometry to detect the neuropeptides present in the species³¹. However, with the recent advancement in next generation sequencing, bioinformatics analyses platforms as well as the availability of transcriptome databases, the pace at which neuropeptides are discovered has accelerated considerably. Within a short timeframe, a list of putative neuropeptides in a vast number of non-model crustacean species has been generated including: A-type allatostatin (AST-A), B-type allatostatin (AST-B), C-type allatostatin (AST-C), bursicon, CHH, crustacean cardioactive peptide (CCAP), diuretic hormone 31 (DH31), diuretic hormone 44 (DH44), eclosion hormone (EH), myosuppressin, orcokinin, short neuropeptide F (SPF), neuroparsin (NP), pyrokinin, pigment-dispersing hormone (PDH), red pigment-concentrating hormone (RPCH), short neuropeptide F (sNPF), SIFamide (SIF), sulfakinin and tachykinin (TK). This approach effectively characterised the neuropeptidome for a number of crustacean species that lack genomic information^{5–9,32}. Still, application of this information for the proper control of reproduction is in its infancy, primarily due to the gaps between *in silico* and *in vivo* functional experiments.

The Australian redclaw crayfish *Cherax quadricarinatus*, has recently received attention as an emerging candidate for sustainable aquaculture production in Australia and abroad³³. FAO predicts that *C. quadricarinatus* would be an attractive freshwater species due to the relative ease of farming and distinct advantages over mariculture³⁴ provided proper guidelines and sustainable management practices are implemented. Moreover, *C. quadricarinatus* serves as a good model organism for the commercially important group of decapod crustaceans; it is hardy, endures a high range of conditions, easy to maintain in the laboratory and its reproductive cycle has been well documented³⁵, with a unique feature of intersex individuals which occur naturally in the population^{36–39}. The molecular toolbox available for *C. quadricarinatus* was recently supplemented with several transcriptomes^{40,41} however, no study has been undertaken to establish the *C. quadricarinatus* neuropeptidome.

This study reports the *in silico* mining and characterization of 53 predicted neuropeptide-encoding genes in *C. quadricarinatus*. Fourteen out of these neuropeptide genes, which were previously linked with reproduction, were screened for tissue spatial expression. Alongside, based on knowledge gained from other crustacean neuropeptide studies, we also highlighted several key candidates for future gene-based studies. Utilizing this molecular toolkit could enhance the aquaculture practice of *C. quadricarinatus* and other crustaceans through comparative studies.

Materials and Methods

Ethical statement. *C. quadricarinatus* are not endangered or protected in Australia where sampling occurred. Animal use and care protocols were approved by the GeneCology Research Centre. No special sampling permission was required.

Animals. Mature individuals of the Australian red claw crayfish *C. quadricarinatus*, ranging in size from 35 to 60 g, were purchased from a local farmer in Queensland, Australia. Prior to dissections, individuals were chilled in ice-cold water for 20 minutes to minimize handling stress. For the eyestalk transcriptome, eyestalks from 15 males and 15 females were pooled for each gender separately. For RT-PCR analysis, 10 female individuals, whose tissues were retrieved, were individually weighed and their carapace lengths measured, as well as the weight of their hepatopancreas and ovary. Crayfish tissues including eyestalk, brain and thoracic ganglia (combined), ovary, hepatopancreas, heart, abdomen muscle and antennal gland were dissected and stored in -80°C until used. Gonadosomatic index was calculated as the ovary weight divided by the total weight times hundred. Ovary was kept for histological examination. The majority of individuals were in their intermolt, assessed by the lack of discernible gastroliths.

Sample preparation and Illumina sequencing. Total RNA was extracted using Trizol Reagent (Invitrogen) following the manufacturer's protocol. All RNA samples were checked on Bioanalyzer 2100 before sending for sequencing. Sequencing using HiSeq 2000 was conducted at BGI, Hong Kong as per manufacturer's protocol (Illumina). Briefly, mRNA was isolated using oligo (dT) beads. All the mRNA resulted from fragmentation process were reverse transcribed into first strand cDNA using reverse transcriptase and random primers. Second strand of cDNA were then synthesized using DNA polymerase I and treated with RNase H. All products were further purified and amplified using PCR to generate the final libraries. All prepared libraries were then sequenced using the Illumina HiSeq 2000, resulting with the final 90 bp paired-end libraries used in the current study. The data was stored in the form of FastQ files.

Bioinformatics analyses. Separately produced FastQ files from males and females were merged together to create one final raw file that represents the sequenced eyestalk transcriptome. FastQ qualities were assessed using FastQC software. (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). For maximum discovery of transcripts, the current study consisted of two different *de novo* assembly approaches. The first approach deployed a cleaning of low quality reads using Trimmomatic⁴² with the command parameter “LEADING = 3, TRAILING = 3, SLIDING WINDOW = 4,15, MIN SIZE = 30, HEADCROP = 6”. Followed by a *de novo* assembly using Trinity_{r20140717}⁴³ with the parameter “-seqType fq -left Cq_EY_1_QCed.fastq -right Cq_EY_2_QCed.fastq -CPU 32 -JM 250G -normalize_reads -min_kmer_cov 2 -min_contig_length 200 -bfly_opts -PasaFly -group_pairs_distance 700 -extended_lock”. The second approach was done mostly using a pre-designed workflow available in the CLC Genomics Workbench v8.0.3 (<https://www.qiagenbioinformatics.com/>). Briefly, raw reads were filtered based on PHRED score (>30), then loaded onto the CLC *de novo* assembly module, running with parameters: word size = 25, bubbles size = 50, minimum contig length = 200 bp. The final reference transcriptomes were then blasted against the NCBI non-redundant (nr) database for annotation of transcripts using BLAST+⁴⁴. For searching of neuropeptides in *C. quadricarinatus*, all annotated sequences were scanned for keywords of previously known neuropeptides for instance “allatostatin”, “neuroparsin”, “corazonin”, or a known conserved amino acid motif such as “PSGFLGR”; list of motif used was taken as described elsewhere¹⁰. These sequences were then re-validated using blastp algorithm. For illustration purposes, cDNA sequences were converted to amino acids using ExPASy translate tool available online (<http://web.expasy.org/translate/>), open reading frames (ORFs) were then chosen and loaded onto a pre-designed pipeline employing SignalP 4.0⁴⁵ to validate they have a signal peptide and TargetP 1.1⁴⁶ to confirm the neuropeptide is in the secretory pathway. Amino acid sequences of neuropeptides were saved as Fasta files for ease of analysis (Supplementary file 1). Amino acid sequences were submitted to the NeuroPred program⁴⁷ to predict peptide cleavage sites. Schematic diagrams of neuropeptide structures were illustrated using the Illustrator for Biological Sequences (IBS) suite⁴⁸. Multiple sequence alignments were done using CLC Genomics Workbench v8.0.3 (<https://www.qiagenbioinformatics.com/>). Multiple alignment files were also imported into MEGA 7.0⁴⁹ for phylogenetic analysis. A maximum likelihood based approach was conducted with 500 bootstrap trials. Genbank accession number of all used sequences can be found in Supplementary file 2.

Tissue distribution using RT-PCR. RT-PCR was performed as previously reported with slight modifications⁵⁰. Out of the 10 individuals sampled above, RNA from different tissues of one mature animal with developed ovary (intermolt stage, weight 52.4 g, GSI = 4.2, oocyte diameter = 2.73 mm) was used to synthesize the complementary DNA libraries using Trizol Reagent (Invitrogen), following the manufacturer’s protocol. In total, 7 tissues including: eyestalk, brain and thoracic ganglia (combined), ovary, heart, abdomen muscle and antennal gland were used for PCR amplification to establish tissue distribution, using 14 neuropeptide gene primers. These neuropeptides were chosen based on literature review highlighting them as candidates for reproduction regulation in other crustacean species. Total RNA was re-extracted using the previously described method followed by cDNA synthesis using Tetro cDNA synthesis kit (Bioline, UK; using approximately 1 µg RNA per tissue). Amplifications were carried out using a touch-down PCR to allow most PCR products to be amplified with minimal non-specific signal. PCR settings were 94 °C for 3 min, followed by 37 cycles of touch down, 94 °C for 30 s, 62–57 °C for 30 s (with 1 °C decrement for each of the first 6 cycles) and 72 °C for 45 s. *Cherax quadricarinatus* Beta-actin (GenBank Accession number AY430093.1) was chosen as the house-keeping gene for this experiment. Primers used for all neuropeptides are listed in Supplementary file 3. Following PCR, products were loaded onto an agarose gel with ethidium bromide (1.5% TBE, EtBr concentration 10 mg/ml), electrophoresed for 30 min in 120 V, 0.4 mA and later visualized under UV-light.

Results

Illumina sequencing and *de novo* assembly. In total, the *C. quadricarinatus* pooled males and female libraries resulted in more than 200 million reads with the designated read length of 90 bp (paired end). All raw reads were submitted to the NCBI SRA under Accession: SRP091408. Read quality statistics including distribution of PHRED score, read length distribution, consistent GC content between samples and N bases in each tissue library were performed (Supplementary file 4). After trimming, the pooled sequenced library had more than 90% reads above the quality threshold (data not shown). Only quality-trimmed reads were then used in all downstream analyses.

In order to maximize the potential to discover neuropeptide transcripts, two different *de novo* assemblers were used. *De novo* assembled transcriptome of the current dataset using either Trinity or CLC *de novo* assembler yielded a large amount of transcripts. A brief summary of the *de novo* assembly statistics is provided in Table 1. All reads were subjected to the BLAST+ program and scanned against the NCBI nr database for hits. Neuropeptide transcripts were then manually chosen based on a number of criteria including N-terminal signal peptide availability, presence of previously known motifs and completeness of transcripts.

Neuropeptide discovery by transcriptome mining: bioinformatics analysis and peptide prediction.

Using a number of available bioinformatics packages, we were able to identify 53 predicted neuropeptide transcripts from the eyestalk of *C. quadricarinatus*, including most of the neuropeptides that were previously identified in other crustacean/insect species (Table 2). A comparative catalogue of neuropeptide sequences generated in the current study and similar characterization of neuropeptides from other crustacean species is also available in Supplementary material 5.

	Trinity	CLC
Number of transcripts	195,402	93,455
Total size of transcripts (bp)	94,219,731	63,041,056
Longest transcripts (bp)	9,234	14,169
Number of transcripts > 1 k bp	16,256	13,607
Number of transcripts > 10 k bp	0	6
Mean transcript size	482	675
N50 transcript length	561	746

Table 1. *De novo* assembly statistics of the reference transcriptome.

Adipokinetic hormone/Corazonin related peptide (ACP), allatostatin A, B (AST-A, B), allatostatin C/Prohormone 1 and allatostatin CC (AST-C, AST-CC), bursicon- α (partial) and CCAP. A putative ACP precursor can be deduced from the *de novo* transcriptome assembly of *C. quadricarinatus* eyestalk (Fig. 1A). The predicted neuropeptide is 12aa long, immediately following a 20aa signal peptide and preceding a dibasic cleavage K₃₃R. Global alignment of *Cq*-ACP with ACP from other related species shows most significant conservation within the ACP mature peptide. We detected a partial transcript (without signal peptide) that encoded for the precursor of *Cq*-AST-A (Fig. 1B), containing 17 predicted peptides flanked by multiple dibasic cleavage sites. The conserved motif recorded was **XYXFLGamide**, which is similar to insect and other crustacean AST-A. One *Cq*-AST-B transcript was found in our *de novo* assembly (Fig. 1C). The neuropeptide precursor has a 22aa signal peptide and six dibasic cleavage sites that, if processed, are predicted to release 5 mature peptides of 11–12aa, and contain the conserved **XWXXXXGXWamide** motif, also observed in insects and other crustaceans⁵¹. Three isoforms of *Cq*-AST-C were found in the eyestalk *de novo* assembly (Fig. 1D). First, an AST-C-like/Prohormone-1 precursor, with a 24aa signal peptide and a 15aa mature peptide, second, a 36aa *Cq*-AST-C was found without a signal peptide, with the conserved **PISCFamide**, and, finally a *Cq*-AST-CC was found with a predicted precursor of 138aa and a 46aa mature peptide. One transcript encoding a *Cq*-bursicon- α was identified in this study, although with no apparent signal peptide or cleavage site following the mature peptide, it is a partial sequence. The deduced amino acid sequence for *Cq*-bursicon- α detected from the *de novo* assembly is SGIFLSCPQILTRAPIDCMCRPCTDVEEGTVLAQEIANFIHDSMPGNVPFLK. A 138aa *Cq*-CCAP precursor was deduced, containing a signal peptide and two cleavage sites (K₄₆R and K₅₈K) that could release a mature peptide **PFCNAFTGCamide** (Fig. 1E).

