



RESEARCH ARTICLE

Comparative genomic analysis of crustacean hyperglycemic hormone (CHH) neuropeptide genes across diverse crustacean species [version 1; referees: 1 approved, 2 approved with reservations]

Wai Hoong Chang , Alvina G. Lai 

Nuffield Department of Medicine, University of Oxford, Oxford, Oxfordshire, OX3 7FZ, UK

v1 First published: 23 Jan 2018, 7:100 (doi: [10.12688/f1000research.13732.1](https://doi.org/10.12688/f1000research.13732.1))
 Latest published: 23 Jan 2018, 7:100 (doi: [10.12688/f1000research.13732.1](https://doi.org/10.12688/f1000research.13732.1))

Abstract

Background: Recent studies on bioactive peptides have shed light on the importance of these compounds in regulating a multitude of physiological, behavioral and biological processes in animals. Specifically, the neuropeptides of the crustacean hyperglycemic hormone (CHH) superfamily is known to control a number of important functions ranging from energy metabolism, molting, osmoregulation to reproduction.

Methods: Given the importance of this peptide family, we employed a conservative approach utilizing extant transcriptome datasets from 112 crustacean species, which not only include important food crop species from the order Decapoda, but also from other lower order crustaceans (Branchiopoda and Copepoda), to identify putative CHH-like sequences.

Results and conclusions: Here we describe 413 genes that represent a collection of CHH-like peptides in Crustacea, providing an important staging point that will now facilitate the next stages of neuroendocrine research across the wider community.

Keywords

Crustaceans, Neuropeptides, Crustacean Hyperglycemic Hormone (CHH), Transcriptomics, Comparative genomics

Open Peer Review

Referee Status:

	Invited Referees		
	1	2	3
version 1 published 23 Jan 2018	 report	 report	 report

- Abigail Elizur**, University of the Sunshine Coast, Australia
Tuan V. Nguyen, University of the Sunshine Coast, Australia
- Chi-Ying Lee**, National Changhua University of Education, Taiwan
Jean-Yves Toullec, Station Biologique de Roscoff, France
- Jos H. M. Schippers**, RWTH Aachen University, Germany

Discuss this article

Comments (0)

Corresponding author: Alvina G. Lai (alvina.lai@ndm.ox.ac.uk)

Author roles: **Chang WH:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Lai AG:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the EMBO Fellowship and the Human Frontier Science Program Fellowship to AGL. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2018 Chang WH and Lai AG. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Chang WH and Lai AG. **Comparative genomic analysis of crustacean hyperglycemic hormone (CHH) neuropeptide genes across diverse crustacean species [version 1; referees: 1 approved, 2 approved with reservations]** *F1000Research* 2018, 7:100 (doi: [10.12688/f1000research.13732.1](https://doi.org/10.12688/f1000research.13732.1))

First published: 23 Jan 2018, 7:100 (doi: [10.12688/f1000research.13732.1](https://doi.org/10.12688/f1000research.13732.1))

Introduction

Crustaceans and insects from the phylum Arthropoda have longstanding histories in peptide biology research, principally in areas related to the roles of peptide hormones in physiology and neuroendocrine signaling. Early discoveries have demonstrated that compounds in the crustacean nervous system were responsible for chromatophore control¹⁻³. Four decades later, it was revealed that a compound known as the red pigment concentrating hormone functions as the first crustacean/invertebrate neuropeptide⁴. Since then, multiple studies have shed light on the highly pleiotropic functions of crustacean neuropeptides implicated in the regulation of a myriad of physiological processes such as light adaptation, molt inhibition, carbohydrate metabolism, reproduction and ion transport⁵⁻⁹.

The crustacean hyperglycemic hormone (CHH) represents a neuropeptide superfamily that is unique to arthropods^{6,10-12}. This superfamily is made up of peptides containing ~70 amino acids originally isolated from the X-organ-sinus-gland system of the decapod *Carcinus maenas*¹³. Given their high degree of structural similarities, and the conservation of six cysteine residues, the molt-inhibiting hormone (MIH) and gonad-inhibiting hormone (GIH) were considered as part of this family collectively known as CHH/MIH/GIH. To date, at least 150 CHH peptides have been isolated and characterized, mainly in decapods through comparative studies on endocrinology^{5-7,14-21}. Although there are reports on CHH peptides in other crustacean taxa such as *Armadillidium vulgare* (Isopoda)^{22,23}, *Daphnia pulex* (Cladocera)²⁴ and *Daphnia magna*¹⁵, investigations beyond decapods have remained scant and the sequences of CHH/MIH/GIH genes in other crustacean taxa have remained elusive.

