RESEARCH ARTICLE

Functional Authentication of a Novel Gastropod Gonadotropin-Releasing Hormone Receptor Reveals Unusual Features and Evolutionary Insight

Scott I. Kavanaugh, Pei-San Tsai*

Department of Integrative Physiology and Center for Neuroscience, University of Colorado, Boulder, Colorado, United States of America

* pei-san.tsai@colorado.edu

Abstract

A gonadotropin-releasing hormone (GnRH)-like molecule was previously identified in a gastropod, Aplysia californica, and named ap-GnRH. In this study, we cloned the full-length cDNA of a putative ap-GnRH receptor (ap-GnRHR) and functionally authenticated this receptor as a bona fide ap-GnRHR. This receptor contains two potential translation start sites, each accompanied by a Kozak sequence, suggesting the translation of a long and a short form of the receptor is possible. The putative ap-GnRHR maintains the conserved structural motifs of GnRHR-like receptors and shares 45% sequence identity with the octopus GnRHR. The expression of the putative ap-GnRHR short form is ubiquitous in all tissues examined, whereas the long form is only expressed in parts of the central nervous system, osphradium, small hermaphroditic duct, and ovotestis. The cDNA encoding the long or the short receptor was transfected into the Drosophila S2 cell line and subject to a radioreceptor assay using ¹²⁵I-labeled ap-GnRH as the radioligand. Further, the transfected cells were treated with various concentrations of ap-GnRH and measured for the accumulation of cAMP and inositol monophosphate (IP1). Radioreceptor assay revealed that only the long receptor bound specifically to the radioligand. Further, only the long receptor responded to ap-GnRH with an increased accumulation of IP1, but not cAMP. Our studies show that despite the more prevalent expression of the short receptor, only the long receptor is the functional ap-GnRHR. Importantly, this is only the second report on the authentication of a protostome GnRHR, and based on the function and the phylogenetic grouping of ap-GnRHR, we suggest that this receptor is more similar to protostome corazonin receptors than chordate GnRHRs.



Citation: Kavanaugh SI, Tsai P-S (2016) Functional Authentication of a Novel Gastropod Gonadotropin-Releasing Hormone Receptor Reveals Unusual Features and Evolutionary Insight. PLoS ONE 11(7): e0160292. doi:10.1371/journal.pone.0160292

Editor: Christian Wegener, University of Würzburg, GERMANY

Received: June 13, 2016

Accepted: July 15, 2016

Published: July 28, 2016

Copyright: © 2016 Kavanaugh, Tsai. This is an open access article distributed under the terms of the <u>Creative Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by National Science Foundation Grant IOS 1352944 (<u>http://www.nsf.gov/awardsearch/showAward?AWD_ID=1352944</u>) to PST. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Gonadotropin-releasing hormone (GnRH) is a member of an ancient peptide superfamily whose origin predated the protostome-deuterostome split [1, 2]. In vertebrates, the most prominent role of GnRH is the stimulation of reproductive maturation and fertility. Orthologs of vertebrate GnRH have recently been discovered in protostomes, but their physiological roles are not clearly defined [3-5]. Physiological studies using native GnRH in molluscan models suggest protostome forms of GnRH have become highly diversified and are multi-functional molecules responsible for a wide range of regulation including neural control [6], motor activity [6, 7], feeding [7], and reproduction [8–10]. Further, the presence of additional molecules that are orthologous to protostome GnRH, including adipokinetic hormone (AKH), corazonin (CRZ), and AKH/CRZ-related peptide (ACP), have been documented [11, 12]. These peptides share common ancestry with protostome GnRH [2, 11, 13] and are present in diverse invertebrate phyla to serve taxon-specific functions [11].

Recent data mining studies suggest that protostome GnRH, AKH, CRZ, and ACP bind to receptors that are structurally related to each other, and these receptor may be difficult to distinguish functionally based on sequence homology alone [2, 11]. For example, several protostome receptors identified through cloning or transcriptome analysis have been initially named GnRH receptor (GnRHR) based on their sequence homology to vertebrate GnRHRs, but one was later found to be an AKH receptor [14, 15] and another was found to be more likely to bind to AKH than protostome GnRH [16]. In an era of whole genome and transcriptome mining that facilitates the identification of molecules based on homology, the functional deorphanization of these receptors becomes the next essential step in conferring the biological importance of the ligand-receptor pair.

In 2008, a GnRH-like molecule was isolated from a gastropod mollusk, *Aplysia californica*, and named ap-GnRH [<u>11</u>, <u>17</u>]. ap-GnRH was an unambiguous ortholog of octopus GnRH (oct-GnRH) [<u>18</u>] and was produced exclusively in the central nervous system (CNS) to modulate neural activity and motor behavior without stimulating reproduction [<u>6</u>, <u>7</u>]. Despite the extensive functional characterization of ap-GnRH [<u>5</u>], the receptor for ap-GnRH (ap-GnRHR) has not been identified. The presence of a related *Aplysia* AKH (ap-AKH) [<u>12</u>], which may possess a receptor structurally similar to ap-GnRHR, also posed a challenge to the search for ap-GnRHR.

The goal of the present study was to isolate the full-length sequence of the putative ap-GnRHR and authenticate it as a functional GnRHR in *A. californica*. To date, only one protostome GnRHR, octopus GnRHR (oct-GnRHR), has been functionally authenticated [19], illustrating the paucity of such studies. Towards this goal, we first analyzed expressed sequence tags and whole genome shotgun sequences of *A. californica* to identify partial sequences for a putative ap-GnRHR. Based on these sequences, we cloned a full-length G-protein-coupled receptor (GPCR) and validated it as a bona fide ap-GnRHR by a radioreceptor assay and the activation of the inositol phosphate (IP) pathway. Interestingly, at least two isoforms of ap-GnRHR transcript were identified, and the ubiquitously expressed form failed to produce a functional ap-GnRHR, further highlighting the importance of receptor validation.

Materials and Methods

Animals

Wild-caught *A. californica* were purchased from Alacrity Marine Biological Services (Redondo Beach, CA) and maintained in an aquarium with recirculating artificial sea water filtered through biological and chemical filters and maintained at 18–20°C. Animals were fed a diet of

Romaine lettuce daily. *A. californica* were anesthetized by an injection of 1/2 body mass of icecold isotonic MgCl₂ before sacrifice. No institutional approval was required for the use of invertebrates such as *A. californica*, but the authors adhered to standards of humane treatment to minimize pain, stress, and suffering of experimental animals.

Isolation of total RNA

Total RNA was isolated from *A. californica* central and peripheral tissues using the Trizol reagent (Invitrogen, Carlsbad, CA) and stored at -80°C until use.

Cloning of ap-GnRHR

Initial *in silico* analyses of *A. californica* expressed sequence tags and whole genome shotgun sequences were performed to identify trochozoan GnRHRs or GnRHR-like receptors. Four partial *A. californica* sequences were identified based on their sequence similarity with the N-terminus of oct-GnRHR. Based on these partial sequences, we designed primers to amplify a larger segment of the target gene and subsequently isolated its 3' end using the 3' rapid amplification of cDNA ends (RACE). Once the complete 3' was identified, primers were designed to capture the 5' end of the sequence using 5' RACE.

Both 5' and 3' RACE procedures were performed using the GeneRacer kit (Invitrogen, Grand Island, NY) according to manufacturer's instructions. Abdominal ganglia total RNA was used to generate the cDNA template for 5' and 3' RACE. The 3' RACE was accomplished using AcGnRHR 13F (5'-CCC TCT ACT TGT CAA CGT ACA-3') as the gene-specific forward primer and GeneRacer 3' Nested Primer (5'-CGC TAC GTA ACG GCA TGA CAG TG-3') as the reverse primer. The 5' RACE utilized GeneRacer 5' Nested Primer (5'-GGA CAC TGA CAT GGA CTG AAG GAG TA-3') as the forward primer and AcGnRHR 21R (5'-TTC TTC TCA TGG GAT CAA GAA TTG-3') as the gene-specific reverse primer. Amplification was performed using 2.5 units of Platinum Pfx DNA Polymerase (Invitrogen), 1X Pfx amplification buffer, 0.3 mM dNTP, 1 mM MgSO₄, and varying amounts of the cDNA template in 50 µl reactions. Cycling conditions included initial melting at 95°C for 1 min, 35 cycles of amplification (45 s at 95°C, 30 s at 55°C, 1 min at 72°C), followed by a 10-min final extension at 72°C. The reactions were then A-tailed by incubating the PCR products with 10 µl 1 mM dATP and 2.5 units BioReady rTaq DNA Polymerase (BullDog Bio, Portsmouth, NH) at 72°C for 20 minutes. Resulting amplicons were cloned into pGEM T-Easy (Promega, Madison, WI) and sequenced at Sequetech Corporation (Mountain View, CA). The overlapping sequences were combined and used to design primers that encompass the entire open-reading frame (ORF). Because our initial analysis suggested this receptor might have two translational start sites, we designed primers to amplify both the long and the short receptor isoforms for in vitro functional studies. For this, the abdominal ganglia cDNA was amplified with the forward primers AcGnRHR Full F1 (5'-CAG GAT TCA GTC CCA TGG ACG GAA CAG-3') (for the long receptor) and AcGnRHR Full F2 (5'-CGA CGG CTC CAA ACA TGA TCT CCA AC-3') (for the short receptor) in conjunction with the reverse primer AcGnRHR Full R2 (5'-GGA GGG GTA TTT TTA GAT GAA AGG AGG-3'). Amplification conditions were identical to the RACE PCR described above. Resulting amplicons were cloned into pGEM T-Easy and verified by sequencing.

