Comparative Biochemistry and Physiology, Part D 8 (2013) 17–23

Contents lists available at SciVerse ScienceDirect

Comparative Biochemistry and Physiology, Part D

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

Complete mitochondrial genome of *Concholepas concholepas* inferred by 454 pyrosequencing and mtDNA expression in two mollusc populations

Gustavo Núñez-Acuña, Andrea Aguilar-Espinoza, Cristian Gallardo-Escárate *

Laboratory of Biotechnology and Aquatic Genomics, Interdisciplinary Center for Aquaculture Research, University of Concepción, PO Box 160-C, Concepción, Chile

ARTICLE INFO

Article history:
Received 2 August 2012
Received in revised form 20 October 2012
Accepted 21 October 2012
Available online 27 October 2012

Keywords:
Mitochondrial genome
Concholepas concholepas
RNA-Seq
Gene expression

ABSTRACT

Despite the great relevance of mitochondrial genome analysis in evolutionary studies, there is scarce information on how the transcripts associated with the mitogenome are expressed and their role in the genetic structuring of populations. This work reports the complete mitochondrial genome of the marine gastropod *Concholepas concholepas*, obtained by 454 pryosequencing, and an analysis of mitochondrial transcripts of two populations 1000 km apart along the Chilean coast. The mitochondrion of *C. concholepas* is 15,495 base pairs (bp) in size and contains the 37 subunits characteristic of metazoans, as well as a non-coding region of 330 bp. *In silico* analysis of mitochondrial gene variability showed significant differences among populations. In terms of levels of relative abundance of transcripts associated with mitochondrion in the two populations (assessed by qPCR), the genes associated with complexes III and IV of the mitochondrial genome had the highest levels of expression in the northern population while transcripts associated with the ATP synthase complex had the highest levels of expression in the southern population. Moreover, fifteen polymorphic SNPs were identified *in silico* between the mitogenomes of the two populations. Four of these markers implied different amino acid substitutions (non-synonymous SNPs). This work contributes novel information regarding the mitochondrial genome structure and mRNA expression levels of *C. concholepas*.

1. Introduction

In general, the genomic organization of the mtDNA in marine organisms have been studied through long-PCR and random genome sequencing with DNA nebulization and the creation of clone libraries with pUC or Bluescript vectors (Burger et al., 2007). However, recent advances in next-generation sequencing technology (NGS) have increased the velocity of mitochondrial genome characterization of several vertebrate and invertebrate marine species (Jex et al., 2008; Neira-Oviedo et al., 2011). The main advantage of using NGS platforms is the possibility of using the data generated to identify specific genes and to infer its relative abundance under different conditions of study (Jex et al., 2008; Neira-Oviedo et al., 2011), which is of particular interest in comparative genomic ecology studies.

The mitochondrial genome is a macromolecule with a size that in the majority of metazoans ranges between 15 and 20 kb (Boyce et al., 1989; Boore, 1999). In practically all metazoans the mitochondrion is circular, except in some classes of cnydaria that have lineal mitochondrial DNA chromosomes (Bridge et al., 1992). The mitochondrial genome generally has 13 enzymatic protein subunits involved in the process of oxidative phosphorylation, 22 tRNA and 2 ribosomal subunits. As well, it has been reported that there is a noncoding region in several species that can play a regulatory role in the duplication and transcription of the mitochondrial genome, although it is difficult to study its homology in higher taxa because it presents similarities only among closely related taxa (Passamonti et al., 2011). Complete mitochondrial genome analysis increases the resolution of works on the molecular phylogeny of species, allowing the identification of genetic reordering among the higher taxa of species (Rawlings et al., 2010). Mitochondrial genome analysis significantly increases the resolution of philogenetic studies of lower taxa, combining the data of all the genes in a single set (Cunha et al., 2009). Finally, intraspecific analysis of the mitochondrial genome evidences possible mtDNA polymorphism (single nucleotide polymorphisms, SNPs) associated with different geographic locations (Xin et al., 2011). Nevertheless, to our knowledge there have been no studies relating mtDNA polymorphism to patterns of mitochondrial gene expression at the biogeographic level. In this context, important questions remain unanswered, such as: Is it possible to find local patterns of mitochondrial gene expression? Does variation of mitochondrial gene expression imply local adaptations?

Our model of study, *Concholepas concholepas*, is a marine gastropod found along the Chilean coast. It has become a commercially important benthic resource for Chile (Leiva and Castilla, 2001). One of the main characteristics of these species at the ecological level is their long larval development stage in the water column, which can last from almost three months to one year. This extensive stage of planktonic development results in a broad geographic distribution of the species from 6°S to 56°S. Genetic studies based on the coding region of the *cytochrome*

* Corresponding author at: Interdisciplinary Center for Aquaculture Research, University of Concepción, PO Box 160-C, Concepción, Chile. Tel.: +56 41 2203422; fax: +56 41 2207310.
E-mail address: crisgallardo@udec.cl (C. Gallardo-Escárate).

1744-117X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cbd.2012.10.004

oxidase I gene (*COX1*) do not show genetic-population structuring among different *C. concholepas* groups over the range of their biogeographic distribution (Cárdenas et al., 2009). Nevertheless, an unanswered question is whether there are genes in *C. concholepas* that show highly differentiated biogeographic patterns due to environmental factors (e.g. northern and southern Chile). New DNA molecular markers have recently been described for *C. concholepas* as microsatellites (Cárdenas et al., 2006) and ESTs obtained by 454 pyrosequencing (Cárdenas et al., 2011). As well, there are now candidate genes related to thermal, oxidative and immune responses that allow for testing hypotheses about local adaptation at the genomic level.

The objective of the present study was to describe for the first time the mitochondrial genome for *C. concholepas* using 454 cDNA pyrosequencing (RNA-seq). As well, we evaluated the levels of gene expression (*in silico* analysis) of all the mitochondrial genes in two populations of *C. concholepas* (1000 km apart) based on data from RNA-seq. Differences in gene expression were confirmed with real time qPCR analysis.

2. Material and methods

2.1. Samples

Fifty adult individuals of *C. concholepas* (Gastropoda, Muricidae) were collected as samples from two locations: Caleta Chungungo, Punta de Choros (29°16′S–71°31′W) and Los Molinos, Valdivia (39°40′S–73°12′W) in February 2010. Shell length of samples were between 10.2 and 12.7 cm. Approximately 100 mg of gill and muscle tissue was extracted from the samples and conserved in 1 mL of RNAlater stabilization solution (Ambion®, USA) and stored at −80 °C until total RNA extraction.

