Gene 557 (2015) 28–34

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

The eyestalk transcriptome of red swamp crayfish *Procambarus clarkii*

Chiara Manfrin [a], Moshe Tom [b], Gianluca De Moro [a], Marco Gerdol [a], Piero Giulio Giulianini [a], Alberto Pallavicini [a,*]

[a] Department of Life Sciences, University of Trieste, Via Giorgieri 5, Trieste 34127, Italy
[b] Israel Oceanographic and Limnological Research, P.O.B 8030, Haifa 31080, Israel

ARTICLE INFO

Article history:
Received 15 July 2014
Received in revised form 5 November 2014
Accepted 1 December 2014
Available online 3 December 2014

Keywords:
Procambarus clarkii
Red swamp crayfish
Transcriptome
Eyestalk
Illumina sequencing
Neuroendocrinology
Vision

ABSTRACT

The red swamp crayfish (*Procambarus clarkii*, Girard 1852) is among the most economically important freshwater crustacean species, and it is also considered one of the most aggressive invasive species worldwide. Despite its commercial importance and being one of the most studied crayfish species, its genomic and transcriptomic layout has only been partially studied. Illumina RNA-sequencing was applied to characterize the eyestalk transcriptome and identify its most characterizing genes. A collection of 83,170,732 reads from eyestalks was obtained using Illumina paired-end sequencing technology. A de novo assembly was performed with the Trinity assembly software generating 119,255 contigs (average length of 1007 bp) and identifying the first sequenced transcriptome in this species.

The eyestalk is a major site for the production of neurohormones and controls a variety of physiological functions such as osmotic regulation, molting, epidermal color patterns and reproduction. Hence, its transcriptomic characterization is interesting and potentially instrumental to the elucidation of genes which have not been comprehensively described yet. Moreover, the availability of such a large amount of information supported the characterization of molecular families which have never been described before. The *P. clarkii* eyestalk transcriptome reported here provides a resource for improving the knowledge of the still incompletely defined neuroendocrinology of this species and represents an important source of data for all the interested carcinologists.

1. Introduction

The red swamp crayfish (*Procambarus clarkii*, Girard 1852) is considered one of the most invasive species worldwide. Because of its high fitness, adaptability to changing environment, and high fecundity, this species has colonized a wide range of aquatic habitats. Originally native to Mexico and Louisiana (USA), this crayfish is now widely distributed worldwide as a result of the introduction of commercial harvests for the food industry (ISSG, 2012). Its expansion has undoubtedly contributed to a loss of biodiversity in the invaded habitats, especially in the areas where autochthon crayfishes were not able to compete with the new invader. The eyestalk — a group of organs of sight and visual perception — contains the neurosecretory X-organ/sinus gland complex (Ollivaux et al., 2006; Christie et al., 2010; Webster et al., 2012), which produces a superfamily of hormones named Crustacean Hyperglycemic Hormone (CHH) that controls many fundamental physiological functions such as molting, osmoregulation, modulation of glycemia, reproduction (Brown, 1934; Scharrer, 1952; Huberman and Aguilar, 1989; De Kleijn, 1994; Chung and Webster, 2003; Katayama et al., 2013; Turner et al., 2013) and behavioral responses, such as aggression (Aquiloni et al., 2012) and anxiety (Lok et al., 1977).

A thorough analysis of the main hormone products produced in the eyestalk was helpful in identifying variants, hormone-like transcripts and patterns of expression revealing aspects never considered before. The absence of the complete sequenced genome for this species makes the transcriptomic information derived from eyestalk among the only ready-to-use dataset currently available for researchers. To date, in fact, most of the molecular studies regarding the genus *Procambarus* are focused on population genetics (Li et al., 2012; Liu et al., 2013; Shen et al., 2013; Zhu et al., 2013) and immunity (Du et al., 2013; Zhang et al., 2013). In the past years, major efforts were devoted to increase knowledge of crustacean neurohormones and their structures, but unfortunately the genomic information of these animals is still limited (Hopkins, 2012; Webster et al., 2012). For the first time, this study reports the characterization of the eyestalk transcriptome of *P. clarkii*, significantly increasing the molecular knowledge base of this species. Prior to this study, 556 *P. clarkii* nucleotide sequences were filed within public databases, whereas about 1532 belonged to the *Procambarus* genus. Consequently, thanks to the application of molecular techniques to investigate different aspects of environmental biology and the development of new sequencing methods, easily applicable to

Abbreviations: CHH, Crustacean Hyperglycemic Hormone; MIH, Molt Inhibiting Hormone; ITP, Ion Transport Peptide; FPKM, Fragments Per Kilobase per Million mapped reads; ORF, Open Reading Frame; GPCR, G Protein-Coupled Receptor; SARA, Smad Anchor for Receptor Activation.

* Corresponding author at: Department of Life Sciences, University of Trieste, Via L. Giorgieri 5, Trieste, Italy.
E-mail address: pallavic@units.it (A. Pallavicini).

http://dx.doi.org/10.1016/j.gene.2014.12.001
0378-1119/

non-model organisms, it is possible to study topics which up to now have been excluded due to technical limitation. As an example of applied research, transcriptome data mining represents a new way to identify hormones that could be used as autocidal methods to hamper the diffusion of the red swamp crayfish.

2. Materials and methods

2.1. Animal maintenance

Two *P. clarkii* females and two males, collected in Sesto al Reghena (12°50′32″E, 45°51′34″N), were used as biological sources to produce the eyestalk transcriptome. All individuals were adult, in intermolt and non-reproductive stages. Before the eyestalk ablation, they were kept for a week in 120-liter tanks provided with closed circuit for filtered and aerated tap water at ~18 °C. Animals were fed on pellets (Sera granular, Heisenberg, Germany) three times a day.

2.1.1. Ethical note

The following experimental procedures comply with the current applicable laws of Italy, the country where they were performed. No specific permits are required for studies that do not involve endangered or protected invertebrate species. Individuals were maintained in appropriate laboratory conditions to guarantee their welfare and responsiveness. Upon completion of experiments, crayfish were euthanized by hypothermia.

