Marine Genomics 37 (2018) 74–81

Contents lists available at ScienceDirect

Marine Genomics

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

De novo assembly and transcriptome characterization of the freshwater prawn *Palaemonetes argentinus*: Implications for a detoxification response

C. Fernando García[a,1], Nicolas Pedrini[a,1], Arturo Sánchez-Paz[b], Carlos S. Reyna-Blanco[c], Sabrina Lavarias[d], Adriana Muhlia-Almazán[e], Analía Fernández-Giménez[f], Aldana Laino[a], Enrique de-la-Re-Vega[g], German Lukaszewicz[h], Alonso A. López-Zavala[g], Luis G. Brieba[h], Michael F. Criscitello[i], Jesús S. Carrasco-Miranda[j,2], Karina D. García-Orozco[j], Adrian Ochoa-Leyva[c], Enrique Rudiño-Piñera[c], Alejandro Sanchez-Flores[c], Rogerio R. Sotelo-Mundo[j,*]

[a] *Instituto de Investigaciones Bioquímicas de La Plata "Dr. Rodolfo Brenner" (INIBIOLP), CCT-La Plata CONICET-UNLP, Calle 60 y 120, La Plata 1900, Argentina*
[b] *Laboratorio de Referencia, Análisis y Diagnóstico en Sanidad Acuícola, Centro de Investigaciones Biológicas del Noroeste S. C. (CIBNOR), Calle Hermosa 101. Col. Los Ángeles, 83106 Hermosillo, Sonora, Mexico*
[c] *Departamentos de Microbiología Molecular, Medicina Molecular y Bioprocesos, Unidad Universitaria de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología (IBT), Universidad Nacional Autónoma de México (UNAM), Avenida Universidad 2001, Colonia Chamilpa, Cuernavaca 62210, Mexico*
[d] *Instituto de Limnología de La Plata "Dr. Raúl A. Ringuelet" (ILPLA), CONICET CCT La Plata-Universidad Nacional de La Plata (UNLP), Boulevard 120 y 62, La Plata 1900, Argentina*
[e] *Laboratorio de Bioenergética, Centro de Investigación en Alimentación y Desarrollo, A.C. Carretera a Ejido La Victoria, Km 0.6, Hermosillo, Sonora 83304, Mexico*
[f] *Laboratorio de Fisiología de Organismos Acuáticos y Biotecnología, Instituto de Investigaciones Marinas y Costeras (IIMYC), Facultad de Ciencias Exactas y Naturales CONICET, Universidad Nacional de Mar del Plata (UNMdP), Mar del Plata 7600, Argentina*
[g] *Universidad de Sonora, Blvd. Colosio y Reforma s/n, Hermosillo, Sonora 83000, Mexico*
[h] *Laboratorio de Ecotoxicología, Instituto de Investigaciones Marinas y Costeras (IIMYC), Facultad de Ciencias Exactas y Naturales, CONICET, Universidad Nacional de Mar del Plata (UNMdP), Mar del Plata 7600, Argentina*
[i] *Comparative Immunogenetics Laboratory, Department of Veterinary Pathobiology, Texas A & M University, College Station, TX 77843, United States*
[j] *Laboratorio de Estructura Biomolecular, Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD). Carretera a Ejido La Victoria Km 0.6, Apartado Postal 1735, Hermosillo, Sonora 83304, Mexico*

ARTICLE INFO

Keywords:
Palaemonetes Argentinus
Shrimp
Prawn
Transcriptome
Detoxification
Crustacea

ABSTRACT

Palaemonetes argentinus, an abundant freshwater prawn species in the northern and central region of Argentina, has been used as a bioindicator of environmental pollutants as it displays a very high sensitivity to pollutants exposure. Despite their extraordinary ecological relevance, a lack of genomic information has hindered a more thorough understanding of the molecular mechanisms potentially involved in detoxification processes of this species. Thus, transcriptomic profiling studies represent a promising approach to overcome the limitations imposed by the lack of extensive genomic resources for *P. argentinus*, and may improve the understanding of its physiological and molecular response triggered by pollutants. This work represents the first comprehensive transcriptome-based characterization of the non-model species *P. argentinus* to generate functional genomic annotations and provides valuable resources for future genetic studies.

Trinity *de novo* assembly consisted of 24,738 transcripts with high representation of detoxification (phase I and II), anti-oxidation, osmoregulation pathways and DNA replication and bioenergetics. This crustacean transcriptome provides valuable molecular information about detoxification and biochemical processes that could be applied as biomarkers in further ecotoxicology studies.

* Corresponding author at: Laboratorio de Estructura Biomolecular, Centro de Investigación en Alimentación y Desarrollo, A.C. (CIAD), Carretera a Ejido La Victoria Km 0.6, Apartado Postal 1735, Hermosillo, Sonora 83304, Mexico.
E-mail address: rrs@ciad.mx (R.R. Sotelo-Mundo).
[1] Both authors contributed equally as first authors.
[2] Present address: Calbiotech. 1935 Cordell Court. El Cajon, CA 92020, USA.

http://dx.doi.org/10.1016/j.margen.2017.08.009
Received 27 June 2017; Received in revised form 25 August 2017; Accepted 26 August 2017
Available online 03 October 2017
1874-7787/ © 2017 Elsevier B.V. All rights reserved.

1. Introduction

Major policies, actions, and control measures have been undertaken globally over recent years to reduce the disposal of hazardous substances into the aquatic environment. However, pollution remains a major problem for marine ecosystems (Vidas and Schei, 2011), especially in developing countries (Wu, 1999), and poses continuing risks to the health of the organisms inhabiting aquatic environments.

Among crustaceans, the Palaemonidae family presents an unusual diversity (it is the second family with more species), and the shrimp species grouped within this family are critical contributing elements in several essential ecological processes of marine and freshwater environments (Bauer, 2005). Because of its evolutionary history, this group is considered an ideal model for studies addressing physiological adaptations associated with successful limnic invasions by marine organisms (Freire et al., 2003). Taken together the high sensitivity to toxic pollutants, like pesticides, and its wide distribution spectrum, penaeid prawns, have been proposed as useful bioindicators for environmental monitoring of anthropogenic impact (García-de la Parra et al., 2006).

P. argentinus is a widely-distributed decapod species in the coastal region of Argentina, Paraguay, Uruguay and southern Brazil, mainly in freshwater ponds and lakes (Morrone and Lopreto, 1995). Although several recent studies have been performed on the toxicological effects of pollutants over *P. argentinus* and its detoxification response (Bertrand et al., 2015; Lavarias and Garcia, 2015; Griboff et al., 2014; Galanti et al., 2013; Montagna and Collins, 2007; Collins and Cappello, 2006), studies of genes associated with the physiologic and metabolic response of this organism to pollutants are scarce.

In recent years, there has been an increasing interest and enthusiasm in applying molecular tools for understanding the impact of contaminant stressors on the health of aquatic organism, to identify specific molecular, biochemical, metabolic, physiological, and behavioral responses of marine species to pollutants, and to identify potential biomarkers of stress caused by contaminants to aquatic life (Hwang et al., 2017; Diaz de Cerio et al., 2017; Shinn et al., 2015; Gust et al., 2014; Meng et al., 2014; Kim et al., 2012). In this context, the emergence of high-throughput sequencing has undoubtedly expanded our knowledge of non-model species in which to focus future research efforts (Mehinto et al., 2012). Unfortunately, to the best of our knowledge, few high-throughput sequencing studies have described the changes in gene expression of freshwater decapods due to environmental stressors (Manfrin et al., 2015; Harms et al., 2013; Griffitt et al., 2007).

Despite the availability and extensive use of low-cost NGS platforms and the commercial value of several crustacean species, the reconstruction of their genomes has remained a particularly challenging task, mainly because crustaceans possess vast and complex genomes, and because of the presence of repetitive elements (Holland and Skinner, 1977), which interfere during sequence assembly. However, RNA-seq is a powerful alternative approach that allows analysis of genomic coding regions and differential gene expression studies. In the absence of a reference genome, *de novo* transcriptomes can deliver thousands of transcripts in a single experiment (Robertson et al., 2010). Once the individual nucleotide sequence of a transcribed gene is known, quantification experiments for gene expression of specific genes can be performed to determine changes related to certain conditions such as xenobiotic exposure, hypoxia, or other environmental parameters by reverse transcription coupled to quantitative PCR. Transcriptome annotation also provides insights about the proteins and metabolic routes present in the organisms, limited of course by the specific gene expression at the point where RNA was obtained.

Therefore, the main purpose of this study was to characterize the transcriptome profile of *P. argentinus* through RNA-seq techniques to identify gene signatures associated with relevant metabolic pathways related to detoxification. This study provides a solid foundation for future comparative studies, and for research on the functional role of particular genes involved in the detoxification response of this crustacean. To the best of our knowledge, this is the first *de novo* transcriptome study in *P. argentinus* to date.

Table 1
Genome and environmental features of the biosample.

Item	Description
Investigation_type	Eukaryote
Project_name	PRJNA309860
Collected_by	Carlos Fernando Garcia
Collection_date	5-August-2013
Latitude_longitude	34.9600 S; 57.7767 W
Depth	0.5 m
Temperature	15 °C
Salinity	0 psu
Environment	Fresh water
Biotic_relationship	Free living
Sequencing technology	Illumina GAIIx
Assembly	Trinity V.20140717
Biome	ENVO:01000297
Feature	ENVO:00000022
Material	ENVO:00000063
Geolocation_name	El Pescado < comma > Argentina
Assembly method	Trinity release 2014
Assembly name	Palaemon argentinus Transcriptome

2. Materials and methods

2.1. Sample collection

P. argentinus adults were collected in “El Pescado” (34°57′0.36″S; 57° 46′0.36″W), a freshwater watercourse in La Plata River, Argentina in 2013 during the pre-reproductive season. The biosample information is shown in Table 1. The samples were taken to the laboratory and immediately submerged in ice-cold RNAlater reagent (Sigma-Aldrich). Samples were stored at − 20 °C until processing.

2.2. Sample preparation and RNA extraction for RNA-seq

The experimental methods were similar to those described previously (Ghaffari et al., 2014; Ioannidis et al., 2014;). Total RNA was isolated from five adult individuals by using the Quick-RNA MicroPrep kit (Zymo Research) using the manufacturer's protocol. RNA quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies). Only RNA samples with an RNA integrity number (RIN) above 7.0 were used for Illumina RNA-seq library preparation. RNA-seq libraries were generated using the TruSeq RNA Sample Prep Kit (Illumina) according to the manufacturer's protocol, followed by sequencing using an Illumina GAIIx platform for 72 paired-end (PE) cycles following the manufacturer's protocol.

2.3. Bioinformatic analysis

The *P. argentinus* sequence reads obtained from the Illumina platform were reconstructed using the Illumina-based Trinity Assembler 2014 (release r20140717) (Grabherr et al., 2011), executed with the Pasafly parameters to reduce the number of reported isoforms. To deduce the protein products by conceptual translation, the software Transdecoder v2.0.1 was used with default parameters. Putative mitochondrial transcripts were identified by using Transdecoder with the arthropod mitochondrial genetic code. The resulting protein sequences were compared to the *Palaemon serenus* (NC_027601) mitochondrial proteins.

2.4. Functional annotation

Annotation was conducted using the Blast v2.2.30 + algorithm (Camacho et al., 2009) against the UniProtKB/Swiss-Prot database. A functional domain annotation was performed using the hmmscan from HMMER v3.1b1 suite (Eddy, 2011) to search against the Pfam-A database. SignalP v4.1 and TMHMM v2.0c programs were used to predict signal peptide and transmembrane regions, respectively. Results were integrated by the Trinotate v2.0.1 (Grabherr et al., 2011) pipeline and generated additional analyses such as Clusters of Orthologous Groups (COG) identifiers (Tatusov et al., 2000) and Gene Ontology terms (GO) (Ashburner et al., 2000). From the integrated results, the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000), euKaryotic Ortholog Groups (KOG) (Tatusov et al., 2000), and Gene Ontology (GO) (Harris et al., 2004) count histograms were elaborated using *ad-hoc* Perl scripts. Blast2GO was also used to annotate function (Conesa et al., 2005; Conesa and Gotz, 2008).

The annotation process consisted of a multi-approach search using programs including BlastP and BlastX against Swissprot/Unifer databases, hmmscan against PFAM-A database, signalP and TMHMM for signal peptide and transmembrane helices regions, respectively. All these results were integrated using the Trinotate package (Finn et al., 2014) with the SQLite database included. From this database, an annotation table was generated (Suppl. Table 1) with additional data from Eggnog and Gene Ontology. All the sequences mentioned in the text are included in Fasta format in supplemental materials.

3. Results

To obtain a wider understanding of the gene content, biological processes, and overview of the gene expression in the freshwater prawn *P. argentinus*, high-quality paired-end RNA-seq libraries were constructed. After processing the data with the Trinity pipeline, 101,251,880 total paired reads led to 24,378 transcripts. The assembly and annotation statistics are shown in Table 2. All the deduced amino acid sequences mentioned in the results and discussion are included as a text file in supplemental materials.

The transcriptome shotgun assembly was deposited at GenBank/DDBJ/EMBL under the accession number GEFN01000000.1 (BioProject: PRJNA309860, BioSample: SAMN04441069, Sequence Read Archive: SRR3124666). A taxonomic analysis based on Blast best hits revealed that the Nevada dampwood termite *Zootermopsis nevadensis* and water flea *Daphnia pulex* were the species with the highest number of similar proteins to *P. argentinus*, among other invertebrates (Fig. 1).

Table 2
RNA-seq assembly and annotation statistics.

Total number of paired reads	101,251,880
Total number of transcripts	24,378
Total number of bases in transcripts	12,265,959
GC content	39.42%
N50/L50	621/5583
Median transcript length (bp)	343
Average transcript length (bp)	504
Total number of predicted proteins (possible ORFs)	15,476
Total number of transcripts with one or more ORFs	12,424
Total number of proteins with annotation	10,238
Total number of transcripts with one or more annotated open reading frames (ORF)	7383
Completeness of transcripts based on translation (% of total predicted proteins)[a]	
Complete	24%
5′partial	37%
3′partial	7%
Internal	32%

5′partial = stop codon found but not a start codon (Met);
3′partial = start codon found (Met) but not a stop codon;
Internal = start and stop codons not found.
[a] Complete = start and stop codons found.

