Marine Genomics 18 (2014) 89–91

Contents lists available at ScienceDirect

Marine Genomics

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

Genomics/Technical resources

Next-generation transcriptome characterization in three *Nacella* species (Patellogastropoda: Nacellidae) from South America and Antarctica

Gonzalo Fuenzalida [a,d], Elie Poulin [b], Claudio Gonzalez-Wevar [b], Cristian Molina [e], Leyla Cardenas [c,*]

[a] UPMC Univ Paris 06, UMR 7144, Equipe ABICE, Station Biologique de Roscoff, 29680 Roscoff, France
[b] Laboratorio de Ecología Molecular, Instituto de Ecología y Biodiversidad (IEB), Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Las Palmeras, 3425 Santiago, Chile
[c] Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Isla Teja S/N, Casilla 567, Valdivia, Chile
[d] CNRS, UMR 7144, Adaptation et Diversité en Milieu Marin, Station Biologique de Roscoff, 29680 Roscoff, France
[e] AUSTRAL-omics, Facultad de Ciencias, Universidad Austral de Chile, Isla Teja S/N, Casilla 567, Valdivia, Chile

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 14 June 2014
Accepted 14 June 2014
Available online 27 June 2014

Keywords:
Pyrosequencing
Comparative transcriptome
Patellogastropods
Antarctica
South America

ABSTRACT

The southern tip of South America and Antarctica are particularly interesting due to many genera and also species currently sharing between both areas. The genus *Nacella* (Patellogastropoda: Nacellidae) is distributed in different regions of South America and Antarctica living preferentially on rocks and boulders and grazing on algae, diatoms and bacterial films. We described the transcriptomes of three *Nacella* species, *Nacella concinna* (Strebel, 1908), inhabiting the Antarctic Peninsula; *Nacella magallanica* (Gmelin, 1791), from Patagonia and *Nacella clypeater* (Lesson, 1831), from central Chile. In total, we obtained over 20,000 contigs with an average length of 583 bp. Homologous protein coding genes (PCGs) for mitochondrial genome of the three species were characterized and a database of molecular markers was also generated. This study represents the first publicly available report on pyrosequencing data for patellogastropod species, and provides an important comparative resource for studies in ecophysiology and evolutionary adaptation in marine invertebrate species.

1. Introduction

The genus *Nacella* (Patellogastropoda: Nacellidae) is currently distributed in different Provinces of the Southern Ocean including Antarctica, the Kerguelen Archipelago, the Antipodean Province, Central Chile and Patagonia (González-Wevar et al., 2010). The limpet *Nacella clypeater* inhabits from southern Peru down to 42° S in Chile. *Nacella magallanica* is found in Patagonia from Puerto Montt to Cape Horn in the Pacific and all along the Atlantic coast up north to the Rio Negro Province in Argentina, including the Falkland/Malvinas Islands. Finally, *Nacella concinna* is the only representative of the genera inhabiting in ice-free rocky areas along maritime Antarctica (Antarctic Peninsula and associated islands) and peri-Antarctic islands (i.e. South Georgia, Gough and Bouvet). Recent molecular studies have shown that the Antarctic limpet was separated from its South American relatives since the end of the Miocene without any evidence of recent or recurrent gene flow events between these regions (González-Wevar et al., 2010).

Antarctic organisms adapt to their environment by changing their physiology, ecology and genomic architecture (Peck and Clark, 2012). Several studies developed mainly in fishes concluded that cold adaptation includes a variety of evolutionary changes such as loss of genes, change in gene expression, genomic rearrangements and evolutionary innovation (Peck and Clark, 2012). In marine invertebrates, adaptation to cold and the genetic basis involved are poorly understood. Only few recent works are intended to describe the transcriptome architecture of some invertebrate species. In *Laternula elliptica*, an infaunal stenothermal bivalve mollusk with a circumpolar distribution, Clark et al. (2010) described their transcriptome focusing on the shell deposition and repair in mollusks. For the Antarctic krill *Euphausia superba*, a keystone species in the Antarctic food chain, two works are describing the transcriptomic architecture placing the attention on genes associated with stress and neuropeptide hormones (Clark et al., 2011; Toullec et al., 2013). In the Antarctic brittle star *Ophionotus victoriae*, the transcriptome was described to characterize the genes involved in regeneration (Burns et al., 2013). In patellogastropods, only one mitochrondrial genome is available (NCBI DQ238599) and only recently the draft genome of *Lottia gigantea* was released (NCBI KB199650). In terms of the available sequence data for nacellid species, there are 667 sequences described in the NCBI database, corresponding mostly to Cytochrome Oxidase I, analyzed in a phylogeographic study (González-Wevar et al., 2013). Thus, here we describe the head transcriptome in three limpet species inhabiting in South America and Antarctica with the aim to generate useful genomic information to study the molecular basis on adaptation in marine invertebrate species.

