Identification, presence, and possible multifunctional regulatory role of invertebrate gonadotropin-releasing hormone/corazonin molecule in the great pond snail (*Lymnaea stagnalis*).

[1]István Fodor, [1]Zita Zrinyi, [2,3]Péter Urbán, [3]Róbert Herczeg, [4]Gergely Büki, [5]Joris M. Koene, [6]Pei-San Tsai and [1]Zsolt Pirger*

[1]Adaptive Neuroethology Research Group, Department of Experimental Zoology, Balaton Limnological Institute, Centre for Ecological Research, 8237 Tihany, Hungary
[2]Microbial Biotechnology Research Group, Szentágothai Research Centre, University of Pécs, 7624 Pécs, Hungary
[3]Bioinformatics Research Group, Szentágothai Research Centre, University of Pécs, 7624 Pécs, Hungary
[4]Department of Medical Genetics, Medical School, University of Pécs, 7624 Pécs, Hungary
[5]Department of Ecological Science, Faculty of Science, Vrije Universiteit, Amsterdam, the Netherlands
[6]Department of Integrative Physiology and Center for Neuroscience, University of Colorado, Boulder, CO 80309-0354, United States

*Address correspondence to Dr. Pirger Zsolt, Adaptive Neuroethology, Department of Experimental Zoology, Balaton Limnological Institute, Centre for Ecological Research, 8237, Tihany, Hungary
E-mail: pirger.zsolt@okologia.mta.hu

Abstract

In the last years, the interpretation of gonadotropin-releasing hormone (GnRH) neuropeptide superfamily has changed tremendously. One main driver is the investigation of functions and evolutionary lineage of previously identified molluscan GnRH molecules. Emerging evidence suggests not only reproductive, but also diverse biological effects of these molecules and proposes they should most likely be called corazonin (CRZ). Clearly, a more global understanding necessitates further exploration of species-specific functions and structure of invGnRH/CRZ peptides. Towards this goal, we have identified the full-length cDNA of invGnRH/CRZ peptide in an invertebrate model species, the great pond snail *Lymnaea stagnalis*, termed ly-GnRH/CRZ, and characterized the transcript and peptide distribution in the central nervous system (CNS) and peripheral organs. Our results are consistent with previous data that molluscan GnRHs are more related to CRZs and serve diverse functions. For this, our findings support the notion that peptides originally termed molluscan GnRH are multifunctional modulators and that nomenclature change should be taken into consideration.

Keywords: gonadotropin-releasing hormone [GnRH]/corazonin [CRZ], coding sequence, phylogenetic tree, peptide localization, mollusk, *Lymnaea stagnalis*

1. INTRODUCTION

The gonadotropin-releasing hormone (GnRH) is an ancient neuropeptide superfamily whose origin predates the protostome-deuterostome split (Lindemans et al., 2011; Plachetzki et al., 2016; Sakai et al., 2017). Currently, the GnRH superfamily consists of five families: 1) vertebrate GnRH; 2) adipokinetic hormone (AKH); 3) corazonin (CRZ); 4) AKH/CRZ-related peptide (ACP), and 5) invertebrate GnRH/CRZ (invGnRH/CRZ; previously called invGnRH) (Plachetzki et al., 2016; Sakai et al., 2017; Tsai, 2018). Through 700 million years of metazoan evolution, these peptides underwent several changes, with only the vertebrate GnRH becoming highly specialized in reproductive activation by stimulating the synthesis and release of luteinizing hormone and follicle-stimulating hormone (Holland and Sower, 2010). For invertebrates, in contrast, AKH, CRZ, ACP and invGnRH/CRZ are shown to be multifunctional modulators of different physiological and behavioral processes (Sakai et al., 2017). Among these peptides, the nomenclature of invGnRH/CRZ molecules has led to some confusion regarding their evolutionary lineage and function (Tsai, 2018). In other words, invGnRH/CRZ molecules, initially named invGnRH (e.g., oct-GnRH, ap-GnRH), may not assume a specialized reproductive role as seen in vertebrate GnRH. Indeed, emerging evidence suggests diverse biological effects of invGnRH/CRZ that include not only reproductive, but also motor, behavioral, and neuromodulatory functions (Iwakoshi et al., 2004; Minakata et al., 2009; Tsai et al, 2010; Sun and Tsai, 2011; Treen et al., 2012; Jung et al., 2014. Nagasawa et al., 2015; Kavanaugh and Tsai, 2016). A more global understanding of invGnRH/CRZ evolution necessitates further exploration of species-specific functions and structure of invGnRH/CRZ molecules.

The overarching goal of the present study was to identify and describe the invGnRH/CRZ molecule in a freshwater pond snail, *Lymnaea stagnalis*, named ly-GnRH/CRZ. For decades, *L. stagnalis* has been a widely used model species for the study of invertebrate neurobiology (Benjamin, 2008; Feng et al., 2009; Kemenes and Benjamin, 2009; Pirger et al., 2014, Rivi et al., 2019), neurodegeneration processes (Maasz et al., 2017, Ford L, et al., 2017), aging (Hermann et al., 2007; Pirger et al., 2014), environmental risk assessment (Bouetard et al., 2014; Charles et al., 2016; Pirger et al., 2018)(Zrinyi et al., 2017; Amorim et al., 2019), and the evolutionary consequence of hermaphroditism (reviewed in Koene, 2010; 2017). *L. stagnalis* possesses unique features such as a well-defined central nervous system (CNS), clearly-defined behaviors, and accessible anatomy that can facilitate the characterization of invGnRH/CRZ. To accomplish our goal, we first identified the coding sequence of prepro-ly-GnRH/CRZ. Subsequently, reverse-transcription (RT) polymerase chain reaction (PCR) and

immunohistochemistry (IHC) were performed to examine the anatomical distribution and cellular localization of ly-GnRH/CRZ transcript and peptide, respectively. Our data revealed the presence of an invGnRH/CRZ preprohormone in *L. stagnalis*. Further, the mRNA and peptide distribution patterns were consistent with the role of ly-GnRH/CRZ as a neuropeptide with diverse functions.

2. MATERIALS AND METHODS

2.1. Experimental animals

Adult (6 months old) specimens of the pond snail, *L. stagnalis*, originating from laboratory-bred stocks (ÖK BLI, Tihany, Hungary) were randomly selected for use in our experiments. Snails were kept in large holding tanks containing oxygenated low-copper artificial snail water at a constant temperature of 20°C (±1.5 °C) and on light:dark regime of 12 h:12 h. Snails were fed lettuce *ad libitum* three times a week. All procedures on snails were performed according to the protocols approved by the Scientific Committee of Animal Experimentation of the Balaton Limnological Institute (VE-I-001/01890-10/2013). Efforts were made to minimize the number of animals used in the experiments.