This work was aimed to provide information on the transcripts for molecular expression studies or to produce recombinant annotated proteins and advance the biochemical toxicology in invertebrates. Individual nucleotide sequences mentioned in the text may be retrieved from the following website: https://www.ncbi.nlm.nih.gov/Traces/wgs/GEFN01?display=contigs&page=1. The *P. argentinus* transcriptome shotgun assembly (TSA) project has the accession number GEFN00000000, and consists of sequences GEFN01000001-GEFN01024378. Since *P. argentinus* is commonly used in ecotoxicological studies, gene families involved in detoxification were identified.

Twenty-seven transcript sequences encoding proteins of the cytochrome P450 superfamily (CYP) with sizes from 205 nt (GEFN01015458.1) to 1519 nt (GEFN01001415.1) were found. After a BLAST search against all invertebrate non-insect P450s in Nelson's database (http://blast.uthsc.edu/) (Nelson, 1999), and based on sequence homology to best hits, a high correspondence to CYPs from the freshwater flea, shrimp or crabs was found.

Two sequences encoding enzymes of the glutathione transferase (GSTs, EC 2.5.1.18) supergene family involved in the phase II detoxification system were found in the assembly. An 855 nt sequence (GEFN01007361.1) showed similarity with the soluble cytosolic mu-GST subclass reported in the crustaceans *Lepeophtheirus salmonis* (GenBank acc. No: ADD38533). In addition, sequence GEFN01023048.1) showed high identity to a delta glutathione transferase from *Palaemon carinicauda* (GenBank acc. No. AGZ89666.1). Finally, sequence GEFN01000783.1 showed high homology with a C-terminal fragment of the alpha helical domain of a theta-GST cytosolic subclass found in the crayfish *Procambarus clarkii* (AEB54656.1) and the shrimp *Penaeus monodon* (APP91325.1).

A detailed analysis of the *P. argentinus* transcriptome revealed a number of genes encoding antioxidant enzymes involved in defense against oxidative stress. Five sequences were annotated as superoxide dismutases (SODs) in the *P. argentinus* transcriptome. A small transcript of 278 nt (GEFN01021101.1) for a Cu-Zn SOD similar to Filobasidiella/Cryptococcus was similar to the Cu-Zn SOD of the Chinese mitten crab *Eriocheir sinensis* (AEP17493.1), to whom it had a 53% identity.

The largest transcript was 1817 nt logn (GEFN01017685.1), similar to Cu-Zn chloroplast SOD. This large SOD has a large identity with the C-terminal domain of two SOD isoforms from *Marsupenaeus japonicus* (BAP28201.1 with 76%) and BAP28203.1 with 63% identity). The whole sequence was 31% identical to an unannotated gene (GenBank EFX88223.1) from *Daphnia pulex* with 31% overall identity in the length of the protein.

In the present study, 63 unigenes annotated as proteins involved in the electron transport chain and oxidative phosphorylation pathways were identified in the *P. argentinus* transcriptome. Most of these sequences are nuclear-encoded partial or complete proteins.

Besides its key role on the respiratory chain, and its participation on the inherent, natural, continuous and harmful formation of reactive oxygen species (ROS), ATP-synthase is targeted by toxins and chemical compounds (Hong and Pedersen, 2008). Some xenobiotics are respiratory chain uncouplers (Boelsterli, 2007), therefore molecular information into the aerobic energy production machinery is fundamental for molecular toxicology. In the current RNA-seq data analysis of the *P. argentinus* transcriptome, two transcripts encoding the endogenous ATPase inhibitor IF1 protein (Chimeo et al., 2015) were identified (GEFN01024353.1 and GEFN01024354.1). The small transcript encodes a peptide that is 66% identical to L. *vannamei* IF1 mitochondrial inhibitory protein. IF1 regulates the complex V, is a highly hydrophobic protein, and contains the conserved IATP (mitochondial ATPase inhibitor) domain that controls ATPase activity under oxidative stress conditions (van Raaij et al., 1996). IF1 maintains the ATP levels and the cellular homeostatic state. Additional sequence analysis of the *P.*

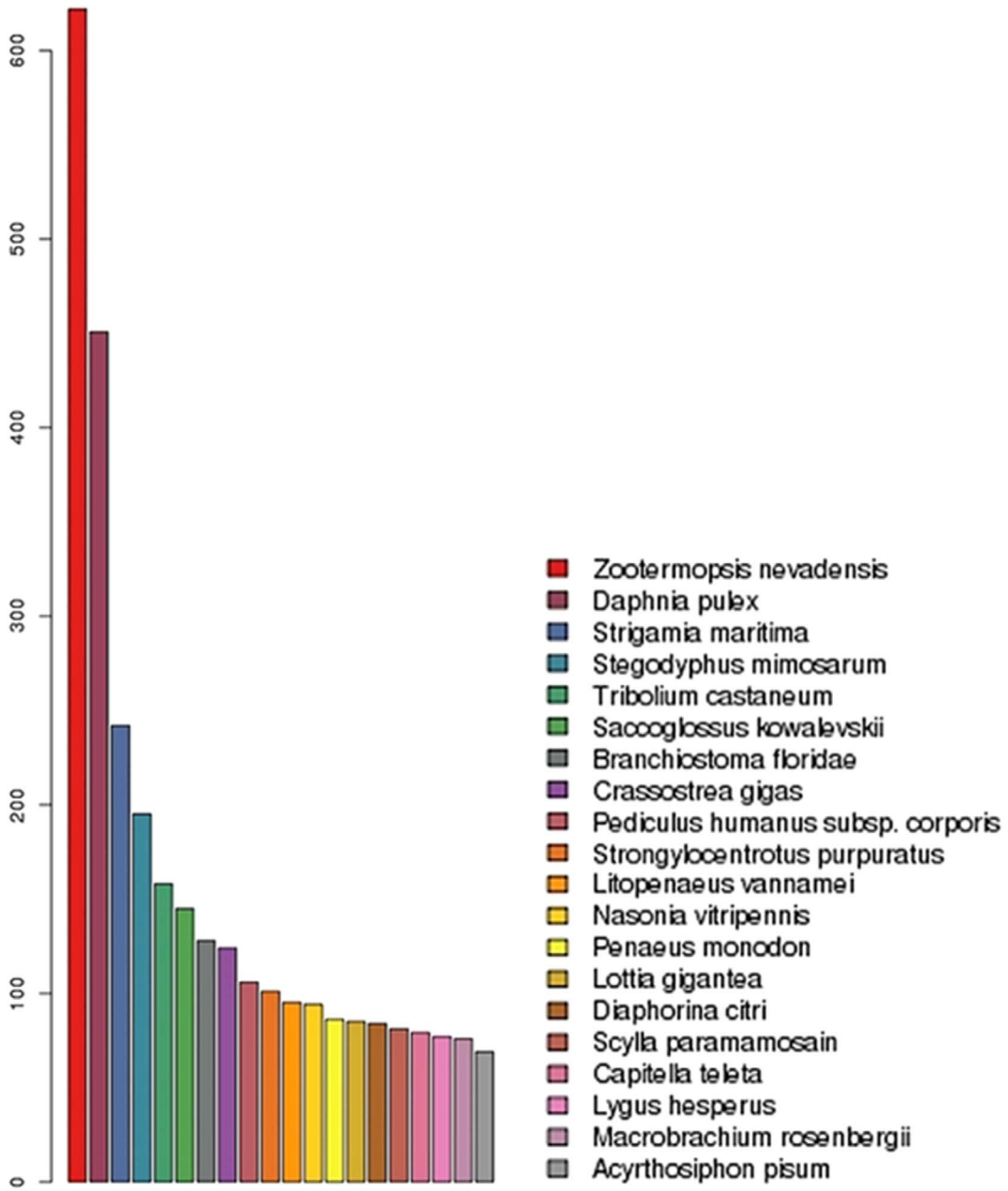

Fig. 1. Blast hits abundance based on taxonomical classification.

argentinus transcriptome related to phase I and II detoxification, enzymatic antioxidant defense, osmoregulation, DNA replication, bioenergetics, and digestion is presented in Supplementary Materials.

4. Discussion

Studies about the analyses of transcriptomes of non-model organisms contribute, among other things, to elucidate the functions of the different genes comprising its genomes (Raherison et al., 2015). However, despite their unparalleled biodiversity, our knowledge on transcriptomes from marine invertebrate species remains scarce, which has hampered the possibility to explore proteins with novel functions. In this study, we have found functional redundancy in a number of transcripts belonging to different gene families. Thus, for example, 27 transcript sequences encoding proteins of the cytochrome P450 superfamily were found. This functional redundancy may provide a selective advantage to the organism either by buffering the effects of neutral loss-of-function mutations over evolutionary time, or by subfunctionalization and neofunctionalization of particular tasks (Lynch and Force, 2000; Ohno, 1970), which increases the response plasticity under stress conditions. This functional redundancy suggests that the transcriptome of *P. argentinus* is highly dynamic in response to changing cell states, environmental conditions, or stressors.

Moreover, it is noteworthy to mention that the highest abundance of Blast hits of our sequences showed high similarity with the Nevada dampwood termite *Zootermopsis nevadensis*. Although it may appear controversial, there are similitudes between the decapod and the termite neuropeptidomes that would need to be further studied into an evolutionary perspective (Veenstra, 2016). Noteworthy is that for the *de novo* transcriptome of the freshwater shrimp *Paratya australiensis* (Decapoda: Atyidae), the termite *Z. nevadensis* appears as one of the best hits in their species transcriptome comparison (Bain et al., 2016).

From our perspective, the fact that the transcriptome of *P. argentinus* shares some similarity with other non-crustacean Arthropods

Table 3
List of CYPs partial expressed transcripts, identified from the *P. argentinus* transcriptome with homologs in GenBank. P450s were identified by using the BlastX tool to search against non-redundant protein sequences.

Original order	Sequence ID	Locus	Sequence length (nt)	Name and best hit	Identity (%)	Aln length (aa)
1	cds.comp234_c0_seq1	GEFN01018194	261	CYP2-clan-fragment1, CYP330A1_*Carcinus maenas* (shore crab)	48.28	58
2	cds.comp738_c0_seq1	GEFN01017694	257	CYP3217-fragment2, CYP3217A-fragment1_Hyalella_azteca	67.00	52
3	cds.comp3323_c0_seq1	GEFN01019097	548	CYP3213F-fragment2, CYP3213A8_Hyalella_azteca	55.06	178
4	cds.comp4034_c0_seq1	GEFN01017675	1010	CYP370C1, CYP370A10_*Daphnia pulex*	42.77	325
5	cds.comp4034_c0_seq2	GEFN01017676	850	CYP370C1. CYP370B2_Daphnia pulex	36.68	199
6	cds.comp6449_c0_seq1	GEFN01007680	421	CYP355E-fragment1, CYP355A2_shrimp (Litopenaeus)	52.67	150
7	cds.comp7192_c0_seq1	GEFN01022480	537	extra-macrochaetae-like, CYP3044B10_Brachionus_rotundifloris	32.69	52
8	cds.comp7195_c0_seq1	GEFN01022480	537	CYP4V-fragment4, CYP4V21_Orconectes limosus (crayfish)	74.77	111
9	cds.comp7882_c0_seq1	GEFN01021267	784	CYP4V-fragment3, CYP4V16_Carcinus maenas (shore crab)	81.63	196
10	cds.comp8133_c0_seq1	GEFN01020629	311	CYP3213A-fragment9, CYP3213A-fragment6_Hyalella_azteca	55.34	103
11	cds.comp8133_c1_seq1	GEFN01019097	548	CYP3213F1, CYP3213A8_Hyalella_azteca	50.59	340
12	cds.comp8480_c0_seq1	GEFN01022022	291	CYP4V-fragment5, CYP4V18_*Litopenaeus vannamei* (Pacific white shrimp)	74.07	81
13	cds.comp9484_c0_seq1	GEFN01022757	435	CYP3217-fragment1, CYP3217A-fragment1_Hyalella_azteca	53.00	134
14	cds.comp10618_c0_seq1	GEFN01000904	323	CYP355E1, CYP355A1_shrimp (Litopenaeus)	52.03	296
15	cds.comp10618_c1_seq1	GEFN01000904	323	CYP355-fragment, CYP355A1_shrimp (Litopenaeus)	62.38	101
16	cds.comp10704_c0_seq1	GEFN01001415	1519	CYP4V41, CYP4V20_Macrobrachium nippinense (fresh water shrimp)	87.89	355
17	cds.comp10704_c0_seq2	GEFN01001414	1503	CYP4V41, CYP4V20_Macrobrachium nippinense (fresh water shrimp)	87.32	355
18	cds.comp10732_c3_seq1	GEFN01023848	786	CYP3213E1, CYP379A1_green shore crab Carcinus maenas	56.25	96
19	cds.comp10732_c3_seq3	GEFN01023848	786	CYP3213E1, CYP379A1_green shore crab Carcinus maenas	56.25	96
20	cds.comp10732_c3_seq4	GEFN01023847	1299	CYP3213E1, CYP3213A8_Hyalella_azteca	44.44	405
21	cds.comp14724_c0_seq1	GEFN01007778	684	CYP3213F-fragment1, CYP3213A1_Hyalella_azteca	54.76	210
22	cds.comp15764_c0_seq1	GEFN01015294	811	CYP379-frgament1, CYP379B7_lobster	57.96	157
23	cds.comp18760_c0_seq1	GEFN01010697	309	CYP4V-fragment1, CYP4V16_Carcinus maenas (shore crab)	76.47	102
24	cds.comp18980_c0_seq1	GEFN01008408	343	CYP4V-fragment2, CYP4V20_Macrobrachium nippinense (fresh water shrimp)	73.91	115
25	cds.comp19393_c0_seq1	GEFN01015458	205	CYP379B-fragment1, CYP379B1 *Petrolisthes cinctipes* (crab)	59.00	44
26	cds.comp19795_c0_seq1	GEFN01015321	258	CYP3012-fragment1, CYP3012B2_Limulus_polyphemus	49.09	55
27	cds.comp19982_c0_seq1	GEFN01012645	215	CYP3213-fragment2, CYP3213A7_Hyalella_azteca	45.16	62

(*Zootermopsis nevadensis*, *Strigamia maritima*, *Nasonia vitripenis*, among others), as well as other invertebrates as mollusks, accurately reflects the complex dynamics of transcriptome evolution in this species. However, further investigation is required to fully understand the evolutionary origin of these sequences.

