Marine Genomics xxx (xxxx) xxx–xxx

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

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.