CCHamide (CCH), crustacean female sex hormone-like (CFSH-like), DH31, corazonin, DH44, eclosion hormone, elevenin. We identified two *Cq*-CCH transcripts in the current *de novo* transcriptome assembly of *C. quadricarinatus* (Fig. 2A). The two transcripts encode for 215aa and 42aa proteins, respectively, each with a signal peptide. *Cq*-CCH-1 has two cleavage sites (K₃₂ and G₄₈K), that if processed would release the mature peptide SCSQFGHSCFGAHamide, while the predicted *Cq*-CCH-2 mature peptide contains GGCLNYGHSLGHAHamide, which has a C-terminal K₃₉R cleavage site. The conserved motif for *Cq*-CCHamide is **GHSCXGAHamide**. A putative CFSH-like transcript was detected in the eyestalk of *C. quadricarinatus*. *Cq*-CFSH-like precursor is 264aa in length, with a signal peptide and a K₁₄₉R cleavage site, followed immediately with the mature peptide (Fig. 2B). A predicted corazonin precursor (*Cq*-Crz) includes a 25aa signal peptide, followed immediately by the *Cq*-Crz mature peptide **pQTFQYSRGWTNamide** and a R₃₇KR cleavage site. Multiple sequence alignments show high similarity of *Cq*-Crz mature peptide with Crz of *M. rosenbergii*, *P. clarkii*, *S. verreauxi* and insects including *Bombyx mori* and *Nilaparvata lugens* (Fig. 2C). A *Cq*-DH31 precursor was deduced to be containing a precursor 23aa signal peptide, a dibasic K₇₀R and R₁₀₄R cleavage site, that if processed would release a 31aa mature DH31 peptide (Fig. 2D). A DH44 transcript encodes a precursor of 84aa, with a 23aa signal peptide and 31aa DH44 mature peptide processed from two dibasic cleavage sites (R₃₈K and K₆₂R) (Fig. 2E). A full-length transcript of eclosion prohormone can be deduced to be consisting of 82aa, including a 26aa signal peptide and a single mature peptide flanked C-terminally by a cleavage site (K₇₉R) (Fig. 2F). A predicted elevenin precursor was identified in our study at 129aa in length, with a 29aa signal peptide and a 17aa mature peptide VDCRKFVFAPVCRGIIA (Fig. 2G).

Crustacean Hyperglycemic Hormone (CHHs) families. In total, three CHH isoforms and a CHH-like transcript were deduced from the eyestalk *C. quadricarinatus de novo* assembly (Fig. 3A). *Cq*-CHH-1, *Cq*-CHH-2 and *Cq*-CHH-3 all consist of a 29aa signal peptide, a CHH precursor-related peptide (CPRP) and a CHH mature peptide of 51aa, 57aa and 57aa, respectively. Moreover, one CHH-like peptide was also detected as containing a CPRP peptide and a CHH mature peptide that deviates in its composition from the other 3 CHHs. One *Cq*-MIH transcript and two *Cq*-MIH-like transcripts were identified (Fig. 3B); *Cq*-MIH-1 is identical to a previously described *Cq*-MIH (Accession ACX55057). The two *Cq*-MIH-like transcripts are 78 and 107aa in length, each with a signal peptide and cleavage sites predicted to release a 42aa and 41aa mature peptide, respectively. Among the CHH families, we identified one putative *Cq*-ITP that has a 31aa signal peptide, and a single mature peptide of 87aa. A phylogenetic tree of the *Cq*-CHH superfamily and CHH peptides from related species was constructed (Fig. 4). Based on this tree, *Cq*-CHH is allocated within the same cluster as lobster CHHs (including *Homarus gammarus*, *Nephrops norvegicus*, *Jasus lalandi*), while CHHs from other crayfish species, including *Procambarus* spp., are

Transcript annotation	Transcript size (bp)/ORF size (aa)	Lowest e-value	Closest blast hit	Species
Adipokinetic hormone (AKH)/corazonin-related peptide (ACP)	493/100	3.8e-38	(Veenstra, 2015)	<i>Procambarus clarkii</i>
Allatostatin A (AST-A)	1358/309	1.0e-121	BAE45266.1	<i>Procambarus clarkii</i>
Allatostatin B (AST-B)	1525/281	8.0e-63	ALQ28584.1	<i>Scylla paramamosain</i>
Allatostatin C1 (AST-C)/Prohormone 1	1544/104	1.3e-21	XP_014598613.1	<i>Polistes canadensis</i>
Allatostatin C2 (AST-C2)	286/36	1.0e-18	AIY69122.1	<i>Neocaridina denticulata</i>
Allatostatin CC (AST-CC)	454/136	8.9e-09	BAO00935.1	<i>Nilaparvata lugens</i>
BursiconA (partial)	461/53	2.1e-21	ADY80040.1	<i>Procambarus clarkii</i>
CCHamide 1	1167/215	6.0e-04	ALM30310.1	<i>Chilo suppressalis</i>
CCHamide 2	260/42	1.0e-05	ALM30310.1	<i>Chilo suppressalis</i>
Corazonin	1116/111	2.0e-17	ALA65535.1	<i>Macrobrachium rosenbergii</i>
DH 31	1020/134	2.1e-57	ACX46386.1	<i>Homarus americanus</i>
CRF-DH precursor/DH44	1371/84	7.0e-05	ALG35940.1	<i>Periplaneta americana</i>
Crustacean cardioactive peptide (CCAP)	1359/138	1.0e-64	BAF34909.1	<i>Procambarus clarkii</i>
Crustacean female sex hormone-like (CFSH)	854/264	3.2e-79	(Veenstra, 2015)	<i>Procambarus clarkii</i>
Crustacean hyperglycemic hormone-1 (CHH-1)	504/131	6.8e-44	BAA89003.1	<i>Procambarus clarkii</i>
Crustacean hyperglycemic hormone-2 (CHH-2)	552/135	1.4e-48	AJK31204.1	<i>Lipopenaeus vannamei</i>
Crustacean hyperglycemic hormone-3 (CHH-3)	1937/137	2.0e-63	ABA42180.1	<i>Homarus gammarus</i>
Crustacean hyperglycemic hormone-like	552/135	6.0e-23	AJK31204.1	<i>Lipopenaeus vannamei</i>
Ecdlosion hormone	1230/82	2.0e-24	ALQ28581.1	<i>Scylla paramamosain</i>
Elevenin	600/129	2.0e-03	BAO00952.1	<i>Nilaparvata lugens</i>
GSEFLamide	1507/335	2.1e-66	ALQ28590.1	<i>Scylla paramamosain</i>
FLRFamide	1248/374	4.4e-74	BAE06262.1	<i>Procambarus clarkii</i>
Glycoprotein A2 (GPA2)	861/120	2.1e-50	AKN21239.1	<i>Locusta migratoria</i>
Glycoprotein B5 (GPB5)	1321/168	1.6e-45	KDR08334.1	<i>Zootermopsis nevadensis</i>
HIGSLYRamide (partial)	724/241	1.3e-07	ALQ28601.1	<i>Scylla paramamosain</i>
Insulin-like peptide	743/214	5.0e-157	AIU40992.1	<i>Cherax quadricarinatus</i>
Kinin	735/173	5.0e-27	ALQ28594.1	<i>Scylla paramamosain</i>
Molt inhibiting hormone (MIH-1)	1331/107	5.0e-72	ACX55057.1	<i>Cherax quadricarinatus</i>
Molt inhibiting hormone-like-1 (MIH-like-1)	703/78	5.0e-41	AIZ05253.1	<i>Procambarus clarkii</i>
Molt inhibiting hormone-like-2 (MIH-like-2)	755/107	4.0e-34	P83636.2	<i>Orconectes limosus</i>
Ion transport protein (ITP)	533/117	9.0e-40	AIZ05253.1	<i>Procambarus clarkii</i>
Myosuppressin	981/99	1.0e-37	BAG68789.1	<i>Procambarus clarkii</i>
Prohormone 4	1101/219	3.0e-97	ALC46992.1	<i>Drosophila busckii</i>
Neuroparsin 1	770/105	1.8e-33	AHG98659.1	<i>Jasus lalandii</i>
Neuroparsin 2	871/101	2.1e-24	ALQ28570.1	<i>Scylla paramamosain</i>
Neuroparsin 3(partial)	368/89	1.03-19	AHX39208.1	<i>Metapenaeus ensis</i>
Neuropeptide F1	468/87	1.4e-17	AEC12204.1	<i>Litopenaeus vannamei</i>
Neuropeptide F2	470/104	6.0e-24	ALQ28586.1	<i>Scylla paramamosain</i>
Orcokinin	900/187	4.9e-70	ACD13197.1	<i>Homarus americanus</i>
Pyrokinin	1418/201	5.7e-32	ALQ28575.1	<i>Scylla paramamosain</i>
Pigment dispersing hormone 1 (PDH-1)	509/79	1.0e-27	JC4756	<i>Penaeus sp.</i>
Pigment dispersing hormone 2 (PDH-2)	393/79	3.0e-14	JC4756	<i>Penaeus sp.</i>
Pigment dispersing hormone 3 (PDH-3)	393/79	6.0e-24	JC4756	<i>Penaeus sp.</i>
Proctolin	697/96	3e-04	AEX08669.1	<i>Rhodnius prolixus</i>
Red pigment concentrating hormone (RPCH)	420/90	3.8e-43	AAV80404.1	<i>Cherax quadricarinatus</i>
Ryamide	462/131	1.0e-10	NP_001280530.1	<i>Tribolium castaneum</i>
Short Neuropeptide F	1617/127	5.3e-39	ALQ28574.1	<i>Scylla paramamosain</i>
SIFamide	619/75	3.8e-24	Q867W1.1	<i>Procambarus clarkii</i>
Sulfakinin	476/116	2.4e-49	ABQ95346.1	<i>Homarus americanus</i>
Tachykinin	1139/235	8.4e-100	BAC82426.1	<i>Procambarus clarkii</i>
Trissin	1052/200	2.0e-09	NP_650471.2	<i>Drosophila melanogaster</i>
Neurophysin - Vasopressin	577/154	1.8e-28	ALQ28600.1	<i>Scylla paramamosain</i>
WXXXRamide (partial)	238/79	2.0e-11	ALQ28592.1	<i>Scylla paramamosain</i>

Table 2. Putative neuropeptide precursors predicted in the eyestalk of *C. quadricarinatus*.

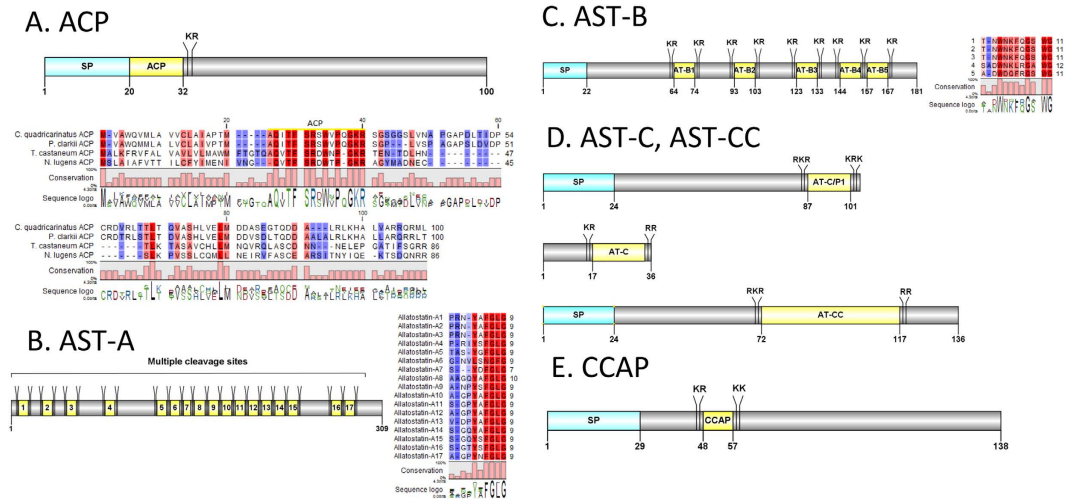


Figure 1. Molecular characterization of *Cherax quadricarinatus* ACP (A), allatostatin-A (B), allatostatin-B (C), allatostatin-C/allatostatin-CC (D) and CCAP (E). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

distributed into another clade. In contrast, *Cq*-MIH groups with crayfish MIHs and is quite distinct from lobster MIHs. The two *Cq*-CHH-like and *Cq*-ITP cluster together as a middle clade, suggesting they are a hybrid form between CHH and MIH.