2.2. Characterization of the mitochondrial genome of C. concholepas

Total RNA was extracted from 100 mg of gills and muscle from each sample independently, using Trizol reagent (Invitrogen, USA) in accordance with the supplier's instructions. The RNA was purified and the concentration and purity were measured with a spectrophotometer (ND-1000, Nanodrop Technologies) and the integrity was visualized with electrophoresis in MOPS/formaldehyde agarose gels at 1.2% staining with ethidium bromide at 0.001%, and then quantified with a 2100 bioanalyzer (Agilent Technologies, USA). Selected extractions presented 260/280 and 260/230 purity indices equal to or greater than 2.0 and showed integral RNA in electrophoresis and bioanalyzer measurements (RIN > 8). Subsequently, sample pools of selected RNAs were prepared (one pool for each population group) and precipitated overnight with 2× volumes of absolute ethanol and 0.1× volume of sodium acetate 0.3 M at −80 °C. Double-stranded cDNA was synthesized from the extracted RNA and pyrosequenced (full plate) in a 454 GS FLX titanium platform by Macrogen Inc. (S. Korea).

The raw data for both samples were trimmed and assembled separately, and combined using CLC Genomics Workbench software (CLC bio, Denmark). The overlap settings used for these assemblies were mismatch cost = 2, insert cost = 3, minimum contig length = 200 bp, similarity = 0.8, and trimming quality score = 0.05. The redundancy of singleton sequences was calculated using the Java application Duplicate Finder (http://bioinfotutlets.blogspot.com/). All contigs and singletons were annotated according to Gene Ontology terms using Blast2Go software (Conesa et al., 2005). From the annotated contigs, we selected those that presented homology with genes present in the mitochondrial genome, and we assembled them using the mitochondrial genome of the muricid *Thais clavigera* (GenBank accession no. DQ159954.1) as a reference (Ki et al., 2010). As well, for regions of the mitochondrial genome that were not covered by EST, specific primers were designed and amplified by PCR and Sanger sequencing (Macrogen Inc., S. Korea) of the amplicons (Table 1). Finally, the sequenced products were aggregated to the assembly to complete the mitochondrial genome of *C. concholepas*. The final assembly was annotated and the beginning and end of each gene was identified. These data were submitted to GenBank (GenBank accession no. JQ446041). As well, tRNAs were identified with CLC Main Workbench software (CLC Bio, Aarhus, Denmark).

Table 1
List of primers for PCR amplification of *C. concholepas* mitochondrial DNA without coverage by 454 pyrosequencing data. In addition, primer sequences for qPCR analysis are included.

Primer name	Gene	Product length	Primer sequence
Cc_ATP8-F	ATP subunit 8	583 bp	5′-GCTGTACCAGGGCGCCTAAA-3′
Cc_ATP8-R			5′-AGCGGAAAACCTGGGACGAC-3′
Cc_12S1-F	s-rRNA	984 bp	5′-TCTTCAGTATTCTCAATGGTCTTATTA-3′
Cc_12S1-R			5′-GTTTAACCGCGGATGCTGGC-3′
Cc_12S2-F	s-rRNA	957 bp	5′-AATGTCCCGCTAACACCAGT-3′
Cc_12S1-R			5′-ACCATGATGCAAAAGGTACAAAGG-3′
Cc_ND1-F	NADH dehydrogenase1	498 bp	5′-ACCCAGAAGATGCAGCCGTC-3′
Cc_ND1-R			5′-GCATCTGCAATAGGCTGGGG-3′
Cc_ND6-F	NADH dehydrogenase 6	536 bp	5′-AGCAATCACATTGTCTTCCTGGT-3′
Cc_ND6-R			5′-AGGCCAGTGGCAATCTGGATA-3′
Cc_CytB-F	Cytochrome *b*	476 bp	5′-TGGAGGCCCCTTACGAGCAA-3′
Cc_CytB-R			5′-GCTTGCTTCACAGGCACCCA-3′
Cc_ND4-F	NADH dehydrogenase 4	565 bp	5′-AGCGGCTCCGCCTAGAGTAA-3′
Cc_ND4-R			5′-CACAGCCAGAAATTAGACAAACCACA-3′
mtCC_qATP6-F	ATP subunit 6	96 bp	5′-TGGGCATACCGCTGTGACTTACT-3′
mtCC_qATP6-R			5′-GTGCAGGGGCACCCATAGGA-3′
mtCC_qCOX1-F	Cytochrome oxidase 1	88 bp	5′-TCAGCTGCTGTAGAGAGTGGGG-3′
mtCC_qCOX1-R			5′-CAACTGAACCACCGGCGTGA-3′
mtCC_qCOX2-F	Cytochrome oxidase 2	118 bp	5′-TCACCGGATAGTTCTTCCAACGCA-3′
mtCC_qCOX2-R			5′-TTAGGCGCCCTGGTACAGCA-3′
mtCC_qCytB-F	Cytochrome *b*	104 bp	5′-GGAAGATGCCCGTTGGAGGC-3′
mtCC_qCytB-R			5′-CCCATAATTTCTGGAGCGCAGGG-3′
mtCC_q16S-F	l-rRNA	115 bp	5′-CGGCAAAACCCAATTCCGCC-3′
mtCC_q16S-R			5′-TTGCACGGTCAGAGTACCGC-3′
mtCC_qNCR-F	Non-coding region	123 bp	5′-CCGTTCGTGCGCCTTAGCTT-3′
mtCC_qNCR-R			5′-ATTGAGGGCGGTGCCCATCT-3′
Contig6815_CC_F	PRKAG	92 bp	5′-TGCAGGAAGCGAAGGATGCG-3′
Contig6815_CC_R			5′-GACAACCATGTGCACCGCCT-3′
Contig18299_CC_F	Hypoxia inducible factor	94 bp	5′-CGCCGCTCATTACGCCAGAA-3′
Contig18299_CC_R			5′-TCCATGAAGGTCATGTCCATCCGA-3′

This final sequence was used as a reference to re-assemble the mitogenome of *C. concholepas* but this time separately for each population. Consensus sequences of each population were obtained and then were aligned using Geneious software (version 5.6, Biomatters Ltd., New Zealand). From this alignment SNP markers were identified from the two populations. Coverage and aminoacid substitution implications were verified in each assembly. Parameters to identify SNPs were minimum coverage = 20 and minimum variant frequency = 0.25.