2.2. RNA extraction, sequencing and de novo transcriptome assembly

Total RNA from 8 frozen eyestalks was extracted by homogenization in TriReagent RNA isolation solution (Sigma-Aldrich) and purified with RNeasy MinElute Cleanup Kit (Qiagen). First dissections were performed removing the retina up to the 4th ganglionic swelling because this region is fully pigmented and negatively impacts the quality of extracted RNA. Therefore, as showed in Fig. 1, the eyestalk tissues and components analyzed in this study included: hypodermic and muscle tissues, 1st–3rd ganglionic swellings (*medulla terminalis*, *medulla interna* and *medulla externa*), sinus gland, and X-organ. In order to obtain high-quality RNA, tissues were dissected and immediately stored in liquid nitrogen. RNA concentration and quality were assessed using Qubit RNA assay (Invitrogen) and Agilent 2100 Bioanalyzer (Agilent Technologies). An average of about 7 μg specimen was extracted, and 5 μg of pooled RNA was sent for sequencing to the Institute of Applied Genomics (IGA) in Udine (Italy). cDNA libraries were created with the TruSeq RNA Sample Prep Kit v2 (Illumina) and then were 100 bp paired-end sequenced using Illumina HiSeq2000. The raw sequencing reads were trimmed according to the quality of bases and the presence of residual Illumina adapters.

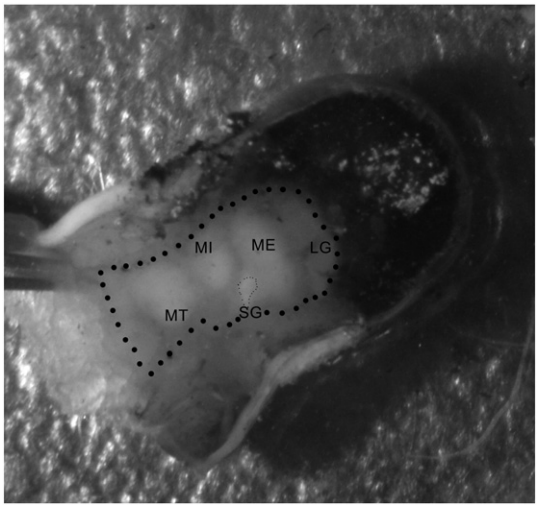

Fig. 1. Internal anatomy of the eyestalk. Marked by the dotted line, the dissected area from which RNA was extracted: LG: lamina ganglionaris, ME: medulla externa, MI: medulla interna, SG: sinus gland, and MT: medulla terminalis.

The de novo assembly of the trimmed reads was performed with Trinity software package (Grabherr et al., 2011), using the default options, allowing the formation of contigs longer than 250 base pairs. The Trinity software is aimed at creating a set of contigs, which is composed of all the possible expressed transcript variants per gene. From this highly redundant contig set we selected as representative the longest transcript with a minimum average coverage of 5 from each gene-related group ('comp' in the Trinity jargon).

2.3. Bioinformatic analysis

The eyestalk non-redundant transcriptome was annotated using the Trinotate suite (http://trinotate.sourceforge.net/). In detail, sequence similarities were detected by BLASTX (Altschul et al., 1990) against the UniProtKB/Swiss-Prot database. Codified peptides were predicted with Transdecoder. Functional domains were identified within the PFAM domain database (Punta et al., 2012) using HMMER (Finn et al., 2011). The prediction of signal peptide and transmembrane domains was performed through SignalP (Petersen et al., 2011) and tmHMM (Krogh et al., 2001), respectively. Finally, transcripts were assigned to eggNOG (Powell et al., 2012) and Gene Ontology (Ashburner et al., 2000) functional categories. We determined the percentage of full-length transcripts using an internal tool of Trinity. Large intergenic non-coding RNAs (lincRNAs) were detected within a selected set of transcripts with length >2 kb and without an Open Reading Frame (ORF) longer than 64 codons using Coding Potential Calculator (CPC) (Kong et al., 2007). RepeatMasker version open-4.0.3 (Smit et al., 2010) was used for detecting interspersed repeats, low complexity DNA sequences, and transposable elements using default conditions. Insertions, deletions and multi-nucleotide variations were identified via CLC Genomics Workbench v.6.03 (Aarhus, Denmark).

Gene expression data was obtained mapping sequence reads on the non-redundant eyestalk assembly using the RNA-seq tool of the CLC Genomics Workbench v.6.03 with the following parameters: minimum allowed length = 0.75, similarity fractions = 0.95 and maximum number of matching contigs = 10. Intact mapped paired-end reads were exclusively used for the calculation of expression values, which are given as Fragments Per Kilobase per Million mapped reads (FPKM).

The phylogenetic analysis was carried out by MEGA 6 (Tamura et al., 2013), initially testing the most suitable substitution model, and then using the Maximum Likelihood algorithm evaluated with 1000 bootstrap replica. All branches supported by less than 50% of replica were collapsed together.

3. Results and discussion

3.1. Transcriptome and assembly characterization

Deep transcriptome sequencing generated 83,170,732 trimmed nucleotide reads of which 80,649,690 are paired-end reads. Reads were filed within the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under the accession number SRR870673. Reads were assembled into 119,255 contigs by Trinity, then reduced to 46,770 as previously described in Materials and methods. Although Trinity assembled contigs were reduced by 60.8%, the 92.32% of reads was back mapped, meaning that only a negligible portion of the sequence data generated was excluded from gene expression analysis. The final non-redundant transcriptome was filed within the NCBI Transcriptome Shotgun Assembly (TSA) database under the accession number GARH01000000.

Table 1
The number of contigs with a positive Trinotate annotation in the selected subset assembly of the eyestalk transcriptome of *P. clarkii*.

Group of transcripts	# of contigs (% on total)
Total contigs	46,770
ORF containing contigs	18,581 (39.73%)
BLASTX resemblance	10,582 (22.63%)
Pfam domain	10,250 (21.92%)
Signal peptide domain	1727 (3.70%)
TmHMM helices	3333 (7.13%)
EggNOG assignment	8687 (18.57%)
Gene Ontology assignment	10,185 (21.78%)

Table 1 summarizes the number of contigs within the selected subset matching the annotation categories provided by Trinotate, such as the presence of signal peptides, transmembrane helices, protein domains, and assignments to orthologous group of genes (eggNOG) and Gene Ontology (GO) categories. The quality of the assembled transcriptome was further confirmed by the fact that about 52% of the transcripts covered more than 80% of their predicted full-length.