Detoxification is a fundamental process for the maintenance of cellular homeostasis against conditions, as environmental toxins and pollutants, which threaten the fitness or survival of living organisms. This response relies on a plethora of complex biochemical processes integrated to avoid harmful effects that may cause cellular dysfunction.

In this study, the *P. argentinus* transcriptome, a species remarkably sensitive to a wide range of pollutants, was investigated using RNA sequencing (RNA-seq), with a particular emphasis on the role of genes associated with xenobiotic phase II detoxification. A total of 27 sequences encoding for different members of the cytochrome P450 superfamily (CYP) were discovered in the *P. argentinus* transcriptome. However, some of these transcripts were not clearly identified by BLASTX to be assigned to a particular P450 family reliably (sequences assigned to a P450 family are enlisted in Table 3). Prominent among these sequences are those inducible by benzo(*a*)pyrene (BaP). BaP, a ubiquitous environmental carcinogenic agent, is a prototypical polycyclic aromatic hydrocarbon (PAH) (Nebert et al., 2013), which enters the marine environment through various routes as industrial discharges, oil spills, airborne fallout, and urban runoffs (Abdel-Shafy and Mansour, 2016; Xiu et al., 2014; Stout and Graan, 2010; Men et al., 2009), and its presence in the marine environment has become an increasingly serious issue in recent years (Ren et al., 2015). As any biotransformation of xenobiotics, biotransformation of BaP is divided into phases I and II. On phase I, exogenous compounds are transformed into a more polar metabolite by unmasking or *de novo* formation of functional groups (e.g. -OH, -NH2, -SH). P450 enzymes are key players in the phase I metabolism (Jancova et al., 2010). On phase II xenobiotics are further converted into water soluble less toxic or inactive metabolites that are more easily excretable (Omiecinski et al., 2011; Jancova et al., 2010). Enzymes such as glutathione S-transferases (GSTs) play a major role in phase II metabolism. Thus, the occurrence of this gene on *P. argentinus* suggests that it may be exploited for detoxification. Also, sequences belonging to CYP3 and CYP4 clans from P450s were found in *P. argentinus*, and they are well recognized as responsible for insecticide resistance (Reddy et al., 2012). In particular, CYP4 has been postulated as a biomarker of xenobiotic exposure (David et al., 2003; Snyder, 1998).

Three sequences with high identities to the μ-, two of the θ- and two microsomal GST subclasses were found in *P. argentinus*. Among the phase II detoxification enzymes of crustaceans, the GSTs are probably the most extensively studied. GSTs are implicated in a wide variety of physiological processes as detoxification of endobiotic and xenobiotic toxins, protection against chemical and oxidative stress, transport of intracellular metabolites and hormones, endogenous metabolites or exogenous chemicals (Salazar-Medina et al., 2010; Zhao et al., 2010; Hansen et al., 2008; Lee et al., 2008; Adewale and Afolayan, 2005; Contreras-Vergara et al., 2004). However, there is limited genomic and proteomic information on GSTs from freshwater crustaceans. Thus, our results may provide a valuable framework to better understand, at the molecular level, the role played by GSTs in the detoxification pathways used by *P. argentinus*.

Roncalli and collaborators have identified 41 GSTs after thorough data mining in several *Calanus finmarchicus* transcriptomes (Roncalli et al., 2015). In this study, we identified 10 GST sequences from samples in the natural environment. It is possible that under exposure to xenobiotics and stress, *P. argentinus* expresses other GSTs not identified in this work. Therefore, RNA-seq would be an ideal tool to further understand the differential expression of the phase II detoxification response in this decapod. Nonetheless, the GSTs from *P. argentinus* identified here are candidate biomarkers for use in pollution monitoring studies of aquatic contamination.

Pollution is strongly tied to most kinds of human activities, and as a consequence aquatic environments have been seriously affected by a myriad of chemical compounds that can react in unexpected ways (Crain et al., 2008). This has exposed the aquatic wildlife to a huge variety of chemical compounds (Benedetti et al., 2015). Environmental pollutants, as heavy metals, PAHs, and polychlorinated biphenyls (PCBs), promote the intracellular formation of ROS, as the superoxide anion ($O_2^{\bullet -}$) (Benedetti et al., 2015; Bayol-Denizot et al., 2000). This anion is implicated in the oxidation of proteins (Brand et al., 2004), lipid peroxidation (Kellogg and Fridovich, 1975), and DNA damage (Keyer et al., 1995). Thus, aquatic organisms must be able to deal with the oxidative stress provoked by an increasing pollutant load. In this study, five sequences were annotated as superoxide dismutases (SODs) in the *P. argentinus* transcriptome. Two transcripts showed high homology to the intracellular forms of the copper/zinc superoxide dismutase (CuZnSOD) of the shrimp *Marsupenaeus japonicas*, and one sequence was homologous to the extracellular CuZnSOD form of the Chinese mitten crab *Eriocheir sinensis*. Superoxide dismutases (SODs) are enzymes that rapidly catalyze the transformation of $O_2^{\bullet -}$ into molecular oxygen (O_2) and hydrogen peroxide (H_2O_2) (Fridovich, 1995). Several studies have demonstrated the role of these enzymes in xenobiotic detoxification (Park et al., 2012; Li et al., 1996a,b; Canada and Calabrese, 1989), and less known is the effect of xenobiotics on the expression of SODs in crustaceans (Ragunathan, 2017; Gorokhova et al., 2013; Chaufan et al., 2006). Thus, this study demonstrates that *P. argentinus* harbors a remarkable arsenal of effective detoxification protective mechanisms against the ROS generated by exposure to toxic compounds.

A total of 63 transcripts encoding key enzymes involved in the electron transport chain and oxidative phosphorylation (OXPHOS) pathways were identified in the annotated *P. argentinus* transcriptome. Several studies have demonstrated that the functionality of the OXPHOS system may be impaired by different chemical compounds (Llobet et al., 2015; Nadanaciva and Will, 2011). The ATP-synthase is a tighlty regulated multimeric complex that synthesizes ATP in a coupled reaction with an electrochemical proton-gradient produced by the electron transport system (Martínez-Cruz et al., 2012). This RNA-seq analysis revealed two transcripts encoding the endogenous ATPase inhibitor IF1 protein. The F_OF_1-ATPase inhibitory factor 1 (IF1) binds to the catalytic F_1 domain of the F_OF_1-ATPase and inhibits the hydrolysis of ATP without affecting its synthesis (Campanella et al., 2008; Bason et al., 2011; Chimeo et al., 2015). Recently, Jimenez et al. (2000) found that the expression of IF1 mRNA in liver of rats was induced by pregnenolone-16α-carbonitrile (PCN), a potent catatoxic steroid that confers resistance to toxic compounds to rats. Apparently, the induction of IF1 mRNA may be part of a complex physiological response to maintain cellular homeostasis during toxic stress. It is worth noting that, although there is no evidence yet of a detoxification role of the *P. argentinus* IF1 protein, and further studies are definitely needed to elucidate its potential contribution to the xenobiotic defense system in *P. argentinus*. This work demonstrates the presence of a complex response toward toxic compounds in this non-model organism. Although not all transcripts are complete, those could be completed using rapid amplification of cDNA ends (RACE). Furthermore, the implicit relevance of this finding is that it opens new research directions for researchers with interest in molecular toxicology.

5. Conclusion

This study is the first integrative analysis of RNA-seq for a palaemonid that describes a large number of transcripts of the freshwater crustacean *P. argentinus* involved in diverse biological pathways. It offers a deeper insight into the molecular mechanisms used by this organism for detoxification, bioenergetics, and osmoregulation. Our findings demonstrate the presence of phase *I and* II detoxification metabolic pathways in *P. argentinus*. Furthermore, as a perspective, we propose the use of several cited genes as molecular markers to perform quantitative expression studies (RT-qPCR) challenging *P. argentinus* to a wide battery of xenobiotics for understanding the impact of environmental perturbation on this species. In future studies, RNA-seq differential expression could be used to evaluate global responses toward exposure to different pollutants and may elucidate the specific enzymes and pathways activated. Also, the proteins related to bioenergetics and osmoregulation are potential subjects of biochemical and structural study to understand the mechanisms behind salinity adaptation of this South American decapod crustacean.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Authors acknowledge support from CONACYT (Mexico's National Council for Science and Technology) for Mexico-Argentina grant C0005-2012-189333 to R. Sotelo-Mundo and CONICET (Argentina's National Council for Scientific and Technical Research) grant COOP-CTES-D680 to C.F. García. CONACYT also supported A.A. Lopez-Zavala under the grant Fondo Consolidacion Institucional (I0007-2015-01, #250973). C.F. García also thanks Argentina's Agencia Nacional de Promoción Científica y Tecnológica for grant PICT 2014-2580. G. Lukaszewicz thanks Mexico's State Department excellence postdoctoral scholarship (Beca de Excelencia AMEXID-SRE). The authors thank Prof. David R. Nelson (The University of Tennessee Health Science Center) for expert advice in naming and annotation of the partial cytochrome P450 transcripts reported in Table 3. We thank the UNAM-CIAD 2016-2017 academic exchange program and Coordinacion de Investigacion from CIAD for support. We thank Lucia Angulo (CIAD) and Jerome Verleyen (IBT-UNAM) for technical support. Authors acknowledge and declare that do not have conflicts of interest in this publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.margen.2017.08.009.

References

Abdel-Shafy, H.I., Mansour, M.S.M., 2016. A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation. Egypt. J. Pet. 25, 107–123.

Adewale, I.O., Afolayan, A., 2005. Organ distribution and kinetics of glutathione transferase from African river prawn, *Macrobrachium vollenhovenii* (Herklots). Aquat. Toxicol. 71, 193–202.

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. The gene ontology consortium. Nat. Genet. 25, 25–29.

Bain, P.A., Gregg, A.L., Kumar, A., 2016. *De novo* assembly and analysis of changes in the protein-coding transcriptome of the freshwater shrimp *Paratya australiensis* (Decapoda: Atyidae) in response to acid sulfate drainage water. BMC Genomics 17, 890.

Bason, J.V., Runswick, M.J., Fearnley, I.M., Walker, J.E., 2011. Binding of the inhibitor protein IF_1 to bovine F_1-ATPase. J. Mol. Biol. 406, 443–453.

Bauer, R.T., 2005. Remarkable shrimps: adaptations and natural history of the carideans. Afr. J. Aquat. Sci. 30, 211–212.

Bayol-Denizot, C., Daval, J.L., Netter, P., Minn, A., 2000. Xenobiotic-mediated production of superoxide by primary cultures of rat cerebral endothelial cells, astrocytes, and neurones. Biochim. Biophys. Acta 1497, 115–126.

Benedetti, M., Giuliani, M.E., Regoli, F., 2015. Oxidative metabolism of chemical pollutants in marine organisms: molecular and biochemical biomarkers in environmental toxicology. Ann. N. Y. Acad. Sci. 1340, 8–19.

Bertrand, L., Monferran, M.V., Metais, I., Mouneyrac, C., Ame, M.V., 2015. MTs in *Palaemonetes argentinus* as potential biomarkers of zinc contamination in freshwaters. Ecol. Indic. 48, 533–541.

Boelsterli, U.A., 2007. Xenobiotic induced oxidative stress: cell injury, signaling and gene regulation. In: Boelsterli, U.A. (Ed.), Mechanistic Toxicology: The Molecular Basis of How Chemicals Disrupt Biological Targets. CRC Press, Boca Raton, pp. 93–147.

