12-1-2020

Cloning of the first cDNA encoding a putative CCRFamide precursor: identification of the brain, eyestalk ganglia, and cardiac ganglion as sites of CCRFamide expression in the American lobster, Homarus americanus

J. Joe Hull
USDA ARS U.S. Arid-Land Agricultural Research Center

Melissa A. Stefanek
USDA ARS U.S. Arid-Land Agricultural Research Center

Patsy S. Dickinson
Bowdoin College

Andrew E. Christie
Pacific Biosciences Research Center

Invertebrate Neuroscience (2020) 20:24
https://doi.org/10.1007/s10158-020-00257-z

SHORT COMMUNICATION

Cloning of the first cDNA encoding a putative CCRFamide precursor: identification of the brain, eyestalk ganglia, and cardiac ganglion as sites of CCRFamide expression in the American lobster, *Homarus americanus*

J. Joe Hull[1] · Melissa A. Stefanek[1] · Patsy S. Dickinson[2] · Andrew E. Christie[3]

Received: 16 January 2020 / Accepted: 18 November 2020 / Published online: 26 November 2020

Abstract

Over the past decade, many new peptide families have been identified via in silico analyses of genomic and transcriptomic datasets. While various molecular and biochemical methods have confirmed the existence of some of these new groups, others remain in silico discoveries of computationally assembled sequences only. An example of the latter are the CCRFamides, named for the predicted presence of two pairs of disulfide bonded cysteine residues and an amidated arginine-phenylalanine carboxyl-terminus in family members, which have been identified from annelid, molluscan, and arthropod genomes/transcriptomes, but for which no precursor protein-encoding cDNAs have been cloned. Using routine transcriptome mining methods, we identified four *Homarus americanus* (American lobster) CCRFamide transcripts that share high sequence identity across the predicted open reading frames but more limited conservation in their 5′ terminal ends, suggesting the *Homarus* gene undergoes alternative splicing. RT-PCR profiling using primers designed to amplify an internal fragment common to all of the transcripts revealed expression in the supraoesophageal ganglion (brain), eyestalk ganglia, and cardiac ganglion. Variant specific profiling revealed a similar profile for variant 1, eyestalk ganglia specific expression of variant 2, and an absence of variant 3 expression in the cDNAs examined. The broad distribution of CCRFamide transcript expression in the *H. americanus* nervous system suggests a potential role as a locally released and/or circulating neuropeptide. This is the first report of the cloning of a CCRFamide-encoding cDNA from any species, and as such, provides the first non-in silico support for the existence of this invertebrate peptide family.

Keywords Decapoda · *Homarus americanus* · In silico transcriptome mining · RT-PCR · Nervous system

Introduction

Recent advances in high-throughput nucleotide sequencing and in the programs used for assembling the resultant data have led to an explosion in the availability and depth of genomic and transcriptomic datasets for many organisms. In silico analyses of these assemblies have provided evidence supporting the existence of a number of previously unknown peptide families (e.g., Christie 2014; Conzelmann et al. 2013; Dircksen et al. 2011; Veenstra et al. 2012). While the presence of some of these groups has been confirmed via molecular cloning, mass spectrometry, and/or biochemical isolation and sequencing (e.g., Bao et al. 2015; Dickinson et al. 2019), others remain in silico discoveries only.

The CCRFamides form one such family, which was originally identified via the mining of an annelid, *Platynereis*

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10158-020-00257-z) contains supplementary material, which is available to authorized users.

✉ J. Joe Hull
joe.hull@ars.usda.gov

[1] Pest Management and Biocontrol Research Unit, US Arid Land Agricultural Research Center, USDA Agricultural Research Services, 21881 North Cardon Lane, Maricopa, AZ 85138, USA

[2] Department of Biology, Bowdoin College, 6500 College Station, Brunswick, ME 04011, USA

[3] Békésy Laboratory of Neurobiology, Pacific Biosciences Research Center, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 1993 East-West Road, Honolulu, HI 96822, USA

dumerilii, transcriptome and named for the predicted presence of two disulfide bonded cysteine pairs and the carboxyl (C)-terminus –RFamide in the putative mature peptide (Conzelmann et al. 2013). In silico analyses suggest that the CCRFamide family is broadly conserved in invertebrates. In addition to the Annelida, genome and transcriptome mining has identified putative CCRFamide-like peptide-encoding sequences in multiple members of the Mollusca and Arthropoda (Conzelmann et al. 2013; Nguyen et al. 2018; Oliphant et al. 2018; Stewart et al. 2014; Veenstra 2010, 2015, 2016a, b). Despite their apparent broad phylogenetic conservation, no CCRFamide precursor has been cloned from any species, and no biochemical or mass spectral data documenting the existence of any putative mature CCRFamide peptide are published (Conzelmann et al. 2013; Nguyen et al. 2018; Oliphant et al. 2018; Stewart et al. 2014; Veenstra 2010, 2015, 2016a, b).