FLRFamide, *GPA2/GBP5*, *GSEFLamide*, *HIGSLYamide*, *insulin*, *kinin* and *myosuppressin*. A predicted *Cq*-FLRFamide precursor consists of a 26aa signal peptide, and processing sites that could release up to 7 mature peptides ranging in size from 8–11aa (Fig. 5A). The FLRFamide motif in *C. quadricarinatus* is **X-Y/F-LRFamide**. We report a putative GPA2/GBP5 (Fig. 5B). The deduced GPA2 is 120aa in length, with an 18aa signal peptide, followed immediately by a 103aa mature peptide. The *Cq*-GBP5 is larger, at 168aa, and consists of a 21aa signal peptide and 147aa GPB5 mature peptide. A *Cq*-GSEFLamide transcript encodes for a 335aa precursor, containing a 24aa signal peptide and followed by 15 mature peptides flanked by multiple cleavage sites (Fig. 5C). The deduced mature peptides vary in length from 7–8aa and each have a conserved **XGSEFLamide** motif. A partial precursor sequence for HIGSLYamide was identified, consisting of 241aa and no signal peptide (Fig. 5D). Multiple dibasic KR/RR cleavage sites suggest the presence of 10 mature peptides, each 8aa in length. The conserved motif within each peptide is **HL/IA/GSLY/HKamide** (Fig. 5D). A complete identical Insulin-like peptide previously characterised in *C. quadricarinatus* was reconfirmed (Accession A1U40992). The neuropeptide is 214aa in length, contains a signal peptide and 2 conserved regions that would be the mature hormone (Fig. 5E). One partial kinin precursor was identified in our *de novo* assembly (lacking the signal peptide) (Fig. 5F). The precursor is 184aa in length and has 4 predicted mature peptides, separated by multiple dibasic cleavage sites. The conserved motif detected is **XFSAWAamide**. A myosuppressin precursor was deduced from the *de novo* assembly (Fig. 5G). The precursor consists of a 29aa signal peptide, a K₈₄R and R₉₇ cleavage site that resulted in an 11aa mature myosuppressin (QDLDHVFLRFamide). *Cq*-myosuppressin has the same motif found in other known crustacean species.

Neuroparsin, *Neuropeptide F*, *Orcokinin*, *pigment dispersing hormone (PDH)*, *pyrokinin*, *prohormone-4*, *proctolin*. Three *Cq*-NP transcripts were found (Fig. 6A), including two full-length sequences with signal peptides of 29aa and 26aa. The predicted mature peptides of 75aa and 80aa are encoded immediately after the signal peptide. A phylogenetic tree analysis of NP precursors was constructed, showing that *Cq*-NP2 clusters strongly with *P. clarkii*-NP3 (Pc-NP3), alongside *S. paramamosain*-NP1 (*Sp*-Np1) and *Sp*-NP4 (Fig. 7). Less well clustered, yet within the same branch, *Cq*-NP3 (partial-length) clusters with *Pc*-NP1. Together these form a clade distinct from *Cq*-NP1, which clusters strongly with *Pc*-NP2, and also cluster with *Sp*-NP2 and *Sp*-NP3. For this tree, the silk moth *Bombyx mori* NP served as an outgroup. Two sequences of *Cq*-NPF were identified in our *de novo* assembly with 87 and 104aa, including a signal peptide of 26 and 20aa, respectively (Fig. 6B). The predicted mature peptide starts immediately right after the signal peptide and both of them have the conserve motif **RPRFamide**. We identified one orckinin precursor 187aa. Orcokinin has a predicted signal peptide of 32aa and multiple cleavage sites that potentially releases up to 8 mature peptides. Of those, there exists a conserved C-terminus motif **XFDEIDRXGFGFXamide** (Fig. 6C), which is similar to other known crustacean species orckinins - NFDEIDRSFGFGFNamide¹⁰. Three PDH precursors were deduced from the eyestalk transcriptome, namely *Cq*-PDH-1, *Cq*-PDH-2 and *Cq*-PDH-3 (Fig. 6D). All have a 22aa signal peptide, followed by a dibasic K₅₇R and R/K₇₈R cleavage site that could release a mature peptide of 18aa. A prohormone-4 transcript was detected in our study, encoding a 219aa precursor that contains a 26aa signal peptide and K₄₀R cleavage site that could release

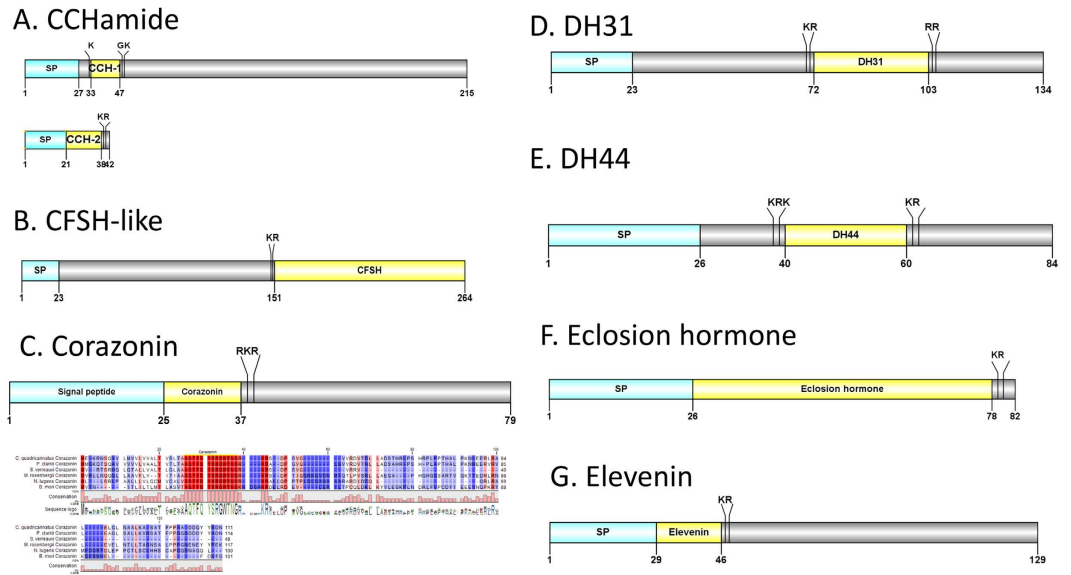


Figure 2. Molecular characterization of *Cherax quadricarinatus* CCHamide (A), CFSH-like (B), corazonin (C), DH31 (D), DH44 (E), eclosion hormone (F) and elevenin (G). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means not similar, dark red means exact amino acid).

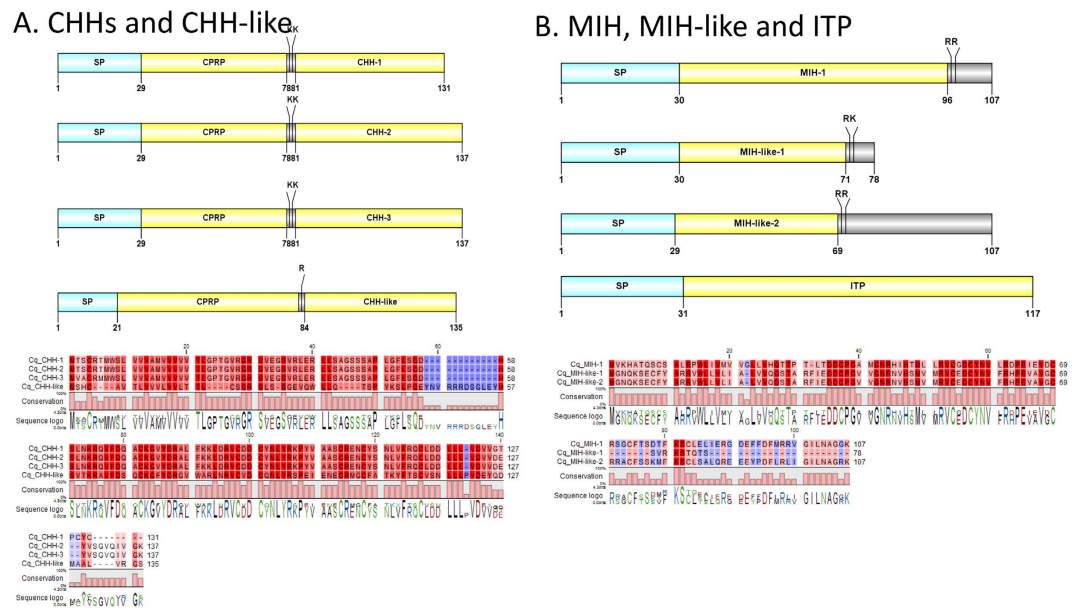


Figure 3. Molecular characterization of *Cherax quadricarinatus* CHHs and CHH-like (A), MIH, MIH-like and ITP (B). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

a 178aa mature peptide (Fig. 6E). A proctolin transcript was identified that encodes a 96aa precursor, encoding a 22aa signal peptide followed immediately by the *Cq*-Proctolin mature peptide (Fig. 6F). *Cq*-Proctolin has a RYLPT motif, similar to other crustacean species¹⁰.

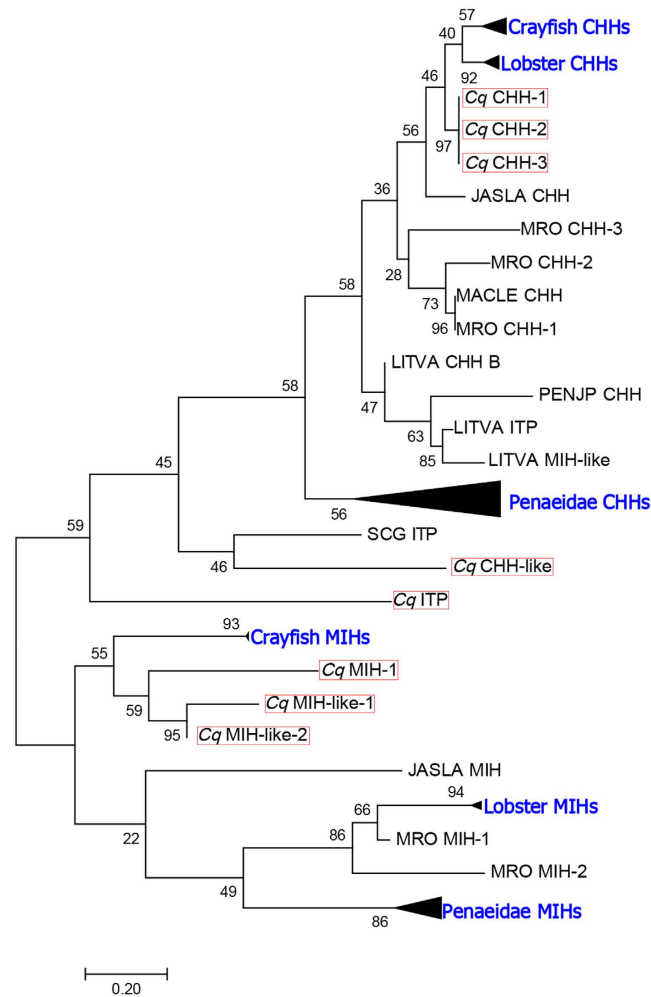
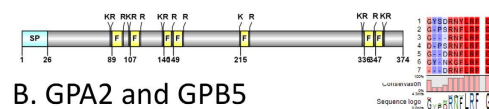


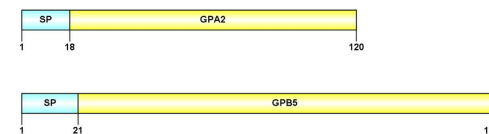
Figure 4. A maximum likelihood method based on the JTT matrix-based model was used to produce the phylogenetic tree using amino acids sequence of CHHs, MIHs and ITPs. 500 bootstrap replicates was used to produce the phylogenetic tree using amino acids sequence of CHHs/MIHs and ITP isoforms (that include signal peptide and the mature sequence). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap values (1–100) are given at each branch. *C. quadricarinatus* CHHs, MIH and ITP are highlighted in red boxes. JASLA: *Jasus lalandi*, LITVA: *Lipopanaeus vannamei*, PENJP: *Panaeus vannamei*, SCG: *Schistocerca gregaria*, MRO: *Macrobrachium rosenbergii*.

Pyrokinin, *RYamide*, *SIFamide*, *short neuropeptide F*, *sulfakinin (SK)*, *tachykinin*, *WXXXRamide*. A partial *Cq*-pyrokinin precursor that lacks a signal peptide, yet contains multiple dibasic cleavage sites that can release 11 highly conserved mature peptides – **ADFAFXPRLamide**, was deduced from the eyestalk transcriptome (Fig. 8A). The 5aa C-terminal region **FXPRLamide** is highly conserved among the pyrokinin/pheromone biosynthesis activating neuropeptide (PBAN) family⁵². A transcript encoding for a 131aa *Cq*-RYamide precursor was deduced from our data (Fig. 8B). The precursor contains a 21aa signal peptide, followed by two mature peptides, QGFYSQRYamide and FIGGSRYamide, which share a conserved **XXXXXXRYamide** motif. One transcript encoding for *Cq*-SIFamide was identified. The prohormone start has a signal peptide of 27aa, followed by the mature peptide with a C-terminal dibasic cleavage site K₄₁R (Fig. 8C). The mature peptide, GYRKPPFNGSIFamide, shares significant sequence similarity with other known SIFamides: GYRKPPFNGSIFamide¹⁰. One transcript was identified that encodes a 127aa sNPF precursor, containing a 25aa signal peptide, followed by three dibasic cleavage sites that could release three mature sNPF peptides (Fig. 8D). The mature peptides share the conserved motif **XPXRLRFamide**. The 116aa *Cq*-sulfakinin is composed of a 22aa signal peptide and two sulfakinin mature peptides GGDYDDYGHLLRRFamide and EFDEYGHMRFamide, which share a common C-terminal **DYGHXRamide** (Fig. 8F). A *Cq*-Tachykinin (TK) transcript was identified, encoding a precursor with a 22aa signal peptide, followed by 6 TK peptides (9aa), each flanked by cleavage sites (Fig. 8G). The predicted TK peptides all share the conserved **APSGFLGMRamide** motif, a feature consistent with other crustacean species TK peptides. A partial transcript of *Cq*-WXXXRamide was detected in our eyestalk *de novo* transcriptome, which encodes a precursor with no signal peptide and multiple cleavage sites that could release up to 6 mature peptides with the C-terminus **WXXXRamide** motif.