Here, we took advantage of the growing number of high-throughput crustacean datasets on public repositories to perform transcriptome mining of the CHH/MIH/GIH superfamily. To this end, we looked at crustacean species from three Classes (Figure 1) and annotated CHH/MIH/GIH genes. This high confidence set of genes identified using our *in silico* framework provides an important basis for understanding neuropeptide biology underpinning physiological adaptations across diverse crustacean species.

Methods

Transcriptome datasets and query sets

We retrieved complete transcriptome datasets for 112 crustacean species available at the time of manuscript preparation from the [European Nucleotide Archive](#). Five non-crustacean arthropod proteomes were retrieved from [Uniprot](#). A complete list of accessions used in this study is provided in [Supplementary Table 1](#). We retrieved a list of query sequences used in subsequent homology searches from [Uniprot](#) and [GenBank](#).

Identification of CHH/MIH/GIH peptides

To identify CHH/MIH/GIH gene orthologs, we used multiple [Basic Local Alignment Search Tool \(BLAST\)](#)-based approaches such as BLASTp and tBLASTn with varying Blocks Substitution matrices based on a previously published workflow²⁵. The BLAST results were filtered by e-value of $< 10^{-6}$, best reciprocal BLAST hits against the GenBank non-redundant (nr) database and redundant contigs having at least 95% identity were collapsed using [CD-HIT](#). We then utilized [HMMER](#) (version 3.1) employing hidden Markov models (HMM) profiles²⁶ to scan for the presence of CHH Pfam domains²⁷ on the best reciprocal nr

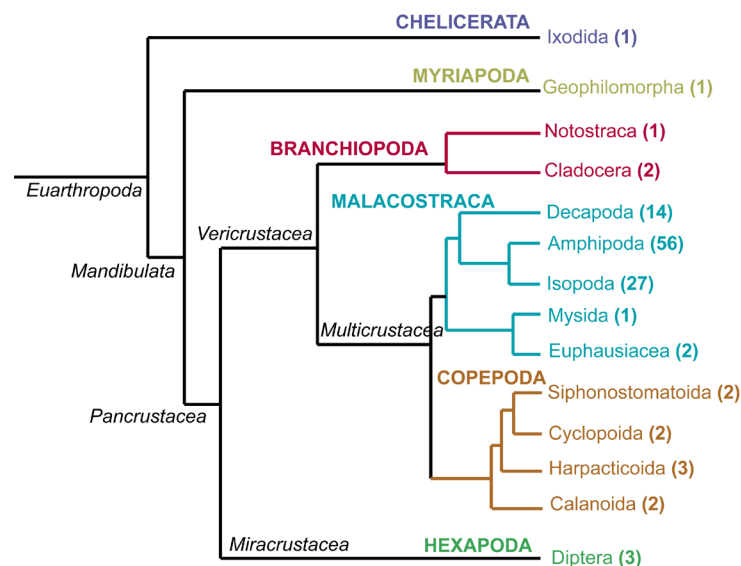


Figure 1. Phylogenetic relationship of Crustacea. The number of species within each taxon is denoted in parentheses.

BLAST hits to compile a final non-redundant set of crustacean CHH/MIH/GIH orthologs. Pfam annotations, associated e-values and fasta sequences are provided in [Dataset 1²⁸](#) and [Dataset 2²⁹](#).

Multiple sequence alignment and phylogenetic tree construction

Multiple sequence alignments of CHH protein sequences were performed using MAFFT (version 7)³⁰. Phylogenetic tree was built from the MAFFT alignment using RAxML WAG + G model to generate best-scoring maximum likelihood trees³¹. Geneious (version 7) was used to generate multiple sequence alignment images as well as graphical representations of the Newick tree³².

Results and discussion

We have annotated CHH/MIH/GIH genes from 112 crustacean transcriptome datasets representing three Classes: Malacostraca (Amphipoda: 56 species, Decapoda: 14 species, Isopoda: 27 species, Euphausiacea: 2 species and Mysida: 1 species), Branchiopoda (3 species), and Copepoda (9 species) ([Supplementary Table 1](#)). We also looked at 5 non-crustacean species from Arthropoda: Insecta (3 species), Arachnida (1 species) and Chilopoda (1 species) ([Supplementary Table 1](#)). Using sequence and motif similarity based approaches, we have conservatively identified a total of 413 genes from these transcriptomes ([Figure 2](#); [Dataset 1²⁸](#) and [Dataset 2²⁹](#)).