2.2. *In silico* searches, cDNA sequencing, and bioinformatics analysis

In silico searches of *L. stagnalis* CNS transcriptome and genome were conducted with TBLASTN (NCBI). First, the coding sequence of ap-GnRH/CRZ (GnRH/CRZ of *Aplysia californica*; NM_001204553.1) was used as the search query and identified a 272 bp long *L. stagnalis* cDNA fragment in the genome consortium data (referred to as Lsta_scaffold164; genome consortium publication in preparation), indicating the presence of GnRH/CRZ in *L. stagnalis*.

For RNA sample collection, the whole CNS was dissected from the snails (n=10) and homogenized using a TissueLyser LT (QIAGEN) in TRI reagent (#93289, Sigma-Aldrich). RNA was isolated with Direct-zol™ RNA MiniPrep (#R2050, Zymo Research) following manufacturer's instructions. The RNA was quantified by Qubit BR RNA Kit (#Q10211, ThermoFisher), and the quality was checked on Agilent Bioanalyzer 2100 using RNA 6000 Nano Kit (#5067-1511, Agilent).

Nanopore sequencing was used for identification of ly-GnRH/CRZ whole transcript. The library was prepared using cDNA-PCR Kit (#SQK-PCS108, Oxford Nanopore Technologies) according to the description of manufacturer. The sample was sequenced on a MinION device

with R9.4.1 flowcells (#FLO-MIN106). Base calling was performed using Guppy v3.2.2 software. Adapters were trimmed with Porechop v0.2.4 (Wick, 2018), and sequences with internal adapters, indicative of chimera reads, were also split with Porechop. Reads were assembled with CLC Genomics Workbench v12.0.3 software *de novo* pipeline (QIAGEN). Consensus sequence was called and manually corrected also within CLC Genomics Workbench. The identified ly-GnRH/CRZ mRNA sequence was submitted to the GenBank Nucleotide database with the accession number MN385595.

2.3. Phylogenetic analysis

The alignment used to generate the maximum likelihood tree consisted of 56 amino acid (AA) sequences of GnRH superfamily preprohormones retrieved from GenBank and the deduced AA sequence of ly-GnRH/CRZ preprohormone (Supplementary information). The alignment was performed using ClustalW with BLOSUM62 substitution matrix in Molecular Evolutionary Genetics Analysis v7 software (Kumar et al., 2016). Alignment was then analyzed to get the best fitting model, which was determined to be LG with gamma-distributed rates. Bootstrapping support for the tree was conducted with 1000 bootstrap replicates.

2.4. RT-PCR analysis of ly-GnRH/CRZ expression in central and peripheral tissues

Total RNA from the CNS (buccal, cerebral, pedal, pleural, parietal, and visceral ganglia) and peripheral organs (heart, and ovotestis) was isolated as presented above. After the RNA preparation, an additional DNase treatment was performed using TURBO DNA-free™ Kit (#AM1907, Thermo Fisher Scientific) following manufacturer's instructions.

The RevertAid H Minus First Strand cDNA Synthesis Kit (#K1631, Thermo Fisher Scientific) was used for RT, applying random hexamer primers and 3 µL volume from the total RNA sample.

PCR primers were designed for the identified coding region of ly-GnRH/CRZ using SnapGene® Viewer software (GSL Biotech, Chicago, IL, version 4.1.7). The applied primer set was as follows: the forward primer for ly-GnRH: 5'- CCT TGT CCT CCT GGC TGT AGT G - 3'; the reverse primer for ly-GnRH: 5'- GAT GTG GCC GGT GTA CGA TGG - 3' (Integrated DNA Technologies, Leuven, Belgium). As a control to ensure the quality of RT, actin was also amplified for all examined tissues. The applied primer set for actin was as follows: the forward primer: 5'- TCC CTT GAG AAG AGC TAC GAG C - 3'; and the reverse primer: 5'- GAG TTG TAG GTG GTT TCG TGG - 3' (Integrated DNA Technologies, Leuven, Belgium).

The PCR reaction was performed in 20 µL reaction volume at 95 °C for 4 min (initial denaturation) followed by 35 cycles of 95 °C for 30 sec (denaturation), 55 °C for 30 sec (annealing) and 72 °C for 10 sec (extension) by a T1 Thermocycler PCR device (Biometra®, Goettingen, Germany). PCR product was checked by agarose gel-electrophoresis using 2% gel (SeaKem LE agarose) and GeneRuler™ 100 bp Plus DNA ladder (#SM0328, Thermo Fisher Scientific).

2.5. Immunohistochemistry

CNS, heart, and ovotestis were dissected from individual snails and pinned out on a Sylgard-coated dish containing 4% paraformaldehyde in 0.1 M phosphate-buffered saline (PBS, pH=7.4) for overnight at 4°C. After washing with PBS, fixed tissues were cryoprotected in 20% glucose solution for 4 hours at room temperature and embedded into cryomatrix (#6769006, Thermo Scientific). A series of 12-14 µm-thick cryostat sections were cut, thaw-mounted onto gelatin-aluminium-coated slides, and processed for IHC as follows: 1) washing with PBS containing 0.25% TritonX-100 for 3 x 10 min; 2) incubation with a rabbit anti-ap-GnRH/CRZ antiserum (#AS203-2, EZbiolab; antigen was a synthetic undecapeptide: CNYHFSNGWYA-amide; Tsai et al., 2010) diluted 1:750 for 24 hours at 4°C; 3) PBS washing for 2 x15 min; 4) incubation with a donkey anti-rabbit IgG conjugated with NorthernLights™ NL557 (#NL004, R&D System) diluted 1:1000 for 2 hours at 4°C. After washing with modified Dulbeco's PBS (DPBS), the nuclei were stained with Hoechst 33342 (1 µg/mL in DPBS; #62249, Thermo Scientific) for 10 min at room temperature. The slides were washed with DPBS and cover-slipped with fluorescent mounting medium (#S3023, Dako). The stained tissues were analyzed with a Leica TCS SP8 DMI laser confocal microscope equipped with a transmitted light detector (BioMarker Ltd, Hungary). Because the primary antiserum was generated against ap-GnRH/CRZ conjugated with hemocyanin (KLH) from giant keyhole limpet (*Megathura crenulata*), to abolish cross-reaction to hemocyanin in tissues of *L. stagnalis* the antiserum was preadsorbed overnight against KLH (#H7017, Sigma). In some cases, the antibody was preadsorbed with 10 µM of synthetic ly-GnRH/CRZ (pQNYHFSNGWYA-amide) overnight at 4°C and applied to the sections.