Brand, M.D., Buckingham, J.A., Esteves, T.C., Green, K., Lambert, A.J., Miwa, S., Murphy, M.P., Pakay, J.L., Talbot, D.A., Echtay, K.S., 2004. Mitochondrial superoxide and aging: uncoupling-protein activity and superoxide production. Biochem. Soc. Symp. 71, 203–213.
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST +: architecture and applications. BMC Bioinf. 10, 421.
Campanella, M., Casswell, E., Chong, S., Farah, Z., Wieckowski, M.R., Abramov, A.Y., Tinker, A., Duchen, M.R., 2008. Regulation of mitochondrial structure and function by the F_1F_0-ATPase inhibitor protein, IF_1. Cell Metab. 8, 13–25.
Canada, A.T., Calabrese, E.J., 1989. Superoxide dismutase: its role in xenobiotic detoxification. Pharmacol. Ther. 44, 285–295.
Chaufan, G., Juárez, A., Basack, S., Ithuralde, E., Sabatini, S.E., Genovese, G., Oneto, M.L., Kesten, E., de Molina, Ríos, M.del C., 2006. Toxicity of hexachlorobenzene and its transference from microalgae (*Chlorella kessleri*) to crabs (*Chasmagnathus granulatus*). Toxicology 227, 262–270.
Chimeo, C., Fernandez-Gimenez, A.V., Campanella, M., Mendez-Romero, O., Muhlia-Almazán, A., 2015. The shrimp mitochondrial F_oF_1-ATPase inhibitory factor 1 (IF_1). J. Bioenerg. Biomembr. 47, 383–393.
Collins, P., Cappello, S., 2006. Cypermethrin toxicity to aquatic life: bioassays for the freshwater prawn *Palaemonetes argentinus*. Arch. Environ. Contam. Toxicol. 51, 79–85.
Conesa, A., Gotz, S., 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int. J. Plant Genomics 2008, 619832.
Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.
Contreras-Vergara, C.A., Harris-Valle, C., Sotelo-Mundo, R.R., Yepiz-Plascencia, G., 2004. A mu-class glutathione S-transferase from the marine shrimp *Litopenaeus vannamei*: molecular cloning and active-site structural modeling. J. Biochem. Mol. Toxicol. 18, 245–252.
Crain, C.M., Kroeker, K., Halpern, B.S., 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecol. Lett. 11, 1304–1315.
David, P., Dauphin-Villemant, C., Mesneau, A., Meyran, J.C., 2003. Molecular approach to aquatic environmental bioreporting: differential response to environmental inducers of cytochrome P450 monooxygenase genes in the detritivorous subalpine planktonic *Crustacea*, *Daphnia pulex*. Mol. Ecol. 12, 2473–2481.
Diaz de Cerio, O., Bilbao, E., Ruiz, P., Pardo, B.G., Martínez, P., Cajaraville, M.P., Cancio, I., 2017. Hepatic gene transcription profiles in turbot (*Scophthalmus maximus*) experimentally exposed to heavy fuel oil n 6 and to styrene. Mar. Environ. Res. 123, 14–24.
Eddy, S.R., 2011. Accelerated profile HMM searches. PLoS Comput. Biol. 7, e1002195.
Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. Nucleic Acids Res. 42, D222–D230.
Freire, C.A., Cavassin, F., Rodrigues, E.N., Torres, A.H., McNamara, J.C., 2003. Adaptive patterns of osmotic and ionic regulation, and the invasion of fresh water by the palaemonid shrimps. Comp. Biochem. Physiol. 136A, 771–778.
Fridovich, I., 1995. Superoxide radical and superoxide dismutases. Annu. Rev. Biochem. 64, 97–112.
Galanti, L.N., Ame, M.V., Wunderlin, D.A., 2013. Accumulation and detoxification dynamic of cyanotoxins in the freshwater shrimp *Palaemonetes argentinus*. Harmful Algae 27, 88–97.
García-de la Parra, L.M., Bautista-Covarrubias, J.C., Rivera-de la Rosa, N., Betancourt-Lozano, M., Guilhermino, L., 2006. Effects of methamidophos on acetylcholinesterase activity, behavior, and feeding rate of the white shrimp (*Litopenaeus vannamei*). Ecotoxicol. Environ. Saf. 65, 372–380.
Ghaffari, N., Sánchez-Flores, A., Doan, R., García-Orozco, K.D., Chen, P.L., Ochoa-Leyva, A., Lopez-Zavala, A.A., Carrasco, J.S., Hong, C., Brieba, L.G., Rudiño-Piñera, E., Blood, P.D., Sawyer, J.E., Johnson, C.D., Dindot, S.V., Sotelo-Mundo, R.R., Criscitiello, M.F., 2014. Novel transcriptome assembly and improved annotation of the whiteleg shrimp (*Litopenaeus vannamei*), a dominant crustacean in global seafood mariculture. Sci Rep 4, 7081.
Gorokhova, E., Löf, M., Reutgard, M., Lindström, M., Sundelin, B., 2013. Exposure to contaminants exacerbates oxidative stress in amphipod *Monoporeia affinis* subjected to fluctuating hypoxia. Aquat. Toxicol. 127, 46–53.
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.
Griboff, J., Morales, D., Bertrand, L., Bonansea, R.I., Monferran, M.V., Asis, R., Wunderlin, D.A., Amé, M.V., 2014. Oxidative stress response induced by atrazine in *Palaemonetes argentinus*: the protective effect of vitamin E. Ecotoxicol. Environ. Saf. 108, 1–8.
Griffitt, R.J., Greig, T.W., Chandler, G.T., Quattro, J.M., 2007. Serial analysis of gene expression reveals identifiable patterns in transcriptome profiles of *Palaemonetes pugio* exposed to three common environmental stressors. Environ. Toxicol. Chem. 26, 2413–2419.
Gust, K.A., Najar, F.Z., Habib, T., Lotufo, G.R., Piggot, A.M., Fouke, B.W., Laird, J.G., Wilbanks, M.S., Rawat, A., Indest, K.J., Roe, B.A., Perkins, E.J., 2014. Coral-zooxanthellae meta-transcriptomics reveals integrated response to pollutant stress. BMC Genomics 15, 591.
Hansen, B.H., Nordtug, T., Altin, D., Booth, A., Hessen, K.M., Olsen, A.J., 2008. Gene expression of GST and CYP330A1 in lipid-rich and lipid-poor female *Calanus finmarchicus* (Copepoda:Crustacea) exposed to dispersed oil. J. Toxicol. Environ. Health 72A, 131–139.
Harms, L., Frickenhaus, S., Schiffer, M., Mark, F.C., Storch, D., Pörtner, H.O., Held, C., Lucassen, M., 2013. Characterization and analysis of a transcriptome from the boreal spider crab *Hyas araneus*. Comp. Biochem. Physiol. 8D, 344–351.
Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., Richter, J., Rubin, G.M., Blake, J.A., Bult, C., Dolan, M., Drabkin, H., Eppig, J.T., Hill, D.P., Ni, L., Ringwald, M., Balakrishnan, R., Cherry, J.M., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S., Fisk, D.G., Hirschman, J.E., Hong, E.L., Nash, R.S., Sethuraman, A., Theesfeld, C.L., Botstein, D., Dolinski, K., Feierbach, B., Berardini, T., Mundodi, S., Rhee, S.Y., Apweiler, R., Barrell, D., Camon, E., Dimmer, E., Lee, V., Chisholm, R., Gaudet, P., Kibbe, W., Kishore, R., Schwarz, E.M., Sternberg, P., Gwinn, M., Hannick, L., Wortman, J., Berriman, M., Wood, V., de la Cruzm, N., Tonellato, P., Jaiswal, P., Seigfried, T., White, R., Consortium, Gene Ontology, 2004. The gene ontology (GO) database and informatics resource. Nucleic Acids Res. 32, D258–D261.
Holland, C.A., Skinner, D.M., 1977. The organization of the main component DNA of a crustacean genome with a paucity of middle repetitive sequences. Chromosoma 63, 223–240.
Hong, S., Pedersen, P.L., 2008. ATP synthase and the actions of inhibitors utilized to study its roles in human health, disease, and other scientific areas. Microbiol. Mol. Biol. Rev. 72, 590–641.
Hwang, J., Suh, S.S., Park, M., Park, S.Y., Lee, S., Lee, T.K., 2017. Differential gene expression patterns during embryonic development of sea urchin exposed to triclosan. Environ. Toxicol. 32, 426–433.
Ioannidis, P., Lu, Y., Kumar, N., Creasy, T., Daugherty, S., Chibucos, M.C., Orvis, J., Shetty, A., Ott, S., Flowers, M., Sengamalay, N., Tallon, L.J., Pick, L., Gunning Hotopp, J.C., 2014. Rapid transcriptome sequencing of an invasive pest, the brown marmorated stink bug *Halyomorpha halys*. BMC Genomics 15, 738.
Jancova, P., Anzenbacher, P., Anzenbacherova, E., 2010. Phase II drug metabolizing enzymes. Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech Repub. 154, 103–116.
Jimenez, B.D., Quattrochi, L.C., Yockey, C.B., Guzelian, P.S., 2000. Identification by differential display of the IF_1 inhibitor peptide of ATP synthase/ATPase as a gene inducible in rat liver by pregnenolone 16alpha-carbonitrile. Life Sci. 67, 1825–1832.
Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.
Kellogg, E.W., Fridovich, I., 1975. Superoxide, hydrogen peroxide, and singlet oxygen in lipid peroxidation by a xanthine oxidase system. J. Biol. Chem. 250, 8812–8817.
Keyer, K., Gort, A.S., Imlay, J.A., 1995. Superoxide and the production of oxidative DNA damage. J. Bacteriol. 177, 6782–6790.
Kim, J.H., Rhee, J.S., Dahms, H.U., Lee, Y.M., Han, K.N., Lee, J.S., 2012. The yellow catfish, *Pelteobagrus fulvidraco* (Siluriformes) metallothionein cDNA: molecular cloning and transcript expression level in response to exposure to the heavy metals Cd, Cu, and Zn. Fish Physiol. Biochem. 38, 1331–1342.
Lavarias, S., Garcia, C.F., 2015. Acute toxicity of organophosphate fenitrothion on biomarkers in prawn *Palaemonetes argentinus* (Crustacea: Palaemonidae). Environ. Monit. Assess. 187, 65.
Lee, K.W., Raisuddin, S., Rhee, J.S., Hwang, D.S., Yu, I.T., Lee, Y.M., Park, H.G., Lee, J.S., 2008. Expression of glutathione S-transferase (GST) genes in the marine copepod *Tigriopus japonicus* exposed to trace metals. Aquat. Toxicol. 89, 158–166.
Li, Y., Kuppusamy, P., Zweier, J.L., Trush, M.A., 1996a. Role of Cu/Zn-superoxide dismutase in xenobiotic activation. I. Chemical reactions involved in the Cu/Zn-superoxide dismutase-accelerated oxidation of the benzene metabolite 1,4-hydroquinone. Mol. Pharmacol. 49, 404–411.
Li, Y., Kuppusamy, P., Zweier, J.L., Trush, M.A., 1996b. Role of Cu/Zn-superoxide dismutase in xenobiotic activation. II. Biological effects resulting from the Cu/Zn-superoxide dismutase-accelerated oxidation of the benzene metabolite 1,4-hydroquinone. Mol. Pharmacol. 49, 412–421.
Llobet, L., Toivonen, J.M., Montoya, J., Ruiz-Pesini, E., López-Gallardo, E., 2015. Xenobiotics that affect oxidative phosphorylation alter differentiation of human adipose-derived stem cells at concentrations that are found in human blood. Dis. Model. Mech. 8, 1441–1455.
Lynch, M., Force, A., 2000. The probability of duplicate gene preservation by subfunctionalization. Genetics 154, 459–473.
Manfrin, C., Tom, M., De Moro, G., Gerdol, M., Giulianini, P.G., Pallavicini, A., 2015. The eyestalk transcriptome of red swamp crayfish *Procambarus clarkii*. Gene 557, 28–34.
Martínez-Cruz, O., Sánchez-Paz, A., García-Carreño, F.L., Jimenez-Gutierrez, L.R., Navarrete del Toro, M.A., Muhlia-Almazan, A., 2012. In: Clark, K. (Ed.), Invertebrates Mitochondrial Function and Energetic Challenges. Bioenergetics. Intech, Rijeka, Croatia, pp. 181–218.
Mehinto, A.C., Martyniuk, C.J., Spade, D.J., Denslow, N.D., 2012. Applications for next-generation sequencing in fish ecotoxicogenomics. Front. Genet. 3, 62.
Men, B., He, M., Tan, L., Lin, C., Quan, X., 2009. Distributions of polycyclic aromatic hydrocarbons in the Daliao River Estuary of Liaodong Bay, Bohai Sea (China). Mar. Pollut. Bull. 58, 818–826.
Meng, X., Tian, X., Liu, M., Nie, G., Jiang, K., Wang, B., Wang, L., 2014. The transcriptomic response to copper exposure by the gill tissue of Japanese scallops (*Mizuhopecten yessoensis*) using deep-sequencing technology. Fish Shellfish Immunol. 38, 287–293.
Montagna, M.C., Collins, P.A., 2007. Survival and growth of *Palaemonetes argentinus* (Decapoda; Caridea) exposed to insecticides with chlorpyrifos and endosulfan as active element. Arch. Environ. Contam. Toxicol. 53, 371–378.
Morrone, J.J., Lopreto, E.C., 1995. Parsimony analysis of endemicity of freshwater Decapoda (Crustacea: Malacostraca) from Southern South America. Neotropica 41, 3–8.
Nadanaciva, S., Will, Y., 2011. New insights in drug-induced mitochondrial toxicity. Curr. Pharm. Des. 17, 2100–2112.
Nebert, D.W., Shi, Z., Gálvez-Peralta, M., Uno, S., Dragin, N., 2013. Oral benzo[a]pyrene:

understanding pharmacokinetics, detoxication and consequences-*Cyp1* knockout mouse lines as a paradigm. Mol. Pharmacol. 84, 304–313.
Nelson, D.R., 1999. Cytochrome P450 and the individuality of species. Arch. Biochem. Biophys. 369, 1–10.
Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag, Berlin.
Omiecinski, C.J., Vanden-Heuvel, J.P., Perdew, G.H., Peters, J.M., 2011. Xenobiotic metabolism, disposition, and regulation by receptors: from biochemical phenomenon to predictors of major toxicities. Toxicol. Sci. 120S1, S49–S75.
Park, S.Y., Nair, P.M., Choi, J., 2012. Characterization and expression of superoxide dismutase genes in *Chironomus riparius* (Diptera, Chironomidae) larvae as a potential biomarker of ecotoxicity. Comp. Biochem. Physiol. 156C, 187–194.
van Raaij, M.J., Orriss, G.L., Montgomery, M.G., Runswick, M.J., Fearnley, I.M., Skehel, J.M., Walker, J.E., 1996. The ATPase inhibitor protein from bovine heart mitochondria: the minimal inhibitory sequence. Biochemistry 35, 15618–15625.
Ragunathan, M.G., 2017. Vicissitudes of oxidative stress biomarkers in the estuarine crab *Scylla serrata* with reference to dry and wet weather conditions in Ennore Estuary, Tamil Nadu, India. Mar. Pollut. Bull. 116, 113–120.
Raherison, E.S.M., Giguère, I., Caron, S., Lamara, M., MacKay, J.J., 2015. Modular organization of the white spruce (*Picea glauca*) transcriptome reveals functional organization and evolutionary signatures. New Phytol. 207, 172–187.
Reddy, B.P., Labbé, P., Corbel, V., 2012. Culex genome is not just another genome for comparative genomics. Parasit. Vectors 5, 63.
Ren, X., Pan, L., Wang, L., 2015. Toxic effects upon exposure to benzo[a]pyrene in juvenile white shrimp *Litopenaeus vannamei*. Environ. Toxicol. Pharmacol. 39, 194–207.
Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S.D., Mungall, K., Lee, S., Okada, H.M., Qian, J.Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, Y.S., Newsome, R., Chan, S.K., She, R., Varhol, R., Kamoh, B., Prabhu, A.L., Tam, A., Zhao, Y., Moore, R.A., Hirst, M., Marra, M.A., Jones, S.J., Hoodless, P.A., Birol, I., 2010. *De novo* assembly and analysis of RNA-seq data. Nat. Methods 7, 909–912.
Roncalli, V., Cieslak, M.C., Passamaneck, Y., Christie, A.E., Lenz, P.H., 2015. Glutathione S-Transferase (GST) gene diversity in the crustacean *Calanus finmarchicus*-contributors to cellular detoxification. PLoS One 10, e0123322.
Salazar-Medina, A.J., García-Rico, L., García-Orozco, K.D., Valenzuela-Soto, E., Contreras-Vergara, C.A., Arreola, R., Arvizu-Flores, A., Sotelo-Mundo, R.R., 2010. Inhibition by Cu^{2+} and Cd^{2+} of a mu-class glutathione S-transferase from shrimp *Litopenaeus vannamei*. J. Biochem. Mol. Toxicol. 24, 218–222.
Shinn, C., Santos, M.M., Lek, S., Grenouillet, G., 2015. Behavioral response of juvenile rainbow trout exposed to an herbicide mixture. Ecotoxicol. Environ. Saf. 112, 15–21.
Snyder, M.J., 1998. Cytochrome P450 enzymes belonging to the CYP4 family from marine invertebrates. Biochem. Biophys. Res. Commun. 249, 187–190.
Stout, S.A., Graan, T.P., 2010. Quantitative source apportionment of PAHs in sediments of little Menomonee River, Wisconsin: weathered creosote versus urban background. Environ. Sci. Technol. 44, 2932–2939.
Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28, 33–36.
Veenstra, J.A., 2016. Similarities between decapod and insect neuropeptidomes. PeerJ 4, e2043.
Vidas, D., Schei, P.J., 2011. The world ocean in globalisation: Challenges and responses for the Anthropocene epoch. In: Vidas, D., Schei, P.J. (Eds.), The World Ocean in Globalisation. Martinus Nijhoff Publishers, Norway (550 pp).
Wu, R.S.S., 1999. Eutrophication, water borne pathogens and xenobiotic compounds: environmental risks and challenges. Mar. Pollut. Bull. 39, 11–22.
Xiu, M., Pan, L., Jin, Q., 2014. Bioaccumulation and oxidative damage in juvenile scallop Chlamys farreri exposed to benzo[a]pyrene, benzo[b]fluoranthene and chrysene. Ecotoxicol. Environ. Saf. 107, 103–110.
Zhao, D., Chen, L., Qin, C., Zhang, H., Wu, P., Zhang, F., 2010. A delta-class glutathione transferase from the Chinese mitten crab *Eriocheir sinensis*: cDNA cloning, characterization and mRNA expression. Fish Shellfish Immunol. 29, 698–703.