Reverse-transcription polymerase chain reaction (RT-PCR) analysis of ap-GnRHR expression

Total RNA from bag cell neurons, abdominal ganglia, cerebral ganglia, pedal/pleural ganglia, buccal ganglia, tail, osphradium, small hermaphroditic duct, heart and ovotestis was harvested

from individual animals, pre-treated with RQ1 DNase, reverse transcribed using Superscript III first-strand cDNA synthesis kit, and amplified with one of the following two sets of genespecific primers: (1) AcGnRHR Full F1 as the forward primer and AcGnRHR 12R (5'-GTA CTT GAC GAA CTT GCA CAT GAC-3') as the reverse primer, and (2) AcGnRHR 12F (5'-CAA CCA CGA GTC TAA TAA AGA CGA-3') as the forward primer and AcGnRHR 12R as the reverse primer. AcGnRHR Full F1 encompassed the first start codon of the putative ap-GnRHR, whereas AcGnRHR Full F1 encompassed the first start codon of the putative ap-GnRHR, whereas AcGnRHR 12F targeted transmembrane domain (TM) 1, and AcGnRHR 12R targeted the end of extracellular loop (EL) 2 and beginning of TM 3. The first set of primers should amplify only the long receptor isoform, whereas the second set of primers would amplify both the long and the short isoforms. Amplification was performed in 25 µl-reactions using BioReady rTaq DNA Polymerase. As a control to ensure the quality of RT, *A californica* actin (ap-Actin) was amplified for all tissues examined using gene-specific primers as described previously [12, 17]. The RT-PCR was repeated 3 times to ensure consistency.

Phylogenetic analysis

The alignment used to generate the maximum likelihood tree consisted of 334 amino acid sequences retrieved from GenBank and the deduced ap-GnRHR amino acid sequence (GenBank accession # AHE78444.1). The alignment was produced using ClustalW in MEGA6 [20], analyzed, and the tree constructed using maximum likelihood method based on the JTT matrix-based model [21] with 1,000 bootstrap replicates.

Construct development, cell culture, and transfection

The ap-GnRHR ORFs previously subcloned into pGEM T-Easy vector were used as PCR templates to generate the coding regions of both receptor isoforms for *in vitro* receptor expression. For this, PCR was performed using Platinum Pfx DNA Polymerase with the following primers linked to selective restriction sites: AcSal1ResR (5'-GGT ATT TGT CGA CTT AGA TGA-3') combined with AcGnRHR-L-XbalF (5'-CAG GAT CTA CAC CCA TGG ACG G-3') to generate the long receptor, or AcSal1ResR combined with AcGnRHR-S-XbalF (5'- GAC GGC TCT AGA CAT GAT CTC C-3') to generate the short receptor. The PCR conditions were identical to RACE. Resulting amplicons were subcloned into pGEM T-Easy, and upon verification by sequencing, digested with XbaI and SalI and subcloned into the pAWG *Drosophila* expression vector, which expressed green fluorescent protein (GFP) as a marker (Drosophila Genomics Resource Center, Bloomington, IN). pAWG plasmid containing the long or the short form of the receptor was named pAWG-L or pAWG-S, respectively.

Drosophila melanogaster S2 cells (Drosophila Genomics Resource Center, Stock #6) were cultured in Schneider's Insect Medium (Sigma, Milwaukee, WI) containing 10% charcoal-stripped fetal bovine serum (FBS) (BioWorld, Dublin, Ohio) and propagated according to the supplier's instructions. For the transfection experiment, S2 cells were seeded in 24 well plates at the density of 10^5 cells/well with 1 ml culture medium and incubated overnight. On Day 2, the culture medium was removed, and the attached cells were incubated with a transfection solution (0.12M CaCl₂, 25mM HEPES, 0.75mM Na₂HPO₄, 140 mM NaCl, pH 7.1) containing 0.5 µg/ml pAWG, pAWG-L, or pAWG-S for 17 h before replacing the transfection solution with the regular culture medium. All functional studies were performed 4-days after transfection, a time point of optimal GFP signal.

Radioreceptor assay

Competitive binding properties of ap-GnRH and ap-AKH to the long and short isoforms of ap-GnRHR were examined using S2 cells transfected with pAWG-L or pAWG-S. Synthetic

ap-GnRH (pQNYHFSNGWYA-amide; GenScript, Piscataway, NJ) was radioiodinated and purified as described previously [$\underline{6}$, $\underline{7}$]. Transfected S2 cells were treated with 10 nM [125 I]-ap-GnRH without or with increasing concentrations (10^{-12} M to 10^{-6} M) of unlabeled ap-GnRH or ap-AKH (pQIHFSPDWGT-NH₂; GenScript) in an assay buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, pH 7.4). The reaction (200 µl total volume) was incubated for 3.5 h on ice in a 4°C refrigerator. Following, cells were gently washed with the assay buffer, collected in a solubilizing solution of 0.5M NaOH /1% SDS, and counted on a Packard Cobra II gamma counter. All samples were run in triplicates in 3 independent experiments.

Inositol 1-phosphate and cAMP assays

The accumulation of inositol 1-phosphate (IP1) was measured by the IP-One ELISA (Cisbio, Bedford, MA) according to the manufacturer's instructions. S2 cells transfected as before with pAWG, pAWG-L or pAWG-S were treated without or with 10^{-12} M to 10^{-6} M of Ap-GnRH in a stimulation buffer provided by the kit for 3 h at 28°C. All treatments were performed in triplicates in 3 independent experiments. Proctolin (RYLPT; GenScript) was used at 10^{-6} M as a positive control for the assay. Proctolin receptor is endogenously expressed in S2 cells and its activation coupled to the activation of the IP pathway [22].

The accumulation of cAMP was measured by the Cyclic AMP Competitive ELISA Kit (ThermoFisher Scientific, Waltham, MA) using the acetylated protocol according to the manufacturer's instructions. Transfected S2 cells were treated without or with 10^{-12} M to 10^{-6} M of Ap-GnRH for 1 hr at 28°C. All treatments were performed in triplicates in 3 independent experiments. Forskolin (5 μ M) was used to show maximum stimulation, and 3-Isobutyl-1-methylxanthine (1 mM) was used to prevent cAMP degradation.

Data analysis

When suitable, data were plotted and EC_{50} calculated using four-parameters logistics-based non-linear regression curve fit feature in the GraphPad Prism5 software (GraphPad, San Diego, CA). Linear regression was used in most cases when a dose effect was not observed.

Results

The full-length ap-GnRHR spans 1531 bp and contains two potential translation start sites, each accompanied by a Kozak consensus sequence [23] (Fig 1). The first start codon generates a 1419-bp ORF encoding a longer 473-amino acid GPCR (ap-GnRHR-L). The second start codon generates a 1191-bp ORF encoding a shorter 397-amino acid GPCR (ap-GnRHR-S). A truncated polyadenylation signal AATA [24] is located 7 bases upstream of the poly (A) tail. The general structural framework predicts this receptor as a Class A GPCR.

The deduced amino acid sequence of ap-GnRHR-L was compared with the related GPCRs from representative taxa (Fig 2). Comparison of the entire receptors, including the N-terminus, revealed that this receptor exhibited 40% identity with *Manduca sexta* (tobacco hornworm) CRZ receptor (Ms-CRZR), 45% with *Octopus vulgaris* (common octopus) GnRHR (oct-GnRHR), 27% with *Ciona intestinalis* (sea squirt) GnRHRI (Ci-GnRHRI), 26% with Ci-GnRHRII, 31% with *Petromyzon marinus* (sea lamprey) GnRHR (Pm-GnRHR) and 28% with *Mus musculus* (house mouse) GnRHR (Mm-GnRHR). The cloned receptor has 8 potential N-glycosylation sites (N-X-^S/_T) at positions 33, 38, 64, 71, 81, 90, 96, and 352, with the first 7 sites clustered in the extracellular N-terminus. The receptor also has 6 potential protein kinase A phosphorylation sites at positions 50, 303, 344, and 450. As with most GnRH receptors, a conserved DRXXX^I/_V motif was seen in the region connecting TM 3 and intracellular