2.3. Gene expression analysis of the mitochondrial genome in C. concholepas

The consensus assembly of the mitochondrial genome generated in the previous step was used as a reference for the RNA-seq expression analysis. Using CLC Genomic Workbench software (version 4.7), the reads for each of the populations were separately assembled against the mitochondrial genome. Then the number of reads per kilobase per million mapped reads (RPKM) was obtained with the same software. This normalizes the number of reads to the size of assembled contigs and allows for assessing the transcripts that are over-expressed among different groups (Mortazavi et al., 2008). Kal's statistical analysis test was used to compare Northern and Southern population in terms of the $\log_2$ fold change ($P<0.0005$, FDR corrected), and also expression levels of RPKM were visualized in a heat-map (sample clustering) from normalized expression values.

The transcripts that presented differential expression between the two populations were validated by qPCR of 20 individuals from each group with the comparative CT method. As well, we assessed the expression of nuclear genes involved in the transcription of mitochondrial genes such as succinate dehydrogenase (*SD*), AMP-activated protein kinase gamma 2 subunit (*PRKAG2*) and hypoxia inducible factor (*HIF*). Specific primers were designed for the qPCR of an amplicon of approximately 100 bp in size, with an annealing temperature of around 60 °C, a high GC content and the absence of dimers and other secondary structures (Table 1). Four genes obtained from the pyrosequencing data were assessed for the reference gene: *GAPDH*, *Elongation factor 1*, *Elongation factor 2*, and a non-coding region of the mitochondrial genome (*NCR*). NormFinder (Andersen et al., 2004) was used to select the reference gene with the highest stability value. The analysis of relative abundance of each of the mitochondrial genes was normalized with NCR as endogenous control. The qPCR analyses were conducted in a thermocycler in real time StepOnePlus™ (Applied Biosystems®, Life Technologies, USA), using the Maxima® SYBR Green/ROX qPCR Master Mix kit (Fermentas) in accordance with the supplier's instructions. The amplification conditions were: enzymatic activation at 95 °C for 10 min, 35 cycles at 95 °C for 15 s and at 60 °C for 1 min. Finally a melting curve was programmed to verify the presence of a unique product and the absence of contamination in the negative controls. Once the data was obtained, the Pfaffl formula (Pfaffl, 2001) was applied to the Ct to obtain relative abundance values. Finally, a one-way ANOVA and a Kruskal–Wallis test were conducted to assess significant differences among the genes and populations of *C. concholepas*, considering a threshold of $p<0.05$. The relative expression analysis by qPCR was conducted with qBasePlus 2 software (Biogazelle, Ghent, Belgium).

3. Results and discussion

3.1. Characterization of the mitochondrial genome of C. concholepas

A total of 22,141 contigs with an average size of 411 bp were obtained from the *de novo* assembly generated with 454 pyrosequencing. Subsequently, 3844 contigs were successfully annotated, which is 17.54% of the total. This value is higher than reported earlier by Cárdenas et al. (2011), who developed another cDNA library with a smaller number of contigs (19,218), with an average length of 160 bp. The annotated transcripts that were associated with mitochondrial genome were extracted for assembly using the mitochondrial genome of *T. clavigera* as a reference. The regions that were resequenced through the Sanger method (direct sequencing of PCR products of remaining regions) owing to the lack of genomic coverage had sizes of 566 bp (*ATP8*), 647 bp (first *s-rRNA* region), 813 bp (second *s-rRNA* region), 481 bp (*ND1*), 502 bp (*ND6*), 462 bp (*CytB*) and 534 bp (*ND4*) (data not shown). These sequences were then reassembled with the rest of the mitochondrial genome of the species, thus obtaining the complete mitochondrion of *C. concholepas*.

The final size of the complete mitochondrial genome of the species was 15,495 bp, which is similar to that of several other species of the gastropod class (see examples of other species in Cunha et al., 2009). The mitochondrial genome of *C. concholepas* is composed of 37 characteristic genes, 13 coding subunits, 22 coding units, 22 transfer RNA (tRNA), and 2 ribosome subunits, 16S and 12S (l-rRNA and s-rRNA respectively). The number of genes found for this species is in accordance with the number of conserved genes in the mitochondrial genome (Wolstenholme, 1992; Boore, 1999). A total of 539 non-coding bp nucleotides (NCR) were observed in 20 intergenic regions. The largest non-coding region was 330 bp long and located between *COXIII* and $tRNA^{Phe}$ (Fig. 1), suggesting a putative control region. However, this region has an A + T nucleotide composition of 53.3%, which is low compared to other species reported earlier (Sun et al., 2011). Of all the identified genes, 29 encoded in the positive strain and 8 encoded in the negative strain (Table 2). The composition of the nucleotide bases had a higher proportion of thymine (36.6%), then adenine (27.3%), guanine (18.7%), and finally cytosine (17.4%). The final GC content was 36.0%, while the final AT content of the coding regions in the positive strain was 62.8%, which is within the range for this composition of metazoans (Bandyopadhyay et al., 2006). Four sequencing overlaps were found among the identified genes. These occurred between the *COX2* and $tRNA^{Asp}$, $tRNA^{Trp}$ and $tRNA^{Gly}$, and the $tRNA^{Gly}$ and $tRNA^{Glu}$ coding regions, and finally between the coding regions of the *ND1* and $tRNA^{Pro}$ genes. Thus all the genes that encode proteins begin with the ATG initiation codon, although there are several species with mitochondrial genes that begin with other "unorthodox" initiation codons. The end codons were TAA and TAG, which are typical of metazoan mitochondrial genome (Wolstenholme, 1992; Bandyopadhyay et al., 2006). We found 22 coding regions for tRNAs, including Serine-2 and Leucine-2, which have been reported for other species (Groenenberg et al., 2012) (Supplemental Fig. 1). The inferred secondary structures of the latter two tRNAs have variations from those of other species. This could be due to the algorithm that was used, or because the tRNAs of the species are in fact different.

NGS provides an alternative methodology to obtain the complete mitochondrial genome in diverse species. Initially it was possible to obtain the mitochondrial genome of a species with the 454 GS20 platform (Roche) after enriching the genomic area of interest with long-range PCR (Jex et al., 2008). The mitochondrial genome of a perciform was subsequently obtained with the same platform and without the need for enrichment, although amplifications by PCR were required owing to the lack of genomic coverage (Cui et al., 2009). It was later possible to obtain the rhinoceros mitochondrial genome with the 454 GS-FLX platform (Roche), which allowed for sequencing with significantly greater coverage than in other studies (Willerslev et al., 2009). Recently, the "shotgun" method has been used in the same platform to enrich genomic regions containing the mitochondrial genome (Feldmeyer et al., 2010). Other types of platforms are now being used, among them Illumina GAII, which eliminates the need for prior or post-sequencing PCR amplification (Groenenberg et al., 2012). There are few reports of mitochondrial genome studies based on transcriptome approximations (EST libraries) (Yasuike et al., 2012). In this context our study is the first to combine 454 pyrosequencing and gene expression analysis to study the mitochondrial genome of *C. concholepas*.