Insertions, deletions, multi-nucleotide variations and repeated elements in the eyestalk transcriptome are given as Supplementary information (Table S1). Among the repeated elements, worth of notice were DNA elements such as hAT-Charlie and TcMar-Tigger variants which accounted for about 0.5% of the eyestalk transcriptome. Data regarding repeated elements highlight a significant number of transposable elements in the *P. clarkii* transcriptome which might have been acquired as functional components of the eukaryotic genomes and have many diverse biological functions including transcriptional gene regulation, post-transcriptional gene silencing, heterochromatin remodeling, suppression of transposon activity, and antiviral defenses (Jurka, 1998; Labreuche and Warr, 2013).

At present, only few studies have analyzed large intergenic noncoding RNAs (lincRNAs) in crustaceans. LincRNAs are emerging as key regulators of diverse cellular processes and intensive efforts are currently invested in their characterization and comprehension of their molecular mechanisms in different cell types (Ilott and Ponting, 2013). This work provides the first list of lincRNAs expressed in a Decapod species (provided in List S2). Two lincRNAs were found among the 100 most expressed transcripts and a total of 294 lincRNAs was identified in the *P. clarkii* transcriptome. These lincRNAs do not show sequence similarities with those identified in *Daphnia pulex*, the evolutionarily closest organism to *P. clarkii*.

3.2. Eyestalk transcriptome description

The complete list of the translated protein transcripts discussed here below is reported in Table S3, along with their Trinotate outputs and putative translated protein sequence, whereas the complete list of nucleotide transcripts is available at NCBI TSA (Accession number: GARH01000000). To provide a short description of the eyestalk transcriptome, we first described the most abundant transcripts. As the eyestalk is an important site of neurohormone production, we focused our attention on transcripts belonging to endocrine pathways.

3.2.1. Description of highly expressed transcripts

The 100 most abundant transcripts, showing the highest Fragments Per Kilobase per Million mapped (FPKM) values, are reported in Table S4 along with relevant annotations.

Among these transcripts, we identified several sequences which are typically expressed in muscle fibers and are related to eye contraction and movement. They included six sequences related to myosin, three related to actin, one tropomyosin and two encoding for troponin T and troponin I. A flightin-like protein was also detected; this protein was first discovered in *Drosophila melanogaster* as part of the indirect flight muscles (IFMs), and has been very recently identified also in crustacean species (Xue et al., 2013). With regard to muscle contraction and energy metabolism, the most frequent transcripts included the enolase and fructose-bisphosphate aldolase mRNAs. Beside transcripts linked to muscle contraction, we also detected a rhodopsin transcript: this protein is commonly involved in the mechanism of light perception and is a biological pigment in photoreceptor cells. Along with highly-expressed arrestin proteins, rhodopsin plays a role in the desensitization of G Protein-Coupled Receptor (GPCR) cascades that mediate neurotransmission (Brady et al., 2012). Rhodopsin and proteins related to light perception are further discussed in a dedicated paragraph.

Abundant transcripts included eight highly-expressed cuticular proteins related to the exoskeleton, probably located in the eyestalk hypodermis (Tom et al., 2014). Transcripts linked to the endocrinal function of the eyestalk, such as the male reproductive-related LIM protein 1 and one protein containing the insect pheromone-binding domain (PF03392), were detected. The latter could be related to the main hormone products present in the eyestalk and controlling important physiological mechanisms, such as molting and reproduction activity.

Furthermore, two crustins and a WAP domain-containing protein were detected. These proteins are antimicrobial peptides (AMPs), fundamental to the innate immune system of the crustaceans. The presence of highly transcribed AMPs in the eyestalk could emerge from circulating hemocytes or lead to the hypothesis of an involvement of eyestalks in the production of immunity-related molecules in crustaceans, as suggested by Liu et al. (2011).

Lastly, 11 sequences did not show any similarity to those stored in public databases. Among them, seven sequences encode for small peptides with an average length of 100 AA and show a secretory nature due to the presence of a signal peptide. Fig. 2 reports their expression level compared to that of the housekeeping Elongation factor 1 alpha transcript (EF 1-alpha) and to CHH, MIH and rhodopsin, which represent the typical and specific products of the eyestalk. A secondary peptide prediction showed different structural organizations, in fact three out of seven (Procl_ES_9_0, Procl_ES_21_1 and Procl_ES_30_0) presented beta-sheet structures, three (Procl_ES_25_0, Procl_ES_56_0 and Procl_74_0) had a mix of beta-sheets and alpha helixes, whereas only one (Procl_ES_8_0) was made exclusively of alpha helixes. The sequences of these unknown secreted proteins are reported in Table S3. Undoubtedly, it is worth to further analyze transcripts which could play a role in hormonal control and/or have inter-individual communication functions.

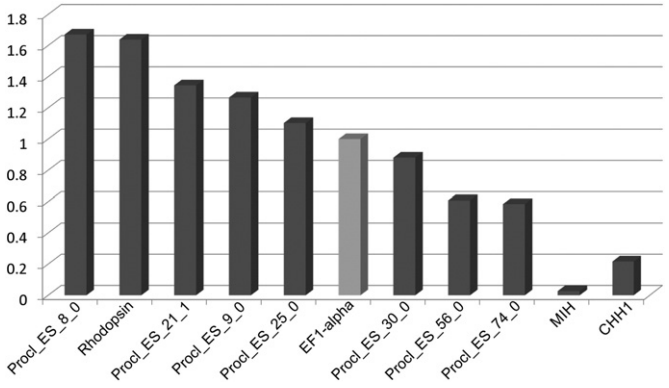

Fig. 2. Expression level of some eyestalk-characterizing transcripts and seven highly abundant unknown secreted proteins. EF1-alpha: Elongation factor 1 alpha (contig Procl_ES_23_0), rhodopsin (contig Procl_ES_23_0), MIH: Molt Inhibiting Hormone (contig Procl_ES_629_0), CHH1_Procl: Crustacean Hyperglycemic Hormone. Y-axis shows the expression level normalized on that of the elongation factor EF1-alpha.

3.3. Photoreceptors

Vision, or light perception, is undoubtedly the main function of the eyes. Retinal photoreceptors are structures involved in the synchronization of the circadian rhythm of sensitivity to light in crayfish. In the eyestalk transcriptome we detected elements belonging to the rhodopsin family (also named 'family A'), such as opsin, a retinylidene protein found in retinal photoreceptor cells, melanopsin, that is involved in circadian rhythms and pupillary reflex, and the compound eye opsin BCRH2. The proteins of the opsin family present a seven-transmembrane structure (PF00001) similar to that of other GPCRs, except for a lysine residue, i.e., a retinal-binding site located in the seventh helix. We also detected a visual pigment-like receptor peropsin which is used in imaging support systems and presents a codified protein which has conserved amino acid residues important for opsins in light detection and a retinal-photoisomerase-like molecular property (Nagata et al., 2009).