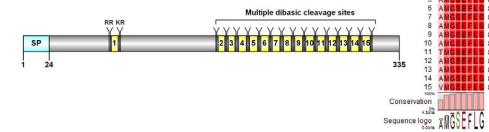
A. FMRFamide/FLRFamide



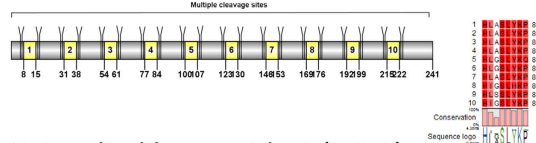
B. GPA2 and GPB5



C. GSEFLamide



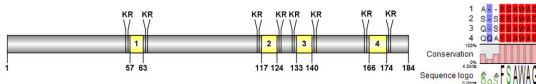
D. HIGSLYamide



E. Insulin-like peptide 1 (ILP-1)



F. Kinin

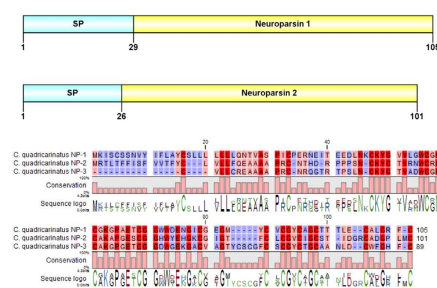


G. Myosuppressin

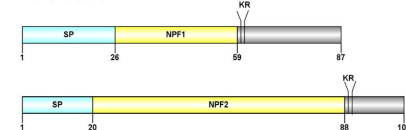


Figure 5. Molecular characterization of *Cherax quadricarinatus* FLRFamide (A), GPA2 and GPB5 (B), GSEFLamide (C), HIGSLYamide (D), Insulin-like peptide (E), Kinin (F), Myosuppressin (G). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

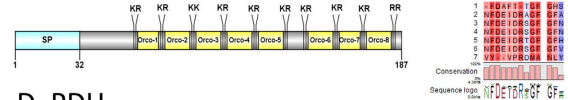
A. Neuroparsin



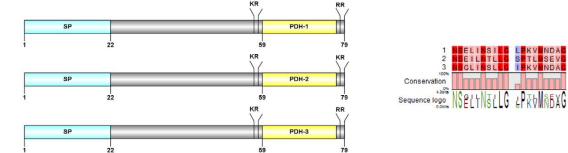
B. NPF



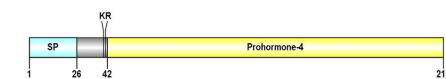
C. Orcokinin



D. PDH



E. Prohormone-4



F. Proctolin



Figure 6. Molecular characterization of *Cherax quadricarinatus* Neuroparsin (A), NPF (B), PDHs (C), FRLFamide (D), Prohormone-4 (E) and Proctolin (F). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

Red pigment concentrating hormone (RPCH), vasopressin-neurophysin. A single *Cq*-RPCH transcript was deduced from the eyestalk transcriptome (Fig. 9). The RPCH precursor consists of a 21aa signal peptide, followed immediately by a 9aa mature peptide **QLNFSPGWamide** and a dibasic cleavage site at K₃₁R. The predicted mature peptide is very similar to known RPCH in other crustacean species (Fig. 9). A vasopressin-neurophysin preprohormone containing a 20aa signal peptide, followed by a vasopressin peptide and a neurophysin mature peptide was also identified (Fig. 10). Multiple sequence alignment shows significant identity of *Cq*-vasopressin-neurophysin with other closely related insect and crustacean species.

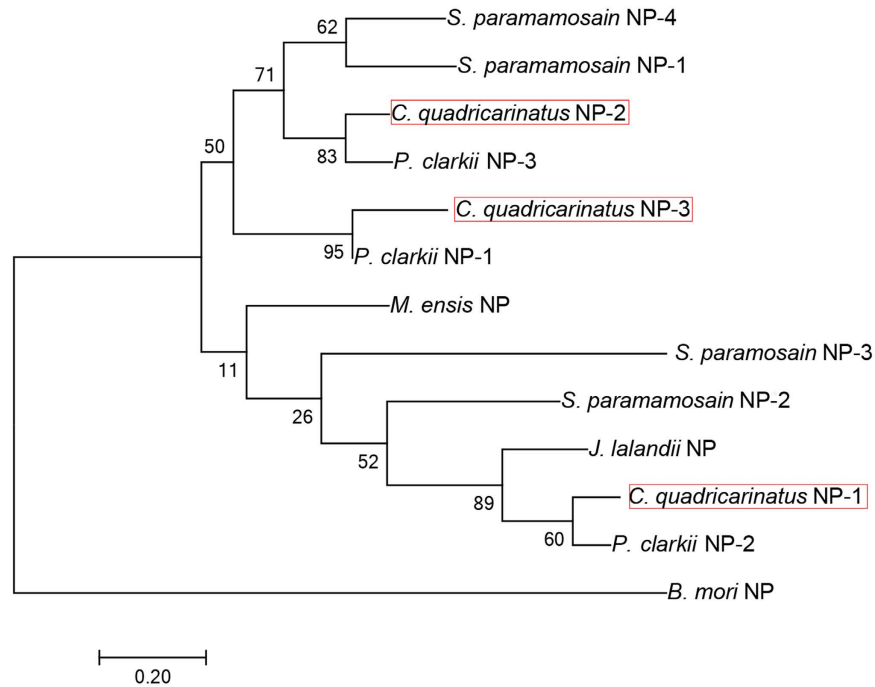


Figure 7. Molecular phylogenetic analysis of Neuroparsin isoforms by Maximum Likelihood method based on the JTT matrix-based model. 500 bootstrap replicates was used to produce the phylogenetic tree using amino acids sequence of prepro-neuroparsin (that include signal peptide and the mature sequence) The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap values (1–100) are given at each branch. *C. quadricarinatus* Neuroparsins are highlighted in red boxes.

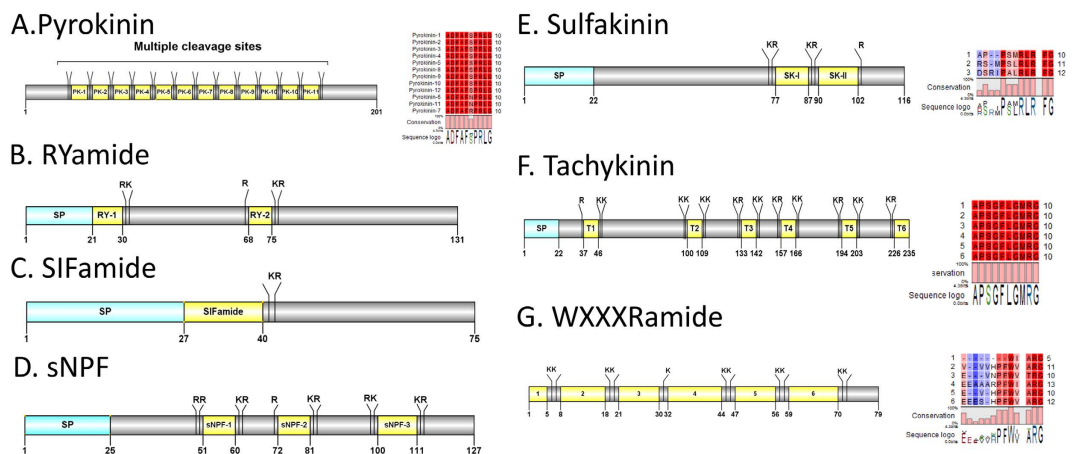


Figure 8. Molecular characterization of *Cherax quadricarinatus* Pyrokinin (A), RYamide (B), SIFamide (C), sNPF (D), Sulfakinin (E), Tachykinin (F) and WXXXRamide (G). Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

GnRH superfamily. In our study, we used the mature peptide sequence of *Cq*-ACP, *Cq*-RPCH, *Cq*-Crz, and other related species ACP, RPCH, Crz as well as GnRH to construct a phylogenetic tree for the GnRH superfamily (Fig. 11). The tree branched into 4 distinct groups ACP, RPCH, Crz, GnRH/GnRH-like (highlighted in red). The analysis indicated that RPCH, ACP and Crz mature peptides are more closely related to each other compared with GnRH, located within the same clade, while GnRHs are well clustered within their own clade.

Tissue distribution using RT-PCR. An RT-PCR experiment was designed to map the expression of 14 of the neuropeptide genes in 7 different *C. quadricarinatus* tissues, including supraesophageal ganglia (referred to as the brain) and thoracic ganglia (combined and referred to as the central nervous system; CNS), eyestalk,

RPCH

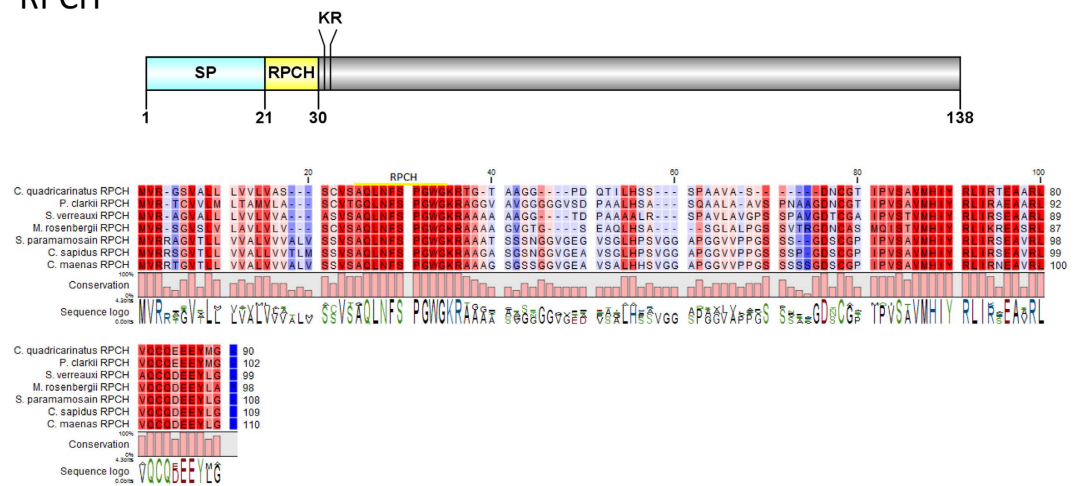


Figure 9. Molecular characterization of *Cherax quadricarinatus* RPCH. Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

Vasopressin-neurophysin

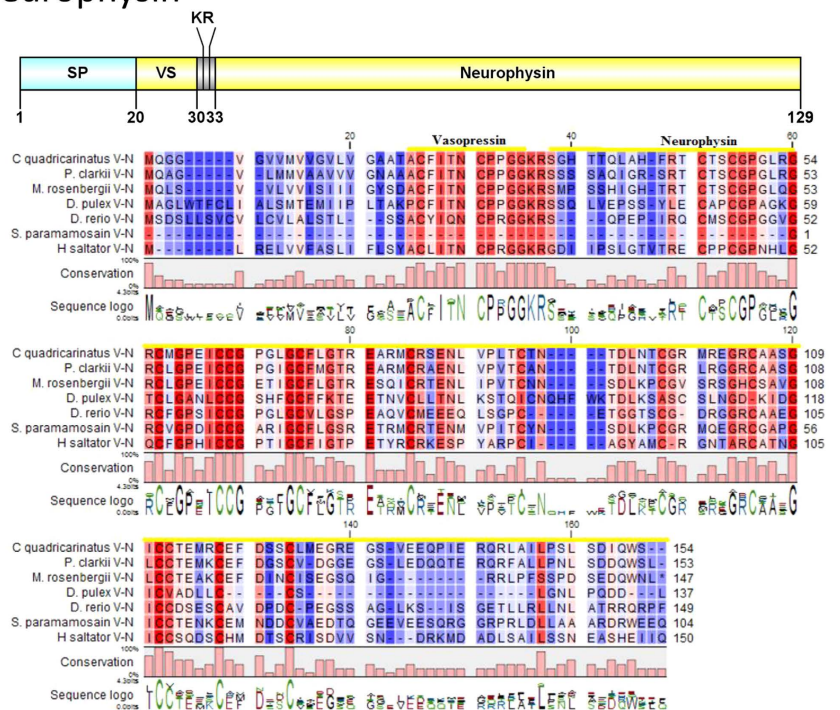


Figure 10. Molecular characterization of *Cherax quadricarinatus* Vasopressin-Neurophysin. Schematic diagrams show structure of neuropeptide precursors, including signal peptide (SP), the mature peptide (yellow) and putative cleavage sites. Precursor sequence alignments are shown with site of the mature peptide highlighted in yellow. Conserved amino acids are shown in a gradient from blue to red (blue means nearly similar, dark red means exact amino acid and white is no conservation).