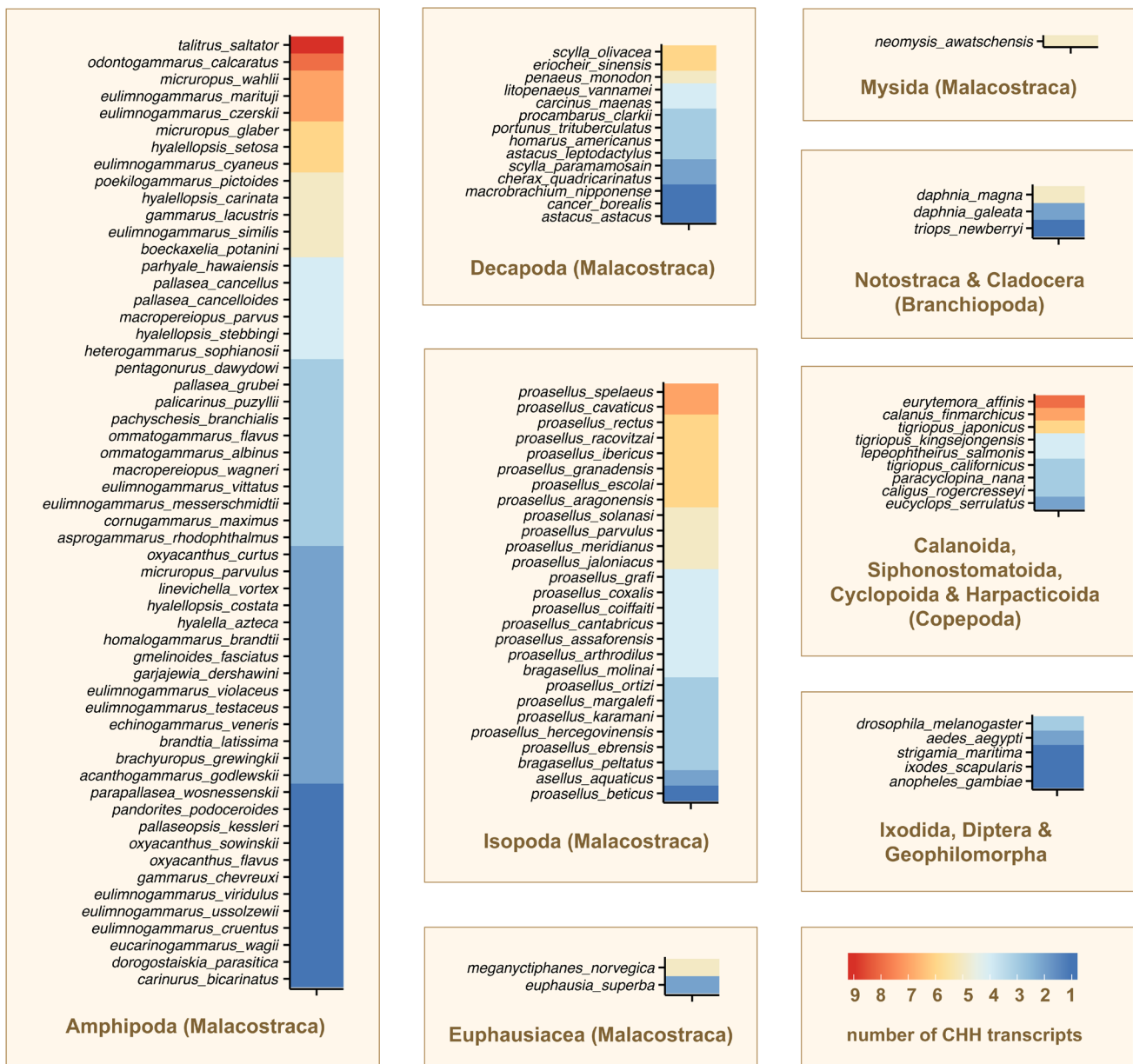


Figure 2. CHH/MIH/GIH genes in Crustacea. Heat maps denote the number of CHH/MIH/GIH genes identified from each crustacean species. CHH/MIH/GIH genes from five non-crustacean species within Arthropoda are also shown.

Multiple sequence alignment analyses on representative CHH/MIH/GIH sequences revealed the presence of a conserved set of six cysteine residues (Figure 3), likely contributing to the formation of disulfide bonds³³. Comparison of insect sequences from *Drosophila melanogaster*, *Anopheles gambiae* and *Aedes aegypti* demonstrated sequence identities of at least 46% (Supplementary Table 2). Within crustacean taxa, a range of sequence identities were observed: Branchiopoda (~25% to 93%), Copepoda (~12% to 30%) and Malacostraca (~10% to 98%) (Supplementary Table 2). This is reflected in the phylogeny where CHH/MIH/GIH sequences from related individuals form distinct clusters (Figure 4). It was previously reported that multiple gene duplications of CHH family peptides occurred in the decapod lineage leading to a high degree of genetic polymorphism¹⁵, hence providing an explanation for our current observation. Two separate clusters of CHH genes exhibiting antagonistic patterns of expression were identified in the decapod *Metapenaeus ensis*, posited to represent an ancient gene duplication event³⁴. Although it is not possible to pinpoint the genomic loci of CHH sequences identified from this study, it is likely that paralogous copies offer mechanisms for evolving new functions through functional divergence. CHH-like genes arising from duplication of the ancestral copy are subjected to reduced selective pressure and therefore

may lose their hyperglycemic activity to adopt more specialized roles¹⁵. Further biochemical studies will be required to unravel the functions of the novel genes identified from this study.

