Analysis of the Central Nervous System Transcriptome of the Eastern Rock Lobster *Sagmariasus verreauxi* Reveals Its Putative Neuropeptidome

**Tomer Ventura**¹, Scott F. Cummins², Quinn Fitzgibbon², Stephen Battaglene², Abigail Elizur¹

¹ Faculty of Science, Health, Education and Engineering, GenECology Research Centre, University of the Sunshine Coast, Sunshine Coast, Queensland, Australia, ² Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

**Abstract**

Neuropeptides have been discovered in many arthropod species including crustaceans. The nature of their biological function is well studied and varies from behavior modulation to physiological regulation of complex biochemical processes such as metabolism, molt and reproduction. Due to their key role in these fundamental processes, neuropeptides are often targeted for modulating these processes to align with market demands in commercially important species. We generated a comprehensive transcriptome of the eyestalk and brain of one of the few commercially important spiny lobster species in the southern Hemisphere, the Eastern rock lobster *Sagmariasus verreauxi* and mined it for novel neuropeptide and protein hormone-encoding transcripts. We then characterized the predicted mature hormones to verify their validity based on conserved motifs and features known from previously reported hormones. Overall, 37 transcripts which are predicted to encode mature full-length/partial peptides/proteins were identified, representing 21 peptide/protein families/subfamilies. All transcripts had high similarity to hormones that were previously characterized in other decapod crustacean species or, where absent in crustaceans, in other arthropod species. These included, in addition to other proteins previously described in crustaceans, prohormone-3 and prohormone-4 which were previously identified only in insects. A homolog of the crustacean female sex hormone (CFSH), recently found to be female-specific in brachyuran crabs was found to have the same levels of expression in both male and female eyestalks, suggesting that the CFSH female specificity is not conserved throughout decapod crustaceans. Digital gene expression showed that 24 out of the 37 transcripts presented in this study have significant changes in expression between eyestalk and brain. In some cases a trend of difference between males and females could be seen. Taken together, this study provides a comprehensive neuropeptidome of a commercially important crustacean species with novel peptides and protein hormones identified for the first time in decapods.

**Introduction**

The Eastern rock lobster *Sagmariasus verreauxi* is one of a few closely related species which constitute the spiny lobster fishery industry in the Southern Hemisphere [1]. Identifying the molecular components which govern fundamental processes in this species might thus prove useful in further enhancing the aquaculture industry of this taxonomic group. Neuropeptides and protein hormones have long been suggested as targets for crustacean aquaculture enhancement [2,3]. They govern a wide array of physiological and behavioral processes and have been studied extensively in crustaceans [4]. Neuropeptides are translated from larger precursors (usually known as prepro-peptides) which include a signal peptide at their N-terminus. The signal peptide directs the prepro-peptide translation into the rough endoplasmic reticulum, where the signal peptide is being cleaved off, leaving the pro-peptide which is then further processed prior to the secretion of the mature peptide [4].

The list of putative neuropeptide sequences from different crustacean species has considerably increased over the past few years with the employment of bioinformatic mining in publicly available databases [5], *de novo* transcriptome assemblies [6–9] and mass spectrometry [10–13]. With the expansion of the crustacean neurohormone database, identification of the conserved features of the mature neurohormones further enables mining of novel neurohormones through *de novo* transcriptomes of crustacean species where neurohormones were not previously identified. Comparisons with other arthropod species where neuropeptidomes have been characterized [14–21] enable insights into species’ life history as in the case of the parasitic wasp *Nasonia vitripennis* [14] and the social honeybee *Apis mellifera* [15] and evolution, as in the case of the fruit fly *Drosophila sp.* [19] and the silk moth *Bombyx mori* [20].

With the recent rapid advancement in transcriptome sequencing capabilities, it becomes increasingly affordable to establish comprehensive transcriptomes of non-model organisms. We collected RNA from several key tissues that are known to be the primary sites of neuropeptide production and secretion in...
### Table 1. Alphabetical list of predicted peptide precursors with transcript and ORF size and best BLAST hit.