Comparative analysis of the mitochondrial genome de *C. concholepas* to other closely related species showed greater similarity to the mitochondrion of *T. clavigera* (85.0%), followed by *Rapana venosa* (84.9%,

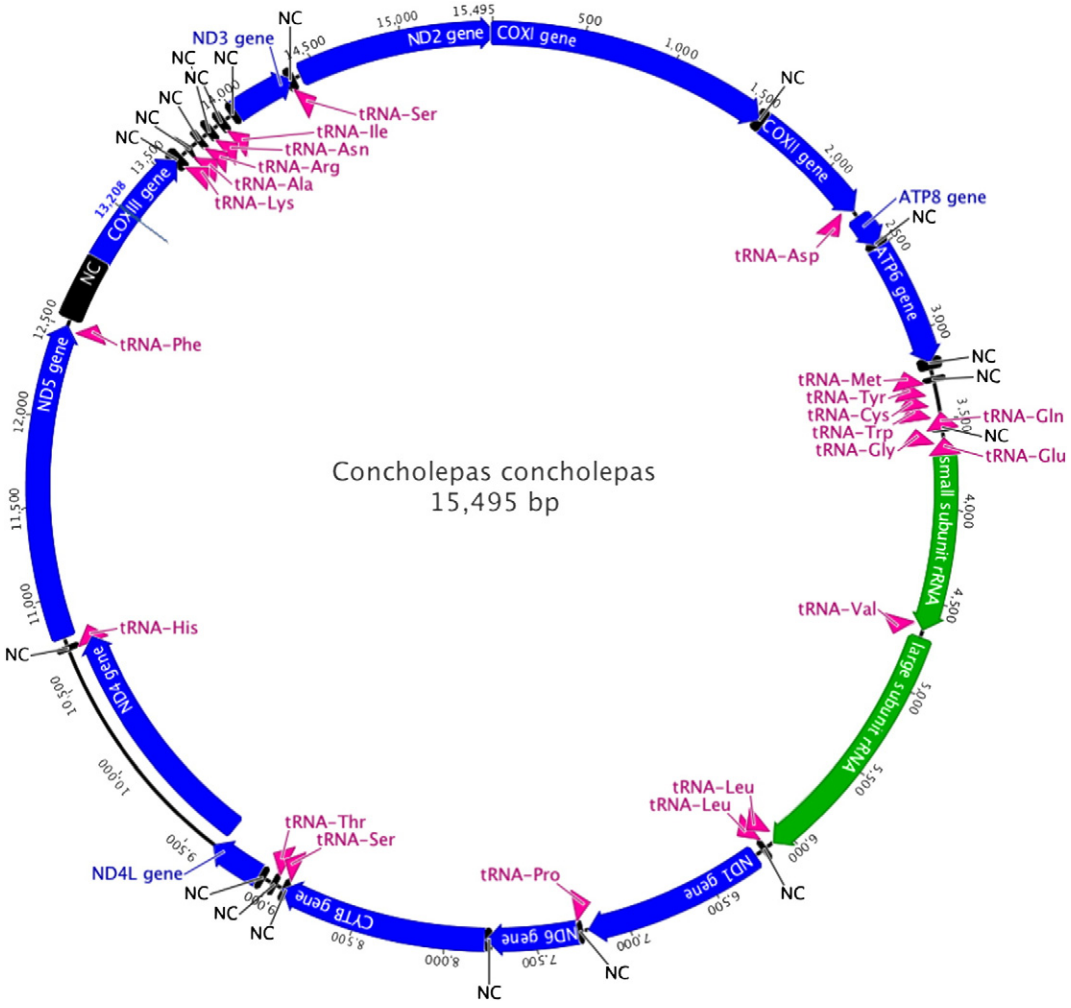

Fig. 1. The mitochondrial genome of *Concholepas concholepas* (15,495 bp). Protein genes are mainly transcribed in a clockwise direction, except for those indicated by the arrowheads. NC corresponds to non-coding intergenic regions. (GenBank accession no. JQ446041).

GenBank accession no. NC_011193). Both species belong to the muricidae family. In relation to other gastropod species, *C. concholepas* mtDNA had molecular similarity to *Lophiotoma cerithiformis* (80.1% of similarity, GenBank accession no. NC_008098) (Bandyopadhyay et al., 2006), *Conus textile* (79.3% of similarity, GenBank accession no. NC_008797), and *Ilyanassa obsoleta* (78.3% of similarity, GenBank accession no. NC_007781) (Simison et al., 2006). The *COX3* gene presented the lowest levels of similarity to this species, with a similarity range of 53.3% and 56.9%. The most conserved gene was *COX2*, which presents a range of similarity between 78.3% and 85.0%. The gene arrangement of the mitochondrial genome of *C. concholepas* is identical to *T. clavigera*, *I. obsoleta* and *L. cerithiformis*. This reflects a high degree of conservation among the mitochondrial genomes described for this group (Cunha et al., 2009).

Regarding intraspecific differences, fifteen SNP polymorphisms were identified between the mitochondrial genome of each population (Table 3). From these SNPs, four were non-synonymous substitutions, involving amino acid changes in *ND1*, *CytB* and *ND2* genes. *ND1* gene evidenced Leucin (L) and Isoluecin (I) substitution, both of them being neutral and hydrophobic amino acids; *Cyt B* had a change between glutamatic acid (E) and aspartic acid (D), which are negatively charged amino acids. Further, *ND2* gene had two non-synonymous SNPs, one of them implied the change between cystein (C) and tyrosin (Y), both neutral and hydrophilic, and the other substitution was between tryptophan (W) and arginine (R), which implied a change between an aromatic amino acid (W) and a positively charged basic amino acid (R). This suggests that the latter SNPs described could be critical for functional changes at the protein level, as it implied substitutions between two amino acids with different chain properties. Other synonymous substitutions were found in *COX1*, *ATP6*, *ND1*, *CytB* and *ND5*.

Studies of the mitochondrial genome in gastropods have been conducted for several purposes. The mitochondrial genomes of some species have been characterized to conduct more robust philogenetic studies (Xin et al., 2011) or for barcode analysis based on more than one mtDNA marker, which is known as multigene barcoding (Zou et al., 2012). However, to date there have not been any studies relating structural mtDNA variations to possible variations in levels of the gene expression of the mitochondrial genome at the population level. The present work assesses whether a possible polymorphism in the gene expression of mtDNA in *C. concholepas* has a biogeographic correlation.