In the study presented here, we report the cloning of the first cDNA encoding a putative CCRFamide precursor. Using primers designed from a transcriptome-derived sequence, the complete open reading frame of an American lobster, *Homarus americanus*, CCRFamide precursor was cloned from supraoesophageal ganglion (brain) cDNAs. Using RT-PCR, we show that in addition to the brain, the prepro-CCRFamide is also expressed in at least two other portions of the lobster nervous system, specifically the eyestalk ganglia, the location of the neuroendocrine X-organ-sinus gland system (e.g., Christie 2011), and the cardiac ganglion (CG), which drives the rhythmic contractions of the heart (e.g., Cooke 2002). We further provide evidence consistent with tissue specific alternative splicing of the CCRFamide transcript.

Materials and methods

Transcriptome mining and prediction of mature peptide structure

In silico transcriptome mining

Searches for putative *H. americanus* CCRFamide-encoding transcripts were conducted using a well-established protocol (e.g., Christie et al. 2015, 2017). Specifically, the database of the online program tblastn (National Center for Biotechnology Information, Bethesda, MD; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was set to Transcriptome Shotgun Assembly (TSA) and restricted to data from *H. americanus* nervous system-specific transcriptomes, *i.e.*, BioProject Nos. PRJNA300643 (brain, abdominal nerve cord, CG and stomatogastric nervous system; Northcutt et al. 2016), PRJNA338672 (eyestalk ganglia-specific; Christie et al. 2017), PRJNA379629 (brain-specific; Christie et al. 2018a) and PRJNA412549 (CG-specific; Christie et al. 2018b). A CCRFamide precursor predicted from crayfish, *Procambarus clarkii,* transcriptomic data (Veenstra 2015) was used as the query sequence for all BLAST searches. In total, four transcriptomic reads (GFDA01062303, GEBG01038526, GFUC01020658 and GFUC01020659) were identified.

Peptide structural prediction

The putative mature structures of a *H. americanus* CCRFamide isoform and several CCRFamide precursor-related peptides were predicted using a well-established workflow (e.g. Christie et al. 2015, 2017). In brief, all hits returned by a given BLAST search were translated using the ExPASy Translate tool (http://web.expasy.org/translate/) and assessed for completeness. Proteins listed as full-length exhibit a start methionine and are flanked on their C-terminus by a stop codon, while those listed as C-terminal fragments are missing the start methionine. Next, each precursor was assessed for the presence of a signal peptide using SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/; Bendtsen et al. 2004). Prohormone cleavage sites were identified based on information presented in Veenstra (2000). The sulfation state of tyrosine residues was predicted using Sulfinator (https://www.expasy.org/sulfinator/; Monigatti et al. 2002), while disulfide bonding between cysteine residues was assessed using DiANNA (http://clavius.bc.edu/~clotelab/DiANNA/; Ferrè and Clote 2005). C-terminal amidation was predicted by homology to known arthropod peptides. Protein and peptide alignments were done using either the online program MAFFT version 7 (http://mafft.cbrc.jp/alignment/software/; Katoh and Standley 2013) or MUSCLE (Edgar 2004) as implemented in Geneious Prime 2020.1.2 (https://www.geneious.com).

Molecular cloning

Total RNAs were purified from freshly dissected individual *H. americanus* brains using TRI Reagent (Life Technologies, Carlsbad, CA) and Direct-zol RNA MiniPrep spin columns (Zymo Research, Irvine, CA, USA) as described in Christie et al. (2018a). First-strand cDNAs were generated from 500 ng of DNase I-treated total RNA using Superscript III Reverse Transcriptase (Life Technologies) and custom made random pentadecamers (IDT, San Diego, CA, USA). The complete open reading frame (ORF) for the putative *H. americanus* CCRFamide precursor was PCR amplified using primers (sense—5′-ATGGTGAGTCGTGTGGGC; antisense—5′-CTAGTCGAGGAGTAGTTCATAACA) designed to encompass the consensus predicted start and stop sites (771 bp) in conjunction with Sapphire Amp Fast PCR Master Mix (Takara Bio USA Inc., Mountain View, CA). PCR was performed on a Biometra TRIO multiblock thermocycler (Biometra,