| Hormone        | Transcript | Transcript size | ORF size | Comments                                           | Best BLASTP result (Protein name [species] accession number) | E-Value  |
|----------------|------------|-----------------|----------|---------------------------------------------------|-------------------------------------------------------------|----------|
| Allatostatins  | >Unigene56418_All 955 | 248  | A-type prepro-allatostatin, partial (N terminus)   | allatostatin precursor protein [Panulirus interruptus] BA64528 | 1.00E-115 |
|                | >Unigene36127_All 462 | 154  | A-type prepro-allatostatin, partial (middle)       | allatostatin precursor protein [Panulirus interruptus] BA64528 | 1.00E-64  |
|                | >Unigene45628_All 1797 | 93   | A-type prepro-allatostatin, partial (C terminus)   | allatostatin precursor protein [Panulirus interruptus] BA64528 | 6.00E-45  |
|                | >Unigene40422_All 704 | 152  | B-type prepro-allatostatin, partial (N terminus)   | B-type preproallatostatin II [Pandalopsis japonica] AFV91539 | 4.00E-21  |
|                | >Unigene25318_All 1537 | 135  | B-type prepro-allatostatin, partial (C terminus)   | B-type preproallatostatin II [Pandalopsis japonica] AFV91539 | 6.00E-44  |
|                | >CL2090.Contig2_All 3784 | 141  | C-type prepro-allatostatin                         | C-type preproallatostatin [Pandalopsis japonica] AFV91540 | 1.00E-33  |
| Bursicon a subunit | >Unigene59348_All 1228 | 142  | prepro-Bursicon a 2                               | bursicon [Procambarus clarkii] ADY90040 | 3.00E-79  |
| Coicturedon     | >CL593.Contig3_All 210 | 49   | prepro-corazonin, partial                         | corazonin preprohormone [Daphnia pulex] ACJ35606 | 4.00E-06  |
| CCAP (crustacean cardioactive peptide) | >Unigene1674_All 1107 | 139  | prepro-CCAP                                        | crustacean cardioactive peptide [Homarus gammarus] AB46292 | 4.00E-62  |
| CHH (crustacean hyperglycemic hormone) | >CL7809.Contig1_All 1021 | 135  | prepro-CHH isoform B 1                            | prepro-crustacean hyperglycemic hormone isoform B [Nephrops norvegicus] AAQ22392 | 1.00E-60  |
|                | >CL7809.Contig3_All 1045 | 133  | prepro-CHH isoform B 2                            | hyperglycemic hormone B [Homarus gammarus] ABA42180 | 8.00E-57  |
|                | >CL7809.Contig4_All 1576 | 112  | prepro-CHH isoform B 3, partial (C terminus)       | crustacean hyperglycemic hormone-like peptide precursor [Procambarus clarkii] ADZ98836 | 1.00E-40  |
|                | >Unigene30324_All 1453 | 126  | prepro-CHH isoform B 4, partial (C terminus)       | prepro-crustacean hyperglycemic hormone-like peptide precursor [Nephrops norvegicus] AAQ22392 | 8.00E-60  |
| MIH/GIH (molt/gonad-inhibiting hormone) | >Unigene47171_All 679 | 115  | prepro-MIH/GIH isoform A 1                        | prepro-gonad-inhibiting hormone isoform A [Macrobrachium nipponense] AEJ54622 | 4.00E-27  |
|                | >Unigene60521_All 1232 | 114  | prepro-MIH/GIH isoform A 2                        | prepro-gonad-inhibiting hormone isoform A [Macrobrachium nipponense] AEJ54622 | 2.00E-26  |
|                | >Unigene58466_All 820 | 111  | prepro-MIH/GIH isoform A 3                        | vitellogenesis inhibiting hormone [Homarus gammarus] ABA42181 | 3.00E-45  |
| CFSH (crustacean female sex hormone) | >Unigene48118_All 1067 | 278  | prepro-CFSH                                        | crustacean female sex hormone, partial [Carcinus maenas] AEJ72264 | 2.00E-08  |
| DH (calcinotide-like diuretic hormone) | >CL8244.Contig2_All 1918 | 135  | prepro-DH class 2                                 | prepro-calcitond-like diuretic hormone [Homarus americanus] ACX46386 | 2.00E-69  |
| Eclosion hormone | >CL2590.Contig2_All 1584 | 82   | prepro-Eclosion hormone 1                         | eclosion hormone [Amphibalanus amphitrite] AKF81936 | 2.00E-14  |
|                | >Unigene55076_All 757 | 86   | prepro-Eclosion hormone 2                         | Eclosion hormone [Acromyrmex echinatior] EG68318 | 4.00E-13  |
| FLP (Myosuppressin) | >Unigene55051_All 819 | 100  | prepro-FLP                                       | prepro-myosuppressin [Homarus americanus] ACX46385 | 2.00E-40  |
| Follistatin     | >CL3958.Contig2_All 686 | 133  | Follistatin-like                                  | follistatin-like, partial [Nematostella vectensis] ABF61774 | 2.00E-15  |
|                | >Unigene49446_708 | 204  | Follistatin-like, partial (N terminus)             | hypothetical protein DAPPUDRAFT_303124 [Daphnia pulex] EFX97722 | 3.00E-41  |
| Hormone                  | Transcript          | Transcript size | ORF size | Comments                          | Best BLASTP result (Protein name [species] accession number)                          | E-Value     |
|--------------------------|---------------------|-----------------|----------|-----------------------------------|--------------------------------------------------------------------------------------|-------------|
| Myostatin                | >CL113.Contig2_All  | 1831            | 419      | Myostatin                        | MSTN [Penaeus monodon] ADO34177                                                     | 0           |
| NPY (neuropeptide Y)     | >Unigene30121_All  | 1287            | 104      | prepro-NPF                       | neuropeptide Y [Lymnaea stagnalis] CAB63265                                          | 3.00E-09    |
| Neuroparsin              | >Unigene5705_All   | 1217            | 103      | prepro-Neuroparsin               | neuroparsin 1 precursor [Schistocerca gregaria] CAC38869                             | 3.00E-12    |
|                          | >CL2744.Contig6_All| 1176            | 102      | prepro-Neuroparsin 2             | neuroparsin 1 precursor [Rhodnius prolixius] ACZ96369                               | 7.00E-11    |
| Neuroparsin              | >Unigene692_All    | 1343            | 205      | prepro-Orcokinin                 | prepro-orcokinin II [Homarus americanus] AC13197                                     | 2.00E-104   |
| Orcokinin                | >CL7594.Contig2_All| 430             | 79       | prepro-PDH                       | pigment dispersing hormone related peptide precursor 79 - penaeid shrimp [Penaeus sp.] JC4756 | 2.00E-29    |
| PDH (pigment dispersing hormone) | >CL7594.Contig3_All | 603             | 79       | prepro-PDH                       | pigment dispersing hormone related peptide precursor 79 - penaeid shrimp [Penaeus sp.] JC4756 | 1.00E-23    |
| Prohormone-3             | >CL1958.Contig1_All| 2238            | 196      | prohormone-3                     | prohormone-3 [Apis mellifera] XP_001122204                                          | 1.00E-44    |
| Prohormone-4             | >Unigene19311_All  | 807             | 143      | prohormone-4                     | prohormone-4-like [Acyrthosiphon pisum] XP_001951503                                | 3.00E-66    |
| RPCH (red pigment concentrating hormone) | >Unigene2547_All | 1158            | 99       | prepro-RPCH                      | red pigment concentrating hormone [Macrobrachium rosenbergii] A0/4675                 | 2.00E-26    |
| Sulfakinin               | >Unigene25008_All  | 902             | 115      | prepro-Sulfakinin                | preprosulfakinin [Homarus americanus] ABQ955346                                     | 7.00E-53    |
| Tachykinin               | >CL7656.Contig2_All| 2181            | 226      | prepro-Tachykinin                | preprotachykinin B [Panulirus interruptus] BAD06363                                  | 2.00E-143   |

doi:10.1371/journal.pone.0097323.t001
crustaceans and generated a comprehensive transcriptome of *S. verreauxi*. These tissues included the eyestalk, where the X-organ-sinus gland (XOSG) neuroendocrine complex resides, the thoracic ganglia and brain. From the transcriptomic data obtained, we compiled a list of the putative neuropeptides and protein hormones and characterized them via comparisons to previously reported neuropeptides to predict the processing of prepro-peptides into mature neuropeptides. The conserved motifs were identified and highlighted, providing a database that might prove useful for further identification of neuropeptides in closely related species.

**Results**

**Allatostatins**

Three transcripts were identified to putatively encode partial *type A allatostatin* precursors representing the N-terminus, middle region and C-terminus, with 248, 154 and 93 amino acids (aa), respectively (Table 1 and Fig. 1). The precursor N-terminus has a predicted signal peptide of 27 aa, followed by 10 predicted neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 1A), while in the C-terminus there are 5 predicted neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 2B). The 13 predicted neuropeptides are 9–14 aa in length with the conserved motif XXDWXXXXXXGXWamide (Fig. 2C). BLAST identified 7 of the above 13 neuropeptides in *type B allatostatin* of the caridean shrimp *Pandalus japonicus*, while the other 6 appear to be novel (Table 2). Both transcripts were found to have comparable expression levels with significantly higher expression of the N-terminus in the eyestalk, compared to the brain (Table 3).

One transcript was identified to putatively encode a complete *type C allatostatin* precursor with 141 aa, starting with a signal peptide of 22 aa, followed by 3 putative neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 3A). The 22 predicted neuropeptides are 14–15 aa in length with the conserved motif highly YFGLamide conserved (Fig. 3D). Using BLAST of the mature neuropeptides individually, they were shown to have either high similarity, or, for most, exact identity to other type A allatostatins, primarily from decapod crustacean species, apart from two who were most similar to insect species.

Most of the Eastern rock lobster putative type A allatostatin neuropeptides (17/22) had highest homology to type A allatostatin of the spiny lobster *Panulirus interruptus* (Table 2). All three type A allatostatin-encoding transcripts were found to have comparable expression levels in the brain, compared to the eyestalk (Table 3).