3. RESULTS

3.1. Nucleotide sequence of ly-GnRH/CRZ preprohormone

Our *in silico* search in the *L. stagnalis* sequence data revealed a 272-bp positive hit (referred to as Lsta_scaffold164) containing a sequence of 32 nucleotides with a 100% match to the query,

thus identifying the potential GnRH/CRZ coding sequence. We termed this sequence ly-GnRH/CRZ. Using nucleotide sequencing, we determined the complete coding sequence. The full-length ly-GnRH/CRZ preprohormone spans 1225 bp (Fig. 1). Based on the data from scientific literature (Iwakoshi et al., 2002; Tsai et al., 2008; Tsai and Zhang 2008), we identified the signal peptide, the mature peptide with the signature tribasic cleavage site and an alpha amidation signal (QNYHFSNGWYAGKKR), and the ly-GnRH/CRZ-associated peptide (Fig. 1). As expected, the mature peptide sequence is highly conserved, while the signal peptide and GnRH/CRZ-associated peptide show high variability.

3.2. Phylogenetic analysis

To update the established evolutionary relationships of the GnRH superfamily with newly annotated sequences including ly-GnRH/CRZ, we performed a phylogenetic analysis. The results of this are presented in Fig. 2, and the segregated clusters from top to bottom are as follows:

1. Two clusters of AKHs from mollusks, arthropods, annelids, and nematodes. Based on data from the literature, all of molluscan, arthropod and nematode AKHs are true AKH molecules. Five manual corrections are required: *Crassostrea gigas* AKH belongs to molluscan AKH group, *Bombyx mori* ACP to the ACP group, *Branchiostoma floridae* GnRH to the GnRH group, while *Capitella teleta* and *Helobdella robusta* (annelids) AKHs are not true AKH molecules but referred to as AKH/GnRHs.
2. An arthropod ACP cluster. Two manual corrections should also be performed here; *Brachionus calyciflorus* AKH has the consensus sequence for arthropod AKH but not ACP cluster, and *Cionia intestinalis* GnRH belongs to GnRH cluster.
3. An arthropod CRZ cluster.
4. A GnRH/CRZ cluster, including the newly determined ly-GnRH/CRZ (arrowhead) and previously identified molluscan GnRH sequences, possesses both GnRH and CRZ properties.
5. A vertebrate GnRH cluster containing all used sequences from chordates and echinoderms.

3.3. Localization of ly-GnRH/CRZ transcript in central and peripheral tissues

RT-PCR revealed wide distribution of ly-GnRH/CRZ transcript (Fig. 3). The ly-GnRH/CRZ transcript (109 bp) can be detected in all CNS ganglia (CG, BG, PeG, PlG, PaG, VG) and peripheral organs examined, including the heart (H) and ovotestis (OT) (Fig. 3, top panel). Actin

(a housekeeping gene) amplification of RNA samples without RT was used as a negative control to test for genomic DNA contamination (Fig. 3, middle panel). The lower panel amplified actin (140 bp) to ensure RNA and RT quality. Negative controls including no RT (middle panel) and no template (NTC) controls did not yield positive signal.

3.4. Localization of ly-GnRH/CRZ peptide in central and peripheral tissues

IHC revealed the presence of ly-GnRH/CRZ peptide in all CNS ganglia (Fig. 4A). Large- (50-80 µm), medium- (30-50 µm) and small-sized (10-30 µm) ly-GnRH/CRZ-immunoreactive (ir) neurons were observed. The distribution of ly-GnRH/CRZ-ir neurons was highly variable in different ganglia (Fig. 4A). In the pleural (PlG) and parietal (PaG) ganglia, only a few ly-GnRH/CRZ-ir neurons were detected, but labeled neurons were more abundant in the buccal (BG), cerebral (CG), and pedal ganglia (PeG) (Fig. 4A). The BG contained several small- and medium-sized ly-GnRH/CRZ-ir neurons largely concentrated around the unlabeled B1 and B2, furthermore, around the immunopositive large-sized B3 and B4 feeding motoneurons (Fig. 4A and B). In the CG, small-sized ly-GnRH/CRZ-ir neurons were detected in the metacerebrum located close to the ventral lobe (vl), in the anterior lobe (al), and a group of medium-sized ly-GnRH/CRZ-ir neurons were observed in the A-cluster (Acl) (Fig. 4A and C). Furthermore, caudo-dorsal cells (CDCs), the reproductive neuroendocrine cells, also displayed immunoreactivity (Fig. 4A and C). The PeG contained several small- and medium-sized immunopositive neurons in the A- (Acl) and E (Ecl) -clusters. The left serotonergic giant cell (LPeD1) and the functionally identified right dopaminergic giant neuron (RPeD1) were also immunopositive (Fig 4A and D). Lastly, some medium- and large-sized ly-GnRH/CRZ-ir neurons were observed in the J cells and M-cluster (Mcl) of the visceral ganglion (VG), and in the B-cluster (Bcl) of the PaG (Fig. 4A and E). Generally, a strong presence of ly-GnRH/CRZ-immunoreactivity was detected in the neuropil regions throughout the CNS (Fig. 4C-E).

Among peripheral tissues, we have investigated the heart and the ovotestis. Confocal projections showed abundant ly-GnRH/CRZ-ir fibers on the surface of the heart atria running perpendicularly to the longitudinal axis of the muscle fibers; however, the heart muscle fibers themselves were not immunopositive (Fig. 5A1 and A2). In ovotestis, immunoreactivity was not observed (Fig. 5B).

The immunostaining was completely abolished when the antiserum was preadsorbed with 10 µM ly-GnRH/CRZ (not shown).

4. DISCUSSION

Previous studies suggested that invGnRH/CRZ molecules, initially named invGnRH, may not assume a specialized reproductive function (Minakata et al., 2009; Tsai et al., 2010; Jung et al., 2014) and may need to be reclassified as CRZ (Hauser and Grimmelikhuijzen, 2014; Plachetzki et al., 2016; Tsai, 2018). The main goal of the present study was to further test this notion by identifying and investigating the invGnRH/CRZ molecule in *L. stagnalis*. The presence of a GnRH-like molecule in *L. stagnalis* has previously been hypothesized (reviewed in Koene, 2010), but remained unconfirmed until the current report. The identified ly-GnRH/CRZ preprohormone sequence contains the mature peptide with the signature tribasic cleavage site and an alpha amidation signal (QNYHFSNGWYAGKKR). However, the predicted mature peptide sequence awaits validation by mass spectrometry. Based on the literature (Nagasawa et al., 2015), we hypothesize that two forms of ly-GnRH/CRZ may be derived from a single prohormone sequence: a mature amidated undecapeptide (pQNYHFSNGWYA-amide) and a non-amidated precursor dodecapeptide (pQNYHFSNGWYAG-OH).

Our phylogenetic analysis presents two challenges commonly seen in peptide hormone analysis. Firstly, the relatively short GnRH superfamily prepropeptide sequences (58–219 AA) contribute to the lower bootstrap values. Secondly, the conserved mature peptide sequence needed for biological activity (e.g., receptor binding) can cause incorrect clustering (Hauser and Grimmelikhuijzen, 2014). Nevertheless, the phylogenetic tree correlated well with previous studies (Hauser and Grimmelikhuijzen, 2014; Plachetzki et al., 2016; Tsai, 2018). The GnRH/CRZ branching suggests sequences originally termed molluscan GnRHs are no more related to vertebrate GnRH than to AKH, ACP or CRZ.