$ \begin{array}{c} 15 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0$					gac	cac	tga	cat	gga	ctg	aag	gag	rtat	ctt	aaa	cat	gat	tga	ıgga	ttc	ata	ato	gtt	aca	gga	tto	agt	ccc	7
$ \begin{array}{c} \mathbf{M} \mathbf{D} \mathbf{G} \mathbf{T} \mathbf{D} \mathbf{S} \mathbf{D} \mathbf{G} \mathbf{S} \mathbf{I} \mathbf{L} \mathbf{Q} \mathbf{R} \mathbf{A} \mathbf{N} \mathbf{Q} \mathbf{A} \mathbf{F} \mathbf{L} \mathbf{A} \mathbf{S} \mathbf{S} \mathbf{G} \mathbf{F} \mathbf{L} \mathbf{Q} \mathbf{L} \mathbf{Z} \\ \begin{array}{c} \mathbf{c} \mathbf{c} \mathbf{c} \mathbf{c} \mathbf{c} \mathbf{c} \mathbf{c} \mathbf$	atg	gac	ada	aca	gac	ago	gat	aaa	rage	ata	ctg	caa	cgo	gca	aat	cag	get	ttt	ctg	gct	tca	agt	tet	add	cct	ttg	rcaa	ctg	15
catatacoctacaattitotocgacaatactagtocaagaaaacgtocatgagatgactatcaccaagagtocatgatcactaccaagaatacgtocatgatcactaccaagaatgacgtocaagagtocaagactatagtocaagactactaccaagaatgacgtocaaactacaacaccacactototaatgtgacgtotgacaacaccaaactacaactacaagaatgacgtotaaaaaggacgactaatggacgaacgacaagacgtocaaaactacaagagtocaaaactacaagagtocaaaactacaagagtocaaaactacaagagtagaacgtotgacaaactagaggacgactaagtgacgacgacggacgacggacg	м	D	G	т	D	s	D	G	s	I	L	Q	R	A	N	Q	A	F	L	A	s	s	s	G	P	L	Q	L	2
P I P Y N F S A Q N S S V H R K R H D V Y S P R S A H H S gatttcattttagottcogacaatatctcgtcacttctcctcacaccgacggctcaaacatgatccacacaca	cca	ata	ccc	tac	aat	tte	tec	act	cad	aac	tca	tca	ato	cac	aga	aaa	cat	cat	gac	ate	tat	tcc	cca	agg	tet	act	cat	cac	24
gattoatttagottocgacaatatotogtoatottotototaacagoggotoaaactgatatocaacaatatototto 32 D F I L A S D N I S S L L N T T A P N M I S N N S S F 33 gagttocagtotaacaggacacaccottotaatgtgacgtotgacgacaacottaacaacaacaactatataaa 40 E F Q F L N T T H P S N V T S E N A P T F T T S L I K 11 acquategitticggaattatgttoggottottgtogtggaatatggogactatagtgogagtgagaggotggoggotgg T I V F G I M F G I S F V G N M A T I V Q M R R L R R R 14 amagocacaacaactototogtaacotggocattggocgattggoggatgagaggtotggogggggggggg	P	I	P	Y	N	F	s	A	8	N	s	s	v	н	R	к	R	н	D	v	Y	s	P	R	s	A	н	н	5
D F I L A S D N I S S L L L N T T A P N H I S N N S S F gagttocagtcotaaacacgacaccottotaatgtgacgtotgaaaacgogocaacottocacacaccacgagtotaataaaaq D F I L A S D N I S S L L L N T T A P N H I S N N S S F gagttocagtcotaaacacgacaccottotaatgtgacgtotgaaaacgogocaacottocacacaccacgagtotaataaaq D F I L N T T H P S N V T S E N A P T F T T T S L I K 111 acgcatogtttocggaattatgttoggcaattatggcgaattatggcgagtotgcagtgaggctgcoggg T I V F G I M F G I S F V G N M A T I V Q M R R L R R R 14 aaaagcaccotcacacacatocatoctoctogtcaactggcgaattatggcgaagtotgcaggtgggggggggg	gat	tto	att	tta	act	tee	aac	aat	ato	ter	tca	ott	ate	oto	aac	aco	aco	ract	cca	220	ato	ato	tee		aac	toa	tee	tta	32
$ \begin{array}{c} ggttocagttoctaaacacqacacaccottotaatgtgacgtotgaaaacqcgcocaacottoacaaccacqagtotaataaaq $	D	F	I	L	A	s	D	N	I	s	s	L	L	L	N	Т	Т	A	P	N	м	I	s	N	N	s	s	F	8
$ \begin{array}{c} \mathbf{E} \ \mathbf{F} \ \mathbf{Q} \ \mathbf{F} \ \mathbf{L} \ \mathbf{N} \ \mathbf{T} \ \mathbf{T} \ \mathbf{H} \ \mathbf{P} \ \mathbf{S} \ \mathbf{N} \ \mathbf{V} \ \mathbf{T} \ \mathbf{S} \ \mathbf{E} \ \mathbf{N} \ \mathbf{A} \ \mathbf{P} \ \mathbf{T} \ \mathbf{F} \ \mathbf{T} \ \mathbf{T} \ \mathbf{T} \ \mathbf{S} \ \mathbf{L} \ \mathbf{I} \ \mathbf{K} \ \mathbf{I} \ \mathbf$	gao	ttc	cac	rttc	cta	aac	aco	aca	lcac	act	tet	aat	ato	raco	tet	gaa	aac	aco	reca	acc	tte	aca	acc	aco	agt	cta	ata	aag	40
acquategttttoggaattatgtteggeatttegttegttegeattatggegaetatagtgeagatgagaegtetgeagaeggeege T I V F G I M F G I S F V G N M A T I V Q M R R L R R R 14 aaaageaccatcaacatetetategeegaettggeegaettggeegaettgyeegaettettetgaegtettettetgtategeeggtggegegee K S T I N T L I V N L A L A D L L V T F F C I A G E A A fggaeccgteacagtgtggeegggaagegeteattggeegaettggeegaettetgeegaegtegee the Transmembrane 2 Transmembrane 2 acgtacateactgtggeeatetgtetgaegaettetgteegaegtetgteegeeettgaegeegeegeegeegeegeegeegeegeegeegeegeeg	Е	F	Q	F	г	N	т	т	н	P	s	N	v	т	s	E	N	A	P	т	F	т	т	т	s	L	I	ĸ	11
$ \begin{array}{c} T & V & F & G & I & M & F & G & I & S & F & V & G & N & M & A & T & I & V & Q & M & R & R & L & R & R & I \\ \hline Transmembrane I & & & & & & & \\ \hline aaaagcaaccatcaacaacctctoatcgtcaacctggcgcattggcgcgtcattggcgtcgtcaggtggtgggggggg$	acg	ate	gtt	ttc	gga	att	atg	ttc	ggc	att	tcg	ttc	gtt	gge	aat	atg	rgeg	ract	ata	gtg	cag	atg	raga	legt	ctg	rcga	icgo	egg	49
Transmembrane 1 aaaaqacactotacatogtoaactogtoaattogtocgtotgtocgtacgtottottotgtatogcogtgaggotgto K S T I N T L I V N L A L A D L L V T F F C I A G E A A Transmembrane 2 tggaccatcacggtacaatggtggggggagggtotgtogcaattogtocaagtacatggaggtgttogcocatcatctgtaca acgtacactoatggtgggagggtggggggggggggggggg	т	I	v	F	G	I	M	F	G	I	s	F	v	G	N	м	A	т	I	v	Q	м	R	R	L	R	R	R	14
analysis of the state of the s		200		Iran	smen	ibrai		a+c	ate		ata		++-		~ +	ata	at a	ata		++-	++-	+ ~+	-		a	~ ~ ~	t		57
$\begin{array}{c} Transmembrane 2 \\ Transmembrane 3 \\ T v T V Q W L A G S V M C K F V X Y M Q V F A L Y L S \\ Transmembrane 3 \\ T v T V Q W L A G S V M C K F V X Y M Q V F A L Y L S \\ T summbrane 3 \\ T v T V Q W L A G S V M C K F V X Y M Q V F A L Y L S \\ T summbrane 3 \\ T v T V Q W L A G S V M C X F V X Y M Q V F A L Y L S \\ T summbrane 3 \\ T v I T V A I S L D R C V A I L D P M R R K S A A Q R V \\ 22 \\ cggaccatgatogtotttgggggggggggggggggggggg$	K	s	T	I	N	T	L	I	v	N	L	A	Ъ	A	D	L	L	v	T	F	F	C	I	A	G	E	A	A	16
$ \begin{array}{c} tggaccgtcacggtacatggtggcaggacggcatgtgcacagttgcgcaggtggcacgtggcacgtggcacgtggcgcacgggggcaggaggagggggggg$					•											— T	rans	mem	bran	e 2 —									
W T V T V Q W L A G S V M C K F V Y M Q V F A L Y L S acqtacatoactgtggccatcagtottgaccgctgtgtcgcattottgacccatgagagagaaatcggcggccacgcgaggt T Y I T V A I S L D R C V A I L D P M R R K S A A Q R V 22 cggaccatgatcgtotttgcgtggatattcagtgcctgtcagtataccccaggcgccacgcgggggcca R T M I V F A W I F S A L F S I P Q P I V F N V L R G P 25 R T M I V F A W I F S A L F S I P Q P I V F N V L R G P 26 acgtcactgttgtgctgcctotggccgtcatgggcacagcctacgggccacatgcagtgcacgccatcgccatcgccatcgccatcgccatggtcacagggggggg	tgg	acc	gto	cacg	rgta	caa	tgg	ttg	laca	gga	age	gto	atg	rtge	aag	ttc	gto	aag	rtac	atg	cag	gtg	tto	gcc	cto	tac	ttg	tca	66
$\begin{array}{c} Transmembrane 3 \\ \hline T Y I T V A I S L D R C V A I L D P M R R K S A A Q R V \\ \hline T Y I T V A I S L D R C V A I L D P M R R K S A A Q R V \\ \hline C ggaccatgatcgtotttgggggatattoagtgcottgtcagtacgaccatgggcaccatgggggggggg$	W	т	v	т	v	Q	W	L	A	G	s	v	м	С	ĸ	F	v	ĸ	Y	м	Q	v	F	A	L	Y	L	s	19
acquare a construction of the set of the se															•					I rar	ismei	mbra	ne 3						
$\begin{array}{c} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 $	acg	tac	ato	act	gtg	gee	ato	agt	T	gac	cgc	tgt	gto	gca 7	att	att	gat	200	atg	aga	aga	aaa	teg	laca	gee	cag	rega P	gtt	22
$\begin{array}{c} cggaaccatgatcgtctttgcgtggatattcagtgccttgttcagtataccccagcccatcgtgttcaatgtcctgcgggggcct \\ R T M I V F A W I F S A L F S I P Q P I V F N V L R G P \\ \hline Transmembrane 4 \\ \hline Transmembrane 4 \\ \hline tcaaagaagtgttttaccagtgtgtcacctcggctcgtacgacagtgcctggcaactgcagtgtacgccatcgccatcgccccccc \\ F K E V F Y Q C V T F G S Y D S A W Q L Q M Y A I A S L \\ \hline S K E V F Y Q C V T F G S Y D S A W Q L Q M Y A I A S L \\ \hline S K E V F V L P L A V M G T A Y G L I F T T I S R K S K E \\ \hline atgctcatgttgtgcacagtattcagatgccggcgaggagagacacgcgcagggtcacgccatcgggtcaactaccaaaagcggggccctgtatg \\ \hline M L M F V L P L A V M G T A Y G L I F T T I S R K S K E \\ \hline atgctcatgtggacagtattcagatgccggcgaggagagaccgggtcacgcgatgggtcaactacaaaacgctgcctctgtatg \\ \hline M L M F V L P A R K T W V T R W V N Y K R C L C M \\ \hline 3 \\ gccttcdgtagcaagcccatccccgaagggcaagtcggtcaactactccggaggaccacgcgcccgggggggg$	-	1	-	-	•	A	<u> </u>	5		D	R	C	•	A	-	-	D	F	D1	R	R	r	5	A	A	×	R	•	22
$\begin{array}{c} \mathbf{R} \mathbf{T} \mathbf{M} \mathbf{I} \mathbf{V} \mathbf{F} \mathbf{A} \mathbf{W} \mathbf{I} \mathbf{F} \mathbf{S} \mathbf{A} \mathbf{L} \mathbf{F} \mathbf{S} \mathbf{I} \mathbf{P} \mathbf{Q} \mathbf{P} \mathbf{I} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{G} \mathbf{P} \mathbf{P} \mathbf{V} \mathbf{F} \mathbf{N} \mathbf{V} \mathbf{L} \mathbf{R} \mathbf{K} \mathbf{S} \mathbf{L} \mathbf{R} \mathbf{K} \mathbf{S} \mathbf{L} \mathbf{L} \mathbf{R} \mathbf{K} \mathbf{S} \mathbf{L} \mathbf{R} \mathbf{K} \mathbf{S} \mathbf{L} \mathbf{R} \mathbf{R} \mathbf{S} \mathbf{L} \mathbf{R} \mathbf{R} \mathbf{R} \mathbf{R} \mathbf{S} \mathbf{R} \mathbf{K} \mathbf{S} \mathbf{K} \mathbf{K} \mathbf{R} $	cgg	acc	ato	Jato	gto	ttt	geg	tgg	rata	ttc	agt	geo	ttg	rtte	agt	ata	ccc	cag	rece	atc	gtg	ttc	aat	gtc	ctg	cgg	iaaa	cet	82