As regards photoreception, 11 transcripts contain the Arrestin-1 domain (PF00339), that binds light-activated phosphorylated rhodopsin. Moreover, a retinal rod rhodopsin-sensitive cGMP 3′,5′-cyclic phosphodiesterase subunit delta domain (PF05351) was reported among the transcripts supporting vision. Moreover, we reported that the presence of four transcripts related to the 5-hydroxytryptamine receptor mediates both excitatory and inhibitory neurotransmission. Several retinol dehydrogenases were also identified and it was assumed that they catalyzed the final step in the biosynthesis of 11-cis retinaldehyde, the universal chromophore of visual pigments. An identified photoreceptor-specific phosphatidate cytidylyltransferase acts in sensory transaction and vision. Other receptors belonging to the rhodopsin family were detected and will be described here below, as their function is ambiguous and not strictly related to vision.

3.4. Melatonin-related pattern

Melatonin was first reported in mammals (Lerner et al., 1958) and thereafter in crustaceans (Kamps et al., 2010). With reference to the latter, this "chronobiotic molecule" has been demonstrated to play a role in several biological processes, from the control of circadian rhythms to glucose homeostasis (Sainath and Reddy, 2010), molting and regeneration (Tilden et al., 1997; Richter et al., 1999), reproduction (Sainath and Reddy, 2011), and anti-oxidant defense (Fanjul-Moles and Prieto-Sagredo, 2003). The presence of melatonin in the eyestalk of *P. clarkii* has already been reported by Agapito et al. (1995) and by Balzer et al. (1997) through radioimmunoassay (RIA). The eyestalk of *P. clarkii* has been also found to contain MT_2-like melatonin receptors; in our transcriptome we individuated a melatonin receptor and three transcripts related to the N-acetyl transferase activity, particularly important for the N-acetylation of the serotonin to melatonin. The presence of mRNAs encoding for melatonin-related transcripts suggested the possibility of a direct production of these proteins within the eyestalk.

3.5. Eyestalk endocrinology

Being the eyestalk a major site for hormone production and control, we dedicated two sections to endocrinology aspects. In particular, the Peptide hormonal signaling section reports the transcripts encoding for proteins containing intracellular or cellular membrane associated with receptor domains, whereas the Hormonal storage and secretion section illustrates transcripts encoding for secreted peptides.

3.5.1. Peptide hormonal signaling

Proteins containing intracellular or cellular membrane associated with receptors were considered potential candidates for understanding the main activities of the eyestalk. According to HMMER analysis, 276 proteins containing receptor-related Pfam domains were identified. Forty-seven transcripts resulted to be G Protein-Coupled Receptors (GPCRs), 18 were transmembrane receptors (Secretin family), and 50 belonged to the low-density lipoprotein receptors family (PF00057).

GPCRs are involved in several biological functions, since they can interact with different ligands of various dimensions, such as light-sensitive compounds, aromatic molecules, pheromones, hormones, and neurotransmitters. The GPCR superfamily encompasses CC chemokine receptors type 6 (family A), which were identified in the *P. clarkii* transcriptome. Chemokine receptors trigger a flux in intracellular calcium ions, which cause chemotaxis activation (Wei et al., 2011; Qin et al., 2013). The detection of transcripts related to the chemoreceptive functions in the sequenced eyestalk samples was deemed to be associated with transcripts of the Bellonci's organ. This organ, which consists of ciliated sensory neurons associated with supporting cells, such as glial, bordering and perilemmal cells, presumably possesses neurosecretory and sensory functions (Stachowitsch, 1992). Moreover, CC chemokine receptors type 6 could play a role in the immune response, in that they are involved in the inflammatory response in vertebrates, whereas no data are available for invertebrates (Esche et al., 2005).

Besides GPCRs, other receptor families were also reported, such as the vasopressin v1a receptor, which has only been described in few arthropods, e.g., *Tribolium castaneum* (Li et al., 2008). Vasopressin is a neurohormone, which is primarily responsible for osmoregulation and is synthesized in the nervous system (Roch et al., 2011). In addition to vasopressin v1a receptors, we identified a diuretic hormone receptor and 12 insulin-like peptides, all of which may have additional osmoregulation functions, in line with the *P. clarkii* ability to survive in dry habitats. We also observed a single transcript encoding for preprotachykinin, a neuropeptide acing as a neuromodulator involved in olfactory processes and in tactile and/or visual sensory systems. Detailed similarity searches did not allow to identify other preprotachykinin transcripts, confirming the hypothesis of a single gene in crustaceans (Yasuda-Kamatani and Yasuda, 2004). We identified the presence of a protein trapped in endoderm-1 (Tre-1), which has been studied in *D. melanogaster*, which represents an orphan receptor of the rhodopsin class and is necessary for primordial germ cell migration (Kamps et al., 2010; Pruitt et al., 2013).

Moreover, a SMAD Anchor for Receptor Activation (SARA) was identified in the eyestalk of *P. clarkii*. SMAD proteins are involved in intracellular communication and mediate the Transforming Growth Factor-beta (TGF-beta) signal from the transmembrane serine–threonine receptor kinases to the nucleus. Two transcripts encoding for Kiss-1 receptor were identified. These receptors have been initially described in the mammalian hypothalamus and have recently been also identified in amphibians (Moon et al., 2009). Kiss-1 receptor is particularly important as it binds to Kisspeptin, the so-called "fertility protein." Whether Gonad Inhibiting Hormone (GIH), Kiss-1 receptor and the two Gonadotropin-releasing hormone receptors co-operate in regulating the reproduction in *P. clarkii* remains unclear, an unanswered question which it would be surely worth examining.

With further reference to the identified GPCRs, five transcripts encoding for a FMRFamide receptor – a widespread neuropeptide in invertebrates – with prominent cardioexcitatory function and characterized by a C-terminal arginine-phenylalanineamide, were isolated. The pyroglutamylated RFamide peptide receptor, also named G-protein coupled receptor 103 (GPR103) and the neuropeptide F receptor (two transcripts were detected in this transcriptome) are mainly involved in the regulation of the energy balance and feeding behavior. FMRFamide causes anorexigenic effects in mice (Kavaliers et al., 1985) and its expression level changes due to gustatory exposure to sugar (Bechtold and Luckman, 2007). Studies carried out on *D. melanogaster* confirmed the role of MRFamide in arthropods and the presence of its receptor, i.e., neuropeptide F.