This is the first study to demonstrate the occurrence and distribution of a snail GnRH/CRZ in the CNS and periphery. Our RT-PCR results show that ly-GnRH/CRZ is ubiquitously present in all central and peripheral tissues examined. However, IHC data show that the simultaneous presence of ly-GnRH/CRZ transcript and peptide occurs predominantly in the CNS. For example, although the ly-GnRH/CRZ transcript can be found in the heart and ovotestis (Fig. 3), the peptide was not observed within the heart muscle proper and absent in the ovotestis (Fig. 5). Similar observations were made in *A. californica* (Tsai et al., 2010; Jung et al., 2014), but the reason for the transcript/peptide mismatch remains to be determined. Proposed explanations include peripheral peptide levels below IHC detection, rapid peripheral peptide turnover or transport from the source tissue, and alternative post-translational processes in the periphery (Jung et al., 2014).

In contrast to *A. californica* in which ap-GnRH/CRZ peptide was found only in pedal and cerebral ganglia (Jung et al., 2014), ly-GnRH/CRZ peptide was detected by IHC in all central

ganglia of *L. stagnalis*, including functionally identified neurons. As such, the possible functions of ly-GnRH/CRZ can be inferred from the distribution and production sites. The B3 and B4 motoneurons in BG (Fig. 4A and B), involved in the execution of feeding (Brierley et al., 1997), showed ly-GnRH/CRZ-ir, suggesting a possible role of ly-GnRH/CRZ in this activity. Furthermore, CV cells in vl of CG (Fig. 4A and C) are well-known for feeding control via the lip motoneurons (McCrohan, 1984), suggesting a role of this neuropeptide in the motor control of feeding. Interestingly, egg-laying regulatory CDCs of cerebral ganglia also showed ly-GnRH/CRZ-ir (Fig. 4A and C). This contrasts with findings that bag cells of *A. californica*, which are homologous to CDCs and also responsible for egg laying (Conn and Kaczmarek, 1989), are negative for ap-GnRH/CRZ (Zhang et al., 2008). In the PeG, ly-GnRH/CRZ-containing neurons overlap with that of the pedal motor serotonergic neurons (A cluster, Fig. 4A and D) involved in the control of foot cilia (McKenzie et al., 1987), thus predicting a role of ly-GnRH/CRZ in locomotion. The suggested roles of ly-GnRH/CRZ in feeding and locomotion are consistent with previous findings showing ap-GnRH/CRZ modulates these activities in *A. californica* (Tsai et al., 2010; Jung et al., 2014). Furthermore, RPeD1, known to be responsible for respiration and heart control (Benjamin, 2008, Scholarpedia), is immunopositive. Since ly-GnRH/CRZ-ir J-cells and M-cluster of the VG are also involved in heart control and respiration, the peptide may also play a role in these processes. Supporting a cardio-respiratory role of ly-GnRH/CRZ, ly-GnRH/CRZ-ir fibers are present on the surface of heart atria muscles (Fig. 5A1 and A2). In this respect, it is interesting to note that CRZ molecule was first discovered and named for its ability to stimulate the heart of the American cockroach (*Periplaneta americana*) (Veenstra, 1989). The first described invertebrate GnRH/CRZ molecule, the oct-GnRH/CRZ, was also identified in the course of bioassay screenings of peptide fractions with cardio-acceleratory effects (Iwakoshi et al., 2002). Additional evidence supporting a multifunctional modulatory role of ly-GnRH/CRZ is the observation that some of ly-GnRH/CRZ-ir neurons overlaps with identified serotonergic (e.g., CV cells, A cluster) and dopaminergic (e.g., RPeD1) neurons responsible for feeding, locomotion, heart control, and respiration (McKenzie et al., 1987; Benjamin and Winlow, 1981; Syed N.I. et al., 1990; Benjamin and Kemenes 2013; Benjamin 2008).

In summary, we conclude that ly-GnRH/CRZ, just like ap-GnRH/CRZ, is a multifunctional neuropeptide involved in diverse central functions including feeding, locomotion, heart control, and reproduction. These finding can be confirmed in the future with physiological studies using purified or synthetic ly-GnRH/CRZ peptide. Our results further suggest the term “invGnRH” (e.g., oct-GnRH, ap-GnRH) should be revised to avoid the connotation that these peptides serve

specialized reproductive function similar to vertebrate GnRHs (Iwakoshi et al., 2004; Minakata et al., 2009; Tsai et al, 2010; Sun and Tsai, 2011; Treen et al., 2012; Jung et al., 2014. Nagasawa et al., 2015; Kavanaugh and Tsai, 2016). Overall, our study reports the presence and peptide distribution of a second invGnRH/CRZ in a gastropod mollusk and adds to a growing body of literature supporting a multifunctional role of invertebrate GnRH/CRZ peptides.

Funding

This work was supported by National Brain Project (No. 2017-1.2.1-NKP-2017-00002), Bolyai Foundation (No. BO/00952/16/8), and National Science Foundation (IOS-1352944).

Figure legends

Figure 1 – Nucleotide and deduced AA sequences of ly-GnRH/CRZ preprohormone (Genbank Accession #MN385595). The nucleotides (upper row) and AAs (lower row) are numbered accordingly. The putative signal peptide region is underlined (solid). The putative mature ly-GnRH/CRZ peptide with an alpha amidation signal is boxed, and the KKR tribasic cleavage site is shaded. The putative ly-GnRH/CRZ-associated peptide is underlined (dashed). The asterisk (*) denotes the stop codon. The nucleotides corresponding to the polyadenylation signal (ATTAAA) are double-underlined.

Figure 2 – Phylogenetic analysis of CRZ, GnRH, GnRH/CRZ, AKH, and ACP preprohormone sequences derived from protostomes including mollusks, annelids, nematodes and arthropods, and deuterostomes including echinoderms and chordates. The bootstrap values (%) from 1000 replicas are indicated at each branch point, the analysis involved 56 sequences. Sequences indicated by an asterisk belong to a neuropeptide group, which is different from the ones in which they cluster in the analysis.

Figure 3 – PCR of cDNA synthesized from the CNS (buccal, cerebral, pedal, pleural, parietal, and visceral ganglia) and peripheral tissues (heart, ovotestis). Top panel: PCR of ly-GnRH/CRZ. Middle panel: PCR of ly-actin using RNA samples that have not been reverse transcribed to check for genomic DNA contamination. Bottom panel: PCR of ly-actin to check the quality of RNA samples. The sizes of PCR products are indicated on the right. CG: cerebral ganglia; BG: buccal ganglia; PeG: pedal ganglia; PlG: pleural ganglia; PaG: parietal ganglia; VG: visceral ganglion; H: heart; OT: ovotestis; NTC: no template negative control.