$\begin{array}{c} \label{eq:spectral_range} \hline Transmembrane 4 \\ \end{tabular} \\ tabu$	R	т	м	I	v	F	А	W	I	F	s	A	г	F	s	I	P	2	P	I	v	F	N	v	L	R	G	P	25
ttoaaagaagtgttttaccagtgtgtcaccttcggctggtacgacagtgccagtgccagtgccagtgccagtgccagtgccactgccatcgccctcccd 91 F K E V F Y Q C V T F G S Y D S A W Q L Q M Y A I A S L 28 atgctcatgttgtgctgcctctggccgtcatggccacagccacgcggccactctcaccactactccagaagtccagggg 99 M L M F V L P L A V M G T A Y G L I F T T I S R K S K E 30 Transmembrane 5 cattcggtgatggacagtattcagatgccggcggggaggaagacctgggtcacgcgatgggtcaactacaaacgctgcctctgtatg 108 H S V M D S I Q M P A R K T W V T R W V N Y K R C L C M 33 gccttctgtagcaagcccatctccagacggtcctccatcatgtccaactactcggaggaccacgccgccagggggccggtgcgg 116 A F C S K P I S R R S S I M S N Y S E D H A A R G P V R 36 agttacctcctccgtaaagccaagcggaagtcgctgatcatgtctttcgtcatcgttttggcctcatggtggtggacggccg 124 S Y L L R K A K R K S L I M S F V I V L A F M V C W T P 39 tactacatcatcttcatttgcatcacctcctggacgatggtcatggccggtcattttaaccaactactcaggtggtctatggacgatg Y Y I I F I C I T F L D K V I D P V I F N Y F S F I G L 42 Transmembrane 7 tccaactccatggtgaacccgatgatttacggaggtttccaagtttgccaagttgcaagttgtgtcaaggttgtcctaaggt 141 S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R 44 ttgactggtaggctcccatgtacttttttgtgcgctgcatggtggtggtggcagtcgtcattttatacccaaggttgttcaaggttgtcaaggtftactatgtactttttttgttgcgctcaaggtggtggtggtggtggtggtggtgtgtgt		•									- Tr	ansn	iemb	rane	4				a de cara de c		1.10.000			•					
F K E V F Y Q C V T F G S Y D S A W Q L Q M Y A I A S L atgeteatgtttgtgetgeetetggeegetatgggeegetatgggeegetatetteeceagaagteeaaggag M L M F V L P L A V M G T A Y G L I F T T I S R K S K E oatteggtgatggaeagtatteagatgeeggegggaggaagaeetgggteaaetaeaaaegetgeetetgtatg H S V M D S I Q M P A R K T W V T R W V N Y K R C L C M geettetgtageaageecaateteeaaaggeggeeggteatetteeggaggaeeeaegeeggggggeeggtggg A F C S K P I S R R S S I M S N Y S E D H A A R G P V R agttaceteeteetgaageecaageeggaagtegtgateatgtetteggeeteaaetaeaaaggetggtgggggeeggtgggg Y L L R K A K R K S L I M S F V I V L A F M V C W T P agttaceteeteetgaaeeeageeggaagtegteatgaeeaggteetgaeatgteetteaaetaeateeteggaggteggte	tte	aaa	gaa	igtg	rttt	tac	cag	tgt	gtc	acc	tte	ggq	stag	rtac	gac	agt	gee	tgg	rcaa	ctg	cag	atg	rtac	gee	ato	gee	tee	ete	91
atgeteatgtttgtgetgeeteatgggeegeetaatgggeetaatetteaeeactateteeagaaagteeaaggag 99 $M \perp M F \vee L P \perp A \vee M G T A Y G L I F T T I S R K S K E 30$ catteggtgatggeaggatteeaggtggeegggggeegggggeeggee	F	ĸ	E	v	F	Y	Q	С	v	т	F	G	s	Y	D	s	A	W	Q	г	Q	м	Y	A	I	A	s	г	28
$\begin{array}{c} M \ L \ M \ F \ V \ L \ P \ L \ A \ V \ M \ G \ T \ A \ Y \ G \ L \ I \ F \ T \ T \ I \ S \ R \ K \ S \ K \ E \\ \hline Transmembrane 5 \\ \hline \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	ato	oto	ate	r+++	ata	ata	cet	oto	acc	ate	ato		aca	acc	tac	aaa	oto	ato	tto	acc	act	ato	tee	ana	330	tee		man	99
$\begin{array}{c} \hline Transmembrane 5 \\ \hline catteggtgatggacagtatteagatgeeggeggaggaggacgeggtggaggaggacggggggaggaggeeggtgggggggggg$	м	L	м	F	v	L	P	L	A	v	м	G	т	A	Y	G	L	I	F	т	т	I	s	R	ĸ	s	ĸ	E	30
catteggtgatggacagtatteagatgeeggegggaggaagaeetgggteaegeggteaeetgeggteaeetgegeteaetgeetetgtatg 108 H S V M D S I Q M P A R K T W V T R W V N Y K R C L C M 33 geettetgtageaageeeateteeagaeggteeteeateatgteeaaetaeteeggaggaeeaegeegeeagggggeeggtgegg 116 A F C S K P I S R R S S I M S N Y S E D H A A R G P V R 36 agttaceteeteeggaageeggagtegetgateatgtettegteeaetgtetggeetteatgtggeegeeggtgggggggg						– Tr	ansn	ıemb	rane	5							-)											
H S V M D S I Q M P A R K T W V T R W V N Y K R C L C M 33 gcottetgtagcaagcccatctccagacggtcctccatcatcgtgggggggg	cat	tcg	gtg	gatg	gac	agt	att	cag	ratg	ccg	gcg	agg	raag	racc	tgg	gto	acg	rcga	ltgg	gtc	aac	tac	aaa	lege	tgo	ato	tgt	atg	108
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	н	s	v	м	D	s	I	Q	м	Р	A	R	ĸ	т	W	v	т	R	W	v	N	Y	ĸ	R	С	L	С	м	33
A F C S K P I S R R S S I M S N Y S E D H A A R G P V R 36 agttacctoctcogtaaagccaagcggaagtcgctgatcatgtottcgtcatcgtttggcctcatggtggtggtggacgccc 124 S Y L L R K A K R K S L I M S F V I V L A F M V C W T P 39 tactacatcatcttcattgcatcaccttcctggacaaggtcattgacccggtcatcttcaactactcagettcatcggactg 133 Y Y I I F I C I T F L D K V I D P V I F N Y F S F I G L 42 tcccaactccatgtgacccgatgatttacggaggtttccaggtggaagttgacaggtgcaatttataatccaaggttgtctaagg 141 S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R 44 ttgactggtaggctcccatgtactttttgtgcgctgcatggtggcagtccgctcgct	gee	tte	tgt	age	aag	race	atc	tee	aga	cgg	tee	tec	ato	atg	tee	aac	tac	teg	rgag	gac	cac	gee	geo	agg	aaa	reeg	gtg	cgg	116
agttacotootoogtaaagcogaagcogotgatoatgtotttogtoatogttttggoottoatggtgtgttggacgooo 124 S Y L L R K A K R K S L I M S F V I V L A F M V C W T P Transmembrane 6 Y Y I I F I C I T F L D K V I D P V I F N Y F S F I G L toccaactocatgotgaaccogatgatttacggagotttocaggtggaagtgggaagttgtattgtoctaagg S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R ttgactggtaggotoccatgtactttttgttgcgotgcatggtggcagtcogototgottatootocttoatotaaaaatac L T G R L P C T F L L R C M V A V R S A Y P P F I * 133 141 141 153 153 153 153 153 153 153 153 153 154 154 154 154 155 155 155 156	A	F	С	S	ĸ	P	I	s	R	R	s	s	I	м	s	N	Y	s	Е	D	H	A	A	R	G	P	v	R	36
S Y L L R K A K R K S L I M S F V I V L A F M V C W T P Transmembrane 6 V C W T P Transmembrane 7 V I F N Y F S F I G L toccaactocatgotgaaccogatgatttacggagotttocagotttgcaaggtgcaattttataatccaaggtttgtoctaagg 141 S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R 44 ttgactggtaggotoccatgtactttttgttgcgotgcatggtggcagtccgototgottatcotocttcatotaaaaatac L T G R L P C T F L L R C M V A V R S A Y P P F I * 47 cccotccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	agt	tac	ato	oto	cgt	aaa	gee	aag	regg	aag	tog	rctg	ato	atg	tet	tto	gto	ato	gtt	ttg	gee	tto	atg	gtg	tgt	tgg	racg	ccc	124
$\begin{array}{c} 133\\ Y \ Y \ I \ I \ F \ I \ C \ I \ T \ F \ L \ D \ K \ V \ I \ D \ P \ V \ I \ F \ N \ Y \ F \ S \ F \ I \ G \ L \ 42\\ \hline & \\ \hline \\ \hline$	s	Y	L	L	R	к	A	к	R	ĸ	s	г	I	м	s	F	v	I	v Trar	L. Ismei	 mbra	r ne 6	м	v	ĉ	W	т	P	39
$\begin{array}{c} Y Y I I F I C I T F L D K V I D P V I F N Y F S F I G L \\ \hline \\ tocaactocatgotgaaccogatgatttacggaggottcocagcttgcaaggtgcaattttataatcocaaggttgtoctaagg 141 \\ \hline \\ S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R \\ \hline \\ ttgactggtaggotcocatgtactttttgtgcgctgcatggtggcagtcogotcgottgcttatcotcottcatctaaaaatac 150 \\ L T G R L P C T F L L R C M V A V R S A Y P P F I * \\ \hline \\ ccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa$	tac	tac	ate	ate	tto	att	tac	ato	acc	tto	oto	aac	aao	rate	att	aac		ato	ato	tto	aac	tac	tto	ago	tto	ate	ada	ata	133
$\begin{array}{c} \text{Transmendrate /} \\ \hline \\ \text{Transmendrate /} \\ \hline \\ \text{S N S M L N P M I Y G A F Q L C K V Q F Y N P R F V L R } \\ \text{ttgactggtaggctcccatgtactttttgtgcgctgcatggtggcagtccgctctgcttatcctccttcatctaaaaatac } \\ \text{L T G R L P C T F L L R C M V A V R S A Y P P F I * } \\ \hline \\ cccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa$	Y	Y	I	I	F	I	c	I	т	F	L	D	к	v	I	D	P	v	I	F	N	Y	F	S	F	I	G	L	42
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	+		+								_					+				***	* - *	- 11a	1151110		ane	-			141
ttgactggtaggotococatgtactttttgttgogotgcatggtggcagtocgototgottatootootttoatotaaaaatac L T G R L P C T F L L R C M V A V R S A Y P P F I * 47 ccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	tee	aac	t d d	atg M	T.	N	P	M	T	v	gga	a get	こしし ひっし	cag	CLL T.	cgc	aag	IGCG V	rcaa	F	v	aat	D D	agg p	F	gtc	CCTA	agg	141
ttgactggtaggeteccatgtactttttgttgegetgeatggtggeagtecgetetgettateeteetteatetaaaaatae 150 L T G R L P C T F L L R C M V A V R S A Y P P F I * 47 ccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaa 153 Eig 1 Nucleatide and deduced amine acid acquerance of the full length nutative on CoPHP (Control Accession # AHE72444 1) The second s	.5		5	24	-		-	24	-	-	3	A		×	-	-	~	,	×	÷.	-		-	-	-	•	-		
L T G R L P C T F L L R C M V A V R S A Y P P F I * 47 ccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	tto	act	aat	ago	rete	cca	tgt	act	ttt	ttg	tta	rege	tgo	ato	gta	gea	gto	ege	tet	get	tat	cct	cet	tte	ato	taa	aaa	tac	150
ccctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	L	т	G	R	г	P	c	т	F	г	г	R	c	м	v	A	v	R	s	A	Y	P	P	F	I	*	=		47
Fig 1 Nucleotide and deduced amine sold esquences of the full length nutative on CnPHP (Conherty Associate # AHE70444 1). Th	aac	tee	aas	laaa	aaa	aaa	aaa	aaa	aaa	aaa	a																		153
			-	tide	ام مر م			I					f	lha f		in arts				~~ ~	UD //	Corl				on 4		7044	4 4) TI