heart, gut, antennal gland, ovary and muscle (Fig. 12). Most neuropeptide amplicons were detected in all the neural-type tissues (eyestalk, brain and thoracic ganglia), except for the *NP-1* isoform that was expressed in the CNS exclusively. Also, the *ACP* gene was expressed in CNS, and with some expression observed in the antennal gland. Seven of the neuropeptide genes tested were found to be expressed in the ovary, including *CCAP*,

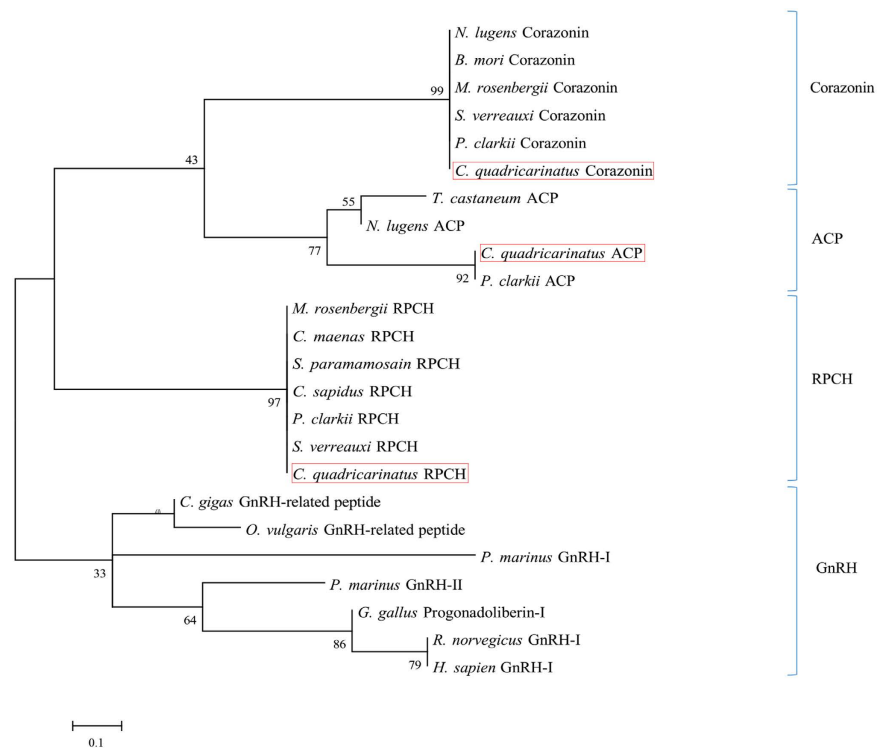


Figure 11. Molecular phylogenetic analysis of GnRH superfamily by Maximum Likelihood method based on the JTT matrix-based model. 500 bootstrap replicates were used to produce the phylogenetic tree using mature amino acids sequence of RPCH, Corazonin, ACP and GnRH. *C. quadricarinatus* neuropeptide are highlighted with red boxes.

myosuppressin, *NP-2*, *NP-3*, *pyrokinin*, *RPCH*, and *SIFamide*. As expected, beta-actin showed up in all tissues tested except negative control.

Discussion

This study presents an RNA sequencing approach to characterise a catalogue of neuropeptide genes in the eyestalk of the Australian red claw crayfish *C. quadricarinatus*. To identify the neuropeptides, two different *de novo* assemblers were employed to maximize transcript discovery as well as completeness. It is now well established that no single assembler is superior in assembly quality when compared with the others, as different assemblers provide different algorithms to address bubble effects, mismatches, and errors, to name a few, that each require different options and optimized parameters^{53–55}. We observed that by having two *de novo* assemblies instead of one, we had greater success in obtaining full-length transcripts. Even with 2 assemblies, for some neuropeptides (e.g. Bursicon, HIGSLYRamide, Kinin, Allatostatin-C), we still could only detect partial-length sequences, possibly due to the fact that crustacean eyestalks are rich in secondary metabolites including phenolic compounds⁵⁶, which can compromise RNA quality.

In the absence of a genome sequence, data mining using transcriptomics data is a powerful tool for neuropeptide discovery in crustacean species^{5–9}. Our *in silico* data mining of both eyestalk assemblies resulted in the identification of 53 neuropeptide-encoding transcripts, of which 14 were further assessed for tissue expression using RT-PCR. A number of neuropeptide genes detected in the current study are very similar to those identified in other *in silico* studies that have investigated related crustacean or insect species. Confirmation of neuropeptide genes in a diverse range of species is important since several arthropod neuropeptides have an unknown function, while others have a function that has only been partially confirmed or appear to have ambiguous roles^{27–30}. A similar recent *in silico* analysis approach in *M. rosenbergii* central nervous system yielded 21 neuropeptide transcripts (with some having multiple isoforms) that encode up to 102 mature peptides⁶. Another transcriptome-based neuropeptide study on *P. clarkii* mined 58 different neuropeptide transcripts, as well as their putative receptors⁷. Analysis of the central nervous system transcriptome in *S. verreauxi* revealed 37 transcripts representing 21 peptide/protein subfamilies⁹. In *H. americanus*, 35 precursors that released 194 distinct neuropeptides were detected from the neural tissues⁸. A very recent comprehensive collection of neuropeptides in a number of decapod species showcase a comprehensive list of characterised transcripts⁵⁷ with high similarity between species, concluding that the gap between insects and crustaceans is narrow as far as the neuropeptide repertoire is concerned. A number of neuropeptides described in the above study by Veenstra were not detected in our dataset including Calcitonins, CFSH, CCRFamide, DILP7-like, Periviscerokin and Trissin. In the current study, we mined a large number of neuropeptides, (comparison can be found in Supplementary Material 5), however it is important to note that in the absence of a sequenced genome, the number of neuropeptides mined, as well as sequencing completeness can

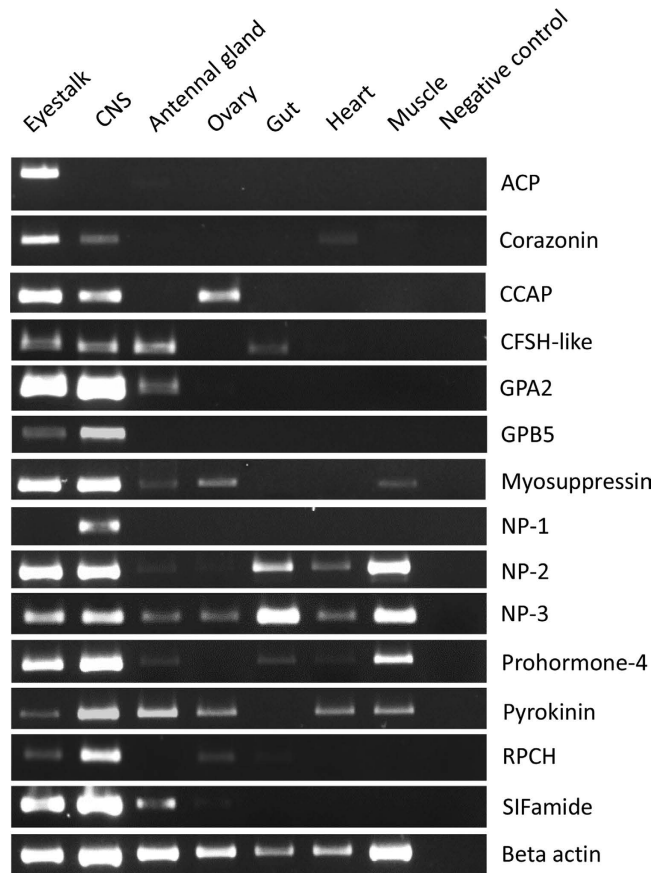


Figure 12. Tissue specific expression of *C. quadricarinatus* neuropeptide genes using RT-PCR. Expression of 14 neuropeptide genes using gene-specific primers, as well as Beta-actin gene. PCR used cDNA derived from tissues of a mature female. Negative control represents no cDNA in PCR. CNS: Central nervous system (Brain + Thoracic ganglia), ACP: Adipokinetic/Corazonin-related peptide, CCAP: Crustacean cardioactive peptide, CFSH-like: Crustacean female sex hormone-like, GPA2: Glycoprotein alpha-2, GPB5: Glycoprotein beta-5, NP: Neuropeptin, RPCH: Red pigment concentrating hormone, sNPF: Short neuropeptide F. All the gels were run under the same experimental conditions and are presented using cropped images. The entire gel photos of all neuropeptide genes can be find in Supplementary material 6.

be elusive, since it is dependent on the timing of expression and tissues chosen – which in our case, is solely based on the eyestalk ganglia, but not the entire CNS of the species.

In the current study, we highlighted a number of neuropeptides that have predicted roles in reproduction and sexual development. In our study, we were able to identify the mature peptide regions for *Cq*-AST-A, B, C, which opens up avenues for functional analysis in *C. quadricarinatus*. AST has a pleiotropic role in insects, for which one function is inhibition of JH synthesis by the corpora allata⁵¹. A precursor named methyl farnesoate (MF) is considered the crustacean equivalent of the insect JH. However, recent findings of the metabolizing enzyme which converts MF to JH – CYP15A1 – in *M. rosenbergii*⁵⁸ and *S. verreauxi*⁵⁹, specifically in the antennal gland (which is linked to the mandibular organ, the production site of MF), supports the notion that JH is an active hormone in crustaceans as well. Additionally, *Diploptera punctata* AST showed a stimulatory effect on MF synthesis in the mandibular organ of the adult crayfish *P. clarkii*⁶⁰. Silencing AST-A reduced egg and testes development in crickets, and the oviposition rate was drastically diminished in both species⁶¹. AST-C may be of particular interest since it has been proven previously to be a key factor in JH synthesis inhibition in different insects^{62,63}. Interestingly, a recent study has shown that local gene duplications in an ancestral arthropod created three allatostatin C genes, namely allatostatin-C, CC and CCC, with all three genes present in decapods⁶⁴. The pattern is similar to our results, where we can predict 3 allatostatin-C isoforms that are distinct from each other. Further research is warranted to investigate differences in functionalities of these genes.

From our results, one bursicon- α partial transcript was deduced from the eyestalk *de novo* assemblies, while we could not retrieve any bursicon- β sequence. This can be elucidated by tissue expression at various physiological stages, since in both *M. rosenbergii* and *Carcinus maenas*, bursicon- β could only be found in the thoracic ganglion but not eyestalk⁶⁵. Bursicon is a neuropeptide found throughout arthropods that was shown to regulate cuticle tanning, hardening and wing expansion of insects following ecdysis⁶⁶. In contrast, much less information is available about bursicon function in crustaceans. However, it is known that injection of recombinant

bursicon- β into female *P. monodon* broodstock causes an increase in vitellogenin gene expression and stimulated ovarian development²⁸.

Crustacean cardioactive peptide (CCAP) is a neuropeptide that is highly conserved between crustacean species, with all containing the PFCNAFTGCamide motif. It has been demonstrated to have a wide range of functions in insects as well as crustaceans including cardiac control⁶⁷, adaptation to environment stressors⁶⁸, stimulation of the oviduct⁶⁹, regulation of gut tissues⁷⁰ and ecdysis^{71,72}. In *S. paramamosain*, up-regulation of CCAP increased haemolymph circulation, and the authors suggested that CCAP facilitates movement of vitellogenin into the ovary during its reproduction phase⁵. In our study, we detected CCAP expression in the CNS, eyestalk and ovary of *C. quadricarinatus*, a similar fashion to that of *M. rosenbergii*⁶, further implicating CCAP role in reproduction.

Crustacean female sex hormone (CFSH) was recently confirmed to be involved in female sexual maturation in the blue crab *C. sapidus*¹⁷. In that study, it was demonstrated that CFSH was highly expressed in females but not males, and has a crucial role in developing the female phenotype¹⁷. In our *de novo* transcriptome, no match for CFSH was detected within the designated threshold. However, we were able to identify a transcript that is very similar to a CFSH-like neuropeptide previously described in *P. clarkii*⁷. That CFSH, whose structure is somewhat related to the mud crab female sex hormone, appears to be expressed exclusively in the eyestalks⁷. From RT-PCR results in our study, we have shown that *Cq*-CFSH-like is not eyestalk exclusive, but is also expressed in the CNS, antennal gland and gut, but not ovary, from one mature individual. Still, the specific role of the CFSH-like peptide in *C. quadricarinatus* and whether it contributes toward reproduction requires further investigation.