Marine Genomics

De novo assembly and transcriptome characterization of the freshwater prawn *Palaemonetes argentinus*: implications for a detoxification response

C.F. García *et al.*

Supplementary materials

S1. Description of the transcriptome assembly

P. argentinus de novo transcriptome assembly

A total of 50,625,940 Illumina paired-end reads from total shrimp RNA yielded a total of ~7.29 Gb. The quality for all read pairs was above Q30 (which is equivalent to the probability of an incorrect base call in 1000 times, and a base call accuracy of 99.9%) and no further quality filtering was required before the transcriptome assembly. Sequence data was *de novo* assembled using the Trinity pipeline (Grabherr et al., 2011). Despite RNA-seq reads provided quality scores, only 50.9% of the transcripts had one or more translations giving a total of 15,476 predicted proteins or open reading frames (ORF) with known biological function, and the remaining reads resulted from non-coding RNAs, including rRNA which, regardless of the mRNA enrichment process in the library construction, there are always presence of it. Of these protein-coding transcripts, 66.15% found a match in at least one of the functional annotation approaches. Regarding the transcript count, those with some annotation information represented 30.24% of the total number of transcripts.

Functional characterization of *P. argentinus* transcriptome.

To develop molecular markers using functional candidate genes, predicted proteins with KEGG and KOG annotation were grouped based on their frequency, and the top

20 functional annotations are shown in supplementary figures S1A and S1B, respectively.

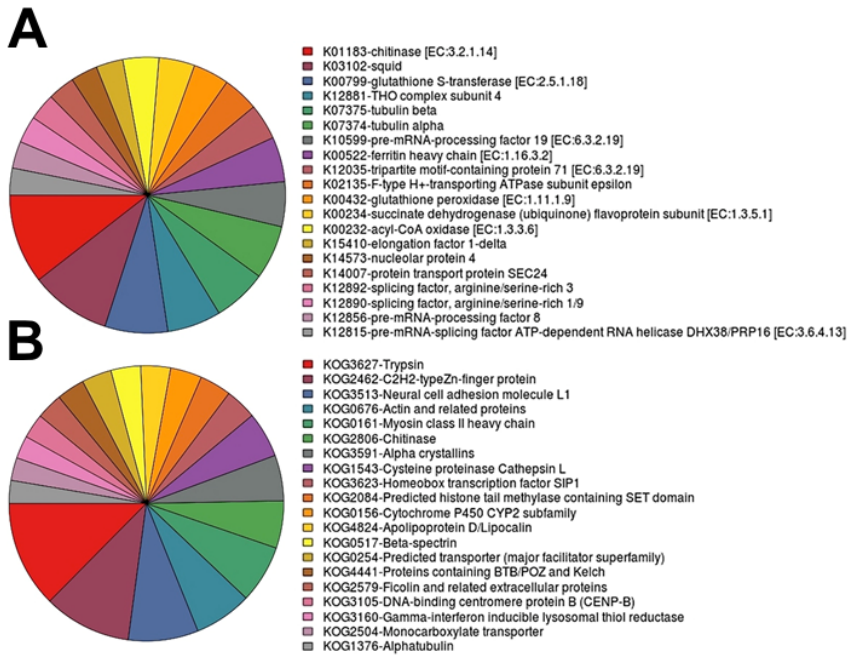

Figure S 1*. Counts for KEGG (A) and KOG (B) functions from annotated transcript products.*

In the case of KEGG functions, three groups contained ~25% of the annotated proteins corresponding to K01183-Chitinase, K03102-Squid and K00799-Glutathione S-transferase activities. As for the KOG annotation, also ~25% of the functions were represented by KOG3627-Trypsin, KOG2462-C2H2-type Zn-finger protein, and KOG3513-Neural cell adhesion molecule L1. Finally, counts from PFAM domain frequencies (Figure S2) revealed functions with high rates such as Zinc-finger double domain (PF13465), RNA recognition motif (PF00076) and reverse transcriptase domain (PF00078) with more than 80% of frequency.

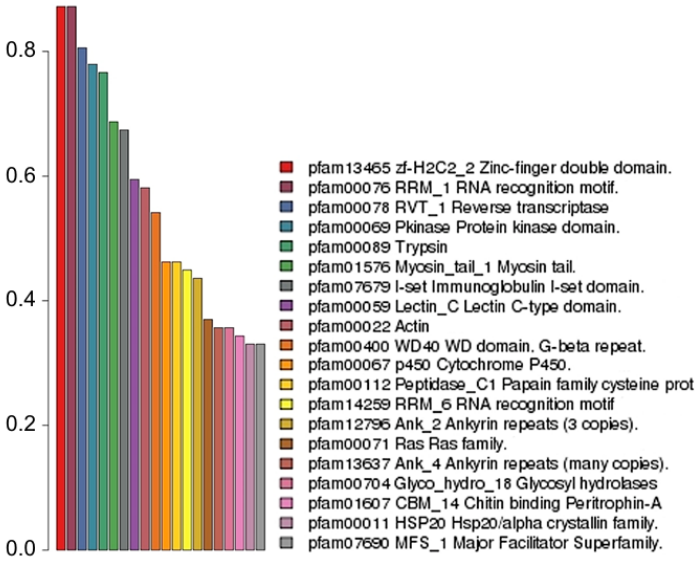

***Figure S 2**. Counts for PFAM domains from annotated transcript products.*

Comparing the PFAM annotation for domain functions of *P. argentinus* (2,160 different PFAM domains) and *L. vannamei* annotations (3,154) (Ghaffari et al., 2014) an overlap of 1,623 domain functions was found, which represents ~75% of the total annotated domains for *P. argentinus*. From the remaining 537 functional domain annotations, 14 domains belonged to ribosomal proteins.

A list of the 20 most expressed transcripts either with or without product annotation is shown in Supp. Table S1 (see at the end of this document). Most of the highly expressed transcripts have translation product annotation related to muscle proteins, which can be explained by the predominant amount of muscle tissue in the shrimp body. In the case of un-annotated transcripts, c10387_g1_i1 (GEFN01000632.1) and c8449_g1_i1 (GEFN01022287.1), they presented a transmembrane prediction and signal peptide region, respectively.

Mitochondrial transcripts

To detect transcripts belonging to the *P. argentinus* mitochondrion, the translation of all the reconstructed transcripts was repeated, although using the arthropod

mitochondrial genetic code. After comparing against mitochondrial genes encoding proteins from the dragonfly *P. serenus*, the mitochondrial genes identified are shown in Supp. Table S2 (Please see at the end of this document).

Enzyme Classification (EC) and KEGG pathway analysis of the transcriptome

The potentially functional enzymes on the *P. argentinus* transcriptome were characterized based on the predictions of Enzyme Commission (EC) numbers for each sequence using Blast2GO software. In total 1,881 sequences were related to EC numbers, and the enzyme classification revealed hydrolases as the largest group of *P. argentinus* enzymes (47.5%, 894 transcripts), followed by transferases (25.1%, 472 transcripts), oxidoreductases (15.3%, 287 transcripts), ligases (5.3%, 99 transcripts), isomerases (4.1%, 77 transcripts) and lyases (2.8%, 52 transcripts) (Supp. Fig. S3).

The number of sequences for each enzyme family is reported in supporting material (Supplementary Figures S4-S10). The 15,476 proteins from shrimp transcriptome were further characterized by KEGG pathway analysis. The predicted enzymes were distributed in 334 KEGG pathways in Supplementary Table S1 (Please see at the end of this document).

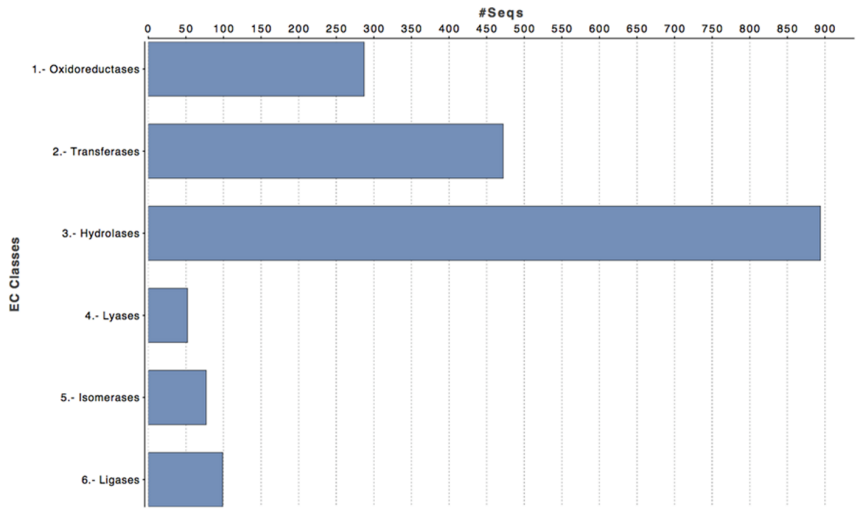

***Figure S 3**. Enzyme Code (EC) distribution.*

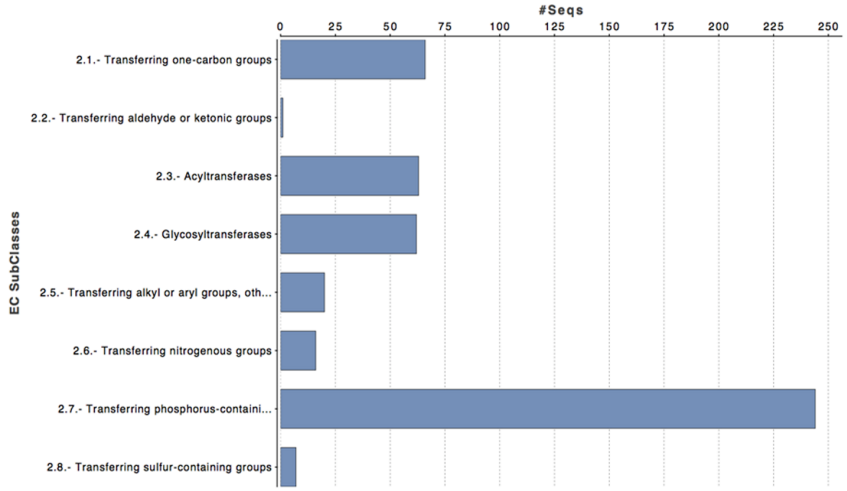

***Figure S 4**. Enzyme code distribution of transferase.*

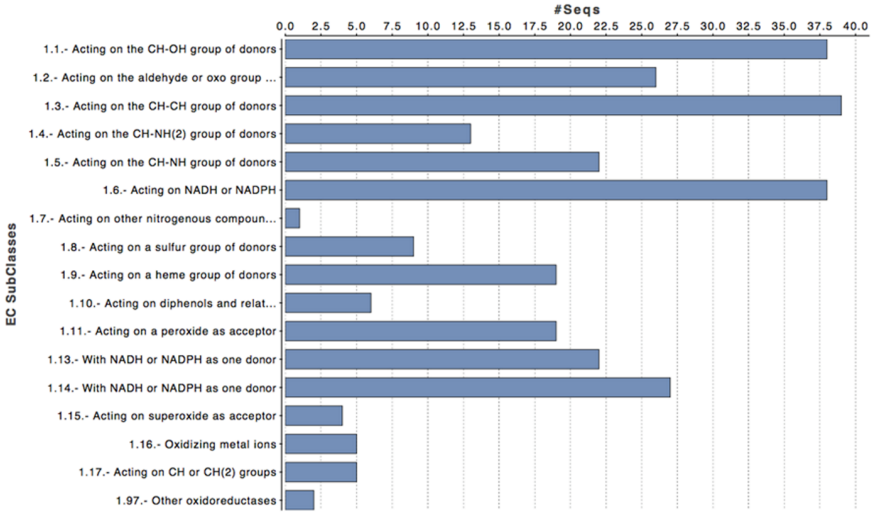

Figure S 5. Enzyme code distribution of oxidoreductases.