Göttingen, Germany) in a 20-µL reaction volume containing 0.5 µL cDNA template and 0.2 µM of each primer. The thermocycler conditions consisted of: 95 °C for 2 min followed by 40 cycles at 95 °C for 20 s, 56 °C for 20 s, 72 °C for 30 s, and a final extension at 72 °C for 5 min. The resulting products were separated on a 1.5% agarose gel using a tris/acetate/EDTA buffer system with SYBR Safe (Life Technologies). The reaction products were subcloned into pCR2.1-TOPO TA (Life Technologies) and sequenced at the Arizona State University DNA Core Laboratory (Tempe, AZ).

RT-PCR expression profiling

To examine the expression of CCRFamide precursor-encoding transcripts, RNAs were purified from four biological replicates of *H. americanus* brain (one brain per replicate), eyestalk ganglia (one eyestalk pair per replicate), CG (ten ganglia per replicate), and cardiac muscle (~25 mg of cardiac muscle per replicate), as described in several recent publications (Christie et al. 2017, 2018a, b). cDNAs were generated as described above from ~100 ng (CG) or 500 ng (brain, eyestalk ganglia, and cardiac muscle) of DNase I-treated total RNAs. RT-PCR was performed in 20-µL reaction volumes containing either 0.5 µL of brain, eyestalk ganglia, or cardiac muscle cDNA templates or 1 µL of CG-derived cDNAs. Primers (sense—5′- GTCTTGGCC AATCACAGC and antisense—5′- GTGAGGACTTCGACG ACT) were designed to generate a 528-bp internal fragment (nt 64-591) common to all of the *H. americanus* CCRFamide precursor transcripts. To assess the expression of potential transcript splice variants, the consensus ORF antisense primer described above was used in conjunction with specific sense primers: variant 1 (GFDA01062303)—5′-ACTTGACTTTCG CACTTCTC; variant 2 (GEBG01038526)—5′-TGAAGGCGT GGGGGTGA; and variant 3 (GFUC01020659)—5′-CCTCTC ACTTCACAAATATTAC. Primers (sense—5′- GGTCGT ACCACCGGTATT and antisense—5′- CATCCTGTCGGC AATTCC) were also designed to amplify nt 439-942 of *H. americanus* actin (FJ217215). PCR was performed with each of the replicated cDNA sets and used thermocycler conditions identical to those described above. The resulting products were visualized on 1.5% agarose gels with images obtained using an AlphaImager Gel Documentation System (ProteinSimple, San Jose, CA) and processed in Photoshop CS6 v13.0 (Adobe Systems Inc., San Jose, CA).

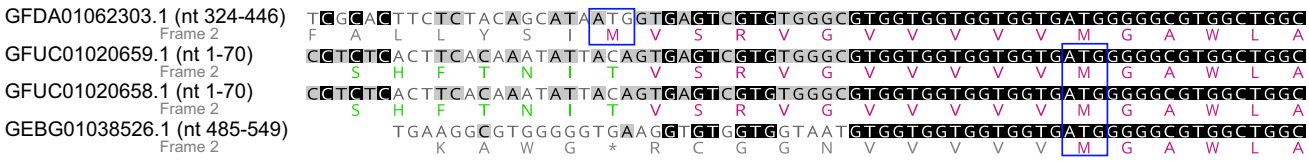
Results

In silico identification of *H. americanus* CCRFamide-encoding transcripts and prediction of CCRFamide precursor proteins and peptides

To identify putative *H. americanus* CCRFamide precursor-encoding transcripts, the sequence of a previously described crayfish, *P. clarkii*, prepro-CCRFamide (Veenstra 2015) was used to search nervous system-specific TSA datasets for sequences encoding putative homologs. The assemblies searched included one produced using multiple nervous system regions as the source of starting RNA (Northcutt et al. 2016), and three region-specific transcriptomes, *i.e.*, those for the eyestalk ganglia (Christie et al. 2017), the brain (Christie et al. 2018a), and the CG (Christie et al. 2018b). While no transcripts encoding putative CCRFamide precursors were found in the CG assembly, four were identified from the other transcriptomes, one each from the eyestalk ganglia and mixed nervous system region datasets (Accession Nos. GFDA01062303 and GEBG01038526, respectively) and two from the brain transcriptome (Accession Nos. GFUC01020658 and GFUC01020659). Although the transcripts are largely identical across a significant portion of the open reading frame (Fig. S1), the regions immediately upstream of the putative start codons differ (Fig. 1a). Only the GFDA01062303 transcript encodes a full-length precursor peptide with a predicted signal peptide; the other transcripts either lack a defined start site (i.e. first in-frame ATG downstream of a stop codon) or have an alternative start relative to GFDA01062303. This apparent transcript specific variation could indicate alternative splicing of the *H. americanus* CCRFamide gene or may reflect assembly issues that arise with repetitive sequences common to untranslated regions (Lima et al. 2017). Significant sequence variation specific to the 5′terminal ends of CCRFamide-encoding transcripts (GGLH01033349.1, GGLH01033499.1, GGLH01033500.1) in *Penaeus monodon* (giant tiger prawn), however, supports the presence of a conserved splicing mechanism.