Crustacean hyperglycemic hormones (CHHs) are an important superfamily of peptides in crustaceans, produced in the X-organ, then transported to the sinus gland, where they are stored and subsequently secreted, together forming the eyestalk XO-SG complex. CHHs have been extensively studied, revealing diverse roles in carbohydrate metabolism, osmoregulation, reproduction and molting [for reviews, see Chan, *et al.*¹³, Christie, *et al.*^{10,12,25}]. In crustaceans, multiple CHH isoforms have been detected in neuropeptidomes. For instance, studies have found 2 isoforms in *S. paramamosain*⁵, 3 isoforms in *M. rosenbergii*⁶ and *P. clarkii*⁷, and 4 isoforms in *H. americanus*⁸ and *S. verreauxi*⁹. Three CHH isoforms were deduced in this study, each with six conserved cysteine residues, which give rise to three intramolecular disulfide bridges¹². Meanwhile, the CHH family member representative *Cq*-MIH is identical to the MIH previously described also in *C. quadricarinatus*⁷³.

Myosuppressin belongs to a group of neuropeptides that are only found in insects or crustaceans. It has previously been demonstrated to exhibit several biological activities, including inhibition of gut contractions⁷⁴, anti-feeding activity⁷⁵, and inhibition of neuropeptide secretion⁷⁶ and has been suggested previously to be a brain/gut peptide⁷⁵. In *C. quadricarinatus*, we detected expression of myosuppressin in the CNS, eyestalk, antennal gland, muscle and most surprisingly, the ovary in a mature female, showing for the first time a potential link between myosuppressin expression and reproduction.

Another group of neuropeptides which are associated with reproduction are neuroparsins (NPs), which have diverse functions in insects, including delayed vitellogenesis⁷⁷. In crustaceans, a number of NP isoforms have been detected in *M. rosenbergii*⁶, *P. clarkii*⁷, *S. paramamosain*⁵ and several decapod species⁵⁷. From the current study, three NP isoforms were deduced; *Cq*-NP-1 and *Cq*-NP-2 have 12 cysteine residues (possibly forming 6 disulfide bridges), while *Cq*-NP-3 is a partial sequence with 16 cysteines. Tissue screening using RT-PCR showed that *Cq*-NP-2 and *Cq*-NP-3 transcripts are widely distributed in most of the tissues we tested, while surprisingly, *Cq*-NP-1 expression is exclusive to the CNS. We propose that this could be a stage-specific Neuroparsin, however further qPCR will be needed to elucidate the expression pattern of this specific neuropeptide. *In vivo* gene silencing of one NP isoform in the sand shrimp *Metapenaeus ensis* caused a significant decrease in hepatopancreas and ovary vitellogenin transcript levels²⁷, supporting a reproduction related role while in *S. paramamosain*, gene expression profiling on Sp-NP1, Sp-NP2 and Sp-NP3 suggests that these neuropeptide have a stimulating effect on early stage vitellogenesis and Sp-NP4 in late stage vitellogenesis⁵.

A single *Cq*-SIFamide transcript was identified in the *C. quadricarinatus* eyestalk, CNS, antennal gland and ovary. Most recently, upregulation of SIFamide was detected during the early vitellogenic stages in the mud crab *S. paramamosain*⁵, while in the prawn *M. rosenbergii*, it has been suggested that SIFamide modulates aggression, in association with adult courtship behaviour⁷⁸. SIFamide is a well conserved neuropeptide in insects and crustaceans⁷⁹. In insects, SIFamide has been proven to have a role in sexual behaviour⁸⁰ and promote sleep in *Drosophila melanogaster*⁸¹.

The PBAN/pyrokinin family is a major group of insect and crustacean neuropeptides that have been implicated in multiple functions during development, mating and reproduction^{52,82,83}. Pyrokinin has now been found in several different species across insects to crustaceans. Our RT-PCR indicated that *Cq*-pyrokinin is present in most of the tissues screened (except gut), suggesting an important regulatory role of pyrokinin in a wide array of biological processes in *C. quadricarinatus*. An effort to silence pyrokinin in *Solenopsis invicta* and *Helicoverpa zea* showed negative impacts post-silencing, including increased mortality, delayed development and reduced levels of sex pheromone production⁸⁴.

A GPA2/GPB5 glycoprotein gene was detected in our *C. quadricarinatus* transcriptome. This ancient glycoprotein was proposed to play a role in maintaining ion balance in the midgut of adult mosquitoes *Aedes aegypti*⁸⁵, development and hydromineral balance in *D. meganogaster*, and reproduction in *B. belcheri*⁸⁶. We detected expression of GPA2/GPB5 in both neural tissues tested, and only GPA2 expression in the ovary. It has been suggested previously that GPB5 binds with another element to form a glycoprotein in the ovary of *P. clarkii*⁷. We propose that there may be a stage-specific expression pattern, where GPA2 and GPB5 are expressed differently between maturation stages and later co-express to form a glycoprotein unit. Therefore, a follow-up qPCR study investigating expression levels of GPA2/GPB5 during different maturation stages might be valuable to assess the potential of this glycoprotein relevant to reproduction processes.

ACP, corazonin, RPCH and the ambiguous scheme for GnRH in crustaceans. ACP is a neuropeptide believed to be the hybrid form of Adipokinetic hormone (AKH) and corazonin⁸⁷. ACP has been demonstrated to act as a neurohormone in the central nervous system of *Rhodnius prolixus*⁸⁸, however we only detected expression of Cq-ACP in the eyestalk of *C. quadricarinatus*. Temporal expression profiling shows that both Rhopr-ACP and Rhopr-ACPR are upregulated after ecdysis, and the author suggested that this neuropeptide may be involved in processes associated with post-ecdysis⁸⁹. Still, little is known about the function of this neuropeptide.

Crz is a well conserved neuropeptide that was originally determined to be a cardioactive neuropeptide (and later named after its function)⁹⁰. The role of Crz, however, is not restricted to cardioactivity. We found Cq-Crz expression in eyestalk, CNS, as well as heart, thus further supporting the classic role of corazonin as a cardioactive neuropeptide. Crz was recently hypothesized to be involved in regulation of stress responses^{91,92}, while Crz administration in *M. rosenbergii* has demonstrated inhibition of spermatogenesis⁹³. Most recently, Crz was suggested to be the decapod equivalent to GnRH, since the GnRH receptor identified from the ovary of the oriental river prawn *M. nipponense* is a corazonin-like receptor⁹⁴. Still, much work needs to be done to confirm the bioactivity of this neuropeptide in reproduction processes.

In *C. quadricarinatus*, expression of RPCH in the CNS, eyestalk, ovary and gut (relatively low) shows a slightly different pattern from that observed in female mud crab *S. olivacea*, where RPCH was found in the eyestalk, brain and ventral nerve cord, but not ovary⁹⁵, however this might reflect the stage from which the tissues were assessed. Global alignment shows high identity throughout the precursor and strict conservation of the mature peptide. RPCH was shown previously to be an ortholog of AKH in insects and APGWamide in mollusks⁹⁶, where it has recently been shown to promote conditioning and spawning in oysters⁹⁷. Early studies of crustacean RPCH suggested a function in pigmentation⁹⁸, however later studies have proven that RPCH can regulate gonadal maturation in *P. clarkii* as determined by both *in vitro* and *in vivo* experiments^{99,100}. In these studies, the authors implied that RPCH acts as a neurotransmitter that triggers the secretion of gonad-stimulating hormone, and can therefore induce ovarian maturation¹⁰⁰. Recently, RPCH was reconfirmed to be involved in ovarian maturation in the mud crab *S. paramamosain*, possibly through a stimulatory effect on the nervous tissues³⁰.

Despite numerous reports regarding the absence of GnRH in crustaceans^{101,102}, there is a trend in discovery of either GnRH-like molecules^{14–16} or GnRH-like receptors⁹⁴. In *C. quadricarinatus* we identified putative transcripts for ACP, corazonin and RPCH, but could not detect any GnRH or GnRH-like transcript in the *de novo* eyestalk transcriptome, perhaps because either (1) the eyestalk does not express these genes, (2) low expression level, or (3) the physiological stage at which the eyestalk RNA was isolated. This result coincides with a recent *in silico* mining investigation that included several decapods, where no trace of a GnRH/GnRH-like transcript could be found⁵⁷. Interestingly, there is an evolutionary connection between GnRH, ACP, RPCH and corazonin at both the receptor and ligand level [(see a canonical review by Roch, *et al.*¹⁰³], supporting the proposition that these factors may share similar functionality.

Other neuropeptides of interest. In addition to the above neuropeptides that were found in the eyestalk and may have reproduction related functions, we have also identified the following neuropeptides, which can possibly serve other physiological aspects. These include diuretic hormones (DHs), a family of neuropeptides that contributes to the maintenance of water homeostasis in insects. In the current study, two distinct DHs (DH31 and DH44) were detected; DH31 was previously suggested to maintain ionic homeostasis of the hemolymph during ovarian development⁵, while the role of DH44 is still vague. Neuropeptide F (NPF) is an invertebrate NPY-like peptide, sharing high sequence similarity and physiological function with NPY¹⁰⁴. NPF/NPY maintain a wide range of physiological processes including feeding behaviour, growth, metabolic homeostasis, stress response and reproduction in insects and mammals (reviewed by Nassel and Wegener¹⁰⁴). For instance, studies have confirmed the role of NPF in regulation of male reproductive processes in the desert locust *Schistocerca gregaria*¹⁰⁵. It has been shown that ecdysis, an important characteristic of arthropods, is controlled by eclosion hormone (EH)¹⁰⁶. In insects, the role of EH is somewhat clear, while in crustaceans, despite numerous reports showing EH transcripts in the eyestalk, information on the role of this neuropeptide is still unclear^{10,107}. Prohormone-4 has previously been characterised in the Eastern spiny lobster (*S. verreauxi*)⁹, *H. americanus* and *C. gigas*¹⁰⁸. The pre-prohormone was first identified in the honeybee¹⁰⁹. Still, whether this neuropeptide in crustaceans plays any role in reproduction requires further investigation. In our *C. quadricarinatus* gene expression experiment, prohormone-4 was found in every tissue tested except the ovary. A complete vasopressin-neurophysin precursor is detected from our *de novo* assembly, however the function of this neuropeptide in crustaceans is currently unclear. This precursor was also detected in *M. rosenbergii*⁶. Another neuropeptide detected in the current *de novo* assembly is WXXXRamide, which is a crustacean orthologue of natalisin, a neuropeptide that was proved to be involved in sexual courtship and fecundity in insects¹¹⁰, but its role in crustaceans has not yet been resolved.

Limitations and insights from the current study. The current dataset is still incomplete in characterising the entire *C. quadricarinatus* neuropeptidome, as it is based solely on the eyestalk ganglia. Nevertheless, when comparing the current study with other recent comprehensive neuropeptidome studies in crustaceans⁵⁷, a near-complete representation of the predicted neuropeptides are found, providing a solid foundation for future studies into the functionality of these neuropeptides in decapods.

The spatial expression investigation demonstrated that our predicted peptides can be detected in the CNS as well as the eyestalk ganglia, thus supporting their roles as neuropeptides. Future gene expression analysis can be designed to examine the spatial and temporal expression of these neuropeptides throughout development and in response to processes such as molting and reproductive maturation.

Conclusion

The current study has provided new information concerning putative neuropeptide genes in *C. quadricarinatus* by mining of eyestalk transcripts, thus building upon accumulating understanding of crustacean neuropeptides involved in key physiological processes. *In silico* data mining combined with RT-PCR experiment resulted in a catalogue of neuropeptides, with an emphasis on those that have been shown to play a role in reproduction in other crustaceans, and likely play a similar role in *C. quadricarinatus*. Results from this study will be useful for both *in vitro* gene-based studies (e.g. design of an effective RNAi experiment to test neuropeptide function) and in applied *in vivo* physiological studies towards control of reproductive processes. This work also lays the foundation necessary for future studies with the aim to manipulating components of the neuropeptidome, which could significantly add to our understanding of reproduction in *C. quadricarinatus* and lead to the development of novel approaches to enhance aquaculture of economically important related species.