Dataset 1. Fasta file for CHH/MIH/GIH sequences in crustaceans and other arthropods

<http://dx.doi.org/10.5256/f1000research.13732.d191194>

Dataset 2. List of Pfam annotated CHH/MIH/GIH genes and associated e-values in crustaceans and other arthropods

<http://dx.doi.org/10.5256/f1000research.13732.d191195>

Conclusions

We have generated a high confidence list of CHH/MIH/GIH sequences from distantly related crustaceans. As a fundamental step in a broader endeavor this data is now available to the wider community to allow detail functional analyses pertinent to the next stages of neuropeptide research. Given the paucity of CHH sequences beyond decapod crustaceans, our analysis forms a promising basis for studies ranging from biochemistry to the evolution of this elusive superfamily.

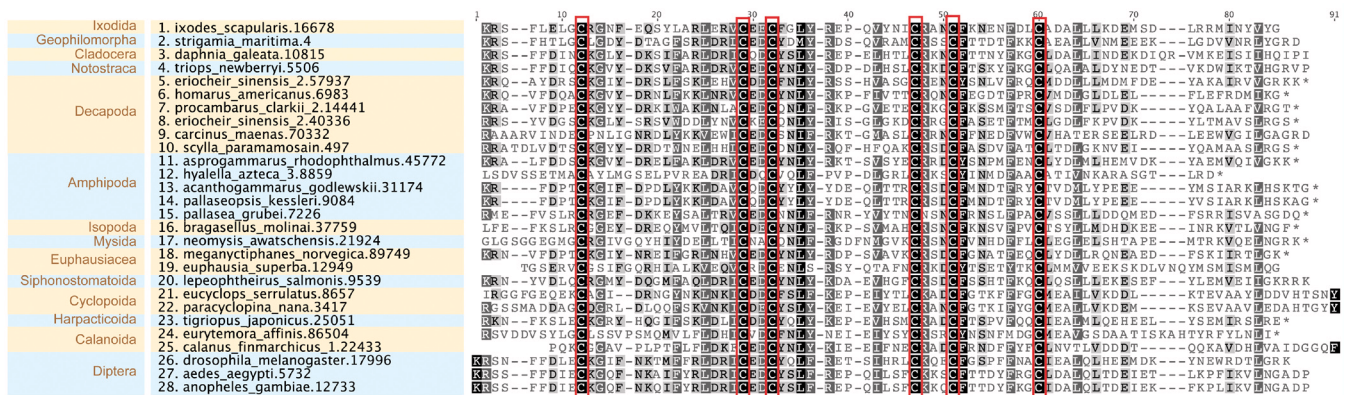


Figure 3. Multiple sequence alignment of representative CHH/MIH/GIH proteins from each taxon. Six conserved cysteine residues are annotated within red boxes.