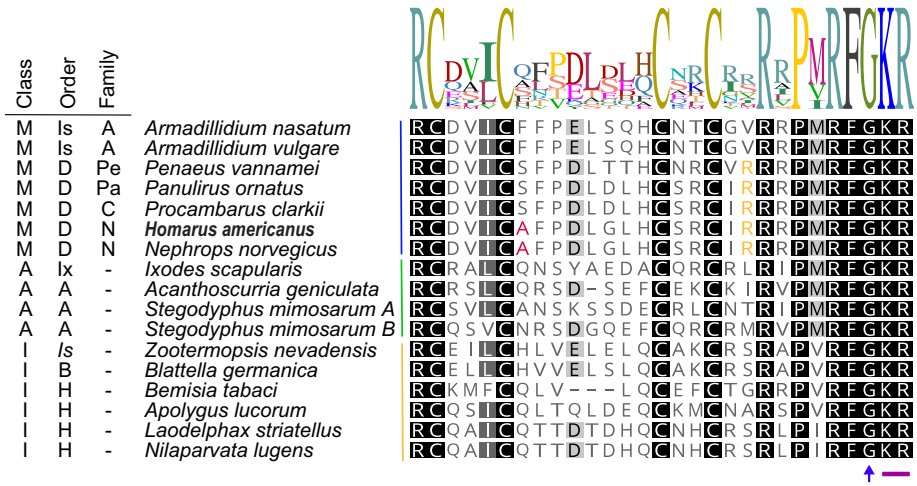
Translation of the identified transcripts revealed three distinct putative CCRFamide precursors: two full-length proteins, one from GFDA01062303 (*H. americanus* prepro-CCRFamide variant 1 [Homam-prepro-CCRFa-v1; 256 amino acids]) and the other from GEBG01038526 (*H. americanus* prepro-CCRFamide variant 2 [Homam-prepro-CCRFa-v2; 245 amino acids]), and one C-terminal fragment (*H. americanus* prepro-CCRFamide variant 3 [Homam-prepro-CCRFa-v3; 262 amino acids]), which was deduced from both GFUC01020658 and GFUC01020659. The three CCRFamide precursors differ only in the region

(A)



(B)

(C)



of their signal peptides and at one conservatively substituted residue (Fig. 1b). A total of six peptides were predicted from the collective set of preprohormones, including one isoform of CCRFamide (Table 1). The putative mature CCRFamide is 36 amino acids in length and is predicted to be C-terminally amidated with disulfide bridging between its first and fifth, and second and fourth cysteines (Table 1). A pair of sulfated tyrosine residues is predicted for CCRFamide precursor-related peptides 2a and 2b (Table 1), which are derived from Homam-prepro-CCRFa-v1/3 and Homam-prepro-CCRFa-v2, respectively (Fig. 1b).

Multiple sequence alignment of HomamCCRFa with CCRFamides from diverse arthropod species reveals