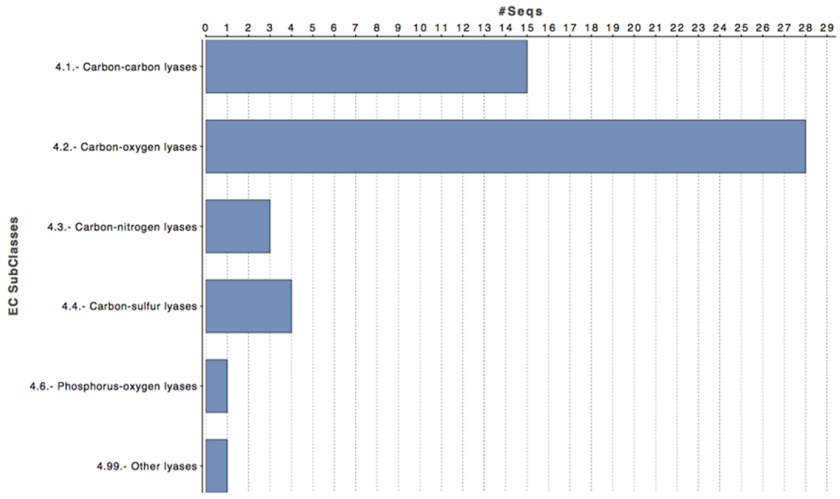

***Figure S 6**. Enzyme code distribution of lyases.*

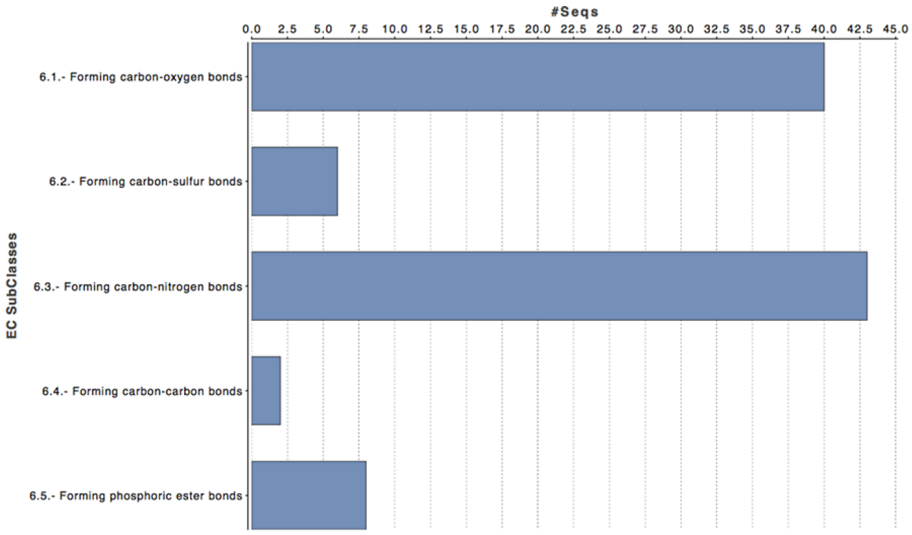

Figure S 7. *Enzyme code distribution of ligases.*

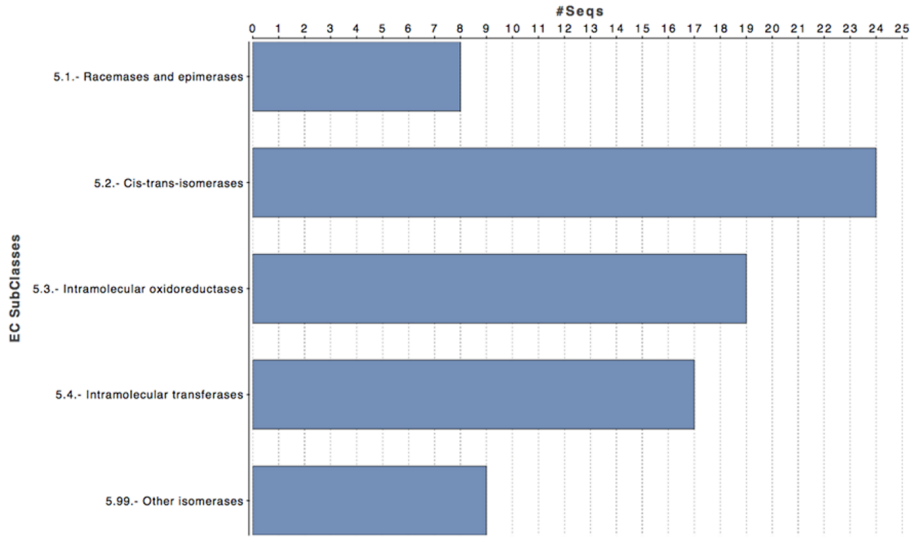

Figure S 8*. Enzyme code distribution of isomerases.*

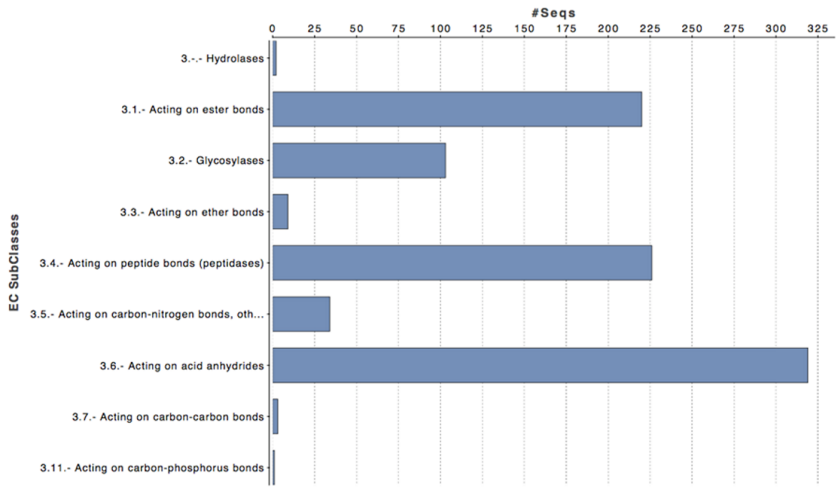

Figure S 9. Enzyme code distribution of hydrolases.

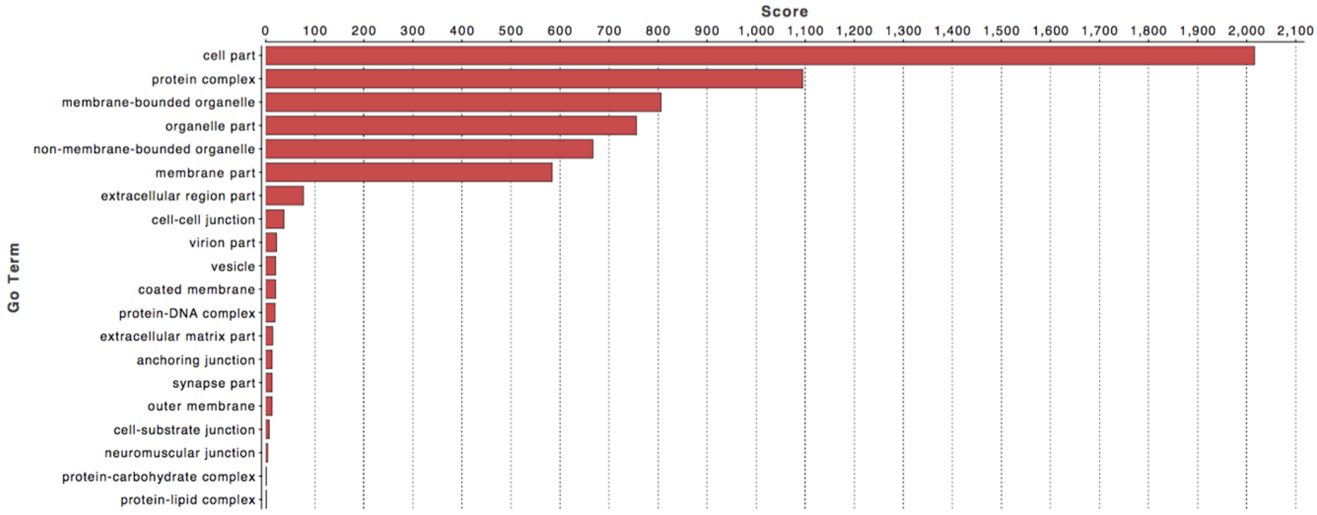

Figure S 10. *Cellular component classification by GO terms.*

S.2. Potential biomarkers in *P. argentinus*.

Several proteins realted to drug metabolism (12 with the drug metabolism-other enzymes, 7 with the metabolism of xenobiotics by cytochrome P450 and 5 with the drug metabolism-cytochrome P450) were found in the transcriptome, indicating a role in xenobiotic detoxification.

Phase I Detoxification- Cytochrome P450 monooxygenases

Cytochrome P450 monooxygenases convert a broad range of lipophilic compounds into more hydrophilic derivatives. They play a critical role both in the synthesis and processing of endogenous metabolites plus degradation of xenobiotic compounds, being a key component of phase I detoxification systems (Feyereisen, 2006).

Three *P. argentinus* CYP sequences, related to a mitochondrial clan, showed high identity percentage to CYP302A1, CYP362A1 and CYP362A2 in *Daphnia pulex*. CYP302A1, a highly conserved Halloween gene, is involved, together with CYP314A1 and CYP315A1 in ecdysone synthesis. Insects share some CYP genes with *D. pulex* in the biosintethic pathway of the molting hormone 20-hydroxyecdysone (CYP302A1,

CYP314A1, CYP315A1). However, insects do not have CYP362A1, CYP362A2 and CYP363A1, suggesting a divergnce within the arthropods and that may have led to novel roles in crustacea (Baldwin et al., 2009).

One of the 13 CYP2-clan sequences found in this transcriptome is similar to the CYP330A1 gene described in the shore crab *Carcinus maenas.* Expression of this gene was induced by ecdysone, ponasterone A, benzo(a)pyrene and phenobarbital in the hepatopancreas of male intermoult crabs (Rewitz et al., 2003), suggesting different roles both in ecdysteroid catabolism and in detoxification of environmental pollutants. Six sequences of *P. argentinus* matched with members of the CYP2L subfamily of the spiny lobster *Panulirus argus.* The last two CYP2-clan sequences found in the *P. argentinus* transcriptome showed homology with subfamily CYP379B of both lobster and crabs, and may be related to a physiological response to environmental stressors (Baldwin et al., 2009).

The CYP3 clan consists of several detoxification enzymes of both xenobiotic and endobiotic compounds. Three sequences belonging to clan 3 were identified in this study, all showed high identity to *D. pulex* CYP360A4, CYP360A5 and CYP361A1, respectively.

The CYP4 family is the most represented in aquatic invertebrates (Snyder, 2007). The seven sequences belonging to the CYP4 family found in *P. argentinus* show high sequence identity with CYP4V16 of the shore crab *C. maenas*. Genes belonging to the CYP3 and CYP4 families were proposed to have important roles in detoxification and are frequently linked in insects with the potential to acquire resistance to chemical insecticides (David et al., 2003).

Phase II detoxification- Glutathione* S-*transferase

Alignments of cds.comp13424 (GEFN01006054.1) and cds.comp23331 (GEFN01016146.1) from *P. argentinus* revealed similitude with the conserved microsomal GST class of the insects *Laodelphax striatela* (AEY80035) and *Papilio xuthus* (BAM18639.1). Microsomal GSTs are membrane-associated proteins that are involved in both eicosanoid and GSH metabolisms (Bresell et al., 2005). The transcript sequences found in *P. argentinus* shared the motif VERVRRXHLNDXENIX, which is described in microsomal GSTs from almost all crustaceans (Supp. Figure S11).

Figure S 11. *Alignment of microsomal GST from P. argentinus and insects. The motif VERVRRXHLNDXENIX from microsomal GSTs is shaded (X is any residue).*

Although in the case of *P. argentinus*, the substitution of a glutamine residue for leucine was observed. This substitution was also found in the crustacean *C. finmarchicus* and was considered species-specific (Roncalli et al., 2015).

Enzymatic antioxidant defense

An inevitable consequence of aerobic metabolism is the production of oxygen reactive species (ROS) that induce oxidative damage including enzyme inactivation, protein

degradation, DNA damage and lipid peroxidation. To prevent cellular damage, enzymatic antioxidant defenses are deployed to scavenge oxygen reactive compounds including superoxide dismutase (SOD, EC 1.15.1.1), catalase (CAT, EC 1.11.1.6) and glutathione peroxidase (GPx, EC 1.11.1.9) (Livingstone, 2003).

SODs are essential enzymes that occur in virtually all oxygen-respiring organisms, constituting the first line of defense against oxidative damage by catalyzing the dismutation of O_2^- to H_2O_2 (Fridovich, 1995).

Regarding MnSOD, in decapod crustaceans there are two types: mitochondrial (mMnSOD) and cytosolic (cMnSOD) isoforms. mMnSODs are found in plants, bacteria, vertebrates, and invertebrates, while cMnSOD is found only in crustaceans (Li et al., 2010). Both isoforms have been identified in crustaceans, including *Callinectes sapidus* (AF264029, AF264030), *Macrobrachium rosenbergii* (DQ073104, DQ157765), *Marsupenaeus japonicas* (GQ181123, GQ478988), *Fenneropenaeus chinensis* (EF427949, DQ205424), *Procambarus clarki* (EU254488, FJ892724), *Cherax quadricarinatus* (JQ040506, JQ763321), and *Penaeus monodon* (AY726542, KC461130). *P. argentinus* showed a high percent identity with two *E. sinensis* MnSODs: one corresponding to cMnSOD (ACV41936.1) (97% identity) and the other to mMnSOD isoform (AGV76057.1) (85% identity). Both sequences in *P. argentinus* have the putative MnSOD signature motif (DVWEHAYY) the same as in the decapod *Exopalaemon carinicauda* (Ren et al., 2015). In addition, mMnSOD in *P. argentinus* have four putative manganese-binding sites (H48, H96, D180, and H184), which are consistent with other mMnSOD sequences from marine decapods (Cheng et al., 2006). The origin of cMnSOD may be a duplication event, as old as the origin of the arthropod phylum) resulting a protein with a new subcellular location (Brouwer et al., 2003).