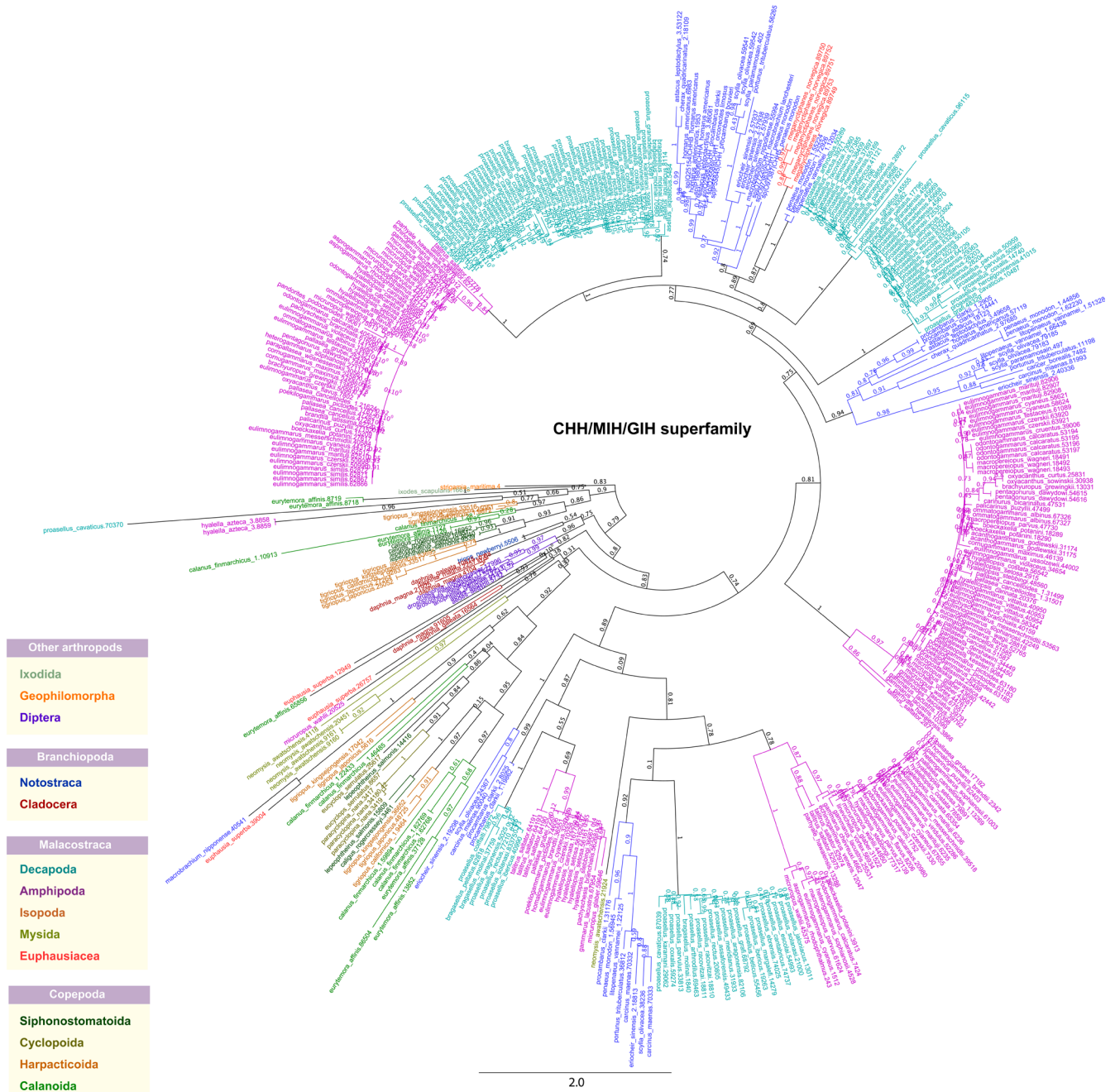


Figure 4. Crustacean CHH/MIH/GIH phylogeny. The tree was constructed using the maximum-likelihood method from an amino acid multiple sequence alignment. The node labels of each taxon are marked with distinctive colors denoted in the figure inset. Bootstrap support values ($n=1000$) are denoted as branch labels.

Data availability

Data supporting the conclusions of this study are provided as [Supplementary Material](#) and [Dataset 1](#) and [Dataset 2](#).

Dataset 1: Fasta file for CHH/MIH/GIH sequences in crustaceans and other arthropods. [10.5256/f1000research.13732.d191194](https://doi.org/10.5256/f1000research.13732.d191194)²⁸

Dataset 2: List of Pfam annotated CHH/MIH/GIH genes and associated e-values in crustaceans and other arthropods. [10.5256/f1000research.13732.d191195](https://doi.org/10.5256/f1000research.13732.d191195)²⁹

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by the EMBO Fellowship and the Human Frontier Science Program Fellowship to AGL.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Supplementary Table 1: List of accession numbers for species used in this study.

[Click here to access the data.](#)

Supplementary Table 2: Protein distance matrix (% identity) constructed from a multiple sequence alignment of CHH/MIH/GIH sequences.

[Click here to access the data.](#)