For the first time in Decapoda, we also isolated an Allatostatin-A receptor, which so far had only been reported for the orders Diplostraca and Siphonostomatoida (de la Fuente et al., 2006). The Allatostatin-A receptor was widely distributed in the animal's brain, including the visual system, central complex and olfactory system (Christie et al., 2010; Polanska et al., 2012).

3.5.2. Hormonal storage and secretion

The eyestalk is a well-known neurohormone producer; in particular, it is the main site of production and storage of the CHH-superfamily (De Kleijn, 1994), which includes the CHH, the Molt Inhibiting Hormone (MIH), the Gonad Inhibiting Hormone (GIH) and the Mandibular Organ-Inhibiting Hormone (MOIH). Only CHH and MIH were expressed in our transcriptome dataset. Based on the amino acid sequences, we followed the classification proposed by Kung et al. (2013), in which two CHH transcripts were identified in hemocytes from *P. clarkii*, namely CHH1 and CHH2, codified by two different genes. As reported by the authors, each gene encodes for two and three different alternative splicing forms, respectively (Kung et al., 2013). Fig. 3 outlines the two CHH genes, each with its alternative splicing forms for both Coding DNA sequences (CDS) and mRNAs. Thanks to the eyestalk transcriptome and the *P. clarkii* hemocytes transcriptome (unpublished data), we were able to show that: 1 — CHH1 is expressed about 6× more than CHH2 in the eyestalk (48.64 and 7.77 FPKM, respectively); 2 — the L-alternative splicing form is very poorly expressed (1 out of 1000 CHH1 transcripts and 1 out of 30 CHH2 transcripts), whereas the truncated form was observed in just 1 read; and 3 — both CHH1 and CHH2 genes are not expressed in hemocytes. The CHH1 sequence found in this assembly perfectly matched the published one, while we extended the CHH2 mRNA sequence towards the 5′-end adding to the N-terminus 6 more amino acids (GenBank accession number KM513658). Within the CHH superfamily, we also reported the full-length transcript of the MIH (contig Procl_ES_629_0). Lastly, an Ion Transport Peptide (ITP) (GenBank accession number KM513659) was detected in the eyestalk transcriptome, an unprecedented finding in *P. clarkii*. ITP belongs to a subgroup of the CHH superfamily and is currently described in insects, collembolans, and branchiopods (Montagné et al., 2010). In order to confirm the ITP identity, we collected CHHs from the subphylum Crustacea in addition to all the ITP sequences filed within the GenBank database, and created a phylogenetic tree. The resulting tree is showed in Fig. 4A. The ITP from *P. clarkii* clustered with ITPs showing a clear distinction from the CHHs, thus confirming its identity. Fig. 4B describes the multiple alignment of the conserved domain, corresponding to the crustacean neurohormone family (PF01147).

4. Conclusions

The increasingly numerous studies on RNA-seq, focusing on non-model organisms, and specifically on crustaceans, undoubtedly point out the great potential of this technique, which could have positive consequences on different aspects, such as a better understanding of the immune and reproduction systems of a species or the management of its development. Moreover, only very recently few NGS-based studies on decapods have been published (Liu et al., 2011; Manfrin et al., 2013; Song et al., 2014) and use data obtained through Illumina sequencing.

The possibility of determining a high number of transcript sequences using in silico assembly of the contigs–a process extremely complex when applying traditional methods–allowed us to investigate from a new and less time-consuming point of view. This approach enabled to combine data of great interest in many aspects, including the definition of tissue-specific transcriptomes, the study of rare or poorly expressed messengers or the determination of different alternative transcripts of the same gene. Our study has increased the amount of genetic information available for *P. clarkii*, even though the sequencing of other tissues would be certainly significant for obtaining a broader overview of the eyestalk-specific transcripts or delineating the function of the transcripts identified in different tissues. By exploiting the genetic knowledge gathered on *P. clarkii* endocrinal system new eradicative solutions can be found for this very invasive and plastic species, able to adapt to very different environmental conditions.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2014.12.001.

Acknowledgments

We would like to especially thank Luca Rodella, who allowed us to use his personal server to perform our analyses more rapidly, and Sara Boscariol, who assisted us in English editing. This study was supported by the European Project LIFE10 NAT/IT/000239 RARITY and by PRIN

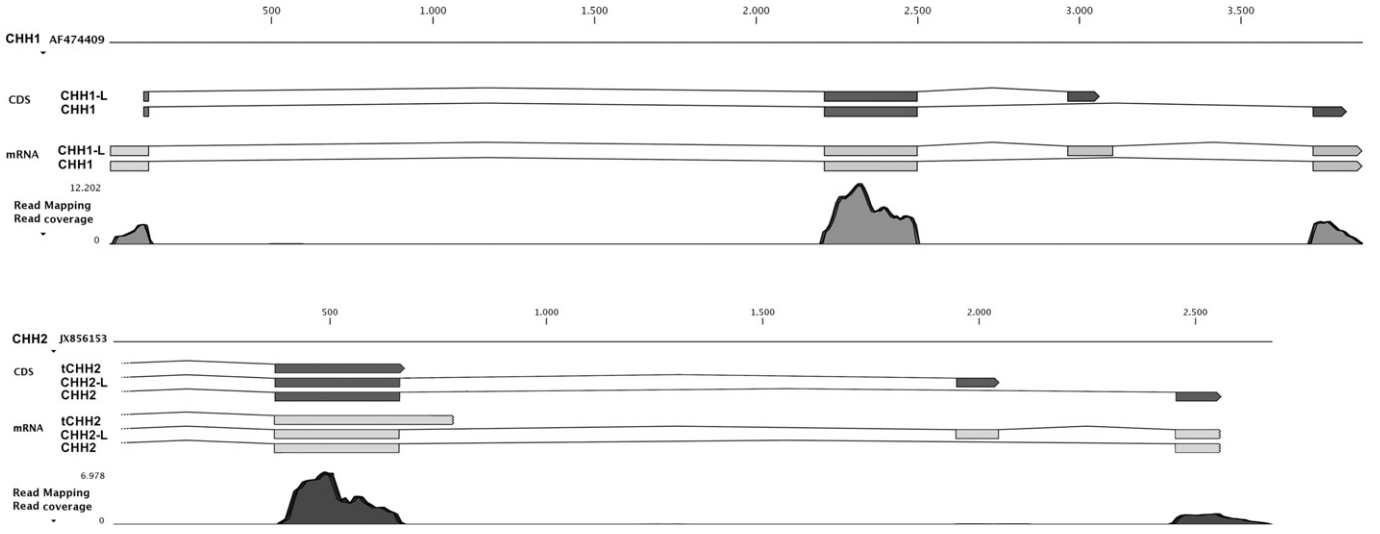

Fig. 3. Scheme of the two CHH genes with their relative alternative splicing forms both for their coding DNA sequences (CDS) and for their mRNAs. The last track for each gene shows the mapping of the *P. clarkii* eyestalk transcriptome reads.