Four sequences encoding for glutathione peroxidases (GPx) were identified in the *P. argentinus* transcriptome, with sequences ranging from approximately 68 to 138 amino acids. After a BLASTp search, the best hit corresponded to GPx from horseshoe crabs, crabs, lobster or prawns. The sequence of *P. argentinus* cds.comp10039 (GEFN01000482.1) showed a high percent identity with GPx-like of *Limulus polyphemus* (XP013772703.1), cds.comp10174 (GEFN01000201.1) with phospholipid GPx of *Macrophthalmus japonicus* (AHJ86274.1), cds.comp9895 (GEFN01024176.1) with GPx of *Palaemon carinicauda* (AGJ03551.1) and cds.comp6948 (GEFN01022279.1) with glutathione GPx of *Procambarus clarkii*. Each of the four sequences has a conserved active site motif "^{151}WNFEKF^{156}". Also in cds.comp10039 it was substituted in the active site, F153 by W; and in cds.comp10174 (GEFN01000201.1), E154 was substituted by T.

It is important to highlight that in cds.comp9895 the presence of the Se-GPx signature motif ^{63}LAFPCNQF^{70} was observed which shared identity with the selenium-dependent subclass GPx (Se-GPx). This isoform is responsible for the reduction of organic and inorganic peroxides, such as hydroperoxide (ROOH) and hydrogen peroxide (H_2O_2). Moreover, the amino acids Q73 and W141 were found that are involved in the fixation of selenium and the putative N-glycosylation site ^{75}NNT^{77} (Duan et al., 2013).

As mentioned above, catalase (CAT) is an important antioxidant enzyme that exists in virtually all oxygen-respiring organisms. The main function of CAT is catalyzing the conversion of H_2O_2 to water and molecular oxygen, but also uses H_2O_2 to oxidize toxins such as phenols, formic acid, formaldehyde and alcohols (Chelikani et al., 2004). Two CAT coding sequences were identified in the *P. argentinus* transcriptome, with translated sequences for approximately 101 (cds.comp20365, GEFN01012618.1) and

240 (cds.comp15814, GEFN01004874.1) amino acids. After a BLAST search against non-redundant protein sequences (nr) using Blastp, the best hit corresponded to different regions of CAT from the prawn *Palaemon carinicauda* catalase-clade-3 (KF367466.1) with high percent identities (96 and 93%, respectively).

It should be noted that GPx plays a significant role in H_2O_2 detoxification in vertebrates, while CAT fulfills this function in invertebrates (Livingstone et al., 1992).

Osmoregulation protein comparison with a marine shrimp

P. argentinus is a widely distributed crustacean species in South America in fresh water and lagoons with variable salinity levels. A key factor for growth, distribution, and survival of crustaceans inhabiting the marine environment is salinity (Anger, 2003).

A study has recently described the adaptation mechanisms and pathways related to salinity stress in *Litopenaeus vannamei* using transcriptomics, and a significantly differential expression of some genes involved in osmotic regulation was observed (Hu et al., 2015). A G-protein (guanine nucleotide-binding protein) activates the K^+, Cl^-, and Na^+ channels through a signaling pathway, while the α-subunit can regulate ion channels indirectly (through second messengers) or directly by physically interacting with the channel protein (Dascal, 2001). In this study, a 72-residue fragment of the guanine nucleotide-binding protein G(q) alpha subunit of *P. argentinus* was found in the transcriptome (cds.comp19945_c0_seq1, GEFN01005018.1). This sequence showed high identity with those reported in other crustaceans such as *Marsupenaeus japonicus* (43%, AB488394.1), *Limulus polyphemus* (41%, NM_001314113.1), *Daphnia pulex* (39%, EFX80373.1) and *L. vannamei* (43%, AY626792.1).

The STE20-like kinases belong to the MAP kinase cascade and participate in chloride transport regulation and cell volume regulation, via phosphorylation of channels

(Strange et al., 2006). An 88-amino acid fragment (cds.comp1882_c0_seq1, GEFN01017124.1) was found in the transcriptome of *P. argentinus*. This fragment shared a 76% pairwise identity to the STE20-like kinase of *L. vannamei* (KP322128.1), and the closest sequence available in GenBank corresponds to that reported for the coelacanth fish (*Latimeria chalumnae* XP_005994681) with a 44% pairwise identity at the amino acid level.

The crustacean gills are multifunctional organs that serve, among other functions, as the primary site of osmoregulatory ion transport. A nearly universal assumption is that the Na/K pump, also called Na^+/K^+ ATPase, functions as the major driving force for ion transport across gills in decapods and other crustaceans (Towle and Weihrauch, 2001). In this study, an 863 amino acid fragment, identified as a Na^+/K^+ ATPase, was found within the transcripts of the grass shrimp *P. argentinus* (cds.comp4027_c0_seq1, GEFN01018493.1). This sequence shared an 85% identity with the a1-subunit of the Na+/K+ ATPase of the crustacean *Artemia parthenogenetica* (AJ269679.2).

In addition, two other sequences putatively encoding for a vacuolar-type H+-ATPase (V-ATPase) (cds.comp4196_c0_seq1, GEFN01018798.1) and a sarco/endoplasmic reticulum calcium-ATPase (SR Ca2+ ATPase) (cds.comp5637_c0_seq1, GEFN01019318.1) were identified from the transcriptome of the grass shrimp *P. argentinus*. The V-ATPase of *P. argentinus* showed an 88.79% identity with that of the white shrimp *Litopenaeus vannamei* (ACM16806.1), while the SR Ca^{2+}ATPase found in this study shared an 80% identity with that reported on *Daphnia pulex* (EFX65740.1). Previous work has demonstrated that a V-ATPase participates together with the Na^+/K^+-ATPase in energizing osmoregulatory ion uptake in the freshwater crayfish (Weihrauch et al., 2001). This information is consistent with the notion that *P. argentinus*, which inhabits low-salinity conditions, possesses a gene encoding for a V-

ATPase. A recent report presents a sarco/endoplasmic reticulum Ca^{2+}ATPase in the white shrimp *P. vannamei*, and it highlights the essential role played by this enzyme in osmotic adjustment (Wang et al., 2013).

DNA replication

DNA replication is an essential biochemical process that must be conducted with high fidelity. Primary components of this process are the DNA polymerases and the current RNA-seq analysis identified several candidate genes encoding enzymes responsible for DNA replication. Transcripts encoding for the DNA polymerase delta (δDNApol) catalytic subunit ranging from 86 (cds.comp23154_c0_seq1, GEFN01003375.1) to 245 amino acids (cds.comp12866_c0_seq1, GEFN01002652.1). This domain is essential for nuclear DNA replication and the elongation and maturation of the Okazaki fragments in the lagging strand. These fragments were most similar to the δDNApol subunits of freshwater crayfish *Procambarus clarkii* (87%) (BAO20827.1) and the mantis shrimp *Oratosquilla oratoria* (87%) (BAO20826.1).

In addition, a transcript encoding for a DNA polymerase epsilon (εDNApol) (cds.comp14989_c0_seq1, GEFN01008396.1) was identified. This transcript, involved in the replication of the leading strand, had a 110 amino acid length. It was most similar to that of the jumping ant *Harpegnathos saltator* (51%) (EFN77324.1) and the fruit fly *Drosophila melanogaster* (64%) (NP_611669.1) (Sanchez Garcia et al., 2009).

Even one of the four subunits of the alpha polymerase-primase complex were identified in the KEGG pathway. This complex is involved in the initiation of the DNA replication because all polymerases lack the capacity to start DNA replication from a single strand, all requires the 3´-end of a primer and here is where the alpha-primase complex is involved (Garg and Burgers, 2005). In the *P. argentinus* transcriptome the α-2 sub unit

was identified with a 210 amino acids length (cds.comp14347_c0_seq1, GEFN01010600.1), and had identity with the Atlantic horseshoe crab *Limulus Polyphemus* (52%) (XP_013771709.1), with the Asian swallow tail *Papilio xuthus* (47%) (KPI92391.1), and with the owl limpet *Lottia gigantea* (47%) (XP_009046367.1).

Bioenergetics

ATP is considered as the "universal energy currency" for most biological processes (Baldwin et al., 2009). Since central metabolic pathways and bioenergetics are essential in conferring to all living organisms, those were analyzed in the *P. argentinus* transcriptome. According to the cellular components classification (Supplementary Figure 8), both membrane bound and non-membrane bound organelle components are highly represented in the transcriptome of this species.

Mitochondrial functions are essential to cellular homeostasis, besides ATP synthesis, for the antioxidant response to oxidative stress, apoptotic mechanisms among others (Tuena de Gomez-Puyou et al., 1999).

The *P. argentinus* transcriptome include partial and full sequences encoding nine enzymes from glycolysis and gluconeogenesis pathways from the α-D-glucose-1-PO_4 conversion to D-glucose-6–PO_4 by phosphoglucomutase (E.C. 5.4.2.2) to all those proteins that participate in the conversion of glucose into pyruvate (Müller et al., 2012). As well, the existence of a transcript encoding an L-lactate dehydrogenase (comp6159_c0_seq1, GEFN01000673.1) confirms the ability of this species to undergo anaerobic metabolism, probably during oxidative stress conditions that are common in the aquatic environment due to periods of low oxygen concentration. This full-length

protein shares length and high identity to other crustacean proteins such as that of the white leg shrimp *Litopenaeus vannamei* (90%, AEC12819), and the water flea *Daphnia pulicaria* (69%, AGR85217), and mosquitoes species such as *Aedes aegypti* (69%, XP001662150) and *Culex quinquefasciatus* (69%; XP001864915).

All of the Krebs cycle enzymes were found as annotated transcripts in the current *P. argentinus* transcriptomic database including the proteins that catalyze each cycle step from pyruvate conversion into acetyl-CoA to the malate dehydrogenase that reversibly catalyzes the oxidation of malate to oxaloacetate. An especially interesting protein is the succinate dehydrogenase or complex II, which is located in the inner mitochondrial membrane and functions in the Krebs cycle and the respiratory chain in mitochondria (Cimen et al., 2010). The complex is comprised in *P. argentinus* by four subunits including the succinate dehydrogenase subunit (cds.comp10307_c0_seq1, GEFN01000551.1).

Cells can accumulate large quantities of phosphagens, which are widely spread in nature. In invertebrates, high-energy phosphorylated compounds, such as phosphoarginine, phospholombricine, phosphoglycocyamine and phosphotaurocyamine (Ellington, 2001).

Despite substrate specificity and source, phosphagen kinases shares well-conserved amino acid sequence (~40% identity) in invertebrates. Most described arginine kinases (AKs) are monomeric (~40 kDa) (Garcia-Orozco et al., 2007; Zhou et al., 1998). However, in the sea cucumber (*Stichopus japonicus*) and in sea anemona (*Anthopleura japonicus)* a homodimeric protein of approximately 80 kDa was found (Suzuki et al., 1999). The current *P. argentinus* transcriptome analysis shows that AK is one of most highly expressed transcripts related to muscle proteins (see Table 2, section A). Sequence alignments show four expressed transcripts in *P. argentinus* transcriptome

annotated as AK protein (cds.comp6637, GB GEFN01018817.1; cds.comp8961, GB GEFN01023367.1; cds.comp9908, GB GEFN01020995.1; cds.comp22727, GB GEFN01015443.1).

Detailed analysis shows two partial transcripts (cds.comp22727, 51 amino acids; and cds.comp6637, 260 amino acids) having 89% and 87 % identity to monomeric AK from *Homarus gammarus* (CAA48654.1) and *L. vannamei* ABI98020), respectively.

Surprisingly, transcript cds.comp9908 is more similar to dimeric AKs from sea cucumber *Apostichopus japonicus* (Q9XY07) and sea anemone *Anthopleura japónica* (BAA22888.1) with 44 % and 46% identity, respectively. Also shows similar values toward octameric-mitochondrial creatine kinase type-U from chicken (45 %, X96403.1) and human (46%, J04469). *P. argentinus* cds.comp9908 transcript show similar identity between both dimeric AKs and creatine kinase, which is consistent with previous reports of multimeric AKs. Overall *P. argentinus* annotated AK transcript is more similar to creatine kinase than monomeric AKs. To the best of our knowledge, this is the first result supporting the presence of monomeric and dimeric AKs in the same crustacean. These findings open the possibility to study phylogenetic and evolutionary aspects related to substrate specificity in guanidine kinases family.

Digestion

Crustaceans produce highly active digestive enzymes that hydrolyze major food items (Fernández-Gimenez, 2013). These are comprised by proteases, lipases and glucanases. In the present work, several trypsin-like sequences ranging from 85 to 266 amino acids in length were identified. However, only two canonical trypsin sequences (cds.comp.5870_c0_seq1; GEFN01018754.1; cds.comp.5870_c1_seq1, GEFN01018755.1) were found. Those were not assembled as a single transcript, so

distinguishing whether they are isoforms or if *P. argentinus* has two trypsin genes is a difficult task.

After a BLASTP search against the non-redundant protein database, high identity hits with *Panulirus argus* trypsin isoform 1a (ADB66711.1), *Litopenaeus vannamei* trypsin (CAA60129.1) and *Daphnia pulex* trypsin (EFX80960.1) were found, and catalytic and active site pocket residues as well.

For carbohydrate digestion, we found two transcripts: cds.comp.10280_c0_se11 (GEFN01000718.1) and cds.comp.10280_c1_seqc1 (GEFN01000716.1) which are similar to *Macrobrachium rosenbergii*, *L. vannamei* and *D. pulex* amylases. The sequence of *P. argentinus* cds.comp10280_c1 has the catalytic site active of amylase and the sequence cds.comp10280_c0 contain the putative Ca^{2+} binding site. The adaptive significance of α-amylase polymorphism in the shrimp is not understood yet, and constitutes at challenge since it is important for aquaculture to know if diets with more carbohydrates than the natural diet can be properly utilized for energy production (Van Wormhoudt and Sellos, 2003).

REFERENCES

Anger, K., 2003. Salinity as a key parameter in the larval biology of decapod crustaceans. Invertebr. Reprod. Dev. 43, 29-45.

Baldwin, W.S., Marko, P.B., Nelson, D.R., 2009. The cytochrome P450 (CYP) gene superfamily in *Daphnia pulex*. BMC Genomics. 10, 169.