* Corresponding author. Tel.: +56 63221673; fax: +56 63221344.
E-mail addresses: gonzalo.fuenzalida.delrio@gmail.com (G. Fuenzalida), epoulin@uchile.cl (E. Poulin), omeuno01@hotmail.com (C. Gonzalez-Wevar), cristian.molina.quiroz@gmail.com (C. Molina), leylacardenas@uach.cl (L. Cardenas).

http://dx.doi.org/10.1016/j.margen.2014.06.004
1874-7787/

Table 1
Statistic for pyrosequencing of three patellogastropods species.

Species	*Nacella clypeater*	*Nacella magallanica*	*Nacella concinna*
Total No of reads	65,389	73,858	148,874
Total size (bases)	20,133,052	22,792,456	53,545,559
Average length (bases)	307	325	359
Total No of contigs	3606	4545	12,603
Average No of reads within contigs	58	71	87
Total size within contig (bases)	1,952,525	2,674,153	7,837,276
Average length of contig (bases)	541	588	621
Largest contig size (bp)	6959	5312	7492
Smallest contig size (bp)	200	203	200
% contigs with blast match	72.60 (2618)	68.94 (3217)	57.73 (7276)

2. Data description

2.1. Animal collections

Samples of adult individuals of the Antarctic limpet *N. concinna* were collected from the intertidal zone during a low tidal period near Base Escudero Station at Fildes bay, King George Island, South Shetland Island (62°10′S, 58°51′W), during the summer of 2012. Adult specimens of *N. magallanica* were obtained from the intertidal zone from Punta Santa Ana, Strait of Magellan (53° 37′S, 70° 54′W) during the summer of 2012. *N. clypeater* individuals were collected from the intertidal zone of La Mision, Valdivia, Chile (39° 46′ S, 73° 23′W) during the summer of 2012. For each species, head tissue extracted from 15 individuals was immediately frozen in liquid nitrogen, and stored at −80 °C. See Supplementary methods for RNA preparation, cDNA library and sequencing.

2.2. Transcriptome analyses

After removing low-quality regions, adaptors and all possible contaminants, we obtained a total of 288,121 high-quality reads containing 96,471,067 bases with an average read length of 356 bases (Table 1). For each species, datasets were assembled using the strategy described in iAssembler (Zheng et al., 2011) by combining Mira and Cap3 Assemblers. Although singletons potentially contained useful sequences with low levels of expression, we excluded them from further analysis. All sequence data were submitted to the NCBI SRA (short read archive) with the accession number SRP041451.

Annotation through Blast2GO pipeline (www.blast2Go.com) was accomplished first by searching for matches in the nr database at NCBI with a low e-value (10^{-6}). Then, the mapping process by aligning the results to the GO database was followed by a final GO annotation step with an e-value cutoff of 10^{-10} and a minimum alignment length of