Two transcripts were identified to putatively encode partial *type B allatostatin* precursors representing the N-terminus and C-terminus, with 270 and 310 aa, respectively (Table 1 and Fig. 2). The N-terminus has a predicted signal peptide of 33 aa, followed by 8 predicted neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 2A), while in the C-terminus there are 5 predicted neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 2B). The 13 predicted neuropeptides are 9–14 aa in length with the conserved motif XXDWXXXXXXGXWamide (Fig. 2C). BLAST identified 7 of the above 13 neuropeptides in *type B allatostatin* of the caridean shrimp *Pandalus japonicus*, while the other 6 appear to be novel (Table 2). Both transcripts were found to have comparable expression levels with significantly higher expression in the brain, compared to the eyestalk (Table 3).

One transcript was identified to putatively encode a complete *type C allatostatin* precursor with 141 aa, starting with a signal peptide of 22 aa, followed by 3 putative neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 3A). The 22 predicted neuropeptides are 14–15 aa in length with no homology between them. The peptide at the precursor C-terminus has two cysteine residues characteristic of other allatostatins (Fig. 3A). Two of the three neuropeptides shared high identity with type C allatostatin peptides from closely related species.
| Hormone | Best BLAST hit | Accession number | Identity |
|---------|----------------|------------------|----------|
| Allatostatin A | HNNYAFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | TPDYAFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | EGMYSFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | GGMYSFGLa |
| | ADLFSFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | SQNYAFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | SKLYSFGLa FGLa-related allatostatin [Nilaparvata lugens] | BAO00953 | QKLYSFGLa |
| | NRQYSFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | SQFYAFGLa type-a prepro-allatostatin (Macrobrachium nipponense) | AEX86939 | 100% identity |
| | PKNYAFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | PTAYSFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | PTPTSFGLa |
| | TASYSFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | SDLYDNDLGRSYDFGL allatostatin precursor protein [Panulirus interruptus] | BAF64528 | SD |
| | SGPSYAFGLa allatostatin precursor protein [Panulirus interruptus] | BAF64528 | 100% identity |
| | GGYAFGLa type-a preproallatostatin II [Pandalopsis japonica] | AFV91539 | ADWSSMGRTWa |
| | PDQLOAPQAVGD Na | Na | Na |
| | GNWDFKHGSWa B-type preproallostatin II [Pandalopsis japonica] | AFV91539 | ANWNKFQGSWa |
| | AEEQAAED Na | Na | Na |
| | ADWNKFHSWa Na | Na | Na |
| | GDEFSPELETTED Na | Na | Na |
| | ANWNKFHSWa B-type preproallostatin II [Pandalopsis japonica] | AFV91539 | ANWNKFQGSWa |
| | GDLDVAEL Na | Na | Na |
| | DWSSLQGTWa B-type preproallostatin I, partial [Pandalopsis japonica] | AFV91539 | GWSLQGGSWa |
| | DWNNHGAWa B-type preproallostatin I, partial [Pandalopsis japonica] | AFV91539 | ANWNHLGAWa |
| | SPDWSLRLGAWa B-type preproallostatin I, partial [Pandalopsis japonica] | AFV91539 | SGDWSLRLGAWa |
| | APDWQFGRSWa B-type preproallostatin I, partial [Pandalopsis japonica] | AFV91539 | DGDWSFGRSWa |
| | VPDEVNEAAHQA Na | Na | Na |
| Allatostatin C | ALGEEQLQEEAAKS Na | Na | Na |
| | MFAPLSPGELPTI C-type preproallostatin [Pandalopsis japonica] | AFV91540 | LFAPLSPPGEIPMT |
| | QIRYHQCFYNPSCF C-type preproallostatin [Pandalopsis japonica] | AFV91540 | QIRYHQCFYPSCF |
| Corazonin | TFQYSRGWTTNa Pro-corazonin [Harpegnathos saltator] | EFN88292 | 100% identity |
| | Crustacean cardioactive peptide | crustacean cardioactive peptide [Homarus gammarus] | ABB46292 | 81% identity in 75% cover |
| | Crustacean female sex hormone | crustacean female sex hormone, partial [Carcinus maenas] | AEI72264 | 26% identity |
| | Crustacean hyperglycemic hormone (CHH) isoform B1 | prepro-crustacean hyperglycemic hormone isoform B [Nephrops norvegicus] | AAQ22392 | 82% identity |
| Hormone                        | Best BLAST hit                                                                 | Accession number | Identity   |
|-------------------------------|-------------------------------------------------------------------------------|-----------------|------------|
| CHH isoform B2                | crustacean hyperglycemic hormone isoform 2 [Rimicaris kairei]                 | ACS35347        | 81% identity |
| CHH isoform B3                | CHH-like protein precursor [Procambarus clarkii]                             | AF474408        | 64% identity |
| CHH isoform B4                | prepro-crustacean hyperglycemic hormone isoform B [Neprops norvegicus]        | AAQ22392        | 85% identity |
| CHH unspecified               | hyperglycemic hormone [Pandalopsis japonica]                                  | AFG16932        | 59% identity |
| Molt inhibiting hormone (MIH) isoform 1 | Molt-inhibiting hormone [Orconectes limosus]                               | P83636          | 55% identity |
| MIH isoform 2                 | Probable molt-inhibiting hormone [Jasus lalandii]                             | P83220          | 70% identity |
| MIH isoform 3                 | Vitellogenesis inhibiting hormone [Homarus gammarus]                          | ABA42181        | 72% identity |
| Diuretic hormone              | prepro-calcitonin-like diuretic hormone [Homarus americanus]                 | ACX46386        | 90% identity |
| Ecdlosion hormone isoform 1   | ecdlosion hormone 2 [Nilaaparvata lugens]                                    | BAC00951        | 62% identity |
| Ecdlosion hormone isoform 2   | ecdlosion hormone 1 [Nilaaparvata lugens]                                    | BAC00950        | 49% identity |
| FLP (myosupressin)            | myosupressin-like-neuropeptide precursor [Procambarus clarkii]               | BAG68789        | 86% identity |
| Follystatin isoform 1          | follistatin-like, partial [Nematostella vectensis]                           | ABF61774        | 54% identity |
| Follystatin isoform 2          | follistatin-related protein 1 isoform 1 [Odobenus rosmarus divergens]        | XP_004403583    | 38% identity |
| Myostatin                     | MSTN [Penaeus monodon]                                                        | ADO34177        | 65% identity |
| Neuropeptide Y                | neuropeptide Y [Lymnaea stagnalis]                                           | CAB63265        | 57% identity |
| Neuropsarin isoform 1          | neuropsarin [Jasus lalandii]                                                  | AHG98659        | 97% identity |
| Neuropsarin isoform 2          | neuropsarin [Jasus lalandii]                                                  | AHG98659        | 48% identity |
| Orcokinin                     | Orcokinin [Procambarus clarkii]                                              | Q9NL83          | 100% identity |
| Pigment dispersing hormone    | Pigment-dispersing hormone [Uca pugilator]                                   | P08871          | NSELINSLGLPKVMNDAa |
| Hormone-3                     | prohormone-3 [Apis mellifera]                                                | XP_001122204    | 43% identity |
| Hormone-4                     | prohormone-4-like [Acythusphion pismu]                                       | XP_001951503    | 89% identity |
| Red pigment concentrating hormone | red pigment-concentrating prohormone [Callinectes sapidus]                  | Q23757          | 63% identity |
| Sulfakinin                    | preprosulfakinin [Homarus americanus]                                        | ABQ95346        | 100% identity |
| Tachykinin                    | preprotachykinin [Procambarus clarkii]                                       | BAC82426        | 100% identity |