Figure 4 – Schematic drawing and representative micrographs showing ly-GnRH/CRZ-ir neuronal elements in different ganglia of CNS. (A) Schematic map (dorsal view) of the distribution of ly-GnRH/CRZ-ir cells in the buccal (BG), cerebral (CG), pedal (PeG), pleural (PlG), parietal (PaG), and visceral (VG) ganglia of *L. stagnalis.* ly-GnRH/CRZ-ir neurons are indicated with black symbols, while outlined areas (black dashed lines) indicate unlabeled but identified (e.g., CGC, VD1) large neurons providing reference points for anatomical orientation. (B) In the right BG small- and medium-sized unidentified ly-GnRH/CRZ-ir cells (arrows), and large identified immunopositive neurons (stars, B3, B4) can be observed. (C) In the left CG several identified (CDCs, arrowheads) and unidentified (arrows) cell bodies show

positive immunosignal. (D, E) Several immunopositive small (10-30 µm), middle (30-50 µm), and large (50-80 µm) neurons (arrows) can be observed in the PeG (A- and E-clusters), VG (M-cluster), and right PaG (B-cluster). Furthermore, intensive immunostaining in the neuropiles, connectives and peripheral nerves can be observed. Abbreviations: dbn, dorsobuccal nerve; pbn, postbuccal nerve; cbc, cerebro-buccal connective; mln, median lip nerve; sln, superior lip nerve; vl, ventral lobe; al, anterior lobe; cc, cerebral commissure; CDCs, caudo-dorsal cells; spn, superior pedal nerve; mpn, median pedal nerve; ipn, inferior pedal nerve; Acl, A-cluster; Ecl, E-cluster; SC, statocyst; lpan, left parietal nerve, cpn, cutaneous pillar nerve; an, anal nerve; in, intestinal nerve; gn, genital nerve; irpan, internal right parietal nerve; erpan, external right parietal nerve; Mcl, M-cluster; Bcl, B-cluster; B1-B4, identified buccal feeding motoneurons; db, dorsal body; np, neuropile; pn, pedal nerve; cpc, cerebro-pedal connective; plpc, pleuro-pedal connective. Scale bars=100 µm.

Figure 5 – Representative ly-GnRH/CRZ IHC of heart atria (A1-A2) and ovotestis (B). Rich ly-GnRH/CRZ-ir fiber bundles (arrows) can be seen on the surface of heart atria muscles. A2 represents a higher magnification detail of the atria muscle. Immunolabeled axon processes (arrows) running perpendicular the longer, longitudinal axis of the muscle fibers. Scale bar=25 µm on A1 and 10 µm on A2. (B) The ovotestis (OT), which is integrated in the hepatopancreatic acini (Hep), is devoid of positive signal. Abbreviations: m, muscle fibers; ge, germ epithelium; po, primary oocyte; so, secondary oocyte; mo, mature oocyte; asterisks, follicular cells; arrows, Sertoli cells; arrowheads, spermatozoon.

References

Benjamin, P.R., 2008. *Lymnaea*. Scolarpedia, p. 4124.

Benjamin, P.R., Kemenes, G., 2010. *Lymnaea* learning and memory. Scolarpedia, p. 4247.

Bigot, L., Zatylny-Gaudin, C., Rodet, F., Bernay, B., Boudry, P., Favrel, P., 2012. Characterization of GnRH-related peptides from the Pacific oyster *Crassostrea gigas*. Peptides 34, 303-310.

Bouetard, A., Cote, J., Besnard, A.L., Collinet, M., Coutellec, M.A., 2014. Environmental versus anthropogenic effects on population adaptive divergence in the freshwater snail *Lymnaea stagnalis*. PLoS One 9.

Charles, S., Ducrot, V., Azam, D., Benstead, R., Brettschneider, D., De Schamphelaere, K., Filipe Goncalves, S., Green, J.W., Holbech, H., Hutchinson, T.H., Faber, D., Laranjeiro, F., Matthiessen, P., Norrgren, L., Oehlmann, J., Reategui-Zirena, E., Seeland-Fremer, A., Teigeler, M., Thome, J.P., Tobor Kaplon, M., Weltje, L., Lagadic, L., 2016. Optimizing the design of a reproduction toxicity test with the pond snail *Lymnaea stagnalis*. Regul. Toxicol. Pharmacol. 81, 47-56.

Di Cristo, C., De Lisa, E., Di Cosmo, A., 2009. Control of GnRH expression in the olfactory lobe of *Octopus vulgaris*. Peptides 30, 538-544.

El Filali, Z., de Boer, P.A., Pieneman, A.W., de Lange, R.P., Jansen, R.F., Ter Maat, A., van der Schors, R.C., Li, K.W., van Straalen, N.M., Koene, J.M., 2015. Single-cell analysis of peptide expression and electrophysiology of right parietal neurons involved in male copulation behavior of a simultaneous hermaphrodite. Invert. Neurosci. 15, 7.

Feng, Z.P., Zhang, Z., van Kesteren, R.E., Straub, V.A., van Nierop, P., Jin, K., Nejatbakhsh, N., Goldberg, J.I., Spencer, G.E., Yeoman, M.S., Wildering, W., Coorssen, J.R., Croll, R.P., Buck, L.T., Syed, N.I., Smit, A.B., 2009. Transcriptome analysis of the central nervous system of the mollusc *Lymnaea stagnalis*. BMC Genomics. 10, 451.

Fernald, R.D., White, R.B., 1999. Gonadotropin-releasing hormone genes: Phylogeny, structure, and functions. Front. Neuroendocrinol. 20, 224-240.

Gasteiger, E., Hoogland, C., Gattiker, A., S., D., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. The Proteomics Protocols Handbook. Humana Press.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95-98.

Kanda, A., Takahashi, T., Satake, H., Minakata, H., 2006. Molecular and functional characterization of a novel gonadotropin-releasing-hormone receptor isolated from the common octopus (*Octopus vulgaris*). Biochem. J. 395, 125-135.

Kemenes, G., Benjamin, P.R., 2009. *Lymnaea*. Curr. Biol. 19, R9-11.

Koene, J.M., 2010. Neuro-endocrine control of reproduction in hermaphroditic freshwater snails: mechanisms and evolution. Front. Behav. Neurosci. 4, 167.

Koene, J.M., 2017. Sex determination and gender expression: Reproductive investment in snails. Mol. Reprod. Dev. 84, 132-143.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870-1874.

Lindemans, M., Janssen, T., Beets, I., Temmerman, L., Meelkop, E., Schoofs, L., 2011. Gonadotropin-releasing hormone and adipokinetic hormone signaling systems share a common evolutionary origin. Front. Endocrinol. 2, 16.