3.2. Analysis of gene expression of mtDNA of C. concholepas

In silico analysis of transcripts expression of the mitochondrial genome showed significant differences between the two populations (north and south) in the variability of expression of some genes. Among the analyzed genes, *COX1* and *CytB* presented the highest level of expression among the northern population (Punta de Choros), while *16S* and *ATP6* presented the highest levels of expression in the southern population (Valdivia). The *COX2* gene had a slightly higher level of expression among individuals from the southern group than those of the northern group. The expression levels of other genes

Table 2
Characteristics of the mitochondrial genome of *Concholepas concholepas*.

Gene	Sense	Position	Length	Start codon	Stop codon	Intergenic nucleotides
COX1	+	1–1533	1533	ATG	TAA	26
COX2	+	1560–2246	687	ATG	TAA	0
tRNA-Asp	+	2244–2312	69			1
ATP8	+	2314–2472	159	ATG	TAA	6
ATP6	+	2479–3174	696	ATG	TAA	37
tRNA-Met	−	3212–3279	68			1
tRNA-Tyr	−	3281–3348	68			0
tRNA-Cys	−	3349–3411	63			0
tRNA-Trp	−	3412–3477	66			0
tRNA-Gln	−	3475–3541	67			9
tRNA-Gly	−	3551–3617	67			0
tRNA-Glu	−	3617–3683	67			0
s-rRNA	+	3684–4650	967			0
tRNA-Val	+	4651–4719	69			0
l-rRNA	+	4720–6084	1365			0
tRNA-Leu	+	6085–6154	70			1
tRNA-Leu	+	6156–6224	69			0
ND1	+	6225–7199	975	ATG	TAG	0
tRNA-Pro	−	7190–7256	67			1
ND6	+	7258–7758	501	ATG	TAA	8
CytB	+	7767–8906	1140	ATG	TAA	8
tRNA-Ser	+	8915–8979	65			2
tRNA-Thr	+	8982–9048	67			9
ND4L	+	9058–9342	285	ATG	AGT	1
ND4	+	9344–10,717	1374	ATG	TAA	1
tRNA-His	+	10,719–10,785	67			0
ND5	+	10,786–12,494	1709	ATG	TTA	0
tRNA-Phe	+	12,495–12,561	67			0
NCR	0	12,562–12,891	330			0
COX3	+	12,892–13,578	687	ATG	TAA	39
tRNA-Lys	+	13,618–13,685	68			3
tRNA-Ala	+	13,689–13,775	67			12
tRNA-Arg	+	13,768–13,836	69			4
tRNA-Asn	+	13,841–13,908	68			15
tRNA-Ile	+	13,924–13,990	67			3
ND3	+	13,994–14,347	354	ATG	TAA	15
tRNA-Ser	+	14,363–14,430	68			0
ND2	+	14,431–15,495	1065	ATG	TAA	0

did not show significant differences between the two locations (Figs. 2 and 3). Additional *in silico* analysis identified five clusters according to the levels of gene expression normalized between the two analyzed populations of *C. concholepas* (Fig. 4). Cluster 1 grouped the *COX2*, *ND3*, *ND2* and *ND6* genes, with a higher tendency of higher expression in the southern group. Cluster 2 contained only one transcript associated with *CytB* gene, which had a higher value of expression in the northern population. Cluster 3 grouped the majority of the genes that show low differentiation between the two locations. Of these, the *COX3*, *ND1*, *ND5*, *ND4L s-rRNA* genes and the non-coding region showed a higher expression in the northern population. The *ATP8* had the opposite tendency, with higher expression in the southern population. The *ND4* gene did not present variations between the two locations. Cluster 4 consisted only of the *COX1* gene with a higher level of expression in the northern population according to normalized values. Finally, cluster 5 grouped the *l-rRNA* and *ATP6* genes with a strong tendency for higher expression in the southern population (Fig. 4).

To validate the *in silico* analysis, we carried out a qPCR analysis, determining the relative abundance of the studied mtDNA genes in *C. concholepas*. The highest levels of transcripts in the northern population were for the *COX1*, *COX2* and *CytB* genes, while the other genes did not present significant differences. The *CytB* gene showed variations between the two locations (Fig. 5). In general, the results of the qPCR analysis were consistent with those obtained from the RNA-seq. However, the *COX2* gene showed an inverse relationship to the *in silico* analysis, in which higher expression of this gene was found in the southern population. A possible explanation of these results is the low differentiation in RPKM values for the *COX2* gene at the population level (Fig. 3), or the results are due to variations of unevaluated gene expression

Table 3
Polymorphic SNPs identified between the mitochondrial genome of two populations. S = synonymous SNPs, NS = non-synonymous SNPs.

Polymorphism	S/NS	Gene	Position	Amino acid	Coverage
T/C	S	COX1	753	Isoleucine (I)	202
G/A	S	ATP6	2616	Valine (V)	60
G/A	S	ATP6	2673	Glycine (G)	76
C/T	S	ATP6	2742	Serine (S)	100
A/G	S	ND1	6311	Glycine (G)	19
A/G	S	ND1	6386	Alanine (A)	46
A/G	S	ND1	6518	Glycine (G)	76
C/A	NS	ND1	7103	Leucine (L)–isoleucine (I)	94
G/A	S	CytB	8216	Leucine (L)	338
A/G	NS	CytB	8417	Glutamic acid (E)–aspartic acid (D)	322
C/T	S	CytB	8594	Leucine (L)	295
T/G	S	ND5	11,433	Leucine (L)	30
A/G	S	ND5	11,946	Glycine (G)	63
T/C	NS	ND2	15,007	Tryptophan (W)–arginine (R)	132
G/A	NS	ND2	15,236	Cysteine (C)–tyrosine (Y)	594

(interaction with the environment). As well, the *l-rRNA* and *ATP6* genes did not show significant differences through the qPCR analysis, although at the *in silico* level these genes show a high degree of correlation and higher gene expression in the southern population of *C. concholepas* (Fig. 4, cluster 5).