◂**Fig. 1** *Homarus americanus* CCRFamide precursor alignments. **a** Nucleotide sequence alignment of the region surrounding the predicted start sites in the four identified *H. americanus* CCRFamide precursor-encoding transcripts. Start codons are shown in blue boxes. Amino acids predicted to comprise a signal peptide are shown in pink, residues upstream of the predicted signal peptide are shown in green, and the untranslated region is shown in grey. **b** MAFFT alignment of putative *Homarus americanus* CCRFamide precursors deduced from nervous system-specific transcriptomic data. PreproCCRFa-v1 corresponds to GFDA01062303, preproCCRFa-v2 to GEBG01038526, and preproCCRFa-v3 to GFUC01020659. In each protein, the signal peptide is shown in pink, residues upstream of the predicted signal peptide are shown in green, the CCRFamide isoform is shown in red, precursor-related peptides are shown in blue, and consensus dibasic cleavage sites are shown in black. In the line immediately below each sequence grouping, the symbol "*" indicates amino acids that are identical in all three proteins, while ":" denotes amino acids that are similar in structure among all three sequences. **c** Multiple sequence alignment of arthropod CCRFamides. Sequences are based on conceptual translation of transcripts from species representing three Classes—Malacostraca (M, blue line), Arachnida (A, green line), and Insecta (I, yellow line). Atypical amino acid usages are shown in red (predominant Nephropidae-based) or orange (decapod specific). Abbreviations are: Order—Isopoda (Is), Decapoda (D), Ixodida (Ix), Araneae (A), Isoptera (*Is*), Blattodea (B), and Hemiptera (H); Family—Armadillidiidae (A), Penaeidae (Pe), Palinuridae (Pa), Cambaridae (C), and Nephropidae (N). The predicted amidation site is indicated by the arrowhead and the consensus C-terminal dibasic convertase cleavage site is indicated by the purple line. The *H. americanus* sequence is shown in bold. Sequence accession numbers are listed in Table S1

significant C-terminal conservation, with most peptides exhibiting a characteristic P-M/V/I-R-F-amide tetrapeptide motif and four equally spaced Cys residues (Fig. 1c). Interestingly, decapod specific atypical amino acid usage was observed, with a charged Arg residue at position −7 (relative to the C-terminus) rather than the polar or small hydrophobic residue present in species from other Orders. In addition, the two Nephropidae in our analysis were distinguished by an atypical Ala at position −20, which, in sequences other than the *Stegodyphus mimosarum* A (African social velvet spider) CCRFamide, is occupied by amino acids with larger sidechains (e.g. Phe, Ser, Gln, Asn, or His).

Cloning of a CCRFamide-encoding cDNA from the brain of *Homarus americanus*

While a number of studies, including the one presented here, have provided in silico evidence for the existence of CCRFamides in invertebrates (Conzelmann et al. 2013; Nguyen et al. 2018; Oliphant et al. 2018; Stewart et al. 2014; Veenstra 2010, 2015, 2016a, b), there has not yet been confirmation of the CCRFamide family using methodologies such as molecular cloning. To generate a complete ORF for the longest of the putative full-length *H. americanus* CCRFamide precursors, *i.e.*, Homam-prepro-CCRFa-v1, primers were designed based on the GFDA01062303 transcript and used to amplify products from brain cDNAs. Sanger sequencing of the resultant amplicons showed them to be > 99% identical in nucleotide sequence to that of the transcript used for primer design. The protein deduced from the cloned cDNA, which was submitted to GenBank under Accession No. MN661244, is identical in amino acid sequence to that of Homam-prepro-CCRFa-v1.

Expression profiling of CCRFamide expression in select portions of the *Homarus americanus* nervous system

Transcriptomic data suggest that the CCRFamide precursor is expressed in the eyestalk ganglia, in addition to the brain, of *H. americanus*. To confirm this expression, and to further investigate whether or not this peptide might also be present in the CG, an offshoot of the central nervous system that controls the rhythmic contractions of the heart (e.g., Cooke 2002), RT-PCR profiling of these portions of the nervous system was conducted; cardiac muscle was used as a negative control for these analyses. PCR products were consistently amplified from multiple brain ($n = 3$ of 4), eyestalk ganglia ($n = 3$ of 4), and CG ($n = 2$ of 4) cDNA samples (Fig. 2); no amplicon was generated from any of the cardiac muscle cDNAs ($n = 3$ independent samples). The inconsistent amplification from CG cDNAs could be an indication that the CCRFamide transcript is conditionally regulated in this tissue.

To assess the validity of the transcript assemblies and to determine if the putative transcript variants exhibit differential expression, sense primers were designed to the variant-specific sequences upstream of the respective start codons. The preproCCRFa-v1 primer consistently generated full-length ORF amplimers of the expected size (800 nt) from brain ($n = 3$ of 4) and eyestalk ganglia ($n = 3$ of 4) but was more limited in CG ($n = 1$ of 4). A preproCCRFa-v2 amplimer (785 nt) was only observed in eyestalk ganglia, albeit faintly ($n = 3$ of 4), and no preproCCRFa-v3 specific product (791 nt) was amplified from any of the cDNA sets (Fig. 2). As with the 500-bp CCRFamide fragment in the CG, the inconsistent amplification of the specific variants could indicate condition-based transcriptional activity.