Bresell, A., Weinander, R., Lundqvist, G., Raza, H., Shimoji, M., Sun, T.H., Balk, L., Wiklund, R., Eriksson, J., Jansson, C., Persson, B., Jakobsson, P.J., Morgenstern, R., 2005. Bioinformatic and enzymatic characterization of the MAPEG superfamily. FEBS J. 1 272, 1688-1703.

Brouwer, M., Hoexum-Brouwer, T., Grater, W., Brown-Peterson, N., 2006. Replacement of a cytosolic copper/zinc superoxide dismutase by a novel cytosolic manganese superoxide dismutase in crustaceans that use copper (haemocyanin) for oxygen transport. Biochem J. 374, 219-228.

Chelikani, P., Fita, I., Loewen, P.C., 2004. Diversity of structures and properties among catalases. Cell Mol. Life Sci. 61, 192-208.

Cheng, W., Tung, Y.H., Chiou, T.T., Chen, J.C., 2006. Cloning and characterisation of mitochondrial manganese superoxide dismutase (mtMnSOD) from the giant freshwater prawn *Macrobrachium rosenbergii*. Fish Shellfish Immunol. 21, 453-466.

Cimen, H., Han, M.J., Yang, Y., Tong, Q., Koc, H., Koc, E.C., 2010. Regulation of succinate dehydrogenase activity by SIRT3 in mammalian mitochondria. Biochemistry. 49, 304-311.

Dascal, N., 2001. Ion-channel regulation by G proteins. Trends Endocrinol. Metab. 12, 391-398.

David, P., Dauphin-Villemant, C., Mesneau, A., Meyran, J.C., 2003. Molecular approach to aquatic environmental bioreporting: differential response to environmental inducers of cytochrome P450 monooxygenase genes in the detritivorous subalpine planktonic Crustacea, *Daphnia pulex*. Mol. Ecol. 12, 2473-2481.

Duan, Y., Liu, P., Li, J., Li, J., Chen, P., 2013. Expression profiles of selenium dependent glutathione peroxidase and glutathione S-transferase from *Exopalaemon carinicauda* in response to *Vibrio anguillarum* and WSSV challenge. Fish Shellfish Immunol. 35, 661-670.

Ellington, W.R., 2001. Evolution and physiological roles of phosphagen systems. Ann. Rev. Physiol. 63, 289-325.

Fernández-Gimenez, A.V., 2013. Digestive physiology of three species of decapod crustaceans of Argentina. J. Shellfish Res. 32, 767-777.

Feyereisen, R., 2006. Evolution of insect P450. Biochem. Soc. Trans. 34, 1252-1255.

Fridovich, I., 1995. Superoxide radical and superoxide dismutases. Annu. Rev. Biochem. 64, 97-112.

Garcia-Orozco, K.D., Aispuro-Hernandez, E., Yepiz-Plascencia, G., Calderon-de la-Barca, A.M., Sotelo-Mundo, R.R., 2007. Molecular characterization of arginine kinase, an allergen from the shrimp *Litopenaeus vannamei*. Int. Arch. Allergy Immunol. 144, 23-28.

Garg, P., Burgersm P.M., 2005. DNA polymerases that propagate the eukaryotic DNA replication fork. Crit. Rev. Biochem. Mol. Biol. 40, 115-128.

Ghaffari, N., Sánchez-Flores, A., Doan, R., García-Orozco, K.D., Chen, P.L., Ochoa-Leyva, A., Lopez-Zavala, A.A., Carrasco, J.S., Hong, C., Brieba, L.G., Rudiño-Piñera, E., Blood, P.D., Sawyer, J.E., Johnson, C.D., Dindot, S.V., Sotelo-Mundo, R.R., Criscitiello, M.F., 2014. Novel transcriptome assembly and improved annotation of the whiteleg shrimp (*Litopenaeus vannamei*), a dominant crustacean in global seafood mariculture. Sci. Rep 4, 7081.

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.

Hu, D.X., Pan, L.Q., Zhao, Q., Ren, Q., 2015. Transcriptomic response to low salinity stress in gills of the Pacific white shrimp, *Litopenaeus vannamei*. Mar. Genom. 24, 297-304.

Li, C., Sun, H., Chen, A., Ning, X., Wu, H., Qin, S,, Xue, Q., Zhao, J., 2010. Identification and characterization of an intracellular Cu, Zn-superoxide dismutase (icCu/Zn-SOD) gene from clam *Venerupis philippinarum*. Fish Shellfish Immunol. 28, 499-503.

Livingstone, D.R., 2003. Oxidative stress in aquatic organisms in relation to pollution and aquaculture. Rev. Med. Vet. 154, 427-430.

Livingstone, D.R., Lips, F., Martinez, P.G., Pipe, R.K., 1992. Antioxidant enzymes in digestive gland of the common mussel, *Mytilus edulis*. Mar. Biol. 112, 265-276.

Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.Y., van der Giezen, M., Tielens, A.G., Martin, W,F., 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol. Mol. Biol. Rev. 76, 444-495.

Ren, X., Pan, L., Wang, L., 2015. Toxic effects upon exposure to benzo[a]pyrene in juve- nile white shrimp Litopenaeus vannamei. Environ. Toxicol. Pharmacol. 39, 194–207.

Rewitz, K., Styrishave, B., Andersen, O.. 2003. CYP330A1 and CYP4C39 enzymes in the shore crab *Carcinus maenas*: sequence and expression regulation by ecdysteroids and xenobiotics. Biochem. Biophys. Res. Commun. 310, 252-260.

Roncalli, V., Cieslak, M.C., Passamaneck, Y., Christie, A.E., Lenz, P.H., 2015. Glutathione S-Transferase (GST) gene diversity in the crustacean *Calanus finmarchicus*-Contributors to cellular detoxification. PLoS One. 10, e0123322.

Sanchez Garcia, J., Baranovskiy, A.G., Knatko, E.V., Gray, F.C., Tahirov, T.H., MacNeill, S.A., 2009. Functional mapping of the fission yeast DNA polymerase δ B-subunit Cdc1 by site-directed and random pentapeptide insertion mutagenesis. BMC Mol. Biol. 10, 82.

Snyder, M.J., 2007. Aquatic P450 species. In: Sigel, A., Sigel, H., Sigel, R.K.O., (Eds.,), The ubiquitous roles of cytochrome P450 proteins: metal ions pp. 97-126.

Strange, K., Denton, J., Nehrke, K., 2006. Ste20-type kinases: Evolutionarily conserved regulators of ion transport and cell volume. Physiology. 21, 61-68.

Suzuki, T., Kamidochi, M., Inoue, N., Kawamichi, H., Yazawa, Y., Furukohri, T., Ellington, W.R., 1999. Arginine kinase evolved twice: evidence that echinoderm arginine kinase originated from creatine kinase. Biochem. J. 340, 671-675.

Towle, D.W., Weihrauch, D., 2001. Osmoregulation by gills of euryhaline crabs: Molecular analysis of transporters. Am. Zool. 41, 770-780.

Tuena de Gomez-Puyou, M., Perez-Hernandez, G., Gomez-Puyou, A., 1999. Synthesis and hydrolysis of ATP and the phosphate-ATP exchange reaction in soluble mitochondrial F_1 in the presence of dimethylsulfoxide. Eur. J. Biochem. 266, 691-696.

Van Wormhoudt, A., Sellos, D., 2003. Highly variable polymorphism of the alpha-amylase gene family in *Litopenaeus vannamei* (Crustacea Decapoda). J. Mol. Evol. 57, 659-671.

Wang, Y., Luo, P., Zhang, L., Hu, C., Ren, C., Xia, J., 2013. Cloning of sarco/endoplasmic reticulum Ca($^{2+}$)-ATPase (SERCA) gene from white shrimp, *Litopenaeus vannamei* and its expression level analysis under salinity stress. Mol. Biol. Rep. 40, 6213-6221.

Weihrauch, D., Ziegler, A., Siebers, D., Towle, D.W., 2001. Molecular characterization of V-type H+-ATPase (B-subunit) in gills of euryhaline crabs and its physiological role in osmoregulatory ion uptake. J. Exp. Biol. 204, 25-37.

Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W.R., Chapman, M.S., 1998. Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. Proc. Natl. Acad. Sci. U. S. A. 95, 8449-8454.

Table S1.- Top 20 most abundant **A**) annotated and **B**) no annotation transcripts. Prot %ID = Amino acid identity percentage with most similar Swiss Prot; FPKM = Fragments per kb of transcript per million mapped reads.

A

Unigene ID	Isoform ID	Swissprot ID	Prot. ID%	Product name	FPKM	GenBank
c8373_g1	c8373_g1_i1	SCP1_ASTLP	85.86	Sarcoplasmic_calcium-binding_protein_1	22887.19	GEFN01022190.1
c9552_g3	c9552_g3_i1	MYSA_DROME	75	Myosin_heavy_chain_muscle	22523.57	GEFN01023696.1
c8909_g2	c8909_g2_i1	MLC1_DROSI	58.39	Myosin_light_chain_alkali	19732.23	GEFN01022858.1
c10023_g6	c10023_g6_i1	ART2_YEAST	63.64	Putative_uncharacterized_protein_ART2	17542.53	GEFN01000046.1
c8909_g1	c8909_g1_i1	MLR_BOMMO	60	Myosin_regulatory_light_chain_2	15143.44	GEFN01022857.1
c10904_g3	c10904_g3_i4	ACT_PROCL	86.67	Actin	13673.66	GEFN01002258.1
c3352_g1	c3352_g1_i1	TNNI_ASTLP	91.67	Troponin_I	11900.54	GEFN01002258.1
c3116_g1	c3116_g1_i1	TNNCA_HOMAM	82.67	Troponin_C_isoform_2A	11007.48	GEFN01016977.1
c7351_g1	c7351_g1_i1	KARG_STIJA	43.91	Arginine_kinase	9996.807	GEFN01020995.1
c9552_g5	c9552_g5_i1	MYSA_DROME	76.92	Myosin_heavy_chain_muscle	9242.349	GEFN01023698.1
c10904_g3	c10904_g3_i1	ACT2_XENTR	85.94	Actin_alpha_cardiac_muscle_2	8133.635	GEFN01002255.1
c10492_g1	c10492_g1_i1	TPM_PANBO	98.94	Tropomyosin	5566.303	GEFN01000821.1
c10904_g3	c10904_g3_i5	ACT_PROCL	87.7	Actin	5075.626	GEFN01002259.1
c10904_g3	c10904_g3_i2	ACT2_XENTR	86.47	Actin_alpha_cardiac_muscle_2	4977.216	GEFN01002256.1
c3915_g2	c3915_g2_i1	COX2_DROLO	69.33	Cytochrome_c_oxidase_subunit_2	4726.941	GEFN01017688.1
c17302_g1	c17302_g1_i1	COX1_ANOGA	73.63	Cytochrome_c_oxidase_subunit_1	4027.057	GEFN01009456.1
c19755_g1	c19755_g1_i1	ACT_MESVI	88	Actin	3912.429	GEFN01012077.1
c3915_g1	c3915_g1_i1	COX3_AEDAE	67.44	Cytochrome_c_oxidase_subunit_3	3553.638	GEFN01017687.1
c9552_g9	c9552_g9_i1	MYSA_DROME	65.22	Myosin_heavy_chain_muscle	2751.925	GEFN01023702.1
c21639_g1	c21639_g1_i1	CYB_DROYA	73.84	Cytochrome_b	2681.274	GEFN01014084.1

B

Unigene ID	Isoform ID	FPKM
c10387_g1	c10387_g1_i1	1.234474e+04
c9552_g7	c9552_g7_i1	1.058587e+04
c6765_g1	c6765_g1_i1	7.327312e+03
c8449_g1	c8449_g1_i1	5.590003e+03
c4714_g2	c4714_g2_i1	5.256098e+03
c10098_g1	c10098_g1_i1	4.891362e+03
c17455_g1	c17455_g1_i1	3.902346e+03
c8372_g4	c8372_g4_i1	3.078985e+03
c4714_g1	c4714_g1_i1	2.256161e+03
c10098_g1	c10098_g1_i2	1.352252e+03
c9378_g1	c9378_g1_i1	8.491003e+02
c17609_g1	c17609_g1_i1	7.398843e+02
c10133_g1	c10133_g1_i1	7.257838e+02
c5326_g1	c5326_g1_i1	6.978764e+02
c5208_g1	c5208_g1_i1	6.802441e+02
c9459_g1	c9459_g1_i1	6.044678e+02
c10280_g1	c10280_g1_i1	5.710552e+02
c7436_g1	c7436_g1_i1	5.257572e+02
c6658_g1	c6658_g1_i1	4.805796e+02
c10805_g2	c10805_g2_i3	4.263200e+02

Table S2. Mitochondrial transcripts. %IDE = Amino acid identity percentage; FPKM = Fragments per kilobase of transcript per million mapped reads.

gene|prot_ID	Product	%IDE	Completeness	FPKM	GenBank
c3915_g2_i1|m.7377	Cytochrome c oxidase subunit II	82.17	complete	4.726941e+03	GEFN01017688.1
c3915_g1_i1|m.7375	ATP synthase F0 subunit 8	62.75	partial	3.553638e+03	GEFN01017687.1
c3915_g1_i1|m.7372	ATP synthase F0 subunit 6	80.80	complete	3.553638e+03	GEFN01017687.1
c3915_g1_i1|m.7371	Cytochrome c oxidase subunit III	81.68	complete	3.553638e+03	GEFN01017687.1
c10976_g1_i1|m.38964	NADH dehydrogenase subunit 5	74.91	partial	7.936385e+02	GEFN01002608.1
c21770_g1_i1|m.62408	NADH dehydrogenase subunit 4	75.85	complete	7.288552e+02	GEFN01014226.1
c21770_g1_i1|m.62409	NADH dehydrogenase subunit 4L	69.89	partial	7.288552e+02	GEFN01014226.1
c21639_g1_i1|m.61979	NADH dehydrogenase subunit 6	64.78	complete	2.681274e+03	GEFN01014084.1
c21639_g1_i1|m.61978	Cytochrome b	89.68	complete	2.681274e+03	GEFN01014084.1
c17316_g2_i1|m.52592	NADH dehydrogenase subunit 1	85.62	complete	1.227179e+03	GEFN01009470.1
c17302_g1_i1|m.52551	NADH dehydrogenase subunit 2	65.62	complete	4.027057e+03	GEFN01009456.1