References

- Brown FA: **The Controlling Mechanism of Chromatophores in Palaemonetes.** *Proc Natl Acad Sci U S A.* 1933; **19**(3): 327–329.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Perkins EB: **Color changes in crustaceans, especially in Palaemonetes.** *J Exp Zool Part A Ecol Genet Physiol.* 1928; **50**(1): 71–105.
[Publisher Full Text](#)
- Kleinholz LH: **Crustacean eye-stalk hormone and retinal pigment migration.** *Biol Bull.* 1936; **70**(2): 159–184.
[Publisher Full Text](#)
- Fernlund P, Josefsson L: **Crustacean color-change hormone: amino acid sequence and chemical synthesis.** *Science.* 1972; **177**(4044): 173–175.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Böcking D, Dirksen H, Keller R: **The crustacean neuropeptides of the CHH/MIH/GIH family: structures and biological activities.** In *Crustacean Nerv Syst.* 2002; **84**–97.
[Publisher Full Text](#)
- Chan SM, Gu PL, Chu KH, *et al.*: **Crustacean neuropeptide genes of the CHH/MIH/GIH family: implications from molecular studies.** *Gen Comp Endocrinol.* 2003; **134**(3): 214–219.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Christie AE, Stemmler EA, Dickinson PS: **Crustacean neuropeptides.** *Cell Mol Life Sci.* 2010; **67**(24): 4135–4169.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Christie AE, Cashman CR, Brennan HR, *et al.*: **Identification of putative crustacean neuropeptides using *in silico* analyses of publicly accessible expressed sequence tags.** *Gen Comp Endocrinol.* 2008; **156**(2): 246–264.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Christie AE, McCoole MD, Harmon SM, *et al.*: **Genomic analyses of the *Daphnia pulex* peptidome.** *Gen Comp Endocrinol.* 2011; **171**(2): 131–150.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Beltz BS: **Crustacean neurohormones.** *Invertebr Endocrinol.* 1988; **2**: 235–258.
[Reference Source](#)
- De Kleijn DP, Van Herp F: **Molecular biology of neurohormone precursors in the eyestalk of Crustacea.** *Comp Biochem Physiol B Biochem Mol Biol.* 1995; **112**(4): 573–579.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Keller R: **Crustacean neuropeptides: structures, functions and comparative aspects.** *Experientia.* 1992; **48**(5): 439–448.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kegel G, Reichwein B, Weese S, *et al.*: **Amino acid sequence of the crustacean hyperglycemic hormone (CHH) from the shore crab, *Carcinus maenas*.** *FEBS Lett.* 1989; **255**(1): 10–14.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Montagné N, Soyez D, Gallois D, *et al.*: **New insights into evolution of crustacean hyperglycaemic hormone in decapods—first characterization in *Anomura*.** *FEBS J.* 2008; **275**(5): 1039–1052.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Montagné N, Desdevises Y, Soyez D, *et al.*: **Molecular evolution of the crustacean hyperglycemic hormone family in ecdysozoans.** *BMC Evol Biol.* 2010; **10**: 62.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chen SH, Lin CY, Kuo CM: ***In silico* analysis of crustacean hyperglycemic hormone family.** *Mar Biotechnol (NY).* 2005; **7**(3): 193–206.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Fanjul-Moles ML: **Biochemical and functional aspects of crustacean hyperglycemic hormone in decapod crustaceans: review and update.** *Comp Biochem Physiol C Toxicol Pharmacol.* 2006; **142**(3–4): 390–400.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Chung JS, Zmora N, Katayama H, *et al.*: **Crustacean hyperglycemic hormone (CHH) neuropeptides family: Functions, titer, and binding to target tissues.** *Gen Comp Endocrinol.* 2010; **166**(3): 447–454.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lacombe C, Grève P, Martin G: **Overview on the sub-grouping of the crustacean hyperglycemic hormone family.** *Neuropeptides.* 1999; **33**(1): 71–80.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Soyez D: **Occurrence and diversity of neuropeptides from the crustacean hyperglycemic hormone family in arthropods. A short review.** *Ann N Y Acad Sci.* 1997; **814**: 319–323.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Van Herp F: **Molecular, cytological and physiological aspects of the crustacean hyperglycemic hormone family.** *Recent Adv Arthropod Endocrinol.* 1998; **53**–70.
[Reference Source](#)
- Grève P, Sorokine O, Berges T, *et al.*: **Isolation and amino acid sequence of a peptide with vitellogenesis inhibiting activity from the terrestrial isopod *Armadillidium vulgare* (Crustacea).** *Gen Comp Endocrinol.* 1999; **115**(3): 406–414.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Martin G, Sorokine O, Van Dorsselaer A: **Isolation and molecular characterization of a hyperglycemic neuropeptide from the sinus gland of the terrestrial isopod *Armadillidium vulgare* (Crustacea).** *Eur J Biochem.* 1993;