nucleotides (lower case letters) and amino acids (upper case letters) are numbered accordingly. The seven predicted transmembrane domains are underlined. The two putative translation start sites are highlighted in green, and the asterisk (*) denotes the stop codon. The nucleotides corresponding to the truncated polyadenylation signal (aata) are double-underlined.

doi:10.1371/journal.pone.0160292.g001

PLOS ONE

loop (IL) 2. Two cysteine residues for the formation of a highly conserved disulfide bridge between EL 1 and EL 2 [25] are located at positions 182 and 260.

Interestingly, ap-GnRHR-L has the longest extracellular N-terminus and IL 3 compared to other similar GPCRs (Fig 2). The TM regions are well conserved between ap-GnRHR and other invertebrate receptors, but the presence of the long IL 3 in ap-GnRHR greatly diminished its similarity to others. The length of extracellular N-terminus in another molluscan GnRHR (oct-GnRHR) is more similar to that of ap-GnRHR-S than ap-GnRHR-L. Like all type II vertebrate GnRHRs [25] and other invertebrate receptors in this superfamily, ap-GnRHR also possesses a C-terminal tail.



	+				+	TM 1					
ap-GnRHR Ms-CRZR oct-GnRHR Ci-GnRHRI Ci-GnRHRII Pm-GnRHRI Mm-GnRHR	R MDGTDSDGSILQRANQAFLASSSGPLQLPIPYNFSAQNSSVHRKRHDVYSPRSAHHDFILASDNISSLLLNTTAPNMISNNSSFEFQFLNTTHPSNVTSENAPTFTTTSLIKTT 										
	IL 1		TM 2	EL 1	TM 3	IL 2					
ap-GnRHR Ms-CRZR oct-GnRHR Ci-GnRHRI Ci-GnRHRII Pm-GnRHRI Mm-GnRHR	GISFVGNMATIVQMRRLRRRK- ILSFFGNLATISIQRGKRGRG VISFFGNTLVIIQIFRIRGSR- CLSMAGNMFVLWSLRGSKSR IFAAISNLTVLTISHNHRKT- LLSTAFNASFLLKLQKWTQKR- :: * ::	STINTLIVNLAJ RARPSWTAIYSLIFQLS] STIQSLILNLAJ HFIMFHLAJ KSHFIMFHLAJ KKGKKLSRMKVLLKHLTI ::*:	LADLLVTFFCIAGEAAWT (ADLLVTIFCIAGEAAWS (ADLMVSFFNILMDIIWS SSNLITIFVMPSDAVWN SSDLITIFVMPSDAVWN (ADLLITFIVMPLDAVWH ANLLETLIVMPLDGWWN ::*::::: *	VTVQWLAGSVMCKF FTVQWYAGNIACKI ATVEWLAGNYMCKI TTMEWLAGDVMCRL TTMEWLAGDVMCRL ITTQWYAGEFACRL ITTQWYAGEFLCKV * :* **. *:	VKYMQVFALYLSTYITV FKFLQMFALYQSTFILV MKYLTVFGLHLSTYITV CQMMKQFGMYASSFMVV CQMMKQFGMYASSFMVV LMFLRLLAMYSSAFITV LSYLKLFSMYAPAFMMV :**	AISLDRCVAILDPMRRKSAAQRVRTMIVFAWI 233 LIGVDRMLAVKYPMKSMATATRSCRLVVIAWU 213 SIALDRCFAILSPMSRSKAPLRVRIMITMAWU 155 VIGADRVTGILSPLPCHSQRKRGYIMVATAWI 174 VIGMDRVTAILSPLTHEGQRKRGYIMVATAWI 164 VISLDRISAILNPLGIGKAKAKNTMLSVAWU 166 VISLDRISAILNPLGIGKAKAKNTMLSVAWU 165 *. ** .: *: :: **					
	TM 4	EL 2	TM 5	ũ		IL 3					
ap-GnRHR Ms-CRZR oct-GnRHR Ci-GnRHRI Ci-GnRHRII Pm-GnRHRI Mm-GnRHR	FALFSIPQETUFNVLRGPFKE LSVULSIPQTVVFRVAKGPFVE LSAIFSIPQAVIFQEQRMFRQ SSLICCLPAGFIFSVASIPTCE TSLLCCIPAGALFSLLTVETCE LSVLLAVPQLFLFHVKSPKGN- LSIVFAGPQLYIFRIYLADGSG * : . * :*	UPYQCUTFGSVDSAW EFYQCUTFGSVDSAW GIPTQCRDSVNALW GIPTYQCIDFNULQDVSI GINUYQCUDFNIVKDRSI -KNFVQCUTHGNFVEQW PTVFSQCUTHCSFPQWW **	DLQMYATASLMLMFVLPL DEQAYTTLSLVFMFVLPL DEQAYTTLSLVFMFVLPL LRPYYFFTMCMSFLLPL LLRPYYFFTMCMSFLLPL LLRPYYFFTMCMSFLLPL HNLLYMFTFVFLFILPL QAFYNFFTFGCLFIIPL * :: *::*	AVMGTAYGLIFTTI VILISTYVSTVRTI IIMVTSYLLILKTI ICTLVSYSLIVCEI ICTLVSYSLIVCEI FIMIFCYCRILLEI LIMLICNAKIIFAL	SRSSKEHSVMDSIQMPA ARSEK VKTSR STMKE	.RKTWVTRWVNYKRCLCMAFCSKPISRRSSIMS 351 VFKPEVRR 295 _QFHDTPISPTSMS 241 250 240 251 251 250 241					
		TM 6	EL	. 3	TM 7						
ap-GnRHR Ms-CRZR oct-GnRHR Ci-GnRHRI Ci-GnRHRII Pm-GnRHRI Mm-GnRHR	NYSEDHAARGPURSYLLRKARR QEKYFTPDMNRRLIDRAKM CYSVNHGQIRTHLFERARK RDRVLMGRRHSVNTASIQRAKN RDRVLMGRRHSVNTASIQRAKN GSISSKEIRLRRSNNNIPKARM DPRKLQLNQSKNNIPRARL : :*:	KSLIMSFVIVLAFMVCW KSLRMSVVIVAAFILCW RSSRMSAVIVAAFILCW RTILMRTLITLTFIVCW RTILMRTLITLTFIVCW RTLKMSIAIVSSFVVCW :: * :.:*:*	PPYYIFICITFLDKV APYYVMMIIFTFLNPDKD PPYYIFLGFAFFQWD SPYYGKGIYDWFIRYEDH SPYYGKGIYDWFQKPAIG PPYYULGIWYWFDRSIVS *** : *	IDPVIFNYFSF QSEELLSGIFF NS-RTVIYFFT TPPDAWDTVMYV -PPAPLDTAMYI RKVVPHFVEEMSLT RVSEPVNHFFFL	IGLSNSMLNPMIYGAFC FGMSNSLVNPVIYGAFT LGTSNCMLNPLIYGAFT VMYLNPVLHPIVFGVFI VMYLNPVLHPIVFGVFF FGLLNACLDPVIYGVFF FAFLNPCFDPLIYGVFS * **::* *	LCKVQFYNPRFVLRLTGR 452 LWPKKKRSHRHSDRESGGHHASLLRRGDN 407 IYKVHRGRSGSANSPSGTRLMIMVNKGRST 352 KEIRGKFKQRLNCARKRFFKQGDFKTVPNAQS 368 KEIRSKFKKTSCVAGLKRKR 345 AHVRREVRRCCRWPRTAAHDRDSSSTPVTGSF 371 L 326					
ap-GnRHR Ms-CRZR oct-GnRHR Ci-GnRHRI Ci-GnRHRII Pm-GnRHRI Mm-GnRHR	LP TTTNRMSGSGRRQLTTGQTITQ SMNYSIASVLNRPRRMSSTSRG VQVRSGSVSRL RYSASSVRSRVPFACGEQPEA	CTFLLRCMVAVR SSIRLTTIRSLRSS CASLTNPHQPVRPSPGIN SFSSYATGATHLNGSSHV SLTSYATGLTQVQYLS TGAHPTPATRLLQRGCLV	SAYPPFI- 	ISLL QPQFFGANRMVAPQ VPSLDESVHTKAPE AFCDASGGGAGGGG	QQLLLSSESAL	- 473 - 435 - 407 - 450 - 401 I 460 -					
Fig 2. Amino	acid sequence alignme	nt of ap-GnRHR wit	th representative G	inRHR-related	receptors. Shaded	amino acids indicate the					