Discussion

In the study presented here, a cDNA for a lobster, *H. americanus*, CCRFamide precursor is reported. This is the first cloning of cDNA encoding a prepro-CCRFamide from any species, and thus represents the only non-in silico evidence for the existence of this invertebrate peptide family. RT-PCR profiling suggests that CCRFamide is broadly expressed in the *H. americanus* nervous system,

Table 1 Peptides predicted from putative *Homarus americanus* CCRFamide precursors

Peptide	Structure
CCRFa	NHSQCDVQALRCDVICAFPDLGLHCSRCIRRRPMRFa
CCRFa-PRP-1	SGEPPRTLSVAATGSPLVAVISHPGPTHLQKTKSQDPGQSGLASSTVQY
CCRFa-PRP-2a	TSGVHDAARPVVSSSLRQLINKVLWNEYTARKTLPSGDRQQEAY-SSQLGRYVSDLQSELGPDY$_{(SO3H)}$DDSTEGVEY$_{(SO3H)}$VSVMRS
CCRFa-PRP-2b	TSGVHDAARPVVSSSLRQLINKVLWNEYTARKTLPSGDRQQEAY-SSQLGRYISDLQSELGPDY$_{(SO3H)}$DDSTEGVEY$_{(SO3H)}$VSVMRS
CCRFa-PRP-3	SPHHTLDNLGSGGVQSDGRFPVE
CCRFa-PRP-4	SGVHEVIEGLMDLLGLDELPVDPTLYGCYELLLD

Numbering of CCRFamide precursor-related peptides is based on their order within the preprohormone

CCRFamide precursor-related peptide 2a was predicted from both prepro-CCRFamide variants 1 and 3, while CCRFamide precursor-related peptide 2b was predicted from prepro-CCRFamide variant 2; the two isoforms differ from one another in valine vs. isoleucine at position 52

In the putative mature CCRFamide isoform, disulfide bridging is predicted between the first and fifth and second and fourth cysteine residues

Peptide abbreviations: CCRFa, CCRFamide; CCRFa-PRP, CCRFamide precursor-related peptide

Abbreviations in peptide structures: Y$_{(SO3H)}$, sulfated tyrosine; C, cysteine residue participating in a disulfide bond; a, carboxyl-terminal amide group

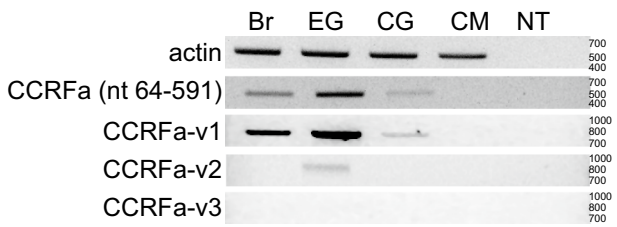


Fig. 2 PCR-based expression profiling of CCRFamide-encoding transcripts in select portions of the *Homarus americanus* nervous system and heart muscle. Primers designed to amplify an internal 528-bp fragment of the CCRFamide precursor as well as variant specific transcripts encompassing the complete ORF were used with supraoesophageal ganglion (brain; Br), eyestalk ganglia (EG), cardiac ganglion (CG), and cardiac muscle (CM) cDNAs. As a control for cDNA integrity, a 503-bp fragment of actin was likewise amplified. Reactions were also performed in the absence of a template (no template; NT). The gel images shown are representative of positive amplification in at least one of four biological replicates

including in the brain, eyestalk ganglia, and CG. While the function(s) of CCRFamide remain unknown, it is highly likely that it serves as a locally released and/or circulating neuromodulator in the lobster. Detection of CCRFamide-encoding transcripts in the eyestalk ganglia suggests that it might be present in the X-organ-sinus gland system, a major neuroendocrine control center in decapod species (Christie 2011). If so, the peptide could be delivered via the circulatory system to a diverse array of potential targets. The presence of CCRFamide in the CG suggests a potential role in modulating cardiac output, a role played by a number of –RFamide-containing peptides in *H. americanus* (Dickinson et al. 2015, 2016; Stevens et al. 2009).