References

1. Van Herp, F. Inhibiting and stimulating neuropeptides controlling reproduction in Crustacea. *Invertebrate Reproduction & Development* **22**, 21–30, doi: 10.1080/07924259.1992.9672253 (1992).
2. Crown, A., Clifton, D. K. & Steiner, R. A. Neuropeptide signaling in the integration of metabolism and reproduction. *Neuroendocrinology* **86**, 175–182, doi: 10.1159/000109095 (2007).
3. Gäde, G. & Hoffmann, K.-H. Neuropeptides regulating development and reproduction in insects. *Physiological Entomology* **30**, 103–121, doi: 10.1111/j.1365-3032.2005.00442.x (2005).
4. De Haes, W. *et al.* Functional neuropeptidomics in invertebrates. *Biochim Biophys Acta* **1854**, 812–826, doi: 10.1016/j.bbapap.2014.12.011 (2015).
5. Bao, C., Yang, Y., Huang, H. & Ye, H. Neuropeptides in the cerebral ganglia of the mud crab, *Scylla paramamosain*: transcriptomic analysis and expression profiles during vitellogenesis. *Sci Rep* **5**, 17055, doi: 10.1038/srep17055 (2015).
6. Suwansa-ard, S. *et al.* *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 (2015).
7. Veenstra, J. A. 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 (2015).
8. Christie, A. E. *et al.* 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 (2015).
9. Ventura, T., Cummins, S. F., Fitzgibbon, Q., Battaglione, S. & Elizur, A. 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 (2014).
10. Christie, A. E., Stemmler, E. A. & Dickinson, P. S. Crustacean neuropeptides. *Cellular and molecular life sciences: CMLS* **67**, 4135–4169, doi: 10.1007/s00018-010-0482-8 (2010).
11. Nagaraju, G. P. Reproductive regulators in decapod crustaceans: an overview. *J Exp Biol* **214**, 3–16, doi: 10.1242/jeb.047183 (2011).
12. Hopkins, P. M. The eyes have it: A brief history of crustacean neuroendocrinology. *General and comparative endocrinology* **175**, 357–366, doi: 10.1016/j.ygcen.2011.12.002 (2012).
13. Chan, S. M., Gu, P. L., Chu, K. H. & Tobe, S. S. Crustacean neuropeptide genes of the CHH/MIH/GIH family: implications from molecular studies. *General and comparative endocrinology* **134**, 214–219, doi: 10.1016/S0016-6480(03)00263-6 (2003).
14. Ngernsoungnarn, A. *et al.* The identification and distribution of gonadotropin-releasing hormone-like peptides in the central nervous system and ovary of the giant freshwater prawn. *Macrobrachium rosenbergii*. *Invertebrate neuroscience: IN* **8**, 49–57, doi: 10.1007/s10158-008-0067-5 (2008).
15. Tinikul, Y. *et al.* Existence and distribution of gonadotropin-releasing hormone-like peptides in the central nervous system and ovary of the Pacific white shrimp, *Litopenaeus vannamei*. *Cell Tissue Res* **343**, 579–593, doi: 10.1007/s00441-010-1112-3 (2011).
16. Guan, Z.-B., Shui, Y., Liao, X.-R., Xu, Z.-H. & Zhou, X. Primary structure of a novel gonadotropin-releasing hormone (GnRH) in the ovary of red swamp crayfish *Procambarus clarkii*. *Aquaculture* **418–419**, 67–71, doi: 10.1016/j.aquaculture.2013.10.010 (2014).
17. Zmora, N. & Chung, J. S. 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 (2014).
18. Ventura, T., Fitzgibbon, Q., Battaglione, S., Sagi, A. & Elizur, A. Identification and characterization of androgenic gland specific insulin-like peptide-encoding transcripts in two spiny lobster species: *Sagmariasus verreauxi* and *Jasus edwardsii*. *General and comparative endocrinology* **214**, 126–133, doi: 10.1016/j.ygcen.2014.06.027 (2015).
19. Ventura, T. *et al.* Expression of an Androgenic Gland-Specific Insulin-Like Peptide during the Course of Prawn Sexual and Morphotypic Differentiation. *ISRN Endocrinol* **2011**, 476283, doi: 10.5402/2011/476283 (2011).
20. Ventura, T., Rosen, O. & Sagi, A. From the discovery of the crustacean androgenic gland to the insulin-like hormone in six decades. *General and comparative endocrinology* **173**, 381–388, doi: 10.1016/j.ygcen.2011.05.018 (2011).
21. Rosen, O. *et al.* A sexual shift induced by silencing of a single insulin-like gene in crayfish: ovarian upregulation and testicular degeneration. *PLoS One* **5**, e15281, doi: 10.1371/journal.pone.0015281 (2010).
22. Manor, R. *et al.* Insulin and gender: an insulin-like gene expressed exclusively in the androgenic gland of the male crayfish. *General and comparative endocrinology* **150**, 326–336, doi: 10.1016/j.ygcen.2006.09.006 (2007).
23. Sharabi, O. *et al.* Identification and Characterization of an Insulin-Like Receptor Involved in Crustacean Reproduction. *Endocrinology* **157**, 928–941, doi: 10.1210/en.2015-1391 (2016).
24. Aizen, J. *et al.* 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 (2016).
25. Webster, S. G., Keller, R. & Dirksen, H. 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 (2012).
26. Subramoniam, T. Mechanisms and control of vitellogenesis in crustaceans. *Fisheries Science* **77**, 1–21, doi: 10.1007/s12562-010-0301-z (2010).
27. Yang, S. P., He, J. G., Sun, C. B. & Chan, S. F. Characterization of the shrimp neuroparsin (MeNPLP): RNAi silencing resulted in inhibition of vitellogenesis. *FEBS open bio* **4**, 976–986, doi: 10.1016/j.fob.2014.09.005 (2014).
28. Sathapondecha, P., Panyim, S. & Udomkit, A. A novel function of bursicon in stimulation of vitellogenin expression in black tiger shrimp *Penaeus monodon*. *Aquaculture* **446**, 80–87 (2015).
29. Siangcham, T. *et al.* The effects of serotonin, dopamine, gonadotropin-releasing hormones, and corazonin, on the androgenic gland of the giant freshwater prawn, *Macrobrachium rosenbergii*. *General and comparative endocrinology* **193**, 10–18, doi: 10.1016/j.ygcen.2013.06.028 (2013).
30. Zeng, H., Bao, C., Huang, H., Ye, H. & Li, S. The mechanism of regulation of ovarian maturation by red pigment concentrating hormone in the mud crab *Scylla paramamosain*. *Anim Reprod Sci* **164**, 152–161, doi: 10.1016/j.anireprosci.2015.11.025 (2016).

31. Koebach, J. & Jackson, K. A. V. Unravelling peptidomes by *in silico* mining. *Peptidomics* **2**, doi: 10.1515/ped-2015-0002, (2015).
32. Christie, A. E. & Chi, M. 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 (2015).
33. Ghanawi, J. & Saoud, I. P. Molting, reproductive biology, and hatchery management of redclaw crayfish *Cherax quadricarinatus* (von Martens 1868). *Aquaculture* **358–359**, 183–195, doi: 10.1016/j.aquaculture.2012.06.019 (2012).
34. FAO. *Cultured aquatic species information programme. Cherax quadricarinatus*, http://www.fao.org/fishery/culturedspecies/Cherax_quadricarinatus/en (2011).
35. Abdu, U. R. L., Yehezkel, G. & Sagi, A. Oocyte development and polypeptide dynamics during ovarian maturation in the red-claw crayfish *Cherax quadricarinatus*. *Invertebrate Reproduction & Development* **37**, 75–83, doi: 10.1080/07924259.2000.9652402 (2000).
36. Parnes, S., Khalaila, I., Hulata, G. & Sagi, A. Sex determination in crayfish: are intersex *Cherax quadricarinatus* (Decapoda, Parastacidae) genetically females? *Genetical research* **82**, 107–116 (2003).
37. Khalaila, I. & Sagi, A. Intersexuality and its Control by the Androgenic Gland in the Crayfish *Cherax quadricarinatus*. *Journal of Reproduction and Development* **43**, 69–70 (1997).
38. López, F. J. V.-L. S. Intersex females in the red claw crayfish, *Cherax quadricarinatus* (Decapoda: Parastacidae). *Revista de Biología Tropical/International Journal of Tropical Biology and Conservation* **55**, doi: 10.15517/rbt.v55i0.5802 (2007).
39. Sagi, A., Khalaila, I., Barki, A., Hulata, G. & Karplus, I. Intersex Red Claw Crayfish, *Cherax quadricarinatus* (von Martens): Functional Males with Pre-vitellogenic Ovaries. *The Biological Bulletin* **190**, 16–23, doi: 10.2307/1542672 (1996).
40. Ali, M. Y., Pavasovic, A., Amin, S., Mather, P. B. & Prentis, P. J. Comparative analysis of gill transcriptomes of two freshwater crayfish, *Cherax cainii* and *C. destructor*. *Mar Genomics* **22**, 11–13, doi: 10.1016/j.margen.2015.03.004 (2015).
41. Ali, M. Y., Pavasovic, A., Mather, P. B. & Prentis, P. J. Analysis, characterisation and expression of gill-expressed carbonic anhydrase genes in the freshwater crayfish *Cherax quadricarinatus*. *Gene* **564**, 176–187, doi: 10.1016/j.gene.2015.03.074 (2015).
42. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120, doi: 10.1093/bioinformatics/btu170 (2014).
43. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644–652, doi: 10.1038/nbt.1883 (2011).
44. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC bioinformatics* **10**, 421, doi: 10.1186/1471-2105-10-421 (2009).
45. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785–786, doi: <http://www.nature.com/nmeth/journal/v8/n10/abs/nmeth.1701.html#supplementary-information> (2011).
46. Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols* **2**, 953–971, doi: 10.1038/nprot.2007.131 (2007).
47. Southey, B. R., Amare, A., Zimmerman, T. A., Rodriguez-Zas, S. L. & Sweedler, J. V. NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res* **34**, W267–272, doi: 10.1093/nar/gkl161 (2006).
48. Liu, W. *et al.* IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics* **31**, 3359–3361, doi: 10.1093/bioinformatics/btv362 (2015).
49. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870–1874, doi: 10.1093/molbev/msw054 (2016).
50. Chandler, J. C. *et al.* 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 (2015).
51. Stay, B. & Tobe, S. S. The role of allatostatins in juvenile hormone synthesis in insects and crustaceans. *Annu Rev Entomol* **52**, 277–299, doi: 10.1146/annurev.ento.51.110104.151050 (2007).
52. Choi, M. Y. & Vander Meer, R. K. Molecular Structure and Diversity of PBAN/pyrokinin Family Peptides in Ants. *Front Endocrinol (Lausanne)* **3**, 32, doi: 10.3389/fendo.2012.00032 (2012).
53. Zhao, Q. Y. *et al.* Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC bioinformatics* **12** Suppl 14, S2, doi: 10.1186/1471-2105-12-S14-S2 (2011).
54. Yang, Y. & Smith, S. A. Optimizing *de novo* assembly of short-read RNA-seq data for phylogenomics. *BMC Genomics* **14**, 328, doi: 10.1186/1471-2164-14-328 (2013).
55. Nguyen, T. V., Jung, H., Nguyen, T. M., Hurwood, D. & Mather, P. Evaluation of potential candidate genes involved in salinity tolerance in striped catfish (*Pangasianodon hypophthalmus*) using an RNA-Seq approach. *Mar Genomics* **25**, 75–88, doi: 10.1016/j.margen.2015.11.010 (2016).
56. Martinez-Perez, F., Zinker, S., Aguilar, G., Valdes, J. & Arechiga, H. Circadian oscillations of RPCH gene expression in the eyestalk of the crayfish *Cherax quadricarinatus*. *Peptides* **26**, 2434–2444, doi: 10.1016/j.peptides.2005.05.018 (2005).
57. Veenstra, J. A. Similarities between decapod and insect neuropeptidomes. *PeerJ* **4**, e2043, doi: 10.7717/peerj.2043 (2016).
58. Ventura, T. *et al.* Post-embryonic transcriptomes of the prawn *Macrobrachium rosenbergii*: multigenic succession through metamorphosis. *PLoS One* **8**, e55322 (2013).
59. Ventura, T., Fitzgibbon, Q. P., Battaglene, S. C. & Elizur, A. Redefining metamorphosis in spiny lobsters: molecular analysis of the phyllosoma to puerulus transition in *Sagmariasus verreauxi*. *Scientific reports* **5** (2015).
60. Kwok, R., Rui Zhang, J. & Tobe, S. S. Regulation of methyl farnesoate production by mandibular organs in the crayfish, *Procambarus clarkii*: a possible role for allatostatins. *J Insect Physiol* **51**, 367–378, doi: 10.3201/eid1105.040755 (2005).
61. Meyering-Vos, M., Merz, S., Sertkol, M. & Hoffmann, K. H. Functional analysis of the allatostatin-A type gene in the cricket *Gryllus bimaculatus* and the armyworm *Spodoptera frugiperda*. *Insect Biochem Mol Biol* **36**, 492–504, doi: 10.1016/j.ibmb.2006.03.008 (2006).
62. Kaneko, Y. & Hiruma, K. Short neuropeptide F (sNPF) is a stage-specific suppressor for juvenile hormone biosynthesis by corpora allata, and a critical factor for the initiation of insect metamorphosis. *Dev Biol* **393**, 312–319, doi: 10.1016/j.ydbio.2014.07.014 (2014).
63. Nouzova, M., Rivera-Perez, C. & Noriega, F. G. Allatostatin-C reversibly blocks the transport of citrate out of the mitochondria and inhibits juvenile hormone synthesis in mosquitoes. *Insect Biochem Mol Biol* **57**, 20–26, doi: 10.1016/j.ibmb.2014.12.003 (2015).
64. Veenstra, J. A. Allatostatins C, double C and triple C, the result of a local gene triplication in an ancestral arthropod. *General and comparative endocrinology* **230–231**, 153–157, doi: 10.1016/j.ygcen.2016.04.013 (2016).
65. Wilcockson, D. C. & Webster, S. G. 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 (2008).
66. Webster, S. G., Wilcockson, D. C., Mrinalini & Sharp, J. H. Bursicon and neuropeptide cascades during the ecdysis program of the shore crab, *Carcinus maenas*. *General and comparative endocrinology* **182**, 54–64, doi: 10.1016/j.ygcen.2012.11.018 (2013).
67. Slama, K., Sakai, T. & Takeda, M. Effect of corazonin and crustacean cardioactive peptide on heartbeat in the adult American cockroach (*Periplaneta americana*). *Archives of insect biochemistry and physiology* **62**, 91–103, doi: 10.1002/arch.20131 (2006).
68. Chung, J. S. *et al.* Identification and developmental expression of mRNAs encoding crustacean cardioactive peptide (CCAP) in decapod crustaceans. *J Exp Biol* **209**, 3862–3872, doi: 10.1242/jeb.02425 (2006).