Fig 2. Amino acid sequence alignment of ap-GnRHR with representative GnRHR-related receptors. Shaded amino acids indicate the transmembrane (TM) domains. The predicted intracellular loops (IL) and extracellular loops (EL) are numbered accordingly. The two putative start sites (M) are each denoted by a "+". Ms-CRZR—corazonin receptor, *Manduca sexta* (AAR14318); Oct-GnRHR—Gonadotropin-releasing hormone receptor, *Octopus vulgaris* (GNRHR_OCTVU); Ci-GnRHRI—gonadotropin-releasing hormone receptor 1, *Ciona intestinalis* (NP_001028997); Ci-GnRHRII—gonadotropin-releasing hormone receptor 2, *Ciona intestinalis* (NP_001028996); Pm-GnRHRI—gonadotropin releasing-hormone receptor, *Petromyzon marinus* (AF439802_1); Mm-GnRH—gonadotropin-releasing hormone receptor isoform *Mus musculus* (NP_001297580). Identical, highly conserved, and less conserved amino acid residues are denoted by an asterisk (*), a colon (:), and a period (.), respectively. Ampersand (&) denotes cysteine residues involved in the disulfide bridge formation.

doi:10.1371/journal.pone.0160292.g002

RT-PCR results suggest ap-GnRHR-L and ap-GnRHR-S are differentially expressed (Fig 3 and S1 Fig). Since ap-GnRHR-S is part of ap-GnRHR-L, it is not possible to amplify only the short receptor without simultaneously detecting the long form. However, primers targeting specifically the long receptor revealed that ap-GnRHR-L was only expressed in the abdominal, cerebral, and buccal ganglia of the central nervous system and a few peripheral tissues including the osphradium (a chemosensory organ), small hermaphroditic duct, and ovotestis (Fig 3 and S1 Fig). Primers targeting both receptor isoforms successfully amplified a transcript in all tissues examined (Fig 3 and S1 Fig). By deduction, our results suggest the expression pattern is ubiquitous for ap-GnRHR-S but restricted for ap-GnRHR-L.

Phylogenetic analysis of 334 predicted amino acid sequences of GnRHR and related receptors (Fig 4) suggests this receptor superfamily is segregated into 3 major clades: (1) deuterostome GnRHR, (2) ACP/AKH receptors (ACPR/AKHR), and (3) CRZ receptors (CRZR). Ap-GnRHR is clustered with oct-GnRHR and several unauthenticated receptors from a limpet (*Lottia gigantea*) and an oyster (*Crossostrea gigas*). These molluscan receptors cluster with



Fig 3. Expression of ap-GnRHR isoforms (top two panels) in different *A. californica* **tissues examined by RT-PCR.** Primers used for the top panel amplify only ap-GnRHR-L; primers used for the second panel amplify both ap-GnRHR-L and ap-GnRHR-S. Negative controls (no RT) were RNA samples that have not been reverse transcribed (third panel). ap-Actin was used as a control to ensure the quality of RNA samples (bottom panel). The sizes of the PCR products are shown on the right. BCN, bag cell neurons; Abg, abdominal ganglia; Ceb, cerebral ganglia; P/P, pedal/pleural ganglia; Buc, buccal ganglia; Tail, ventral tail tissue; Osp, osphradium; SHD, small hermaphroditic duct; Hrt, heart; Ovt, ovotestis;—Cont, negative water control.

doi:10.1371/journal.pone.0160292.g003

multiple insect receptors, including the authenticated fruit fly CRZR [26], to form a CRZR lineage. On the other hand, the chordate GnRHRs form a distinct group that is segregated from all other protostome receptors examined. Interestingly, a putative AKHR was identified from *A*. *californica* genome and belonged in the ACPR/AKHR lineage with other authenticated insect ACPRs and AKHRs (Fig 4).

Radioreceptor assays of S2 cells transfected with ap-GnRHR-L or ap-GnRHR-S show that $[^{125}I]$ -ap-GnRH binds specifically to ap-GnRHR-L, with unlabeled ap-GnRH displacing radioligand binding at EC₅₀ of 2.52 nM (Fig 5). No specific binding was observed in S2 cells transfected with ap-GnRH-S (Fig 5). Further, binding of $[^{125}I]$ -ap-GnRH to either ap-GnRHR-L or ap-GnRHR-S could not be displaced by a related peptide, ap-AKH (Fig 6). ap-GnRH did not stimulate cAMP accumulation in either ap-GnRHR isoform (Fig 7), but increased IP1 accumulation up to 6-fold in S2 cells transfected with ap-GnRHR-L (Fig 8). The quality of cAMP and IP1 assays was validated by the positive effects of forskolin and proctolin (Figs 7 and 8), respectively. S2 cells transfected with the empty vector, pAWG, failed to bind specifically to $[^{125}I]$ -ap-GnRH and did not respond to ap-GnRH in terms of cAMP or IP1 accumulation (Figs 5 and 8).

Discussion

The rapidly emerging field of protostome GnRH has provided valuable insight into the functional and structural evolution of this hormone superfamily [2, 27]. There is now strong evidence to suggest the GnRH superfamily co-evolved with its cognate receptors exclusively in the bilaterian lineages, with both the ligands and receptors absent in non-bilaterians such as sponges, cnidarians, placozoans and ctenophores [1]. These findings placed the origin of the GnRH and GnRHR superfamilies around 570 to 700 million years ago [28], and this evolutionary timeline allowed for the tremendous diversification of ancestral molecules into modern clades consisting of chordate GnRHs, non-chordate orthologs such as AKH, CRZ, ACP, and their cognate receptors [29]. Such diversity underscores the importance of receptor





Fig 4. Maximum likelihood analysis of the phylogeny of GnRHR-related receptors. The bootstrap values (in %) from 1000 replicas are indicated at each branch point. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 334 amino acid sequences and a total of 1171 positions in the final dataset. The invertebrate cardioacceleratory peptide receptors and vertebrate vasopressin receptor-like GPCRs were used as outgroups to root the tree. The detailed nomenclature and sequences are listed as data in <u>S1 Dataset</u>.

doi:10.1371/journal.pone.0160292.g004

deorphanization to define the exclusive ligand-receptor pairing, which has been the primary goal of the present study.

Our results validate a candidate receptor as an authentic ap-GnRHR. This receptor has several features that are novel among most known GnRHRs. The first feature is the presence of two potential start sites that might result in the translation of two endogenous receptor proteins (Figs 1 and 2). Neither start site contains the ideal Kozak sequence ($GCC^A/_CCATGG$) found among vertebrates [23], but both start site variations have been documented as functional in higher eukaryotes. Unlike vertebrates, protostomes appear to lack a strong preference for a G at position #4 of the Kozak sequence (last G in ATGG) [30], suggesting the two start sites in ap-GnRHR may be similarly effective. The second feature is the differential expression of at least two forms of transcript that vary in the 5' region corresponding to the extracellular N-terminus (Fig 3 and S1 Fig). A parsimonious explanation is that two splice variants, one fulllength and one with a 5' truncation, are produced in a tissue-specific manner. This may have functional implications as discussed in a later paragraph. The third feature is the unusually long IL 3 of ap-GnRHR compared to other related GPCRs (Fig 2). The IL 3 in ap-GnRHR is approximately twice as long as oct-GnRHR and is rich in phosphorylation target residues. IL 3 in vertebrate GnRHRs plays a crucial role in the coupling to $G_{q/11}$ [25], but the functional significance of this region in ap-GnRHR is at present unclear. Despite these unusual features, ap-GnRHR-L is a bona fide receptor that binds with high affinity to ap-GnRH to activate a well-defined signaling pathway for vertebrate GnRHRs [25].



Fig 5. Analysis of [¹²⁵I]-ap-GnRH binding to S2 cells transfected with ap-GnRHR-L or ap-GnRHR-S. Binding to the radioligand was displaced by increasing concentrations of unlabeled ap-GnRH in S2 cells transfected with ap-GnRHR-L ($EC_{50} = 2.52$ nM), but not ap-GnRHR-S. Binding data are expressed as the percentage of [¹²⁵I]-ap-GnRH binding to the respective receptor in the absence of unlabeled hormone. Empty vector-transfected cells were used as a negative control. Each data point represents N = 3 assays run in triplicates.

doi:10.1371/journal.pone.0160292.g005



Fig 6. Analysis of [¹²⁵]-ap-GnRH binding to S2 cells transfected with ap-GnRHR-L or ap-GnRHR-S. Binding to the radioligand was not displaced by increasing concentrations of unlabeled ap-AKH in S2 cells transfected with either ap-GnRHR-L or ap-GnRHR-S. Binding data are expressed as the percentage of [¹²⁵]-ap-GnRH binding to the respective receptor in the absence of unlabeled hormone. Each data point represents N = 3 assays run in triplicates.

doi:10.1371/journal.pone.0160292.g006

PLOS ONE

The nomenclature of invertebrate GnRH and GnRHR superfamilies is at present problematic due to several naming attempts [15, 16, 19, 27] as well as our evolving understanding of these molecular families [2, 31]. A phylogenomic analysis of the deep relationships among GnRH-like peptides revealed a strong monophyletic grouping of all chordate GnRHs, but their relationships with protostome GnRH-like molecules such as AKH, CRZ, ACP, and protostome GnRHs are paraphyletic [29]. Our present phylogenetic analysis of GnRHR-like GPCRs mirrors findings on the ligands and further re-classifies protostome GnRHR-like receptors into the ACPR/AKHR and CRZR clades, all of which are paraphyletic to chordate GnRHRs (Fig 4). At some point, the grouping and functional characteristics of protostome GnRHRs should encourage an attempt at a more consistent nomenclature. For example, ap-GnRH is stimulatory to neither gonadotropin secretion nor gonadal function in *A. californica*, but instead controls behavioral and motor coordination [7]. Further, the multifunctional nature of ap-GnRH is similar to other non-GnRH peptides, such as CRZ, AKH and ACP, in this superfamily [11].