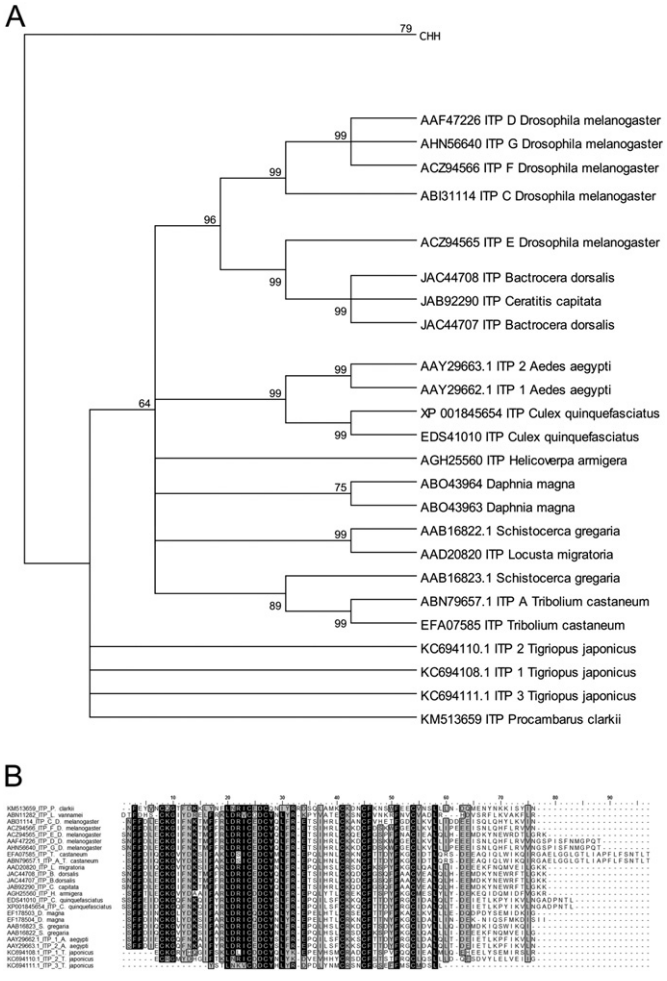

Fig. 4. A. Ion Transport Peptide cladistic tree. It has been rooted on a series of CHH proteins, collapsed in this tree (GenBank accession numbers: ADL27417 (*Litopenaeus vannamei*); AAL79192 (*Procambarus clarkii*); AAL79193 (*P. clarkii*); AFV95077 (*P. clarkii*); AAX09331 (*Pontastacus leptodactylus*); ADZ98836 (*P. clarkii*); CAA56674 (*Orconectes limosus*); Q25589 (*O. limosus*); Q25683 (*P. clarkii*); AFV95082 (*P. clarkii*); AFV95083 *P. clarkii*); AAB32871 (*Homarus americanus*); P19806 (*H. americanus*); Q25154 (*H. americanus*); AAQ24527 (*Penaeus monodon*); AAQ24526 (*P. monodon*); AAQ24525 (*P. monodon*); BAE78493 (*Marsupenaeus japonicus*); AAC36310 (*Macrobrachium lanchesteri*); AF372657 (*Macrobrachium rosenbergii*); AF372657 (*M. rosenbergii*); AAB25454 (*Armadillidium vulgare*); AAQ22391 (*Nephrops norvegicus*); AAQ22392 (*N. norvegicus*); P83485 (*Cherax destructor*); P55845 (*Procambarus bouvieri*); P83800 (*Astacus astacus*); P56687 (*Jasus lalandii*); ACS35347 (*Rimicaris kairei*); AAK28329 (*Bythograea thermydron*); AAS45136 (*Callinectes sapidus*); ACB46189 (*Portunus trituberculatus*); AFM29133 (*Portunus pelagicus*); ABQ41272 (*Cancer productus*); ABQ41269 (*C. productus*); ABQ41270 (*C. productus*); ABQ41271 (*C. productus*); P14944 (*Carcinus maenas*); AAO27805 (*Pachygrapsus marmoratus*); ABA70560 (*Potamon ibericum*); O97383 (*P. monodon*); AAD11813 (*Metapenaeus ensis*); AFV95078 (*P. clarkii*); and BAA89003 (*P. clarkii*)). Le & Gasquel substitution model + gamma distribution + proportion of invariable sites (LG + G + I) were applied for 1000 bootstrap replica. Concerning ITPs, branches are collapsed due to significance less than 50%. B. Multiple alignment of the Crustacean neurohormone domain (PF01147). The amino acids' numbering refers to the *P. clarkii* ITP protein.

2010–2011 20109XZEPR_002 (Italian Ministry of Education, Universities and Research).

References

Agapito, M.T., Herrero, B., Pablos, M.I., Miguel, J.L., Recio, J.M., 1995. Circadian rhythms of melatonin and serotonin-N-acetyltransferase activity in *Procambarus clarkii*. Comp. Biochem. Physiol. A Physiol. 112, 179–185.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Aquiloni, L., Giulianini, P.G., Mosco, A., Guarnaccia, C., Ferrero, E.A., Gherardi, F., 2012. Crustacean hyperglycemic hormone (cHH) as a modulator of aggression in crustacean decapods. PLoS ONE 7, e50047.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25, 25–29.

Balzer, I., Espínola, I.R., Fuentes-Pardo, B., 1997. Daily variations of immunoreactive melatonin in the visual system of crayfish. Biol. Cell. 89, 539–543.