Lindemans, M., Liu, F., Janssen, T., Husson, S.J., Mertens, I., Gade, G., Schoofs, L., 2009. Adipokinetic hormone signaling through the gonadotropin-releasing hormone receptor modulates egg-laying in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA. 106, 1642-1647.

Nagasawa, K., Osugi, T., Suzuki, I., Itoh, N., Takahashi, K.G., Satake, H., Osada, M., 2015. Characterization of GnRH-like peptides from the nerve ganglia of Yesso scallop, *Patinopecten yessoensis*. Peptides 71, 202-210.
Pirger, Z., Zrinyi, Z., Maasz, G., Molnar, E., Kiss, T., 2018. Pond snail reproduction as model in the environmental risk assesment: Reality and doubts, in: Ray, S. (Ed.), Biological Resources of Water. IntechOpen, London, pp. 33-53.
Rodet, F., Lelong, C., Dubos, M.P., Costil, K., Favrel, P., 2005. Molecular cloning of a molluscan gonadotropin-releasing hormone receptor orthologue specifically expressed in the gonad. Biochim. Biophys. Acta. 1730, 187-195.
Sakai, T., Shiraishi, A., Kawada, T., Matsubara, S., Aoyama, M., Satake, H., 2017. Invertebrate gonadotropin-releasing hormone-related peptides and their receptors: An update. Front. Endocrinol. 8, 217.
Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques. 28, 1102, 1104.
Straub, V.A., Staras, K., Kemenes, G., Benjamin, P.R., 2002. Endogenous and network properties of *Lymnaea* feeding central pattern generator interneurons. J. Neurophysiol. 88, 1569-1583.
Sun, B., Tsai, P.S., 2011. A gonadotropin-releasing hormone-like molecule modulates the activity of diverse central neurons in a gastropod mollusk, *Aplysia californica*. Front. Endocrinol. 2, 36.
Tajima, F., 1993. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135, 599-607.
Treen, N., Itoh, N., Miura, H., Kikuchi, I., Ueda, T., Takahashi, K.G., Ubuka, T., Yamamoto, K., Sharp, P.J., Tsutsui, K., Osada, M., 2012. Mollusc gonadotropin-releasing hormone directly regulates gonadal functions: a primitive endocrine system controlling reproduction. Gen. Comp. Endocrinol. 176, 167-172.
Tsai, P.S., 2018. Gonadotropin-releasing hormone by any other name would smell as sweet. Gen. Comp. Endocrinol. 264, 58-63.
Tsai, P.S., Sun, B., Rochester, J.R., Wayne, N.L., 2010. Gonadotropin-releasing hormone-like molecule is not an acute reproductive activator in the gastropod, *Aplysia californica*. Gen. Comp. Endocrinol. 166, 280-288.
Tsai, P.S., Zhang, L., 2008. The emergence and loss of gonadotropin-releasing hormone in protostomes: orthology, phylogeny, structure, and function. Biol. Reprod. 79, 798-805.
Young, K.G., Chang, J.P., Goldberg, J.I., 1999. Gonadotropin-releasing hormone neuronal system of the freshwater snails *Helisoma trivolvis* and *Lymnaea stagnalis*: possible involvement in reproduction. J. Comp. Neurol. 404, 427-437.
Zhang, L., Tello, J.A., Zhang, W., Tsai, P.S., 2008. Molecular cloning, expression pattern, and immunocytochemical localization of a gonadotropin-releasing hormone-like molecule in the gastropod mollusk, *Aplysia californica*. Gen. Comp. Endocrinol.156, 201-209.

```
tttgacaagaaagttgtcggtgtctttgtgtttctgttggtgctgatattgctggggccg    60
tcaaaacttcttgtttagttccgaatactttttagactgatcacctttggtttttttt    120
cgtttcgcctcacaccagcagaaggtcacagcaatcagcaacatgacgtccagcaatctg    180
                                            M  T  S  S  N  L      6
atgtccactttactggtgcttatccttgtcctcctggctgtagtgcacaccacaacagca    240
 M  S  T  L  L  V  L  I  L  V  L  L  A  V  V  H  T  T  T  A     26
caaaactaccacttcagcaacggctggtacgccggcaagaagcgctcttccccgtcgtac    300
 Q  N  Y  H  F  S  N  G  W  Y  A  G  K  K  R  S  S  P  S  Y     46
accggccacatcattggtgacagtagcagtgaggcggtgagaggcccaaggacggcagat    360
 T  G  H  I  I  G  D  S  S  S  E  A  V  R  G  P  R  T  A  D     66
gaggcttgcaatatcagaccggaggccgttctacttatcaacagaataatacaggaagaa    420
 E  A  C  N  I  R  P  E  A  V  L  L  I  N  R  I  I  Q  E  E     86
gtggcgagaatccaaaaagtgtgtactgctgagactccaagtggtctgagggaaatcctc    480
 V  A  R  I  Q  K  V  C  T  A  E  T  P  S  G  L  R  E  I  L    106
gagaatgctgctacaagattggactcggagtcgaaatggtaaaacactccaagaggagag    540
 E  N  A  A  T  R  L  D  S  E  S  K  W  *                      119
acgaagtcgaggagaccttgacagggtagacagtttctcttgcataaggatatagcaaaa    600
gatattacgggcaaacaaggaactaaagcagatagtgtgctgtccttaactcccgtgaat    660
tacaccatattttacgtaggttccaatcatcactcacgttctaaagacagacattcaata    720
gacacagcactgaacttattttgtacgatgttctgaacagttctgaacgctgaatttatc    780
aatgtaaagaaatacgtaatttttgccagataaggactttgtgagagggggcacttgcac    840
gtctgtgtggtgagaatattgcgttgtggtagaggttctatttcaaaatgaaatgttaca    900
aaatgtatcagatctattggatctcatttagaatagtgaaatctgcagctagagaactga    960
ggtttattaattttacatgtacgctttacccttgtgaaaaaaaagttttcccttagcacc   1020
tatacattagattgtagactcattggtagaattttggtagaatgtgtttaattttggcat   1080
agtctgtttcagcttaaaaatttactcttcgaaaactatttacaccaaaaatggaacaaa   1140
tgtgtatttctcgccttctgtcatcttctatagaaaattgcatgagaataaagtcagaac   1200
caccaaaaaaaaaaaaaaaaaaaaa                                      1225
```

Fig1.