Fig 7. ap-GnRH did not stimulate cAMP accumulation in S2 cells transfected with either ap-GnRHR-L or ap-GnRHR-S. Data are expressed as fold stimulation over cAMP levels in the untreated cells. Forskolin was used as a positive control. N = 3 assays run in triplicates.

doi:10.1371/journal.pone.0160292.g007

By virtue of its function and the phylogenetic grouping of its receptor, ap-GnRH should be more appropriately named ap-CRZ and its receptor named ap-CRZR.



Fig 8. ap-GnRH stimulated IP1 accumulation in S2 cells transfected with ap-GnRHR-L, but not ap-GnRHR-S. Data are expressed as fold stimulation over IP1 levels in the untreated cells. Proctolin was used as a positive control in ap-GnRHR-L or ap-GnRHR-S-transfected cells, and empty vector-transfected cells were used as a negative control. N = 3 assays run in triplicates.

doi:10.1371/journal.pone.0160292.g008

Our functional characterization of ap-GnRHR revealed that the expression of ap-GnRHR-L, but not ap-GnRHR-S, resulted in ligand binding and the initiation of intracellular signaling (Figs 5-8). A caveat is that S2 cultures transfected with ap-GnRHR-L may also produce ap-GnRHR-S, since the short receptor sequence is included in the long receptor. The short receptor alone is non-functional, but it may play a role in modulating the signaling properties of the long receptor via receptor hetero-dimerization. Hetero-dimerization among GnRHR splice variants and paralogs is common [32-35], and in the case of tunicate GnRHRs, an orphan paralog with no binding affinity for any tunicate GnRHs can dimerize with another functional tunicate GnRHR to markedly enhance the biological potency of the latter [34, 35]. Future studies using site-directed mutagenesis to alter the second start codon of ap-GnRHR-L should allow us to examine the activation of ap-GnRHR-L in isolation.

Although vertebrate GnRHRs are predominantly coupled to Gq/11 and activate the IP pathway, their coupling to G_s and G_i has also been reported [25]. The preferential coupling of a vertebrate GnRHR to a particular G protein may also change in response to the ligand [36]. Consistent with this notion, two tunicate GnRHRs were reported to activate exclusively the cAMP pathway, but one reported to activate both cAMP and IP pathways when stimulated with a specific tunicate GnRH, reaffirming the versatility of G protein coupling [37]. Our results support the IP pathway as the primary signaling pathway activated by ap-GnRHR and are consistent with findings from oct-GnRHR [19]. That said, cAMP accumulation should be routinely examined for other trochozoan GnRHR-related receptors since ecdysozoan CRZR and AKHR are known to activate both IP and cAMP pathways [38, 39]. In addition, we noted that ap-AKH, an ortholog of ap-GnRH and ecdysozoan AKH [12], failed to bind to ap-GnRHR-L or ap-GnRHR-S (Fig 5), demonstrating a high level of receptor selectivity for specific protostome ligands. In contrast to this selectivity, an amphioxus GnRHR with structural similarity to protostome GnRHR was more promiscuous and responded similarly to amphioxus GnRH, oct-GnRH and an insect AKH [40, 41]. Based on these results, we suggest a high degree of fidelity in the ligand-receptor pairing within the protostome members of the GnRH/ GnRHR superfamilies that may have resulted from long periods of ligand-receptor coevolution.

The expression pattern of a receptor is frequently used to infer its function, but our results suggest a caveat to this approach. In our case, the non-functional ap-GnRHR-S is ubiquitously expressed in all tissues examined, an outcome consistent with our earlier work [5], but the functional ap-GnRHR-L is expressed in far fewer tissues (Fig 3). This finding carries the significant implication that some invertebrate GnRHR-like receptors may not be functional even if expressed at high levels. As such, expression studies should be regarded with caution in the absence of functional data.

The functional ap-GnRHR-L is expressed in both central and peripheral tissues, a finding consistent with previous data suggesting ap-GnRH as a neuropeptide with diverse roles in the control of neuronal, feeding, and motor activities [5–7]. Both ap-GnRHR-L and ap-GnRHR-S are weakly expressed in the ovotestis, an observation that may invite speculations on a reproductive role of this receptor. However, we have not observed any stimulatory effects of ap-GnRH on gonadal histology, egg-laying, and mating behavior in *A. californica* [7]. In fact, ap-GnRH reduced the ability of the neuroendocrine bag cell neurons to secrete an important reproductive hormone, thereby decreasing the probability of egg-laying in these animals [6]. The expression pattern of ap-GnRHR-L is also consistent with the functional diversity reported for ecdysozoan CRZ, AKH, and ACP in neural, behavioral, and metabolic regulation [11]. Reproductive activation by these neuropeptides in ecdysozoans is uncommon, and if observed, is likely secondary to their primary functional repertoire instead of being a dedicated and conserved effect comparable to vertebrate GnRH. Our current analyses of ap-GnRHR, when

combined with the previous phylogenomic analysis [29] and physiological characterization [5– 7], steer us towards the conclusion that ap-GnRHR is more similar to protostome CRZR than chordate GnRHR in structure, function, and evolutionary ancestry.

In sum, we have molecularly isolated a putative ap-GnRHR and functionally validated it as a bona fide ap-GnRHR. This receptor exhibits some unusual features, including the presence of two potential translation start sites and the differential expression of at least two transcript isoforms, but with only one form being functional. This is only the second report on the functional authentication of a trochozoan GnRHR, thus it is unclear if these features are unique to ap-GnRHR or also present in other trochozoan receptors. Additional functional pairing of protostome ligands and receptors within the GnRH and GnRHR superfamilies is needed to shed light on the histories of how these ligands and receptors co-evolved in the last 570–700 million years.

Supporting Information

S1 Dataset. A complete list of amino acid sequences and Accession numbers used to construct the evolutionary tree (Fig 4). (PDF)

S1 Fig. Full images of RT-PCR (presented in Fig 3) with DNA ladder to indicate the size of the amplicons. Amplification of ap-GnRHR-L and ap-GnRHR-S (A), ap-Actin (B), and ap-Actin using RNA samples that have not been reversed transcribed (C) are shown. Tissue abbreviations are indicated in Fig 3 legend. The 12 visible DNA ladder bands are: 1517, 1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, and 100 bp (New England Biolabs, Ipswich, MA). (TIF)

Acknowledgments

We thank Joshua Johnson and Biao Sun for their assistance in animal husbandry and tissue harvest.

Author Contributions

Conceptualization: SIK PST. Formal analysis: SIK PST. Funding acquisition: PST. Investigation: SIK PST. Methodology: SIK PST. Project administration: PST. Resources: SIK PST. Supervision: SIK PST. Validation: SIK PST. Writing - original draft: SIK PST. Writing - review & editing: SIK PST.

References

- Plachetzki DC, Tsai PS, Kavanaugh SI, Sower SA. Ancient Origins of Metazoan Gonadotropin-Releasing Hormone and their Receptors Revealed by Phylogenomic Analyses. General and comparative endocrinology. 2016. doi: 10.1016/j.ygcen.2016.06.007 PMID: 27318276.
- Roch GJ, Busby ER, Sherwood NM. GnRH receptors and peptides: skating backward. General and comparative endocrinology. 2014; 209:118–34. doi: <u>10.1016/j.ygcen.2014.07.025</u> PMID: <u>25107740</u>.
- Minakata H, Shigeno S, Kano N, Haraguchi S, Osugi T, Tsutsui K. Octopus gonadotrophin-releasing hormone: a multifunctional peptide in the endocrine and nervous systems of the cephalopod. J Neuroendocrinol. 2009; 21(4):322–6. Epub 2009/02/13. [pii] doi: <u>10.1111/j.1365-2826.2009.01852.x</u> PMID: <u>19210294</u>.
- Minakata H, Tsutsui K. Oct-GnRH, the first protostomian gonadotropin-releasing hormone-like peptide and a critical mini-review of the presence of vertebrate sex steroids in molluscs. General and comparative endocrinology. 2016; 227:109–14. Epub 2015/09/01. doi: <u>10.1016/j.ygcen.2015.07.011</u> PMID: <u>26319132</u>.