The expression of mtDNA genes has been studied from diverse points of view. For example, the response of the *COX1* gene in zebra

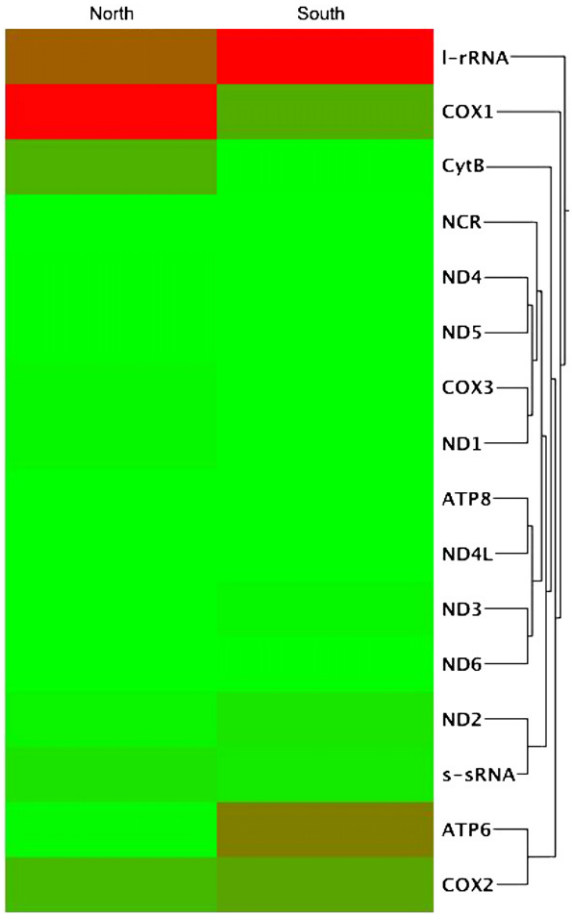

Fig. 2. Heat-map of mtDNA gene expression levels for *Concholepas concholepas*. Expression level was measured in RPKM from normalized values. Red color indicates high expression values and green color low expression values. Distance was inferred by Euclidean distance using average linkage. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

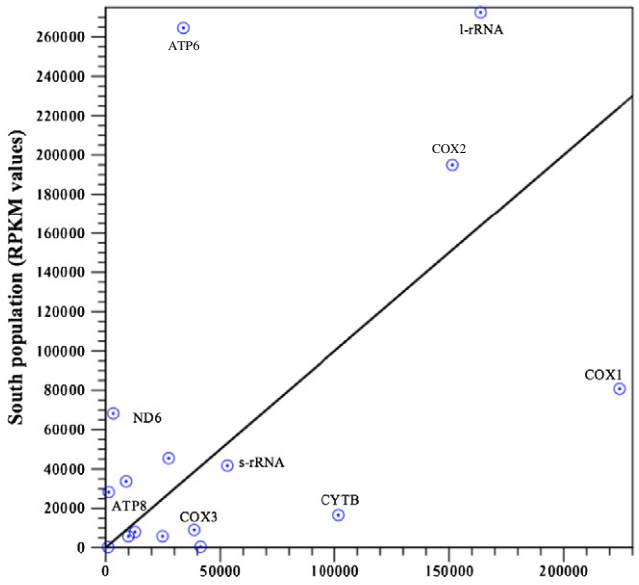

Fig. 3. Scatter plot of RPKM values between South and North populations for mtDNA genes in *Concholepas concholepas*.

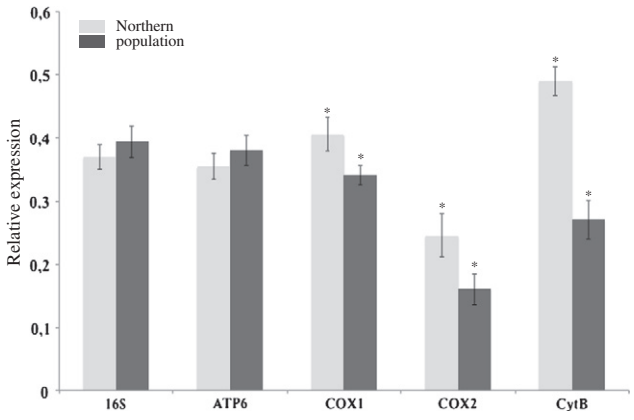

Fig. 5. Quantitative PCR analysis of mtDNA genes in *Concholepas concholepas* populations. * represents significant differences ($p<0.05$).

fish has been associated with the presence of metallic pollutants in the diet (Gonzalez et al., 2005). The levels of expression of *COX1* in marine bivalves and fish have been associated with the presence of cadmium in marine environments (Achard-Joris et al., 2006). Similar behavior of this gene was described in different developmental states of the cephalopod *Sepia officinalis* under conditions of excessive environmental CO_2 (Hu et al., 2011). A study in the fish *Sebastiscus marmoratus* found that *COX2* is over-expressed due to exposure to the pollutant TBT (Li et al., 2010). The *COX2* gene has also been related to the adaptation of the fish *Gadus morhua* to low temperatures (Lucassen et al., 2006). The *ATP6* gene has been associated with stress response induced by excess CO_2 in marine environments, which established a relationship between acidification of marine environments and levels of *ATP6* expression (Stumpp et al., 2011). Future studies will assess how the environment modifies expression patterns of mtDNA genes and relationships at the population level. As well, it is necessary to assess the temporal component of variations in gene expression and determine if there are patterns associated with the biogeography of *C. concholepas*.

Parallel to qPCR analysis of the mtDNA genes, the present study assessed three genes associated with the expression of the mitochondrial genome. There were significant differences in the *PRKAG2* and *SD* genes of the two populations (Fig. 6), with higher levels of expression in the *PRKAG2* gene of the northern population and the *SD* gene of the southern population. The levels of HIF expression do not show statistically significant variations. Studies of hypoxia in the bivalve *Crassostrea virginica* have directly associated *HIF* and *SD* levels to the transcriptome response of the mitochondrial genome in aquatic environments with low levels of dissolved oxygen (Ivanina et al., 2010; Ivanina et al., 2011). Finally, intra-population analysis of mtDNA gene expression showed that the population of *C. concholepas* in northern Chile probably has a higher level of gene transcription, such as *COX1*, *COX2*, *CytB* and *PRKAG2*, while the individuals from the south in general showed lower relative abundance of mtDNA gene transcripts, but higher expression of the *SD* gene.

The debate about how these genes behave in the context of diverse sources of induction at the level of their expression remains unresolved (Arnqvist et al., 2010). In this context, there are studies to determine if there is a mitochondrial and nuclear genetic origin to the changes at a metabolic level in invertebrates (Ellison and Burton, 2008; Ellison and Burton, 2010; Van Wormhoudt et al., 2011). Nevertheless, future studies should be focused on the behavior of these genes at a functional level to establish genotype–environment type relationships to identify the origin of the physiological changes in marine species in response to exogenous and endogenous agents.