- 211(3): 601–607.
[PubMed Abstract](#) | [Publisher Full Text](#)
24. Gard AL, Lenz PH, Shaw JR, *et al.*: **Identification of putative peptide paracrines/hormones in the water flea *Daphnia pulex* (Crustacea; Branchiopoda; Cladocera) using transcriptomics and immunohistochemistry.** *Gen Comp Endocrinol.* 2009; **160**(3): 271–287.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Lai AG, Aboobaker AA: **Comparative genomic analysis of innate immunity reveals novel and conserved components in crustacean food crop species.** *BMC Genomics.* 2017; **18**(1): 389.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Finn RD, Clements J, Eddy SR: **HMMER web server: interactive sequence similarity searching.** *Nucleic Acids Res.* 2011; **39**(Web Server issue): W29–37.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Bateman A, Coin L, Durbin R, *et al.*: **The Pfam protein families database.** *Nucleic Acids Res.* 2004; **32**(Database issue): D138–41.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Chang WH, Lai AG: **Dataset 1 in: Comparative genomic analysis of crustacean hyperglycemic hormone (CHH) neuropeptide genes across diverse crustacean species.** *F1000Research.* 2018.
[Data Source](#)
29. Chang WH, Lai AG: **Dataset 2 in: Comparative genomic analysis of crustacean hyperglycemic hormone (CHH) neuropeptide genes across diverse crustacean species.** *F1000Research.* 2018.
[Data Source](#)
30. Katoh K, Asimenos G, Toh H: **Multiple alignment of DNA sequences with MAFFT.** *Methods Mol Biol.* 2009; **537**: 39–64.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics.* 2014; **30**(9): 1312–3.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Kearse M, Moir R, Wilson A, *et al.*: **Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.** *Bioinformatics.* 2012; **28**(12): 1647–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Christie AE: **Identification of the first neuropeptides from the Amphipoda (Arthropoda, Crustacea).** *Gen Comp Endocrinol.* 2014; **206**: 96–110.
[PubMed Abstract](#) | [Publisher Full Text](#)
34. Gu PL, Yu KL, Chan SM: **Molecular characterization of an additional shrimp hyperglycemic hormone: cDNA cloning, gene organization, expression and biological assay of recombinant proteins.** *FEBS Lett.* 2000; **472**(1): 122–128.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 08 October 2018

doi:10.5256/f1000research.14919.r37870



Jos H. M. Schippers

Institute of Biology, RWTH Aachen University, Aachen, Germany

The manuscript presented by Chang and Lai describes a collection of CHH neuropeptide genes in over 100 crustacean species.

I read with interest this short report and could appreciate that a majority of neuropeptide studies have been limited to decapod crustacean species previously. In addition the presence of neuropeptide transcripts in other divergent crustaceans including those from basal lineages (copepods and branchiopods) are presented.

Like in their previous work (Chang and Lai, 2018, *BMC Genomics*¹), a broad sampling strategy was followed. Here the approach was across multiple crustacean lineages, which allows for drawing conclusions on the evolutionary trajectory of CHH genes and their potential functional role during life histories.

Although the paper describes a single finding, it is a valuable asset, as it provides an inventory of a high-quality list of carefully curated CHH neuropeptide transcripts for future functional studies.

The paper is well-written and concise, and definitely fits with the scope and eligibility of the *F1000Research* guidelines.

As a minor comment: it would be good to describe how the tree in Figure 1 was made in the legend.

References

1. Chang W, Lai A: Mixed evolutionary origins of endogenous biomass-depolymerizing enzymes in animals. *BMC Genomics*. 2018; **19** (1). [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 20 April 2018

doi:[10.5256/f1000research.14919.r30523](https://doi.org/10.5256/f1000research.14919.r30523)



Chi-Ying Lee¹, Jean-Yves Toullec²

¹ Department of Biology, National Changhua University of Education, Changhua, Taiwan

² Sorbonne Universités, UPMC Université Paris 06, UMR 7144 CNRS, Equipe ABICE, Station Biologique de Roscoff, Roscoff, France

The authors collected and analyzed transcriptome datasets for 112 crustacean species (representing three Classes - Malacostraca, Branchiopoda, and Copepoda), as well as those for 5 non-crustacean (arthropod) species (Insecta, Arachnida and Chilopoda) (Supplementary Table 1; Fig. 1), retrieved from public repository.

Collectively, 413 genes are annotated based on sequence and motif similarity as encoding peptides belonging to the so-called CHH/MIH/GIH family. The number of CHH/MIH/GIH genes identified from each crustacean species is shown as heat maps (Fig. 2). Protein sequence, annotation information, etc. are listed (Dataset 1 and 2). Between-species % identity is reported (Supplementary Table 2). Multiple sequence alignment on representative CHH/MIH/GIH sequences is presented in Fig. 3 to show the 6 conserved cysteine residues (Fig. 3). A phylogenetic tree is built based on the multiple sequence alignment data (Fig. 4).

The main drawback of this work is it only results in a 413-sequence inventory of CHH/MIH/GIH family peptides, but with little information (nucleotide length, amino acid sequence).

The authors did not describe how the tree presented in Fig. 1 was built with what sorts of data. If they took a constructed tree from other source, this should be cited. There are obviously other current interpretations of a phylogenetic relationship of Crustacea, in addition to the one shown in the manuscript.