Best BLAST hit shows arthropods that are not decapod crustaceans (underlined) and non-arthropods (italicized and underlined). Identity of proteins is given as percentage and peptides as sequence with non-identical aa underlined (amidation is noted by ‘a’).  
doi:10.1371/journal.pone.0097323.t002
Table 3. Alphabetical list of peptide precursors with RPKM quantity in male and female brain and eyestalk.

| Hormone                          | Transcript                     | Comments                                      | M  | BR  | F  | M  | ES  | F  | BR  | ES  |
|----------------------------------|--------------------------------|-----------------------------------------------|----|-----|----|----|-----|----|-----|-----|
| Allatostatins                    | >Unigene56418_All              | A-type prepro-allatostatin, partial (N terminus) | 34.32 | 35.6 | 16.05 | 13.42 | 34.96 | 14.74 |
|                                 | >Unigene36127_All              | A-type prepro-allatostatin, partial (middle)   | 34.14 | 35.77 | 11.78 | 13.75 | 34.96 | 12.77 |
|                                 | >Unigene45628_All              | A-type prepro-allatostatin, partial (C terminus) | 32.56 | 38.8 | 13.85 | 13.49 | 35.68 | 13.67 |
|                                 | >Unigene40422_All              | B-type prepro-allatostatin, partial (N terminus) | 39.12 | 40.22 | 32.4 | 60.14 | 39.67 | 56.27 |
|                                 | >Unigene25318_All              | B-type prepro-allatostatin, partial (C terminus) | 46.66 | 54.5 | 57.75 | 64.03 | 50.58 | 60.89 |
|                                 | >CL2090.Contig2_All            | C-type prepro-allatostatin                    | 11.08 | 11.04 | 3.22 | 3.42 | 11.06 | 3.32 |
|                                 | >Unigene59348_All              | Insects prohormone-1                         | 527.56 | 565.95 | 379.48 | 353.88 | 546.76 | 366.68 |
|                                 |                                | Bursicon a subunit                            | 2.07 | 2.06 | 0 | 0 | 2.07 | 0.00 |
|                                 |                                | Corazonin                                     | 0.54 | 0 | 14.64 | 20.12 | 0.27 | 17.38 |
|                                 |                                | CCAP (crustacean cardioactive peptide)        | 30.9 | 28.86 | 81.84 | 54.44 | 29.88 | 68.14 |
|                                 |                                | CHH (crustacean hyperglycemic hormone)        | 0.12 | 0.46 | 309.12 | 450.84 | 0.29 | 379.98 |
|                                 |                                | MIH/GIH (molt/gonad inhibiting hormone)       | 0.04 | 0.19 | 127.88 | 238.64 | 0.12 | 2.81 |
|                                 |                                | CFSH (crustacean female sex hormone)          | 0.05 | 0 | 4.41 | 8.33 | 0.03 | 6.37 |
|                                 |                                | DH (calcitonin-like diuretic hormone)         | 78.61 | 70.11 | 66.36 | 61.51 | 74.36 | 63.94 |
|                                 |                                | Eclosion hormone                              | 3.9 | 3.01 | 49.2 | 29.2 | 3.46 | 39.20 |
|                                 |                                | FLP (Myosuppressin)                           | 56.18 | 65 | 58.99 | 48.62 | 60.59 | 53.81 |
|                                 |                                | Follistatin                                    | 0.18 | 0.29 | 0.06 | 0.06 | 0.24 | 0.06 |
|                                 |                                | Myostatin                                      | 4.24 | 5.19 | 13.07 | 13.52 | 4.72 | 13.30 |
|                                 |                                | NPY (neuropeptide Y)                          | 3.08 | 2.91 | 47.64 | 46.86 | 3.00 | 47.25 |
|                                 |                                | Neuroparsin                                    | 428.86 | 665.33 | 462.26 | 370.29 | 547.10 | 416.28 |
|                                 |                                | Orcokinin                                      | 12.06 | 6.44 | 14.2 | 17.83 | 9.25 | 16.02 |
|                                 |                                | PDH (pigment dispersing hormone)               | 72.2 | 54.85 | 52.4 | 48.95 | 63.53 | 50.68 |
|                                 |                                | Prohormone-3                                   | 8.55 | 0.18 | 150.22 | 144.9 | 4.37 | 147.56 |
|                                 |                                | Prohormone-4                                   | 2.98 | 0.33 | 70.95 | 62.33 | 1.66 | 66.64 |
|                                 |                                | RPCH (red pigment concentrating hormone)      | 12.84 | 15.22 | 48.32 | 52.63 | 14.03 | 50.48 |
identified in *P. japonica* (Table 2). Another transcript was identified to putatively encode a complete prohormone-1 with 105 aa, starting with a signal peptide of 25 aa, followed by 1 putative neuropeptide, separated by dibasic proteinase cleavage sites (Fig. 3B). The putative neuropeptide in prohormone-1 shares a conserved motif QCXFNXXXSCF with the last putative peptide in the type C allatostatin (Fig. 3C), and is identical to the neuropeptide encoded by prohormone-1 of insects (Table 2). While like allatostatin type C, prohormone-1 has a significantly higher expression in the brain compared with eyestalk, the overall expression of prohormone-1 is one order of magnitude higher compared to all other allatostatins (Table 3).

**Bursicon alpha subunit**

One transcript was identified to putatively encode a complete bursicon alpha subunit precursor with 142 aa, starting with a 25 aa signal peptide, followed by a predicted C-terminal cysteine knot-like domain of 89 aa which contains ten conserved cysteine residues (Table 1 and Fig. 4). The mature hormone share up to 90% identity with bursicon alpha subunit identified in other decapod crustacean species (Table 2). The level of expression is very low in the brain and not evident in the eyestalk (Table 3).

**Corazonin**

One transcript was identified to putatively encode 49 aa of the N-terminus of the corazonin precursor, starting with a 24 aa long signal peptide followed by a 11 aa conserved peptide (identical to corazonin peptides of insects; Table 2) followed by a carboxyl-peptidase cleavage site (Table 1 and Fig. 5). Corazonin expression was found to be almost exclusive to the eyestalk with slight higher levels in females (Table 3).