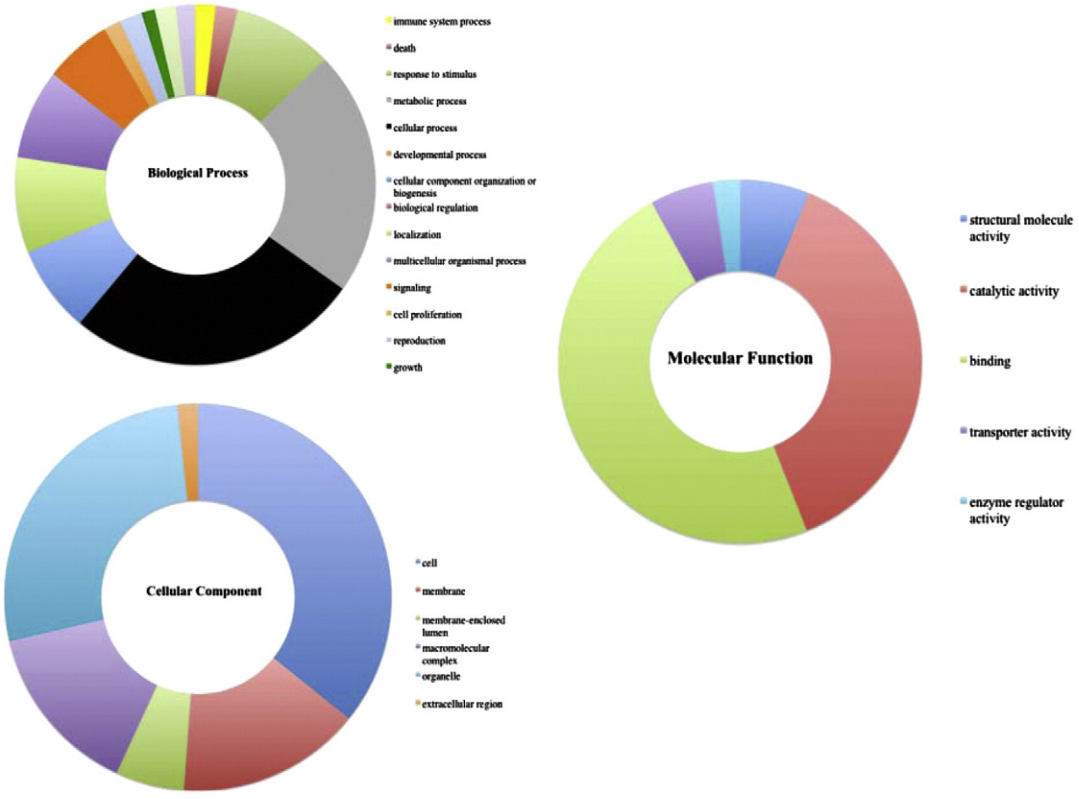

Fig. 1. Gene Ontology Analysis. *N. concinna* was the species with most GO assignments to different biological process including cell process (26%), metabolic process (25%), single-organism process (13%), biological regulation (7%), localization (6%) and response to stimulus (5%). Those assigned to cellular components included cells (35–36%), organelles (27–29%), macromolecular complexes (14%–17%), membranes (14%–16%) and proteins located in the extracellular regions (2%). Finally, those assigned to molecular function were mainly linked to the binding of ATP, zinc ions, and protein (48%–50%), catalytic activities of enzymes (38%–40%), and structural molecular activity (6%–9%) and enzyme regular activity (2%).

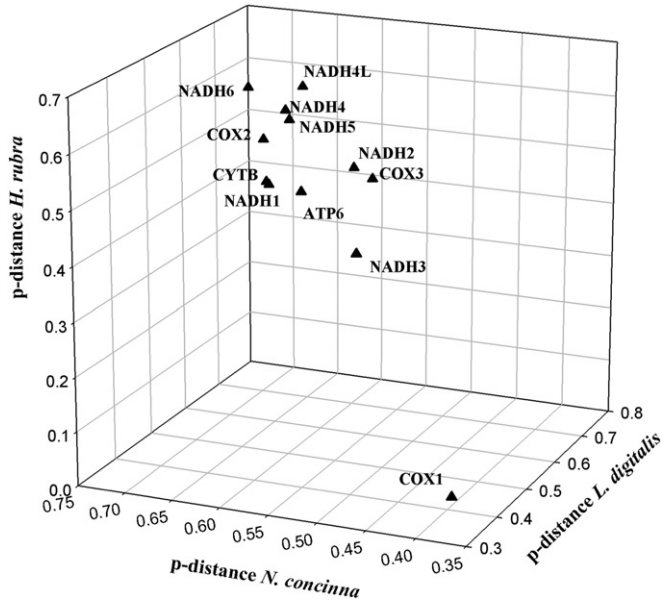

Fig. 2. Divergence of mitochondrial protein-coding genes (PCGs) of *Nacella concinna*. P-distance corresponds to the number of base differences per site from the pairwise comparison among the *N. concinna* with respect to the PCGs from *L. digitalis* (NCBI DQ238599) and the vertigastropod *H. rubra* (NCBI AY588938).

100 bp. The search in the nr database at NCBI resulted in a total of 13,111 matches (Table 1). Interestingly, 29% of the total matches were transcripts from the owl limpet *L. gigantea* followed by the sea slug *Aplysia californica* with 14.5% of the total. *N. concinna* was the species with the most GO assignments to different biological processes including cell process, metabolic process, single-organism process, biological regulation, localization and response to stimulus, among others (Fig. 1). The top 30 commonly expressed sequences with associated BLAST matches in the three nacellid species, are shown in the Supporting Information, File S2.

To identify the homologous protein coding genes (PCGs) for mitochondrial genome present in the three libraries, we performed a comparison alignment with homologous genes of a unique complete mitochondrial genome available up to date for the *Lottia digitalis* and the vetigastropod *Haliotis rubra* (NCBI AY588938) species because it is a group closely related to the patellogastropods (White et al., 2011). The limits of PCG genes were adjusted manually based on the location of adjacent genes and the first start and stop codons in frame. Data were analyzed using Geneious Pro 5.5.6 (Drummond et al., 2010). Through the analysis we were able to identify 12 of the 13 PCGs with only ATP8 absent in the three databases (Supporting Information, File S3). The divergence analysis showed *N. concinna* closer to the vetigastropod *H. rubra* rather than to the patellogastropod *L. digitalis* (Fig. 2).