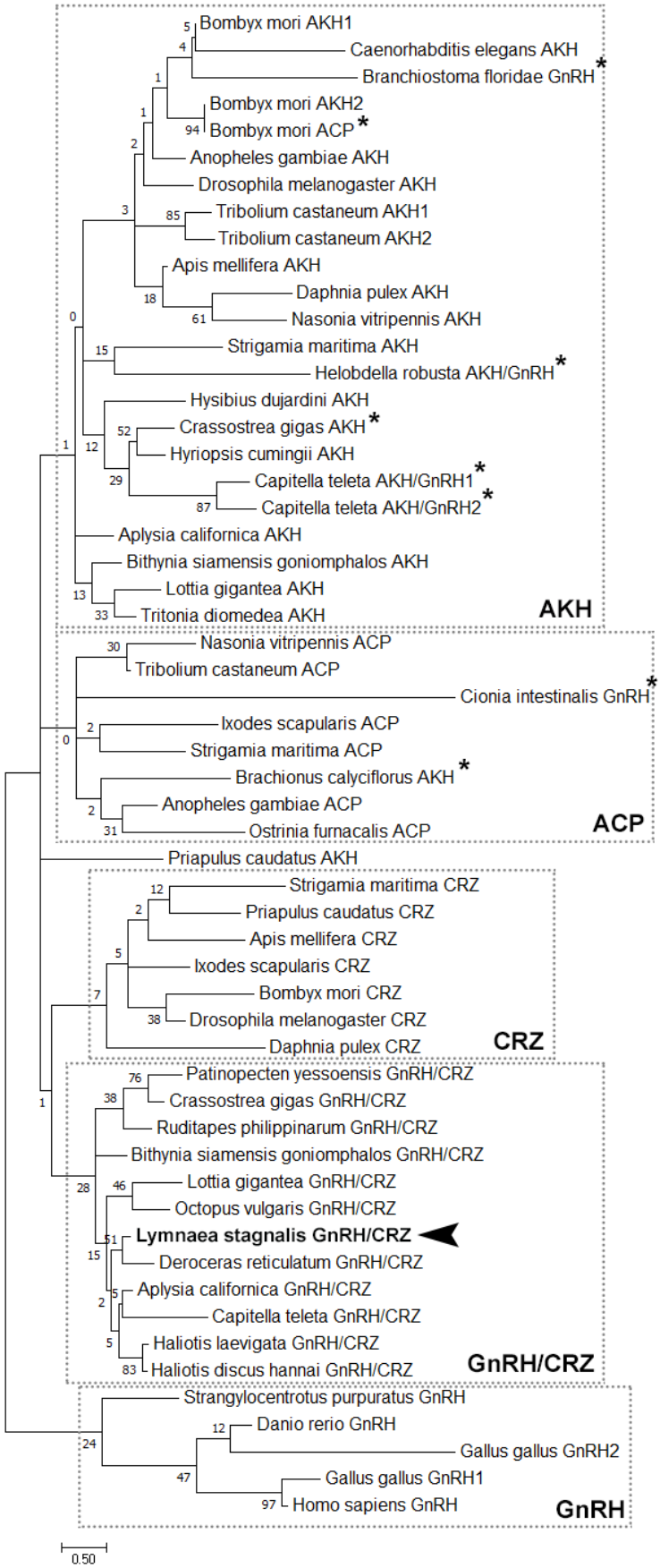
Bombyx mori AKH1
Caenorhabditis elegans AKH
Branchiostoma floridae GnRH
Bombyx mori AKH2
Bombyx mori ACP
Anopheles gambiae AKH
Drosophila melanogaster AKH
Tribolium castaneum AKH1
Tribolium castaneum AKH2
Apis mellifera AKH
Daphnia pulex AKH
Nasonia vitripennis AKH
Strigamia maritima AKH
Helobdella robusta AKH/GnRH
Hysibius dujardini AKH
Crassostrea gigas AKH
Hyriopsis cumingii AKH
Capitella teleta AKH/GnRH1
Capitella teleta AKH/GnRH2
Aplysia californica AKH
Bithynia siamensis goniomphalos AKH
Lottia gigantea AKH
Tritonia diomedea AKH
AKH
Nasonia vitripennis ACP
Tribolium castaneum ACP
Cionia intestinalis GnRH
Ixodes scapularis ACP
Strigamia maritima ACP
Brachionus calyciflorus AKH
Anopheles gambiae ACP
Ostrinia furnacalis ACP
ACP
Priapulus caudatus AKH
Strigamia maritima CRZ
Priapulus caudatus CRZ
Apis mellifera CRZ
Ixodes scapularis CRZ
Bombyx mori CRZ
Drosophila melanogaster CRZ
Daphnia pulex CRZ
CRZ
Patinopecten yessoensis GnRH/CRZ
Crassostrea gigas GnRH/CRZ
Ruditapes philippinarum GnRH/CRZ
Bithynia siamensis goniomphalos GnRH/CRZ
Lottia gigantea GnRH/CRZ
Octopus vulgaris GnRH/CRZ
Lymnaea stagnalis GnRH/CRZ
Deroceras reticulatum GnRH/CRZ
Aplysia californica GnRH/CRZ
Capitella teleta GnRH/CRZ
Haliotis laevigata GnRH/CRZ
Haliotis discus hannai GnRH/CRZ
GnRH/CRZ
Strangylocentrotus purpuratus GnRH
Danio rerio GnRH
Gallus gallus GnRH2
Gallus gallus GnRH1
Homo sapiens GnRH
GnRH
0.50

Fig2.

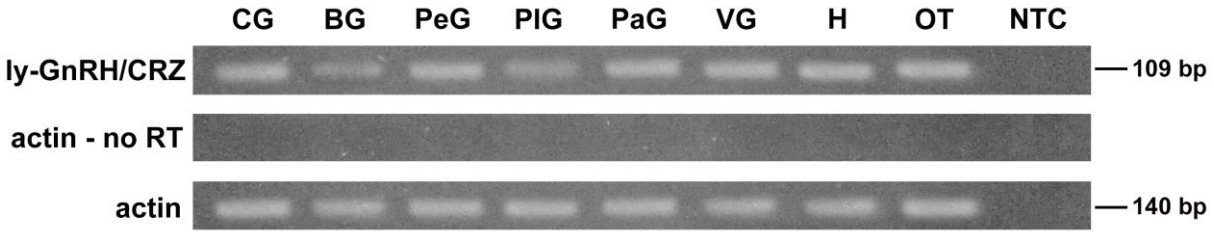
CG
BG
PeG
PIG
PaG
VG
H
OT
NTC
ly-GnRH/CRZ
109 bp
actin - no RT
actin
140 bp

Fig3.