**Crustacean cardioactive peptide (CCAP)**

One transcript was identified to putatively encode a complete 139 aa open reading frame (ORF) of CCAP precursor starting with a 29 aa signal peptide followed by four predicted peptides (10, 9, 52 and 23 aa in length), separated by carboxyl-peptidase cleavage sites. One of those peptides is highly conserved and contains two cysteine residues predicted to form a disulfide bridge and is amidated (Table 1 and Fig. 6). The highest identity level of the entire ORF, excluding the signal peptide was 81%, with another decapod crustacean CCAP, covering 75% of the ORF (Table 2). The transcript encoding CCAP had significantly higher expression in the eyestalk compared with the brain, with a higher expression in male eyestalk (Table 3).

**Crustacean hyperglycemic hormone (CHH)**

Five transcripts were identified to putatively encode three complete and two partial CHH peptide precursors with 112–139 aa (Table 1 and Fig. 7). All three complete sequences start with a predicted signal peptide of 25–26 aa. One partial sequence has part of the signal peptide (16 aa). All 5 sequences have a CHH-conserved domain of 71–73 aa, preceded by a carboxyl-peptidase cleavage site. The 6 cysteine residues predicted to give rise to 3 disulfide bridges are all aligned between the 5 sequences (Fig. 7A–E). Overall the sequence similarity between the CHH domains is high with up to 89% identity between isoforms B1 and B2 (Fig. 7F, G). Compared with previously described CHHs, identity of the mature hormone was between 59%–85% (Table 2). Isoforms B1-3 had the highest expression of all five transcripts, and found almost exclusively in the eyestalk, while isoform B4 had much lower expression (two orders of magnitude) only in the eyestalk. The
unspecified isoform had equivalent expression to that of isoform B4 in both the eyestalk and the brain. Interestingly, all five isoforms had higher levels in females compared with males (Table 3).

Molt/Gonad-inhibiting hormone (MIH/GIH)

Three transcripts were identified to putatively encode three complete MIH/GIH peptide precursors with 111–115 aa (Table 1 and Fig. 8). All three sequences start with a predicted signal peptide of 33–37 aa followed by an MIH-conserved domain of 74 aa. The 6 cysteine residues predicted to give rise to 3 disulfide bridges are all aligned between all 3 sequences (Fig. 8A–C). Overall, the sequence similarity between the MIH domains is lower than the CHH isoforms with 53%–54% identity (Fig. 8D). Compared with previously described MIHs/GIHs, identity of the mature hormone was between 55%–72% (Table 2). All 3 putative MIH transcripts were found to be specifically expressed in the eyestalk with isoform A2 showing highest expression. Similar to CHH, all three MIH isoforms showed higher expression levels in females compared with males (Table 3).

Crustacean female Sex hormone (CFSH)

One transcript was identified to putatively encode a complete CFSH peptide precursor with 278 aa (Table 1 and Fig. 9). The sequence starts with a 22 aa signal peptide and contains 10 conserved cysteine residues predicted to form 5 disulfide bridges.
(Fig. 9), although the overall identity of the mature hormone does not exceed 26% with other decapod crustaceans (Table 2). CFSH was found to be specifically expressed in the eyestalk, with equivalent expression in both males and females (Table 3).

Diuretic hormone (DH)

One transcript was identified to putatively encode a complete DH peptide precursor with 135 aa (Table 1 and Fig. 10). The sequence starts with a 23 aa signal peptide and the active 31-residue DH peptide is released using dibasic proteinase cleavage sites. This peptide shared 90% identity with a clawed lobster DH (Table 2). The transcript is expressed in both brain and eyestalk with a non significant higher level in brain and in males (Table 3).

Eclosion hormone

Two transcripts were identified to putatively encode complete isoforms of the eclosion hormone precursor (Table 1 and Fig. 11) with 82 and 86 aa, each starting with a signal peptide of 26–28 aa, followed by 55–57 aa eclosion hormone domains each containing 6 conserved cysteine residues predicted to form 3 disulfide bridges (Fig. 11A, B). Other than the cysteine residues, the similarity level between the two eclosion hormone domains is intermediate, with 47% identity (Fig. 11C). Compared to other eclosion hormones, identity of S. verreauxi eclosion was 49%–62% with insect eclosion hormones (Table 2). The first isoform had a significantly higher expression in the eyestalk compared with the brain, and higher expression in males compared with females. The second isoform showed only a basal expression in the female eyestalk (Table 3).

Follistatin

Two transcripts were identified to putatively encode a complete (133 aa) and a partial (204 aa) isoforms of the follistatin precursor (Table 1 and Fig. 12), each starting with a signal peptide of 15 aa, followed by identical 23 aa follistatin domains each containing 4 conserved cysteine residues predicted to form 2 disulfide bridges (Fig. 12A, B). In each predicted peptide, the follistatin domain is followed by a 45 aa kazal-type serine protease inhibitor domain whose N-terminus is identical between the isoforms with 5 cysteine residues and the C-terminus contains 2 additional cysteine residues in the partial isoform (Fig. 12C). The shorter, yet complete follistatin-like isoform ends with a 23 aa predicted transmembrane region. The mature hormones showed identity of 38%–54% to a cnidarians and a mammalian species’ follistatins (Table 2). The first transcript had a very low expression in all tissues and the second transcript had very low expression and was exclusively found in the female brain (Table 3).

Myostatin

One transcript was identified to putatively encode a complete 419 aa ORF of a myostatin precursor, starting with a 18 aa signal peptide, followed by a 136 aa TGF-beta propeptide domain, followed by another 96 aa TGF-beta domain (Table 1 and Fig. 13). The mature hormone showed 65% identity with another decapod crustacean myostatin (Table 2). Myostatin showed significantly higher expression in the eyestalk compared to the brain (Table 3).

Myosupressin

One transcript was identified to putatively encode a complete myosupressin peptide precursor with 100 aa (Table 1 and Fig. 14). The sequence starts with a 29 aa signal peptide and the active 10-
residue myosupressin peptide is released using dibasic and arginine proteinase cleavage sites. Overall the prohormone showed 86% identity with myosupressin of the penaeid shrimp *Penaeus monodon* (Table 2). Myosupressin showed similar expression in the eyestalk and the brain (Table 3).

Neuropeptide Y (NPY)

One transcript was identified to putatively encode a complete NPY precursor with 104 aa (Table 1 and Fig. 15). The sequence starts with a 26 aa signal peptide followed by a 36 aa pancreatic hormone/neuropeptide F/peptide YY family domain, which...
showed 57% identity with an NPY from a mollusk (Table 2).

Neuropeptide Y showed significantly higher expression in the eyestalk compared to the brain (Table 3).

Two transcripts were identified to putatively encode complete neuroparsin peptide precursors with 103–102 aa (Table 1 and Fig. 16A,B). Both sequences contain a 93–101 aa neuroparsin domain with very low similarity (44% identity), although all 12 cysteine residues, predicted to form 6 disulfide bridges are aligned (Fig. 16C). Although the similarity between the two isoforms was rather low, both showed similarity to the same neuroparsin of a spiny lobster (97% and 48%; Table 2). The first neuroparsin encoding transcript had higher expression compared with the second transcript. In both cases the expression was not significantly different between tissues, due to high variation between males and females (Table 3).