Finally, expressed sequence tag (EST)-derived simple sequence repeats (SSRs) were identified in the three databases using the MSATCOMMANDER software (Faircloth, 2008). There were 2737 sequences containing microsatellite motifs with enough flanking regions to design primers, 1678 of these were identified in *N. concinna*, 639 were identified in *N. magallanica* and 420 were identified in *N. clypeater* (Supplementary File S4). The most abundant EST-SSRs were trinucleotide repeat motifs with the AAT motif being the most abundant (Table 2).

Table 2
Expressed sequence tag (EST)-derived simple sequences repeat (SSR) founded in the three *Nacella* species.

		N. clypeater	
N. magallanica		*N. concinna*	
Total EST_SSR	420	639	1678
% dinucleotide motif	12.6	17.5	20.3
% trinucleotide motif	81.2	72.3	71.5
% tetranucleotide motif	3.3	7.0	4.7
% pentanucleotide motif	1.9	3.1	2.6
% hexanucleotide motif	1.0	0.0	0.8

Here we report the first comprehensive transcriptome database for *Nacella* species from Antarctica and South America. More than 20,000 putative transcripts were obtained with a large percentage of similarity to known proteins. Thus, the information provided here constitutes a step forward in the knowledge of the genetic characteristics of this particular group recognized as a basal branch of the extant Gastropoda.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.margen.2014.06.004.

Acknowledgments

We thank AUSTRAL_*omics* (www.australomics) for the help in the bioinformatic analysis. The authors also acknowledge the support of the Instituto Antartico Chileno, Grant Inach T22-10. Finally, our thanks are given to the Millennium Nucleus Center for the Study of Multiple-drivers on Marine Socio-Ecological Systems (MUSELS) by MINECON Project NC120086.

References

Burns, G., Thorndyke, M.C., Peck, L.S., Clark, M.S., 2013. Transcriptome pyrosequencing of the Antarctic brittle star *Ophionotus victoriae*. Mar. Genomics 9, 9–15.
Clark, M.S., Thorne, M.A.S., Vieira, F.A., Cardoso, J.C.R., Power, D.M., Peck, L.S., 2010. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing. BMC Genomics 11, 362.
Clark, M.S., Thorne, M.A.S., Toullec, J.-Y., Meng, Y., Guan, L.L., et al., 2011. Antarctic krill 454 pyrosequencing reveals chaperone and stress transcriptome. PLoS ONE 6 (1), e15919. http://dx.doi.org/10.1371/journal.pone.0015919.
Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., et al., 2010. Geneious v5.5. Available at http://www.geneious.com.
Faircloth, B.C., 2008. msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol. Ecol. Resour. 8, 92–94.
González-Wevar, C.A., Nakano, T., Cañete, J.I., Poulin, E., 2010. Molecular phylogeny and historical biogeography of *Nacella* (Patellogastropoda: Nacellidae) in the Southern Ocean. Mol. Phylogenet. Evol. 56, 115–124.
González-Wevar, C.A., Saucède, T., Morley, S., Chown, S.L., Poulin, E., 2013. Extinction and recolonization of maritime Antarctica in the limpet *Nacella concinna* (Strebel, 1908) during the last glacial cycle: toward a model of Quaternary biogeography in shallow Antarctic invertebrates. Mol. Ecol. 22, 5221–5236.
Peck, Ll, Clark, M.S., 2012. Understanding adaptations and responses to change in Antarctica: recent physiological and genomic advances in marine environments. Chapter 9 In: Verde, C., di Prisco, G. (Eds.), Adaptation and Evolution in Marine Environments, Vol. 1, pp. 157–182 (From Pole to Pole 3).
Toullec, J.-Y., Corre, E., Bernay, B., Thorne, M.A.S., Cascella, K., et al., 2013. Transcriptome and peptidome characterisation of the main neuropeptides and peptidic hormones of a euphausiid: the Ice Krill, *Euphausia crystallorophias*. PLoS ONE 8 (8), e71609. http://dx.doi.org/10.1371/journal.pone.0071609.
White, T.R., Conrad, M.M., Tseng, R., Balayan, S., Golding, R., de Frias, A.M., Dayrat, B.A., 2011. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. BMC Evol. Biol. 11, 295.
Zheng, Y., Zhao, L., Gao, J., Fei, Z., 2011. iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. BMC Bioinforma. 12, 453.