- Sun B, Kavanaugh SI, Tsai PS. Gonadotropin-releasing hormone in protostomes: insights from functional studies on Aplysia californica. General and comparative endocrinology. 2012; 176(3):321–6. Epub 2011/12/17. doi: <u>10.1016/j.ygcen.2011.11.030</u> PMID: <u>22172342</u>.
- 6. Sun B, Tsai PS. A gonadotropin-releasing hormone-like molecule modulates the activity of diverse central neurons in a gastropod mollusk, Aplyisa californica. Front Endocrinol. 2011; 2:36.
- Tsai PS, Sun B, Rochester JR, Wayne NL. Gonadotropin-releasing hormone-like molecule is not an acute reproductive activator in the gastropod, Aplysia californica. General and comparative endocrinology. 2010; 166(2):280–8. Epub 2009/10/06. doi: 10.1016/j.ygcen.2009.09.009 PMID: 19800884.
- Iwakoshi-Ukena E, Ukena K, Takuwa-Kuroda K, Kanda A, Tsutsui K, Minakata H. Expression and distribution of octopus gonadotropin-releasing hormone in the central nervous system and peripheral organs of the octopus (Octopus vulgaris) by in situ hybridization and immunohistochemistry. J Comp Neurol. 2004; 477(3):310–23. PMID: <u>15305367</u>.
- Treen N, Itoh N, Miura H, Kikuchi I, Ueda T, Takahashi KG, et al. Mollusc gonadotropin-releasing hormone directly regulates gonadal functions: a primitive endocrine system controlling reproduction. General and comparative endocrinology. 2012; 176(2):167–72. Epub 2012/02/14. doi: <u>10.1016/j.ygcen.</u> <u>2012.01.008</u> PMID: <u>22326349</u>.
- Nagasawa K, Oouchi H, Itoh N, Takahashi KG, Osada M. In Vivo Administration of Scallop GnRH-Like Peptide Influences on Gonad Development in the Yesso Scallop, Patinopecten yessoensis. PLoS One. 2015; 10(6):e0129571. doi: <u>10.1371/journal.pone.0129571</u> PMID: <u>26030928</u>; PubMed Central PMCID: PMCPMC4451010.
- Hauser F, Grimmelikhuijzen CJ. Evolution of the AKH/corazonin/ACP/GnRH receptor superfamily and their ligands in the Protostomia. General and comparative endocrinology. 2014; 209:35–49. Epub 2014/07/25. doi: <u>10.1016/j.ygcen.2014.07.009</u> PMID: <u>25058364</u>.
- Johnson JI, Kavanaugh SI, Nguyen C, Tsai PS. Localization and functional characterization of a novel adipokinetic hormone in the mollusk, Aplysia californica. PLoS One. 2014; 9(8):e106014. doi: <u>10.1371/</u> journal.pone.0106014 PMID: 25162698; PubMed Central PMCID: PMCPMC4146582.
- Lindemans M, Liu F, Janssen T, Husson SJ, Mertens I, Gade G, et al. Adipokinetic hormone signaling through the gonadotropin-releasing hormone receptor modulates egg-laying in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2009; 106(5):1642–7. Epub 2009/01/24. [pii] doi: <u>10.1073/pnas.</u> <u>0809881106</u> PMID: <u>19164555</u>.
- Hauser F, Sondergaard L, Grimmelikhuijzen CJ. Molecular cloning, genomic organization and developmental regulation of a novel receptor from Drosophila melanogaster structurally related to gonadotropin-releasing hormone receptors for vertebrates. Biochemical and biophysical research communications. 1998; 249(3):822–8. Epub 1998/09/10. PMID: 9731220.
- Staubli F, Jorgensen TJ, Cazzamali G, Williamson M, Lenz C, Sondergaard L, et al. Molecular identification of the insect adipokinetic hormone receptors. Proc Natl Acad Sci U S A. 2002; 99(6):3446–51. Epub 2002/03/21. doi: <u>10.1073/pnas.052556499</u> PMID: <u>11904407</u>; PubMed Central PMCID: PMC122543.
- Rodet F, Lelong C, Dubos MP, Costil K, Favrel P. Molecular cloning of a molluscan gonadotropinreleasing hormone receptor orthologue specifically expressed in the gonad. Biochim Biophys Acta. 2005; 1730(3):187–95. Epub 2005/09/10. doi: 10.1016/j.bbaexp.2005.05.012 PMID: 16150500.
- Zhang L, Tello JA, Zhang W, Tsai PS. Molecular cloning, expression pattern, and immunocytochemical localization of a gonadotropin-releasing hormone-like molecule in the gastropod mollusk, Aplysia californica. General and comparative endocrinology. 2008; 156(2):201–9. Epub 2008/01/08. doi: S0016-6480(07)00401-7 [pii] doi: <u>10.1016/j.ygcen.2007.11.015</u> PMID: <u>18178211</u>.
- Iwakoshi E, Takuwa-Kuroda K, Fujisawa Y, Hisada M, Ukena K, Tsutsui K, et al. Isolation and characterization of a GnRH-like peptide from Octopus vulgaris. Biochemical and biophysical research communications. 2002; 291(5):1187–93. PMID: <u>11883942</u>.
- Kanda A, Takahashi T, Satake H, Minakata H. Molecular and functional characterization of a novel gonadotropin-releasing-hormone receptor isolated from the common octopus (Octopus vulgaris). Biochem J. 2006; 395(1):125–35. Epub 2005/12/22. [pii] doi: <u>10.1042/BJ20051615</u> PMID: <u>16367741</u>.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30(12):2725–9. Epub 2013/10/18. doi: <u>10.1093/molbev/mst197</u> PMID: <u>24132122</u>; PubMed Central PMCID: PMCPmc3840312.
- Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992; 8(3):275–82. PMID: <u>1633570</u>.
- Baines RA, Lange AB, Downer RG. Proctolin in the innervation of the locust mandibular closer muscle modulates contractions through the elevation of inositol trisphosphate. J Comp Neurol. 1990; 297 (4):479–86. doi: <u>10.1002/cne.902970402</u> PMID: <u>1974556</u>.

- Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987; 15(20):8125–48. PMID: <u>3313277</u>; PubMed Central PMCID: PMCPMC306349.
- Schutze J, Krasko A, Custodio MR, Efremova SM, Muller IM, Muller WE. Evolutionary relationships of Metazoa within the eukaryotes based on molecular data from Porifera. Proc Biol Sci. 1999; 266 (1414):63–73. doi: <u>10.1098/rspb.1999.0605</u> PMID: <u>10081159</u>; PubMed Central PMCID: PMCPMC1689648.
- Millar RP, Lu ZL, Pawson AJ, Flanagan CA, Morgan K, Maudsley SR. Gonadotropin-releasing hormone receptors. Endocr Rev. 2004; 25(2):235–75. Epub 2004/04/15. PMID: <u>15082521</u>.
- Cazzamali G, Saxild N, Grimmelikhuijzen C. Molecular cloning and functional expression of a Drosophila corazonin receptor. Biochemical and biophysical research communications. 2002; 298(1):31–6. PMID: <u>12379215</u>.
- Tsai PS, Zhang L. The emergence and loss of gonadotropin-releasing hormone in protostomes: orthology, phylogeny, structure, and function. Biol Reprod. 2008; 79(5):798–805. Epub 2008/07/11. [pii] doi: 10.1095/biolreprod.108.070185 PMID: 18614699.
- Ryan JF, Mazza ME, Pang K, Matus DQ, Baxevanis AD, Martindale MQ, et al. Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, Nematostella vectensis. PLoS One. 2007; 2(1):e153. doi: <u>10.1371/journal.pone.0000153</u> PMID: <u>17252055</u>; PubMed Central PMCID: PMCPMC1779807.
- Plachetzki D, Tsai PS, Kavanaugh SI, Sower SA. Ancient origins of metazoan gonadotropin-releasing hormone and their receptors revealed by phylogenomic analyses. General and comparative endocrinology. 2016. doi: <u>10.1016/j.ygcen.2016.06.007</u> PMID: <u>27318276</u>.
- Tatematsu K, Uchino K, Sezutsu H, Tamura T. Effect of ATG initiation codon context motifs on the efficiency of translation of mRNA derived from exogenous genes in the transgenic silkworm, Bombyx mori. Springerplus. 2014; 3:136. doi: <u>10.1186/2193-1801-3-136</u> PMID: <u>25674439</u>; PubMed Central PMCID: PMCPMC4320137.
- Roch GJ, Busby ER, Sherwood NM. Evolution of GnRH: diving deeper. General and comparative endocrinology. 2011; 171(1):1–16. Epub 2010/12/28. doi: <u>10.1016/j.ygcen.2010.12.014</u> PMID: <u>21185290</u>.
- **32.** Cheung TC, Hearn JP. Dimerizations of the wallaby gonadotropin-releasing hormone receptor and its splice variants. General and comparative endocrinology. 2005; 144(3):280–8. doi: <u>10.1016/j.ygcen.</u> 2005.06.008 PMID: <u>16102759</u>.
- Wang L, Bogerd J, Choi HS, Seong JY, Soh JM, Chun SY, et al. Three distinct types of GnRH receptor characterized in the bullfrog. Proc Natl Acad Sci U S A. 2001; 98(1):361–6. doi: <u>10.1073/pnas.</u> <u>011508498</u> PMID: <u>11120886</u>; PubMed Central PMCID: PMCPMC14595.
- Sakai T, Aoyama M, Kusakabe T, Tsuda M, Satake H. Functions of a GnRH receptor heterodimer of the ascidian, Ciona intestinalis. Acta Biol Hung. 2008; 59 Suppl:241–3. doi: <u>10.1556/ABiol.59.2008</u>. Suppl.34 PMID: 18652398.
- Sakai T, Aoyama M, Kusakabe T, Tsuda M, Satake H. Functional diversity of signaling pathways through G protein-coupled receptor heterodimerization with a species-specific orphan receptor subtype. Mol Biol Evol. 2010; 27(5):1097–106. doi: <u>10.1093/molbev/msp319</u> PMID: <u>20026483</u>.
- Millar RP, Pawson AJ, Morgan K, Rissman EF, Lu ZL. Diversity of actions of GnRHs mediated by ligand-induced selective signaling. Front Neuroendocrinol. 2008; 29(1):17–35. doi: <u>10.1016/j.yfme.</u> <u>2007.06.002</u> PMID: <u>17976709</u>; PubMed Central PMCID: PMCPMC2667102.
- Tello JA, Rivier JE, Sherwood NM. Tunicate gonadotropin-releasing hormone (GnRH) peptides selectively activate Ciona intestinalis GnRH receptors and the green monkey type II GnRH receptor. Endocrinology. 2005; 146(9):4061–73. Epub 2005/06/18. doi: 10.1210/en.2004-1558 PMID: 15961566.
- Gade G, Auerswald L. Mode of action of neuropeptides from the adipokinetic hormone family. General and comparative endocrinology. 2003; 132(1):10–20. PMID: 12765639.
- Yang J, Huang H, Yang H, He X, Jiang X, Shi Y, et al. Specific activation of the G protein-coupled receptor BNGR-A21 by the neuropeptide corazonin from the silkworm, Bombyx mori, dually couples to the G(q) and G(s) signaling cascades. J Biol Chem. 2013; 288(17):11662–75. doi: <u>10.1074/jbc.M112.</u> <u>441675</u> PMID: <u>23457297</u>; PubMed Central PMCID: PMCPMC3636857.
- Tello JA, Sherwood NM. Amphioxus: beginning of vertebrate and end of invertebrate type GnRH receptor lineage. Endocrinology. 2009; 150(6):2847–56. Epub 2009/03/07. [pii] doi: <u>10.1210/en.2009-0028</u> PMID: <u>19264870</u>.
- Roch GJ, Tello JA, Sherwood NM. At the transition from invertebrates to vertebrates, a novel GnRHlike peptide emerges in amphioxus. Mol Biol Evol. 2014; 31(4):765–78. doi: <u>10.1093/molbev/mst269</u> PMID: <u>24361996</u>; PubMed Central PMCID: PMCPMC3969558.