Orcokinin

One transcript was identified to putatively encode a complete orcokinin peptide precursor with 205 aa (Table 1 and Fig. 17), starting with a signal peptide of 20 aa, followed by 11 putative neuropeptides, separated by dibasic proteinase cleavage sites (Fig. 17A). The predicted neuropeptides are 8–13 aa in length with NFDEIRDR X GFGF as the most conserved motif (Fig. 17B). All 11 neuropeptides had high homology (5 identical) with orcokinin of either the clawed lobster Homarus americanus or the red swamp crayfish Procambarus clarkii (Table 2). Orcokinin showed higher expression in the male brain compared with the female brain, with similar expression in the eyestalk and the brain (Table 3).

Pigment dispersing hormone (PDH)

Two transcripts were identified to putatively encode complete, highly similar isoforms of PDH precursors (Table 1 and Fig. 18) with 79 aa, both starting with an identical signal peptide of 22 aa, followed by a 23 aa transmembrane region in only one isoform, followed by a carboxy-peptidase cleavage site prior to an 18 aa PDH domain in both isoforms (Fig. 18A, B). Of the 18 aa’s, 15 are identical and the other 3 are similar (Fig. 18C). Both neuropeptides had high homology with previously identified PDH of decapod crustaceans (Table 2). Both of the PDH encoding transcripts showed significantly higher expression in the eyestalk compared with the brain and a higher level in the male brain compared with the female brain (Table 3).

Prohormone-3

One transcript was identified to putatively encode a complete prohormone-3 peptide precursor with 196 aa (Table 1 and Fig. 19). The sequence starts with a 21 aa signal peptide and contains 12 cysteine residues (Fig. 19), all conserved with other insect prohormone-3 sequences, with up to 43% identity in sequence (Table 2). Prohormone-3 encoding transcript showed higher expression in the eyestalk compared to the brain, with higher expression in the male brain compared with the female brain (Table 3).

Prohormone-4

One transcript was identified to putatively encode a partial C-terminus of prohormone-4 peptide precursor with 143 aa (Table 1 and Fig. 20). The highest homology to an insect species was 89% (Table 2). Prohormone-4 encoding transcript showed higher expression in the brain compared to the eyestalk, with higher expression in the male brain compared with the female brain.
expression in the male compared with the female, in both eyestalk and brain (Table 3).

Red pigment concentrating hormone (RPCH)

One transcript was identified to putatively encode a complete RPCH peptide precursor with 99 aa (Table 1 and Fig. 21). The sequence starts with a 21 aa signal peptide followed by the 8-residue RPCH peptide (with 100% identity to peptides of other RPCHs) and RPCH-associated peptide C-terminal domain (Fig. 21). The overall prohormoe shared 63% identity with the blue swimmer crab Callinectes sapidus RPCH (Table 2). Red pigment concentrating hormone encoding transcript showed higher expression in the eyestalk compared to the brain (Table 3).

Sulfakinin

One transcript was identified to putatively encode a complete sulfakinin peptide precursor with 115 aa (Table 1 and Fig. 22). The sequence starts with a 27 aa signal peptide followed by two sulfakinin putative peptides of 10 aa and 13 aa, separated by carboxy-peptidase cleavage sites (Fig. 22). The two peptides had high homology with sulfakinin of H. americanus (Table 2). Sulfakinin encoding transcript showed higher expression in males compared to females both in the brain and the eyestalk (Table 3).

Tachykinin

One transcript was identified to putatively encode a complete tachykinin peptide precursor with 226 aa (Table 1 and Fig. 23). The sequence starts with a 22 aa signal peptide followed by seven...
Figure 12. Follistatin precursors predicted ORFs and conserved peptide. A, B) Complete and a partial follistatin precursor predicted ORFs (derived from CL3958.Contig2_All and Unigene49446_All) each starting with a signal peptide (red) followed by an identical follistatin domain (green) with 4 conserved cysteine residues (yellow), followed by a kazal-type serine protease inhibitor domain (pink) with 5–6 cysteine residues (yellow). The complete, shorter isoform (A) ends with a predicted transmembrane domain (blue). Asterisk indicates the stop codon.

C) Amino acid alignment between the kazal-type domains.

doi:10.1371/journal.pone.0097323.g012
Figure 13. Myostatin precursor predicted ORF. A complete myostatin predicted ORF (derived from CL113.Contig2_All) starting with a signal peptide (red) followed by a TGF-beta propeptide domain (green), followed by another TGF-beta domain domain (pink). Asterisk indicates the stop codon.

doi:10.1371/journal.pone.0097323.g013
identical tachykinin putative peptides of 9 aa each (APSGFLGM-Ramide), separated by carboxy-peptidase cleavage sites (Fig. 23). This peptide was found to be identical to the tachykinin found in *P. clarkii* (Table 2). Tachykinin encoding transcript showed significantly higher expression in the brain compared with the eyestalk (Table 3).

**Discussion**

This study has elucidated the putative neuropeptidome of the previously uncharacterized Eastern rock lobster *S. verreauxi*. Overall 37 partial and complete transcripts were identified which putatively encode 21 peptide families/sub-families (Table 1). These included three partial allatostatin type A transcripts, where one is presumed to represent the N-terminus (Fig. 1A), the other is presumed to represent the middle region (Fig. 1B) and the third is presumed to represent the C-terminus (Fig. 1C). It is conceivable that these three transcripts are part of a one, larger transcript which includes all three, as in most studied arthropod species only one type A allatostatin gene was identified [22], except for blowflies [23]. Overall there are 22 mature peptides of 8 aa predicted to arise from the above three transcripts, each containing the highly conserved YXFGLamide motif (Fig. 1D), found in all arthropods type A allatostatins [22]. Two partial peptides were identified as the putative N-terminus and C-terminus of type B allatostatin precursors (Fig. 2A and B, respectively). The level of conservation between the 13 putative mature peptides encoded by these transcripts was much lower compared with the conservation between the predicted type A allatostatins and six arc novel (Fig. 2C). Two transcripts were identified to encode complete type C allatostatin precursors with very low conservation between the two predicted mature peptides which include the signature cysteine residues of the type C allatostatins (Fig. 3A, B, C). The latter sequence whose best BLAST hit was the predicted prohormone-1 of the honey bee (Table 1) includes the predicted mature peptide which is broadly conserved among crustaceans SYWKQCAFVSCFamide [24]. Most of the mature peptides had very high homology with other arthropods, primarily other decapod crustacean species. Most prominent was the conservation of type A allatostatin-derived peptides with those of the spiny lobster *P. interruptus* and the broadly conserved peptide in prohormone-1 (Table 2).