As with the biological role and target tissue, the cognate CCRFamide receptor has yet to be elucidated. Although most regulatory peptides and neuromodulators signal through class A G protein-coupled receptors (GPCRs) (Tikhonova et al. 2019), disulfide-bridged peptides like CCRFamide frequently act on other signaling systems such as insulin-like peptides and prothoracicotropic hormone (PTTH) with tyrosine kinases (Nässel and Zandawala 2019), eclosion hormone with a guanylate cyclase (Nässel and Zandawala 2019), or, as proposed for the agatoxin-like peptides, with ion channels (Sturm et al. 2016). Thus, specific identification of the CCRFamide receptor is potentially complicated by the breadth of potential molecular targets. Further complicating receptor identification and peptide functionality is the fact that the CCRFamide gene has been lost in a number of the more genetically malleable model organisms (e.g. *Drosophila melanogaster*) typically employed for functional studies (Veenstra 2019). However, given that the neural circuit contained within the CG, the neuromuscular junction between the ganglion and cardiac muscle, and the muscle itself are all potential targets for CCRFamide modulation (e.g., Dickinson et al. 2015, 2016; Fort et al. 2007a, b; Stevens et al. 2009), additional transcriptomic studies may provide insights into CCRFamide function and receptor elucidation.

Acknowledgements Lisa Baldwin is thanked for reading and editing an earlier version of this article. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture; the USDA is an equal opportunity provider and employer.

Funding This work was supported by funds from the National Science Foundation (IOS-1353023; IOS-1354567; IOS-1856307;

IOS-1856433), the National Institutes of Health (INBRE Grant 8P20GM103423-12), the Cades Foundation, and base CRIS funding from the US Department of Agriculture (2020-22620-022-00D).

Compliance with ethical standards

Conflict of interest The authors declared that they have no conflict of interest statement.

References

Bao C, Yang Y, Huang H, Ye H (2015) Neuropeptides in the cerebral ganglia of the mud crab, *Scylla paramamosain*: transcriptomic analysis and expression profiles during vitellogenesis. Sci Rep 5:17055

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: signalP 3.0. J Mol Biol 340:783–795

Christie AE (2011) Crustacean neuroendocrine systems and their signaling agents. Cell Tissue Res 345:41–67

Christie AE (2014) Prediction of the peptidomes of *Tigriopus californicus* and *Lepeophtheirus salmonis* (Copepoda, Crustacea). Gen Comp Endocrinol 201:87–106

Christie AE, Chi M, Lameyer TJ, Pascual MG, Shea DN, Stanhope ME, Schulz DJ, Dickinson PS (2015) Neuropeptidergic signaling in the American lobster *Homarus americanus*: new insights from high-throughput nucleotide sequencing. PLoS ONE 10:e0145964

Christie AE, Roncalli V, Cieslak MC, Pascual MG, Yu A, Lameyer TJ, Stanhope ME, Dickinson PS (2017) Prediction of a neuropeptidome for the eyestalk ganglia of the lobster *Homarus americanus* using a tissue-specific *de novo* assembled transcriptome. Gen Comp Endocrinol 243:96–119

Christie AE, Yu A, Pascual MG, Roncalli V, Cieslak MC, Warner AN, Lameyer TJ, Stanhope ME, Dickinson PS, Hull JJ (2018a) Circadian signaling in *Homarus americanus*: region-specific *de novo* assembled transcriptomes show that both the brain and eyestalk ganglia possess the molecular components of a putative clock system. Mar Genom 40:25–44

Christie AE, Yu A, Roncalli V, Pascual MG, Cieslak MC, Warner AN, Lameyer TJ, Stanhope ME, Dickinson PS, Hull JJ (2018b) Molecular evidence for an intrinsic circadian pacemaker in the cardiac ganglion of the American lobster, *Homarus americanus*—Is diel cycling of heartbeat frequency controlled by a peripheral clock system? Mar Genom 41:19–30

Conzelmann M, Williams EA, Krug K, Franz-Wachtel M, Macek B, Jékely G (2013) The neuropeptide complement of the marine annelid *Platynereis dumerilii*. BMC Genom 14:906

Cooke IM (2002) Reliable, responsive pacemaking and pattern generation with minimal cell numbers: the crustacean cardiac ganglion. Biol Bull 202:108–136

Dickinson PS, Calkins A, Stevens JS (2015) Related neuropeptides use different balances of unitary mechanisms to modulate the cardiac neuromuscular system in the American lobster, *Homarus americanus*. J Neurophysiol 113:856–870

Dickinson ES, Johnson AS, Ellers O, Dickinson PS (2016) Forces generated during stretch in the heart of the lobster *Homarus americanus* are anisotropic and are altered by neuromodulators. J Exp Biol 219:1187–1202

Dickinson PS, Dickinson ES, Oleisky ER, Rivera CD, Stanhope ME, Stemmler EA, Hull JJ, Christie AE (2019) AMGSEFLamide, a member of a broadly conserved peptide family, modulates multiple neural networks in *Homarus americanus*. J Exp Biol 222:jeb194092