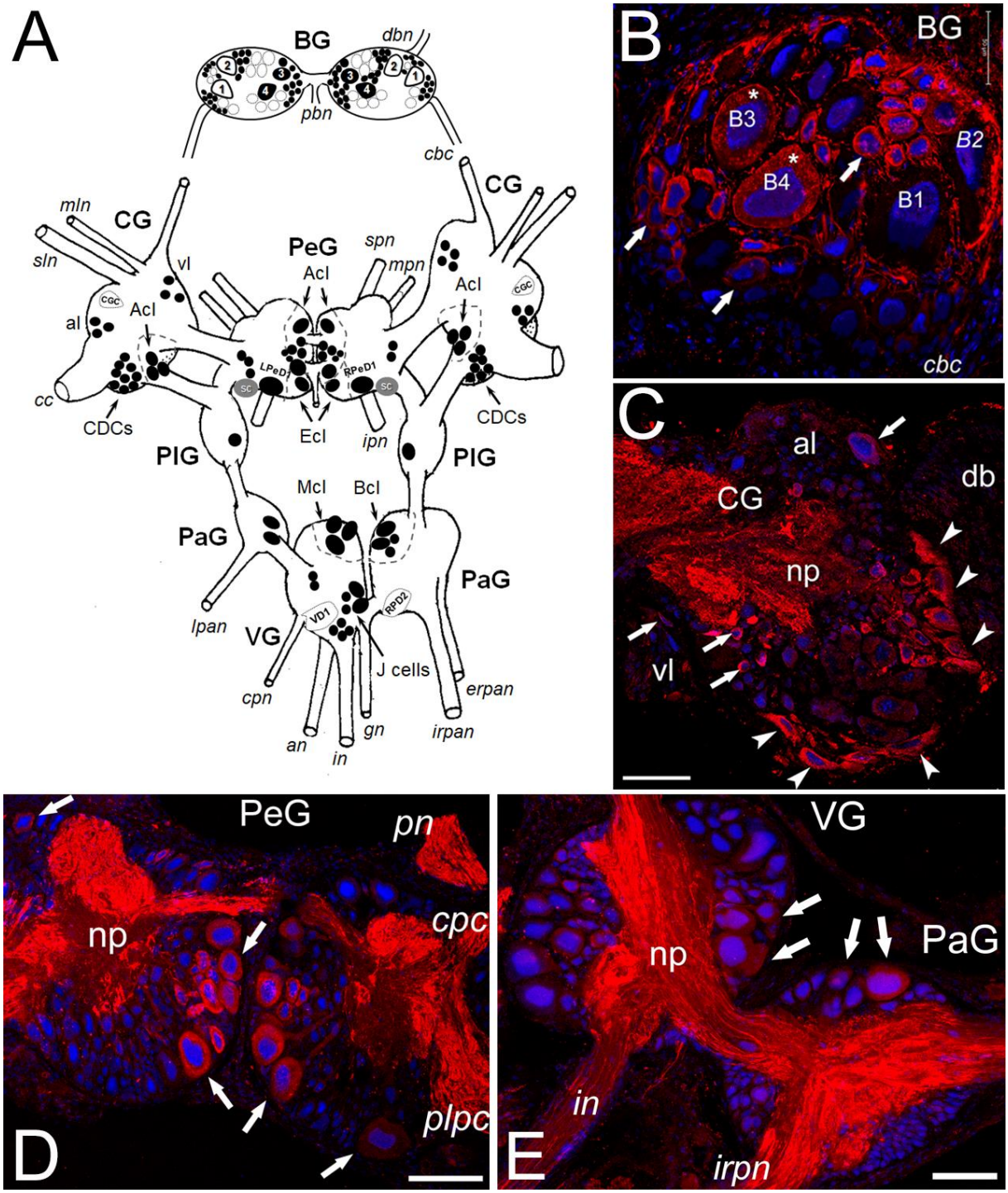
A
BG
dbn
pbn
cbc
CG
mln
sln
vl
PeG
spn
mpn
AcI
al
CGC
cc
CDCs
LPeD1
RPeD1
SC
EcI
ipn
PIG
McI
BcI
PaG
lpan
VG
VD1
RPD2
J cells
cpn
erpan
an
in
gn
irpan
B
B3
B4
B2
B1
C
db
np
D
pn
cpc
plpc
E

Fig4.

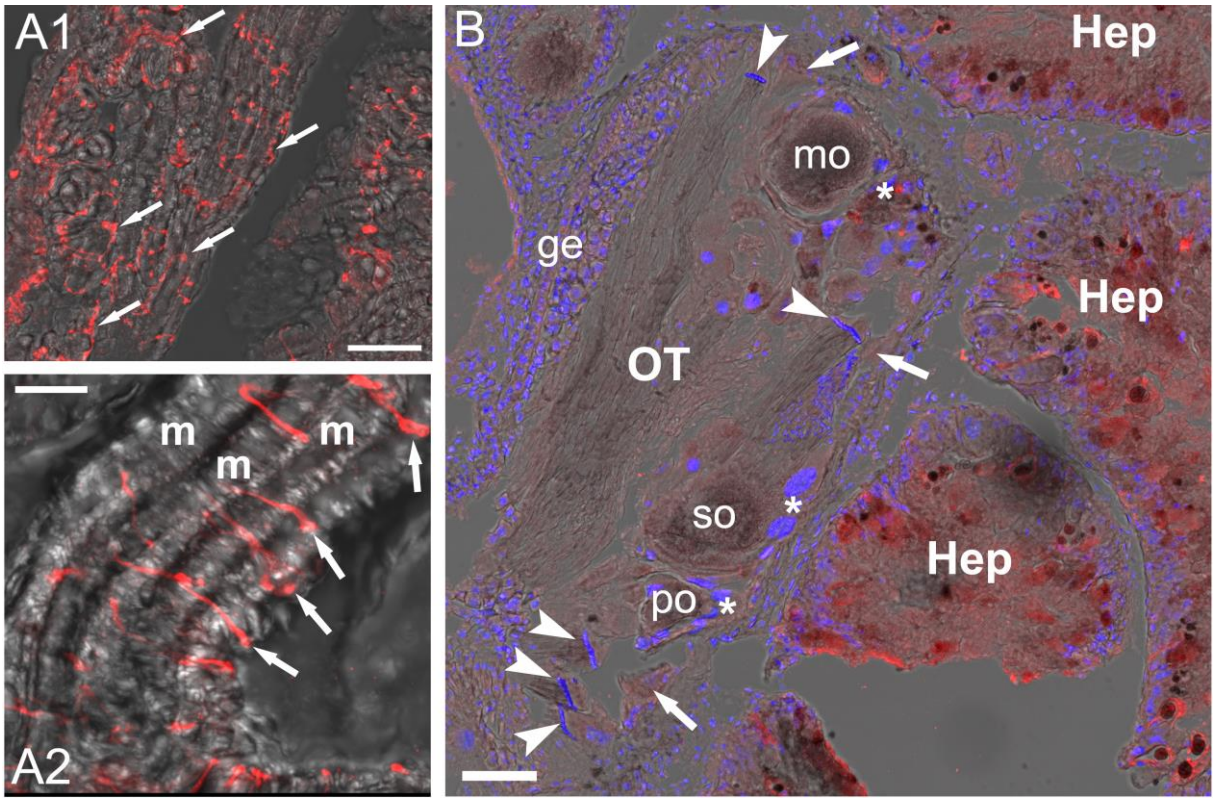

Fig5.