One complete bursicon alpha subunit predicted sequence was identified, containing a signal peptide and a predicted C-terminal cysteine knot-like domain (Table 1, Fig. 4) with 11 cysteine residues well conserved with other crustacean and insect species, 10 of which are hypothesized to form five disulfide bridges [25]. Another transcript is hypothesized to be the N-terminus of a corazonin precursor, comprising a signal peptide, followed by the 11 aa conserved peptide which is the signature of corazonin (QTFQYSRGVTamide) [26], with 139 aa and high similarity to other crustacean sequences (Table 1&2, Fig. 5). Another sequence is predicted to encode the crustacean cardioactive peptide precursor (CCAP), with 139 aa and high similarity to other crustacean sequences (Table 1&2, Fig. 6).

Five sequences were identified to encode four predicted complete and near complete type B CHH precursors (Crustacean hyperglycemic hormones) and another unspecified CHH precursor. The putative peptides were identified to be specific to the eyestalk as expected from CHHs and included a signal peptide (in 4 out of 5 sequences) and a conserved CHH domain (Table 1, Fig. 7). Although the occurrence of splice variance-derived isoforms of CHH is well documented [27], we currently cannot rule out that the high similarity between the 5 sequences identified (up to 99% identity) is due, at least in part, to sequencing/assembly

**Figure 14.** Myosuppressin precursor predicted complete ORF. A complete Myosuppressin peptide precursor (derived from Unigene55051_All) with a signal peptide (red) and a conserved peptide (green) with an amidated glycine (light blue), bordered by carboxy-peptidase cleavage sites. Asterisk indicates the stop codon. doi:10.1371/journal.pone.0097323.g014

**Figure 15.** Neuropeptide Y (NPY) precursor predicted complete ORF. A complete NPY precursor (derived from Unigenene30121_All) starting with a signal peptide (red) followed by a Pancreatic hormones/neuropeptide F/peptide YY family domain (green) with an amidated glycine (light blue). Asterisk indicates the stop codon. doi:10.1371/journal.pone.0097323.g015
errors rather than actual isoforms. Three sequences were identified to putatively encode complete isoforms of Molt/Gonad-inhibiting hormone (MIH/GIH). All predicted isoforms included a signal peptide followed by a conserved MIH/GIH domain with intermediate similarity (up to 54% identity; Table 1, Fig. 8), suggesting these are more reliably representing isoforms, compared with the predicted CHHs. The homology of CHHs and MIHs with others identified in decapod crustaceans was in some
cases higher than the homology between the isoforms themselves (Table 2), consistent with these genes being diverged for a long time. Most CHH and MIH isoforms were found to be expressed predominantly in the eyestalk with three of the CHH isoforms and one MIH isoform that are most abundantly expressed (Table 3). In most isoforms higher expression was found in females, suggesting that the females sampled were more advanced in the molt cycle. Repeating the neuropeptidome analysis with more samples of males and females of distinct molt stages will enable better distinction between neuropeptides whose expression change with

![Figure 17. Orcokinin precursor predicted complete ORF and conserved motif. A) A complete orckokin precursor predicted ORF (derived from Unigene692_All) with signal peptide (red) and 11 predicted orckokin peptides (green), separated by carboxyl-peptidase cleavage sites (underlined) Asterisk indicates the stop codon. B) Orcokin peptides conservation: 11 predicted neuropeptides of 8–13 aa in length with NFDEIRDRKFGFX conserved. doi:10.1371/journal.pone.0097323.g017](#)

![Logo](#)
relation to molt cycle and neuropeptides whose expression change between genders. Another sequence which was found to express specifically in the eyestalk was predicted to encode a complete Crustacean female sex hormone precursor (CFSH; Table 1, Fig. 9). CFSH was recently identified in two brachyuran crabs and was found to be specifically expressed in the female eyestalk. CFSH knock-down was shown to inhibit the appearance of the female reproductive characteristics which accompany the terminal molt in these species (GenBank Accession # ADO00266).

Interestingly, the putative CFSH in S. verreauxi, identified in this study, was found to be specific to the eyestalk although it is present also in male eyestalks with the same level of expression as in females.

One transcript was predicted to encode a complete calcitonin-like diuretic hormone (DH), with high similarity to the one identified in the American lobster H. americanus [28] (Table 1, Fig. 10). Two transcripts were predicted to encode two complete eclosion hormone precursor isoforms (with 47% identity) each starting with a signal peptide and containing 6 conserved cysteine residues within their eclosion hormone domain (Table 1, Fig. 11). Two transcripts were predicted to encode follistatin-like peptides. Although not considered as neuropeptides, these were included here as it might be of interest to further pursue their precise functionality in crustaceans. The N-termini of both predicted isoforms include identical signal peptides, followed by identical follistatin domains, followed by a kazal-type serine protease inhibitor domain whose N-terminus is identical and the C-terminus was different (Table 1, Fig. 12). One isoform includes a predicted transmembrane region and is a complete ORF (Fig. 12A), while the other is longer, without a predicted transmembrane region and a partial ORF (Fig. 12B). One transcript was identified to encode a complete myostatin precursor with the exact same sequence of that identified in the penaeid shrimp P. monodon (Table 1, Fig. 13). Although also not considered a neuropeptide, like follistatin, its function in regulating muscle development in crustaceans is an interesting aspect to

Figure 18. PDH precursor predicted complete ORFs and conserved motif. A, B) Two complete PDH precursor predicted ORFs (derived from CL7594.Contig2_All and CL7594.Contig3) each starting with an identical signal peptide (red), a transmembrane region in one isofrom (dark blue) and a predicted PDH peptide (green), preceded by a carboxy-peptidase cleavage site (underlined) in each predicted isofrom with an amided glycine (light blue). Asterisk indicates the stop codon. C) PDH peptides conservation 15/18 aa are identical with the other 3 similar in characteristics.
doi:10.1371/journal.pone.0097323.g018
Figure 19. Prohormone-3 precursor predicted complete ORF. A complete prohormone-3 peptide precursor (derived from CL1958.Contig1_All) with a signal peptide (red) and 12 cysteine residues (yellow). Asterisk indicates the stop codon.
doi:10.1371/journal.pone.0097323.g019

Figure 20. Prohormone-4 precursor predicted partial ORF. A partial prohormone-4 peptide precursor (derived from Unigene19311_All). Asterisk indicates the stop codon.
doi:10.1371/journal.pone.0097323.g020

Figure 21. RPCH precursor predicted complete ORF. A complete RPCH peptide precursor (derived from Unigene2547_All) starting with a signal peptide (red) followed by a RPCH domain (green) with an amidated glycine (blue). Asterisk indicates the stop codon.
doi:10.1371/journal.pone.0097323.g021
pursue and is thus included here. Recently, an opposite role was assigned to myostatin in *P. monodon* compared with vertebrates [29]. Based on the identical sequence identified in this study, the Eastern rock lobster might serve a good candidate species to revisit this hypothesis. A complete myosupressin precursor was predicted with a signal peptide and high similarity with *H. americanus* myosuppressin (Table 1, Fig. 14).