Bechtold, D., Luckman, S.M., 2007. The role of RFamide peptides in feeding. J. Endocrinol. 192, 3–15.

Brady, P., Elizur, A., Williams, R., Cummins, S.F., Knibb, W., 2012. Gene expression profiling of the cephalothorax and eyestalk in *Penaeus monodon* during ovarian maturation. Int. J. Biol. Sci. 8, 328–343.

Brown, F.A., 1934. The chemical nature of the pigments and the transformations responsible for color changes in *Palaemonetes*. Biol. Bull. 67, 365–380.

Christie, A.E., Stemmler, E.A., Dickinson, P.S., 2010. Crustacean neuropeptides. Cell. Mol. Life Sci. 67, 4135–4169.

Chung, J.S., Webster, S.G., 2003. Moult cycle-related changes in biological activity of moult-inhibiting hormone (MIH) and crustacean hyperglycaemic hormone (CHH) in the crab, *Carcinus maenas*: from target to transcript. Eur. J. Biochem. 270, 3280–3288.

De Kleijn, D.P.V., 1994. Cloning and expression of mRNA encoding prepro-gonad-inhibiting hormone (GIH) in the lobster *Homarus americanus*. FEBS Lett. 353, 255–258.

de la Fuente, J., Almazán, C., Naranjo, V., Blouin, E.F., Meyer, J.M., Kocan, K.M., 2006. Autocidal control of ticks by silencing of a single gene by RNA interference. Biochem. Biophys. Res. Commun. 344, 332–338.

Du, H.H., Hou, C.L., Wu, X.G., Xie, R.H., Wang, Y.Z., 2013. Antigenic and immunogenic properties of truncated VP28 protein of white spot syndrome virus in *Procambarus clarkii*. Fish Shellfish Immunol. 34, 332–338.

Esche, C., Stellato, C., Beck, L.A., 2005. Chemokines: key players in innate and adaptive immunity. J. Investig. Dermatol. 125, 615–628.

Fanjul-Moles, M.L., Prieto-Sagredo, J., 2003. The circadian system of crayfish: a developmental approach. Microsc. Res. Tech. 60, 291–301.

Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39, 29–37.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Hopkins, P.M., 2012. The eyes have it: a brief history of crustacean neuroendocrinology. Gen. Comp. Endocrinol. 175, 357–366.

Huberman, A., Aguilar, M.B., 1989. A neuropeptide with molt-inhibiting hormone activity from the sinus gland of the mexican crayfish *Procambarus bouvieri* (Ortmann). Comp. Biochem. Physiol. B 93, 299–305.

Ilott, N.E., Ponting, C.P., 2013. Predicting long non-coding RNAs using RNA sequencing. Methods 63, 50–59.

ISSG, I.S.S.G., 2012. Species profile: *Procambarus clarkii* (Girard, 1852). The IUCN Red List of Threatened Species. International Union for Conservation of Nature (IUCN) (Global Invasive Species Database Version 2014.2).

Jurka, J., 1998. Repeats in genomic DNA: mining and meaning. Curr. Opin. Struct. Biol. 8, 333–337.

Kamps, A.R., Pruitt, M.M., Herriges, J.C., Coffman, C.R., 2010. An evolutionarily conserved arginine is essential for Tre1 G protein-coupled receptor function during germ cell migration in *Drosophila melanogaster*. PLoS ONE 5, e11839.

Katayama, H., Ohira, T., Nagasawa, H., 2013. Crustacean peptide hormones: structure, gene expression and function. Aqua-BioSci. Monogr. 6, 49–90.

Kavaliers, M., Hirst, M., Mathers, A., 1985. Inhibitory influences of FMRFamide on morphine- and deprivation-induced feeding. Neuroendocrinology 40, 533–535.

Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L., Gao, G., 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 35, 345–349.

Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Kung, P.C., Wu, S.H., Nagaraju, G.P.C., Tsai, W.S., Lee, C.Y., 2013. Crustacean hyperglycemic hormone precursor transcripts in the hemocytes of the crayfish *Procambarus clarkii*: novel sequence characteristics relating to gene splicing pattern and transcript stability. Gen. Comp. Endocrinol. 186, 80–84.

Labreuche, Y., Warr, G.W., 2013. Insights into the antiviral functions of the RNAi machinery in penaeid shrimp. Fish Shellfish Immunol. 34, 1002–1010.

Lerner, A.B., Case, J.D., Takahashi, Y., Lee, T.H., Mori, W., 1958. Isolation of melatonin, the pineal gland factor that lightens melanocytes. J. Am. Chem. Soc. 80, 2587-2587.

Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., Williamson, M., Arakane, Y., Verleyen, P., Schoofs, L., Schachtner, J., Grimmelikhuijzen, C.J.P., Park, Y., 2008. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. Genome Res. 18, 113–122.

Li, Y., Guo, X., Cao, X., Deng, W., Luo, W., Wang, W., 2012. Population genetic structure and post-establishment dispersal patterns of the red swamp crayfish *Procambarus clarkii* in China. PLoS ONE 7, e40652.

Liu, Y., Cui, Z., Song, C., Wang, S., Li, Q., 2011. Multiple isoforms of immune-related genes from hemocytes and eyestalk cDNA libraries of swimming crab *Portunus trituberculatus*. Fish Shellfish Immunol. 31, 29–42.