For annotation information (Dataset 2), it should be more discerning in that peptide should be assigned as CHH, MIH/GIH, MOIH (mandibular organ-inhibiting hormone) or ITP (ion transport peptide), instead of only giving an inclusive description - Crustacean CHH/MIH/GIH neurohormone family.

Results derived from further analysis of the sequence data (Figs, 2, 3, and 4) are not adequately

discussed or lead to conclusions that have been previously established.

For example, the only conclusion made with Fig. 3 is the 6 highly cysteine residues, a signature of the family already extensively described. Figure 4 is left without discussion. Moreover, it is not mentioned which type of peptide (CHH, MIH/HIV, or ITP) the different clusters highlighted in the tree belong to.

Part of the discussion (left column, p. 4) starting with “It was previously reported that multiple gene duplications of CHH family peptides.....” is not particularly relevant to the data being discussed and is again making a point already extensively described and discussed. It should also be mentioned that genomic data are more adequate than transcriptomic data when discussing gene duplication and gene copy number (Fig. 2).

Figure 2 is also left without discussion. In addition, data presented in this figure might be suffering from the criticism that the number of CHH/MIH/GIH genes assigned to each species is likely biased as the number will be absolutely influenced by the type of tissue used for sequencing (eyestalk ganglia, other tissue, or whole animal) and sequencing depth. Accordingly, the relevant information should be given in Supplementary Table 1 and comparison should only be made among species with the same sequenced tissue type (e.g., eyestalk ganglia) and comparable sequencing depth.

Overall, while the methods employed are largely (but not entirely) appropriate for the intended analysis, the goal of the study is not clearly set and this work only produces a sequence inventory without novel finding or solid discussion. Additional analyses expected to yield novel finding should be added to the manuscript. For example, with a vast amount of sequence information that could be extracted from a set of 413 peptides from animals encompassing 3 Classes, would it not be possible to uncover some function-defining sequence characteristics or motif? This piece of information would be useful for functional analysis, which the authors thought should be an important aspect of the next stage of research.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Referee Expertise: Endocrinology

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Referee Report 26 February 2018

doi:10.5256/f1000research.14919.r31165



Abigail Elizur , Tuan V. Nguyen

Genecology Research Center, Faculty of Science, Health, Engineering and Education, University of the Sunshine Coast, Sippy Downs, Qld, Australia

The current study is an interesting “*in silico*” approach that mine for CHH/MIH/GIH neuropeptide in a wide range of crustacean. Introduction is updated and concise. However, there are some minor points that will need to be addressed to provide clarity information for readers. Materials and methods is sound and appropriate, although I raise some particular suggestions to further expand the quality of the MS. Results and discussions are overall good. One major point that needs to be fixed is the phylogenetics tree, whereas the color code system does not correlate between legend and figure.

In addition, I have specific comments, (the fact that it does not appear as numbered pages and lines makes it a bit awkward, but see below):

Page 2: ...known as CHH/MIH/GIH- There are also Mandibular Inhibiting hormone (MOIH) and also ITP (Ion Transport Protein) to be part of the super family. See Ohira (2016)¹.

...from three Classes - Please elaborate on this, are these branchiopoda, malacostraca and copepoda?

(Figure 1) – How was Figure 1 generated ? Is this a phylogenetics based tree built on mitochondrial genome or it is just an illustration?

Genebank nr database... What about TSA (Transcriptome Shotgun Assembly) database? As that will increase the chance that the algorithm will discover more CHH/MIH/GIH? Since you mention in the introduction "Here, we took advantage of the growing number of high-throughput crustacean datasets on public repositories to perform transcriptome mining of the CHH/MIH/GIH superfamily." Adopting the TSA database will be better suited for the MS.

...and redundant contigs... Why 95%?, let's assume that CHH prepropeptide can be roughly more than 120aa. For instance, two distinct CHH in *Homarus americanus* CHH-A (P19806) and CHH-B (Q25154) have 127/134 aa that are similar (~95%). Will they be collapsed with CD-HIT?

Page 4

Within crustacean taxa... Similarity was pretty high in Branchiopoda and Malacostraca (93% & 98%), but significantly lower in Copepoda. Can you propose an explanation for that?

It was previously reported that... In light of this, Veenstra (2016)² has some very intensive mining of CHH in multiple decapod species. Why not provide a section that compares between the number of CHH/MIH/GIH predicted from your dataset with the above study?

Figure 4 - Color coding system is misleading, for instance Decapoda was assigned as teal in legends, but actually blue in the phylogenetics tree.

References

1. Ohira T: Crustacean Hyperglycemic Hormone. 2016. 403-1 [Publisher Full Text](#)
2. Veenstra JA: Similarities between decapod and insect neuropeptidomes. *PeerJ*. 2016; **4**: e2043 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research