69. Donini, A., Agricola, H. & Lange, A. B. Crustacean cardioactive peptide is a modulator of oviduct contractions in *Locusta migratoria*. *J Insect Physiol* **47**, 277–285 (2001).
70. Sakai, T., Satake, H., Minakata, H. & Takeda, M. Characterization of crustacean cardioactive peptide as a novel insect midgut factor: isolation, localization, and stimulation of alpha-amylase activity and gut contraction. *Endocrinology* **145**, 5671–5678, doi: 10.1210/en.2004-0722 (2004).
71. Phlippen, M. K., Webster, S. G., Chung, J. S. & Dirksen, H. Ecdysis of decapod crustaceans is associated with a dramatic release of crustacean cardioactive peptide into the haemolymph. *J Exp Biol* **203**, 521–536 (2000).
72. Gammie, S. C. & Truman, J. W. Ecdysis hormone provides a link between ecdysis-triggering hormone and crustacean cardioactive peptide in the neuroendocrine cascade that controls ecdysis behavior. *J Exp Biol* **202**, 343–352 (1999).
73. Pamuru, R. R. *et al.* Stimulation of molt by RNA interference of the molt-inhibiting hormone in the crayfish *Cherax quadricarinatus*. *General and comparative endocrinology* **178**, 227–236, doi: 10.1016/j.ygcen.2012.05.007 (2012).
74. Marciniak, P., Kuczer, M. & Rosinski, G. New physiological activities of myosuppressin, sulfakinin and NVP-like peptide in *Zophobas atratus* beetle. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology* **181**, 721–730, doi: 10.1007/s00360-011-0563-5 (2011).
75. Vilaplana, L., Pascual, N., Perera, N., Leira, D. & Belles, X. Antifeeding properties of myosuppressin in a generalist phytophagous leafworm, *Spodoptera littoralis* (Boisduval). *Regul Pept* **148**, 68–75, doi: 10.1016/j.regpep.2008.02.002 (2008).
76. Yamanaka, N. *et al.* Neuropeptide receptor transcriptome reveals unidentified neuroendocrine pathways. *PLoS One* **3**, e3048, doi: 10.1371/journal.pone.0003048 (2008).
77. Girardie, J., Boureme, D., Couillaud, F., Tamarelle, M. & Girardie, A. 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 (1987).
78. Vazquez-Acevedo, N. *et al.* GYRKPFFNGSIFamide (Gly-SIFamide) modulates aggression in the freshwater prawn *Macrobrachium rosenbergii*. *Biol Bull* **217**, 313–326, doi: 10.1086/BBLv217n3p313 (2009).
79. Verleyen, P., Huybrechts, J. & Schoofs, L. SIFamide illustrates the rapid evolution in Arthropod neuropeptide research. *General and comparative endocrinology* **162**, 27–35, doi: 10.1016/j.ygcen.2008.10.020 (2009).
80. Terhzaz, S., Rosay, P., Goodwin, S. F. & Veenstra, J. A. The neuropeptide SIFamide modulates sexual behavior in *Drosophila*. *Biochem Biophys Res Commun* **352**, 305–310, doi: 10.1016/j.bbrc.2006.11.030 (2007).
81. Park, S., Sonn, J. Y., Oh, Y., Lim, C. & Choe, J. SIFamide and SIFamide receptor defines a novel neuropeptide signaling to promote sleep in *Drosophila*. *Mol Cells* **37**, 295–301, doi: 10.14348/molcells.2014.2371 (2014).
82. Torfs, P. *et al.* Pyrokinin neuropeptides in a crustacean. Isolation and identification in the white shrimp *Penaeus vannamei*. *Eur J Biochem* **268**, 149–154, doi: 10.1046/j.1432-1327.2001.01858.x (2001).
83. Jurenka, R. & Nusawardani, T. The pyrokinin/pheromone biosynthesis-activating neuropeptide (PBAN) family of peptides and their receptors in Insecta: evolutionary trace indicates potential receptor ligand-binding domains. *Insect Mol Biol* **20**, 323–334, doi: 10.1111/j.1365-2583.2010.01065.x (2011).
84. Choi, M. Y., Vander Meer, R. K., Coy, M. & Scharf, M. E. Phenotypic impacts of PBAN RNA interference in an ant, *Solenopsis invicta*, and a moth, *Helicoverpa zea*. *J Insect Physiol* **58**, 1159–1165, doi: 10.1016/j.jinsphys.2012.06.005 (2012).
85. Paluzzi, J. P., Vanderveken, M. & O'Donnell, M. J. The heterodimeric glycoprotein hormone, GPA2/GPB5, regulates ion transport across the hindgut of the adult mosquito, *Aedes aegypti*. *PLoS One* **9**, e86386, doi: 10.1371/journal.pone.0086386 (2014).
86. Tando, Y. & Kubokawa, K. Expression of the gene for ancestral glycoprotein hormone beta subunit in the nerve cord of amphioxus. *General and comparative endocrinology* **162**, 329–339, doi: 10.1016/j.ygcen.2009.04.015 (2009).
87. Hansen, K. K. *et al.* Discovery of a novel insect neuropeptide signaling system closely related to the insect adipokinetic hormone and corazonin hormonal systems. *The Journal of biological chemistry* **285**, 10736–10747, doi: 10.1074/jbc.M109.045369 (2010).
88. Patel, H., Orchard, I., Veenstra, J. A. & Lange, A. B. The distribution and physiological effects of three evolutionarily and sequence-related neuropeptides in *Rhodnius prolixus*: Adipokinetic hormone, corazonin and adipokinetic hormone/corazonin-related peptide. *General and comparative endocrinology* **195**, 1–8, doi: 10.1016/j.ygcen.2013.10.012 (2014).
89. Zandawala, M., Haddad, A. S., Hamoudi, Z. & Orchard, I. Identification and characterization of the adipokinetic hormone/corazonin-related peptide signaling system in *Rhodnius prolixus*. *The FEBS journal* **282**, 3603–3617, doi: 10.1111/febs.13366 (2015).
90. Veenstra, J. A. Isolation and structure of corazonin, a cardioactive peptide from the American cockroach. *FEBS Lett* **250**, 231–234, doi: 10.1016/0014-5793(89)80727-6 (1989).
91. Veenstra, J. A. Does corazonin signal nutritional stress in insects? *Insect Biochem Mol Biol* **39**, 755–762, doi: 10.1016/j.ibmb.2009.09.008 (2009).
92. Boerjan, B., Verleyen, P., Huybrechts, J., Schoofs, L. & De Loof, A. In search for a common denominator for the diverse functions of arthropod corazonin: a role in the physiology of stress? *General and comparative endocrinology* **166**, 222–233, doi: 10.1016/j.ygcen.2009.09.004 (2010).
93. Poljaroen, J., Tinikul, Y., Phoungpetchara, I. & Kankoun, W. The effects of biogenic amines, gonadotropin-releasing hormones and corazonin on spermatogenesis in sexually mature small giant freshwater prawns, *Macrobrachium rosenbergii* (De Man, 1879). *Aquaculture* **321**, 121–129 (2011).
94. Du, Y.-X., Ma, K.-Y. & Qiu, G.-F. 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 (2015).
95. Kornthong, N. *et al.* Characterization of red pigment concentrating hormone (RPCH) in the female mud crab (*Scylla olivacea*) and the effect of 5-HT on its expression. *General and comparative endocrinology* **185**, 28–36, doi: 10.1016/j.ygcen.2013.01.011 (2013).
96. Martinez-Perez, F., Becerra, A., Valdes, J., Zinker, S. & Arechiga, H. A possible molecular ancestor for mollusk APGWamide, insect adipokinetic hormone, and crustacean red pigment concentrating hormone. *Journal of molecular evolution* **54**, 703–714, doi: 10.1007/s00239-001-0036-7 (2002).
97. In, V. V. *et al.* Reproductive neuropeptides that stimulate spawning in the Sydney Rock Oyster (*Saccostrea glomerata*). *Peptides* **82**, 109–119, doi: 10.1016/j.peptides.2016.06.007 (2016).
98. Fernlund, P. & Josefsson, L. Crustacean color-change hormone: amino acid sequence and chemical synthesis. *Science (New York, N.Y.)* **177**, 173–175 (1972).
99. Fingerman, M. Roles of neurotransmitters in regulating reproductive hormone release and gonadal maturation in decapod crustaceans. *Invertebrate Reproduction & Development* **31**, 47–54, doi: 10.1080/07924259.1997.9672562 (1997).
100. Sarojini, R., Nagabhushanam, R. & Fingerman, M. A neurotransmitter role for red-pigment-concentrating hormone in ovarian maturation in the red swamp crayfish *Procambarus clarkii*. *J Exp Biol* **198**, 1253–1257 (1995).
101. Hauser, F. & Grimmlikhuijzen, C. J. Evolution of the AKH/corazonin/ACP/GnRH receptor superfamily and their ligands in the Protostomia. *General and comparative endocrinology* **209**, 35–49, doi: 10.1016/j.ygcen.2014.07.009 (2014).
102. Tsai, P. S. Gonadotropin-releasing hormone in invertebrates: structure, function, and evolution. *General and comparative endocrinology* **148**, 48–53, doi: 10.1016/j.ygcen.2005.09.016 (2006).
103. Roch, G. J., Busby, E. R. & Sherwood, N. M. Evolution of GnRH: diving deeper. *General and comparative endocrinology* **171**, 1–16, doi: 10.1016/j.ygcen.2010.12.014 (2011).
104. Nassel, D. R. & Wegener, C. A comparative review of short and long neuropeptide F signaling in invertebrates: Any similarities to vertebrate neuropeptide Y signaling? *Peptides* **32**, 1335–1355, doi: 10.1016/j.peptides.2011.03.013 (2011).

105. Van Wielendaele, P. *et al.* Neuropeptide F regulates male reproductive processes in the desert locust. *Schistocerca gregaria*. *Insect Biochem Mol Biol* **43**, 252–259, doi: 10.1016/j.ibmb.2012.12.004 (2013).
106. Horodyski, F. M. Neuroendocrine control of insect ecdysis by eclosion hormone. *Journal of Insect Physiology* **42**, 917–924, doi: 10.1016/0022-1910(96)00035-2 (1996).
107. Truman, J. W. Hormonal control of insect ecdysis: endocrine cascades for coordinating behavior with physiology. *Vitam Horm* **73**, 1–30, doi: 10.1016/S0083-6729(05)73001-6 (2005).
108. Zhang, G. *et al.* The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**, 49–54, doi: 10.1038/nature11413 (2012).
109. Hummon, A. B. *et al.* From the genome to the proteome: uncovering peptides in the Apis brain. *Science (New York, N.Y.)* **314**, 647–649, doi: 10.1126/science.1124128 (2006).
110. Jiang, H. *et al.* Natalisin, a tachykinin-like signaling system, regulates sexual activity and fecundity in insects. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E3526–3534, doi: 10.1073/pnas.1310676110 (2013).

Acknowledgements

The current work was supported through a USC International PhD scholarship to Tuan Viet Nguyen and a Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme (612296-DeNuGReC).

Author Contributions

Conceived and designed the experiments: TVN, SFC, AE, TV. Performed the experiments: TVN, TV. Analysed the data: TVN. Wrote the paper: TVN. All authors proofread and accepted the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Nguyen, T. V. *et al.* Transcriptomic characterization and curation of candidate neuropeptides regulating reproduction in the eyestalk ganglia of the Australian crayfish, *Cherax quadricarinatus*. *Sci. Rep.* **6**, 38658; doi: 10.1038/srep38658 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016