One complete predicted neuropeptide Y (NPY) precursor was identified with a conserved active peptide sequence (Table 1, Fig. 15) and two predicted neuroparsin complete peptide precursors were identified with 12 conserved cysteine residues in each, but with rather intermediate similarity between them (Table 1, Fig. 16). Another predicted neuropeptide, orcokinin was identified that included a highly conserved motif of NFDEIRDRXGFGFX within its 11 predicted mature peptides (Table 1, Fig. 17). Two isoforms of the pigment dispersing hormone (PDH) precursor were identified with intermediate similarity overall. The predicted mature peptide shows high similarity between the two sequences (15/18 aa identical). Two sequences were predicted to encode complete prohormone-3 and prohormone-4 precursors (Table 1, Fig. 19, 20). Both have been characterized solely in insects, apart from one prohormone-4 like peptide identified in the copepod *Acartia pacifica* (GenBank accession number AGN29584), hence this is the first report of the two hormones in decapods.

A predicted red pigment concentrating hormone (RPCH) precursor was identified with a signal peptide and RPCH domain (Table 1, Fig. 21). Another sequence is predicted to encode a complete sulfakinin precursor with a signal peptide and two mature peptides separated by peptidase cleavage sites (Table 1, Fig. 22). Finally, one sequence was identified to putatively encode a complete tachykinin precursor with a signal peptide followed by seven identical tachykinin peptides, separated by peptidase cleavage sites (Table 1, Fig. 23). The tachykinin putative sequence had high similarity to the one identified in the spiny lobster *P. interruptus*.

Materials and Methods

Animals

*Sagmariasus verreauxi* individuals were maintained at Institute for Marine and Antarctic Studies under previously described parameters [30]. Prior to dissections, animals were anesthetized on ice for at least 20 min.
Sample Preparation and Sequencing

Total RNA from eyestalks and brains of two mature *S. verreauxi* males and two mature females were isolated separately with the Trizol Reagent (Invitrogen), according to the manufacturer’s instructions, followed by next generation sequencing by BGI (HongKong Co. Ltd) as per manufacturer’s protocol (Illumina, San Diego, CA). Briefly, poly (A) mRNA was isolated using oligo (dT) beads and the addition of fragmentation buffer for shearing mRNA into short fragments (200–700 nt) prevented priming bias during the synthesis of cDNA using random hexamer-primers. The short fragments were further purified using QiaQuick PCR extraction kit and resolved, with EB buffer for ligation, with Illumina Paired-end adapters. This was followed by size selection (~200 bp). PCR amplification and Illumina sequencing using an Illumina Genome Analyzer (HighSeq 2000, Illumina, San Diego, CA), performing 90 bp-paired end sequencing. The sequence reads were stored as FASTQ files. Overall, at least 4 Gb of cleaned data (at least 45 million reads) was generated for each of the four samples sequenced, which included pooled eyes of two males and two females, pooled brains of two males and two females.

Bioinformatics analyses

Cleaning of low quality reads, assembly and annotation were done by BGI, using unpublished algorithms (BGI, HongKong Co. Ltd), Trinity [31] and Blast2GO [32], respectively. We validated that the reads obtained by BGI are clean using FASTQ/A Trimmer (http://hannonlab.cshl.edu/fastx_toolkit/index.html), which gave an output of over 99.99% of the reads untrimmed. The list of annotated sequences was scanned for key words, including names and abbreviations of previously known neuropeptides as well as general key words such as ‘hormone’. Multiple sequence alignment of the predicted neuropeptide hormones as well as general key words such as ‘hormone’, including names and abbreviations of previously known neurohormones was performed with ClustalW [33], followed by a Neighbor Joining Phylogram (for the CHH sequences) generated via MEGA 5.0 [34] with 1000 bootstrap trials. The multiple sequence alignment file was then exported to TexShade [35] for highlighting the conserved sequence motifs. Signal peptide was predicted using SignalP 4.1 server [36]. Domain prediction was done either via SMART [37] or by comparison with references of other crustacean neuropeptide sequences. The re-validated clean FASTQ files were re-assembled using default parameters in CLC Genomics Workbench v4 (CLC Bio) and validated the assembled transcripts corresponding the neuropeptides using BLAST. Digital Gene Expression was computed using CLC Genomics Workbench v4 (CLC Bio), with default parameters with the exception of 0.9 similarity fraction instead of 0.8. Resulting BAM files were deposited in the sequence read archive (http://www.ncbi.nlm.nih.gov/sra) as biosample SAMN02419461. BAM files were then uploaded onto Partek Genomics Suite (Partek GS) where quantification was performed, yielding reads per kilobase per million reads (RPKM). The quantified data was analyzed using ANOVA, performed in Partek GS, with contrast between values in eye and brain for each neuropeptide. The threshold for statistical significance was set to p<0.05. Since there was only one male and one female sample for each tissue, no statistical analysis was applicable to compare males and females.

Acknowledgments

The Authors would like to thank Ms. Bromwyn Rotgans for compiling a list of previously known neuropeptides in crustaceans. This list served for verifying the identified putative peptides.

Author Contributions

Conceived and designed the experiments: TV AE SFC. Performed the experiments: TV. Analyzed the data: TV. Contributed reagents/materials/analysis tools: SCB QPF. Wrote the paper: TV.
crustacean C-type allatostatin-like peptide with both neuromodulatory and cardioactive properties. Journal of Experimental Biology 212: 1140–1152.

25. Luo C-W, Dewey EM, Sudo S, Ever J, Hsu SY, et al. (2005) Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2. Proceedings of the National Academy of Sciences of the United States of America 102: 2020–2025.

26. Veenstra JA (1989) Isolation and structure of corazonin, a cardioactive peptide from the American cockroach. FEBS Letters 250: 231–234.

27. Chen S-H, Lin C-Y, Kuo CM (2004) Cloning of two crustacean hyperglycemic hormone isoforms in freshwater giant prawn (Macrobrachium rosenbergii): evidence of alternative splicing. Marine Biotechnology 6: 83–94.

28. Christie AE, Stevens JS, Bowers MR, Chapline MC, Jensen DA, et al. (2010) Identification of a calcitonin-like diuretic hormone that functions as an intrinsic modulator of the American lobster, Homarus americanus, cardiac neuromuscular system. The Journal of experimental biology 213: 118–127.

29. De Santis C, Wade NM, Jerry DR, Preston NP, Glencross BD, et al. (2011) Growing backwards: an inverted role for the shrimp ortholog of vertebrate myostatin and GDF11. The Journal of experimental biology 214: 2671–2677.

30. Jensen MA, Carter CG, Adams LR, Fitzgibbon QP (2013) Growth and biochemistry of the spiny lobster Sagmariasus verreauxi cultured at low and high density from hatch to puerulus. Aquaculture 376–379: 162–170.

31. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29: 644–652.

32. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talón M, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

33. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673–4680.

34. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731–2739.

35. Beitz E (2000) TExshade: shading and labeling of multiple sequence alignments using LaTeX2e. Bioinformatics 16: 135–139.

36. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8: 785–786.

37. Schulz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. Proceedings of the National Academy of Sciences 95: 5857–5864.