Liu, G., Zhou, L., Li, X., Lu, D., 2013. Population genetic structure of the invasive red swamp crayfish in China revealed by ITS1 variation. Biochem. Genet. 1–12.
Lok, C.K., Kiat, N.S., Koh, T.K., 1977. An autocidal ovitrap for the control and possible eradication of *Aedes aegypti*. Southeast Asian J. Trop. Med. Public Health 8, 56–62.
Manfrin, C., Tom, M., De Moro, G., Gerdol, M., Guarnaccia, C., Mosco, A., Pallavicini, A., Giulianini, P.G., 2013. Application of D-crustacean hyperglycemic hormone induces peptidases transcription and suppresses glycolysis-related transcripts in the hepatopancreas of the Crayfish *Pontastacus leptodactylus* — results of a Transcriptomic Study. PLoS ONE 8, e65176.
Montagné, N., Desdevises, Y., Soyez, D., Toullec, J.Y., 2010. Molecular evolution of the crustacean hyperglycemic hormone family in ecdysozoans. BMC Evol. Biol. 10.
Moon, J.S., Lee, Y.R., Oh, D.Y., Hwang, J.I., Lee, J.Y., Kim, J.I., Vaudry, H., Kwon, H.B., Seong, J.Y., 2009. Molecular cloning of the bullfrog kisspeptin receptor GPR54 with high sensitivity to Xenopus kisspeptin. Peptides 30, 171–179.
Nagata, T., Koyanagi, M., Tsukamoto, H., Terakita, A., 2009. Identification and characterization of a protostome homologue of peropsin from a jumping spider. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. 196, 51–59.
Ollivaux, C., Vinh, J., Soyez, D., Toullec, J.Y., 2006. Crustacean hyperglycemic and vitellogenesis-inhibiting hormones in the lobster *Homarus gammarus*: implications for structural and functional evolution of a neuropeptide family. FEBS J. 273, 2151–2160.
Petersen, T.N., Brunak, S., Von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.
Polanska, M.A., Tuchina, O., Agricola, H., Hansson, B.S., Harzsch, S., 2012. Neuropeptide complexity in the crustacean central olfactory pathway: Immunolocalization of A-type allatostatins and RFamide-like peptides in the brain of a terrestrial hermit crab. Mol. Brain 5.
Powell, S., Szklarczyk, D., Trachana, K., Roth, A., Kuhn, M., Muller, J., Arnold, R., Rattei, T., Letunic, I., Doerks, T., Jensen, L.J., Von Mering, C., Bork, P., 2012. eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Res. 40, D284–D289.
Pruitt, M.M., Lamm, M.H., Coffman, C.R., 2013. Molecular dynamics simulations on the Tre1 G protein-coupled receptor: exploring the role of the arginine of the NRY motif in Tre1 structure. BMC Struct. Biol. 13.
Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A., Finn, R.D., 2012. The Pfam protein families database. Nucleic Acids Res. 40, D290–D301.
Qin, S., Klamar, C.R., Fallert Junecko, B.A., Craigo, J., Fuller, D.H., Reinhart, T.A., 2013. Functional characterization of ferret CCL20 and CCR6 and identification of chemotactic inhibitors. Cytokine 61, 924–932.
Richter, K., Peschke, E., Peschke, D., 1999. Effect of melatonin on the release of prothoracicotropic hormone from the brain of *Periplaneta americana* (Blattodea: Blattidae). Eur. J. Entomol. 96, 341–345.
Roch, G.J., Busby, E.R., Sherwood, N.M., 2011. Evolution of GnRH: diving deeper. Gen. Comp. Endocrinol. 171, 1–16.
Sainath, S., Reddy, P.S., 2010. Melatonergic regulation of hemolymph sugar levels in the fresh water crab, *Oziotelphusa senex senex*. J. Exp. Zool. 313.
Sainath, S.B., Reddy, P.S., 2011. Effect of selected biogenic amines on reproduction in the fresh water edible crab, *Oziotelphusa senex senex*. Aquaculture 313, 144–148.
Scharrer, B., 1952. Neurosecretion. XI: the effects of nerve section on the intercerebralis–cardiacum–allatum system of the insect *Leucophaea maderae*. Biol. Bull. 261–272.
Shen, H., Braband, A., Scholtz, G., 2013. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. Mol. Phylogenet. Evol. 66, 776–789.
Smit, A.F.A., Hubley, R., Green, P., 2010. RepeatMasker Open-3.0. software package. http://repeatmasker.org.
Song, Y.N., Shi, L.L., Liu, Z.Q., Qiu, G.F., 2014. Global analysis of the ovarian microRNA transcriptome: implication for miR-2 and miR-133 regulation of oocyte meiosis in the Chinese mitten crab, *Eriocheir sinensis* (Crustacea:Decapoda). BMC Genomics 1, 157.
Stachowitsch, M., 1992. The Invertebrates: An Illustrated Glossary. Wiley-Liss Inc., New York.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tilden, A.R., Rasmussen, P., Awantang, R.M., Furlan, S., Goldslein, J., Palsgrove, M., Sauer, A., 1997. Melatonin cycle in the fiddler crab *Uca pugilator* and influence of melatonin on limb regeneration. J. Pineal Res. 23, 142–147.
Tom, M., Manfrin, C., Chung, S.J., Sagi, A., Gerdol, M., De Moro, G., Pallavicini, A., Giulianini, P.G., 2014. Expression of cytoskeletal and molt-related genes is temporally scheduled in the hypodermis of the crayfish *Procambarus clarkii* during premolt. J. Exp. Biol. 217, 4193–4202.
Turner, L.M., Webster, S.G., Morris, S., 2013. Roles of crustacean hyperglycaemic hormone in ionic and metabolic homeostasis in the christmas island blue crab, *Discoplax celeste*. J. Exp. Biol. 216, 1191–1201.
Webster, S.G., Keller, R., Dircksen, H., 2012. The CHH-superfamily of multifunctional peptide hormones controlling crustacean metabolism, osmoregulation, moulting, and reproduction. Gen. Comp. Endocrinol. 175, 217–233.
Wei, W., Gao, L., Zhang, F.F., Cui, L.X., Xie, X., 2011. Expression and function analysis of chemokine receptor CCR6 in HEK293 cells. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi 27, 33–35.
Xue, J., Zhang, X.Q., Xu, H.J., Fan, H.W., Huang, H.J., Ma, X.F., Wang, C.Y., Chen, J.G., Cheng, J.A., Zhang, C.X., 2013. Molecular characterization of the flightin gene in the wing-dimorphic planthopper, *Nilaparvata lugens*, and its evolution in Pancrustacea. Insect Biochem. Mol. Biol. 43, 433–443.
Yasuda-Kamatani, Y., Yasuda, A., 2004. APSGFLGMRamide is a unique tachykinin-related peptide in crustaceans. Eur. J. Biochem. 271, 1546–1556.
Zhang, X.W., Liu, Y.Y., Mu, Y., Ren, Q., Zhao, X.F., Wang, J.X., 2013. Overexpression of a C-type lectin enhances bacterial resistance in red swamp crayfish *Procambarus clarkii*. Fish Shellfish Immunol. 34, 1112–1118.
Zhu, B.F., Huang, Y., Dai, Y.G., Bi, C.W., Hu, C.Y., 2013. Genetic diversity among red swamp crayfish (*Procambarus clarkii*) populations in the middle and lower reaches of the Yangtze River based on AFLP markers. Genet. Mol. Res. 12, 791–800.