4. Conclusions

This study reports the mitochondrial genome of *C. concholepas*, which can be used in evolutionary and ecological studies for the species. As well, it is the first study to characterize the mitochondrial genome by 454 pyrosequencing of RNA samples, and in parallel *in silico* and qPCR analysis of differential expression of the transcripts associated with the mitochondrial genome. However, the major contribution of this work is to show that state-of-the-art molecular tools can be used to contribute in one aspect of the understanding of the physiological responses of

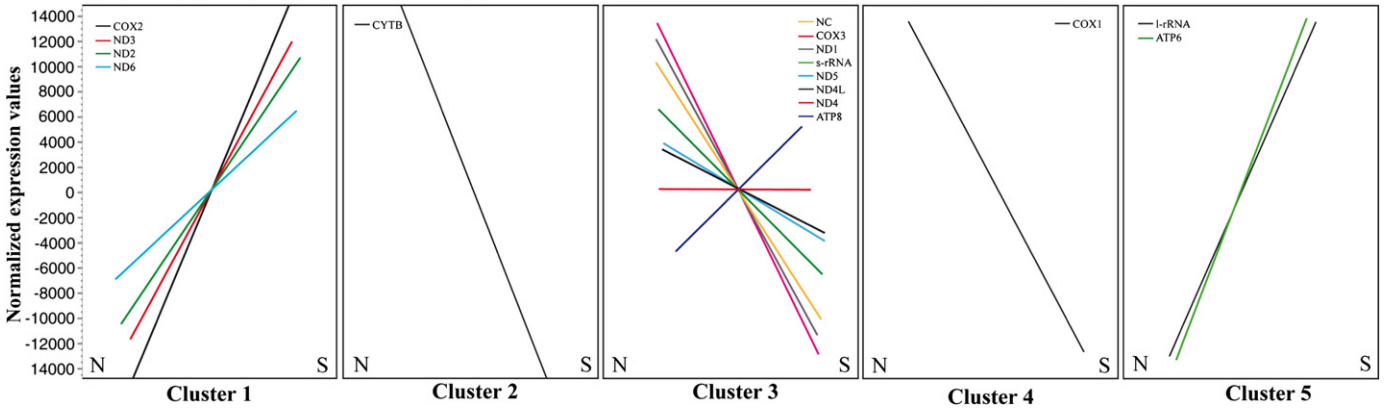

Fig. 4. Clusters analysis for mtDNA gene expression of *Concholepas concholepas*. Normalized expression values are inferred by South and North populations.

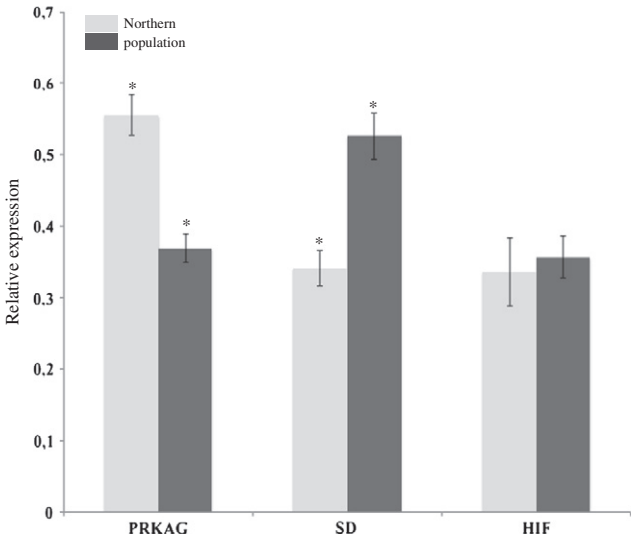

Fig. 6. Quantitative PCR analysis of nuclear genes involved in mtDNA gene expression of *Concholepas concholepas* populations. (SD) Succinate dehydrogenase, (PRKAG2) protein kinase, AMP-activated, gamma 2 non-catalytic subunit and (HIF) hypoxia inducible factor. * represents significant differences ($p<0.05$).

commercial aquatic species in different environments through the analysis of gene expression of transcripts associated with this genome. This represents necessary information to understand the relationship between the levels of genes associated with the mitochondrion in populations separated spatially.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cbd.2012.10.004.

Acknowledgments

This work was supported by the FONDEF-CONICYT, Chile (Project D09I1065).

References

Achard-Joris, M., Gonzalez, P., Marie, V., Baudrimont, M., Bourdineaud, J.-P., 2006. Cytochrome *c* oxydase subunit I gene is up-regulated by cadmium in freshwater and marine bivalves. Biometals 19, 237–244.

Andersen, C.L., Jensen, J.L., Ørntoft, T.F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245–5250.

Arnqvist, G., Dowling, D.K., Eady, P., Gay, L., Tregenza, T., Tuda, M., Hosken, D.J., 2010. Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. Evolution 64, 3354–3363.

Bandyopadhyay, P.K., Stevenson, B.J., Cady, M.T., Olivera, B.M., Wolstenholme, D.R., 2006. Complete mitochondrial DNA sequence of a Conoidean gastropod, *Lophiotoma* (Xenuroturris) *cerithiformis*: gene order and gastropod phylogeny. Toxicon 48, 29–43.

Boore, J., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Boyce, T., Zwick, M., Aquadro, C., 1989. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. Genetics 123, 825–836.

Bridge, D., Cunningham, C., Schierwater, B., DeSalle, R., Buss, L., 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. Proc. Natl. Acad. Sci. U. S. A. 89, 8750–8753.

Burger, G., Lavrov, D.V., Forget, L., Lang, B.F., 2007. Sequencing complete mitochondrial and plastid genomes. Nat. Protoc. 2, 603–614.

Cárdenas, L., Daguin, C., Castilla, J.C., Viard, F., 2006. Isolation and characterization of 11 polymorphic microsatellite markers for the marine gastropod *Concholepas concholepas* (Brugière, 1789). Mol. Ecol. Notes 7, 464–466.

Cárdenas, L., Castilla, J.C., Viard, F., 2009. A phylogeographical analysis across three biogeographical provinces of the south-eastern Pacific: the case of the marine gastropod *Concholepas concholepas*. J. Biogeogr. 36, 969–981.

Cárdenas, L., Sánchez, R., Gomez, D., Fuenzalida, G., Gallardo-Escárate, C., Tanguy, A., 2011. Transcriptome analysis in *Concholepas concholepas* (Gastropoda, Muricidae): mining and characterization of new genomic and molecular markers. Mar. Genomics 4, 197–205.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676.

Cui, Z., Liu, Y., Li, C.P., You, F., Chu, K.H., 2009. The complete mitochondrial genome of the large yellow croaker, *Larimichthys crocea* (Perciformes, Sciaenidae): unusual features of its control region and the phylogenetic position of the Sciaenidae. Gene 432, 33–43.