Dircksen H, Neupert S, Predel R, Verleyen P, Huybrechts J, Strauss J, Hauser F, Stafflinger E, Schneider M, Pauwels K, Schoofs L, Grimmelikhuijzen CJ (2011) Genomics, transcriptomics, and peptidomics of *Daphnia pulex* neuropeptides and protein hormones. J Proteome Res 10:4478–4504

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Ferrè F, Clote P (2005) DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res 33:W230–W232

Fort TJ, Brezina V, Miller MW (2007a) Regulation of the crab heartbeat by FMRFamide-like peptides: multiple interacting effects on center and periphery. J Neurophysiol 98:2887–2902

Fort TJ, García-Crescioni K, Agricola HJ, Brezina V, Miller MW (2007b) Regulation of the crab heartbeat by crustacean cardioactive peptide (CCAP): central and peripheral actions. J Neurophysiol 97:3407–3420

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Lima L, Sinaimeri B, Sacomoto G, Lopez-Maestre H, Marchet C, Miele V, Sagot MF, Lacroix V (2017) Playing hide and seek with repeats in local and global de novo transcriptome assembly of short RNA-seq reads. Algorithms Mol Biol 12:1–19

Monigatti F, Gasteiger E, Bairoch A, Jung E (2002) The Sulfinator: predicting tyrosine sulfation sites in protein sequences. Bioinformatics 18:769–770

Nässel DR, Zandawala M (2019) Recent advances in neuropeptide signaling in Drosophila, from genes to physiology and behavior. Prog Neurobiol 179:101607

Nguyen TV, Rotllant GE, Cummins SF, Elizur A, Ventura T (2018) Insights into sexual maturation and reproduction in the Norway lobster (*Nephrops norvegicus*) via in silico prediction and characterization of neuropeptides and G protein-coupled receptors. Front Endocrinol (Lausanne) 9:430

Northcutt AJ, Lett KM, Garcia VB, Diester CM, Lane BJ, Marder E, Schulz DJ (2016) Deep sequencing of transcriptomes from the nervous systems of two decapod crustaceans to characterize genes important for neural circuit function and modulation. BMC Genom 17:868

Oliphant A, Alexander JL, Swain MT, Webster SG, Wilcockson DC (2018) Transcriptomic analysis of crustacean neuropeptide signaling during the moult cycle in the green shore crab, *Carcinus maenas*. BMC Genom 19:711

Stevens JS, Cashman CR, Smith CM, Beale KM, Towle DW, Christie AE, Dickinson PS (2009) The peptide hormone pQDLDHVFLRFamide (crustacean myosuppressin) modulates the *Homarus americanus* cardiac neuromuscular system at multiple sites. J Exp Biol 212:3961–3976

Stewart MJ, Favrel P, Rotgans BA, Wang T, Zhao M, Sohail M, O'Connor WA, Elizur A, Henry J, Cummins SF (2014) Neuropeptides encoded by the genomes of the Akoya pearl oyster *Pinctata fucata* and Pacific oyster *Crassostrea gigas*: a bioinformatic and peptidomic survey. BMC Genom 15:840

Sturm S, Ramesh D, Brockmann A, Neupert S, Predel R (2016) Agatoxin-like peptides in the neuroendocrine system of the honey bee and other insects. J Proteom 132:77–84

Tikhonova IG, Gigoux V, Fourmy D (2019) Understanding peptide binding in class A G protein-coupled receptors. Mol Pharm 96:550–561

Veenstra JA (2000) Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch Insect Biochem Physiol 43:49–63

Veenstra JA (2010) Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. Gen Comp Endocrinol 167:86–103

Veenstra JA (2015) The power of next-generation sequencing as illustrated by the neuropeptidome of the crayfish *Procambarus clarkii*. Gen Comp Endocrinol 224:84–95

Veenstra JA (2016a) Neuropeptide evolution: chelicerate neurohormone and neuropeptide genes may reflect one or more whole genome duplications. Gen Comp Endocrinol 229:41–55

Veenstra JA (2016b) Similarities between decapod and insect neuropeptidomes. PeerJ 4:e2043

Veenstra JA (2019) Coleoptera genome and transcriptome sequences reveal numerous differences in neuropeptide signaling between species. PeerJ 7:e7144

Veenstra JA, Rombauts S, Grbić M (2012) In silico cloning of genes encoding neuropeptides, neurohormones and their putative G-protein coupled receptors in a spider mite. Insect Biochem Mol Biol 42:277–295