Cunha, R.L., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evol. Biol. 9, 210.

Ellison, C.K., Burton, R.S., 2008. Genotype-dependent variation of mitochondrial transcriptional profiles in interpopulation hybrids. Proc. Natl. Acad. Sci. U. S. A. 105, 15831–15836.

Ellison, C.K., Burton, R.S., 2010. Cytonuclear conflict in interpopulation hybrids: the role of RNA polymerase in mtDNA transcription and replication. J. Evol. Biol. 23, 528–538.

Feldmeyer, B., Hoffmeier, K., Pfenninger, M., 2010. The complete mitochondrial genome of *Radix balthica* (Pulmonata, Basommatophora), obtained by low coverage shot gun next generation sequencing. Mol. Phylogenet. Evol. 57, 1329–1333.

Gonzalez, P., Dominique, Y., Massabuau, J.C., Boudou, A., Bourdineaud, J.P., 2005. Comparative effects of dietary methylmercury on gene expression in liver, skeletal muscle, and brain of the zebrafish (*Danio rerio*). Environ. Sci. Technol. 39, 3972–3980.

Groenenberg, D., Pirovano, W., Gittenberger, E., Schilthuizen, M., 2012. The complete mitogenome of *Cylindrus obtusus* (Helicidae, Ariantinae) using Illumina Next Generation Sequencing. BMC Genomics 13, 114.

Hu, M.Y., Tseng, Y.-C., Stumpp, M., Gutowska, M.A., Kiko, R., Lucassen, M., Melzner, F., 2011. Elevated seawater PCO_2 differentially affects branchial acid–base transporters over the course of development in the cephalopod *Sepia officinalis*. Am. J. Physiol. 300, R1100–R1114.

Ivanina, A.V., Sokolov, E.P., Sokolova, I.M., 2010. Effects of cadmium on anaerobic energy metabolism and mRNA expression during air exposure and recovery of an intertidal mollusk *Crassostrea virginica*. Aquat. Toxicol. 99, 330–342.

Ivanina, A.V., Froelich, B., Williams, T., Sokolov, E.P., Oliver, J.D., Sokolova, I.M., 2011. Interactive effects of cadmium and hypoxia on metabolic responses and bacterial loads of eastern oysters *Crassostrea virginica* Gmelin. Chemosphere 82, 377–389.

Jex, A.R., Hu, M., Littlewood, D.T.J., Waeschenbach, A., Gasser, R.B., 2008. Using 454 technology for long-PCR based sequencing of the complete mitochondrial genome from single *Haemonchus contortus* (Nematoda). BMC Genomics 9, 11.

Ki, J.-S., Lee, Y.-M., Jung, S.-O., Horiguchi, T., Cho, H.-S., Lee, J.-S., 2010. Mitochondrial genome of *Thais clavigera* (Mollusca: Gastropoda): affirmation of the conserved, ancestral gene pattern within the mollusks. Mol. Phylogenet. Evol. 54, 1016–1020.

Leiva, G.E., Castilla, J.C., 2001. A review of the world marine gastropod fishery: evolution of catches, management and the Chilean experience. Rev. Fish Biol. Fish. 11, 283–300.

Li, B., Wang, C., Yu, A., Chen, Y., Zuo, Z., 2010. Identification of differentially expressed genes in the brain of *Sebastiscus marmoratus* in response to tributyltin exposure. Toxicology 99, 248–255.

Lucassen, M., Koschnick, N., Eckerle, L.G., Pörtner, H.-O., 2006. Mitochondrial mechanisms of cold adaptation in cod (*Gadus morhua* L.) populations from different climatic zones. J. Exp. Biol. 209, 2462–2471.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 1–8.

Neira-Oviedo, M., Tsyganov-Bodounov, A., Lycett, G.J., Kokoza, V., Raikhel, A.S., Krzywinski, J., 2011. The RNA-Seq approach to studying the expression of mosquito mitochondrial genes. Insect Mol. Biol. 20, 141–152.

Passamonti, M., Ricci, A., Milani, L., Ghiselli, F., 2011. Mitochondrial genomes and Doubly Uniparental Inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (Bivalvia Mytilidae). BMC Genomics 12, 442.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45.

Rawlings, T.A., MacInnis, M.J., Bieler, R., Boore, J.L., Collins, T.M., 2010. Sessile snails, dynamic genomes: gene rearrangements within the mitochondrial genome of a family of caenogastropod molluscs. BMC Genomics 11, 440.

Simison, W.B., Lindberg, D.R., Boore, J.L., 2006. Rolling circle amplification of metazoan mitochondrial genomes. Mol. Phylogenet. Evol. 39, 562–567.

Stumpp, M., Dupont, S., Thorndyke, M.C., Melzner, F., 2011. CO(2) induced acidification impacts sea urchin larval development II: gene expression patterns in pluteus larvae. Comp. Biochem. Physiol. Part A 160, 320–330.

Sun, M., Shen, X., Liu, H., Liu, X., Wu, Z., Liu, B., 2011. Complete mitochondrial genome of *Tubulipora flabellaris* (Bryozoa: Stenolaemata): the first representative from the class Stenolaemata with unique gene order. Mar. Genomics 4, 159–165.

Van Wormhoudt, A., Roussel, V., Courtois, G., Huchette, S., 2011. Mitochondrial DNA introgression in the European abalone *Haliotis tuberculata tuberculata*: evidence for experimental mtDNA paternal inheritance and a natural hybrid sequence. Mar. Biotechnol. 13, 563–574.

Willerslev, E., Gilbert, M.T., Binladen, J., Ho, S., Campos, P., Ratan, A., Tomsho, L., da Fonseca, R., Sher, A., Kuznetsova, T., Nowak-Kemp, M., Roth, T., Miller, W., Schuster, S., 2009. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. BMC Evol. Biol. 9, 95.

Wolstenholme, D., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. 141, 173–216.

Xin, Y., Ren, J., Liu, X., 2011. Mitogenome of the small abalone *Haliotis diversicolor* Reeve and phylogenetic analysis within Gastropoda. Mar. Genomics 1–10.

Yasuike, M., Leong, J., Jantzen, S.G., von Schalburg, K.R., Nilsen, F., Jones, S.R.M., Koop, B.F., 2012. Genomic resources for sea lice: analysis of ESTs and mitochondrial genomes. Mar. Biotechnol. 14, 155–166.

Zou, S., Li, Q., Kong, L., 2012. Multigene barcoding and phylogeny of geographically widespread muricids (Gastropoda: Neogastropoda) along the coast of China. Mar. Biotechnol. 14, 21–34.