Marine Genomics 23 (2015) 37–40

Contents lists available at ScienceDirect

Marine Genomics

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

Genomics/Technical resources

Transcriptome analysis of gill tissue of Atlantic cod *Gadus morhua* L. from the Baltic Sea

Magdalena Małachowicz, Agnieszka Kijewska, Roman Wenne *

Institute of Oceanology PAS, 81-712 Sopot, Poland

ARTICLE INFO

Article history:
Received 23 March 2015
Received in revised form 15 April 2015
Accepted 15 April 2015
Available online 24 April 2015

Keywords:
Atlantic cod
Baltic
454 pyrosequencing
Transcriptome

ABSTRACT

The Atlantic cod (*Gadus morhua* L.) is one of the most ecologically and economically important marine fish species in the North Atlantic Ocean. Using Roche GS-FLX 454 pyrosequencing technique 962,516 reads, representing 379 Mbp of the Baltic cod transcriptome, were obtained. Data was assembled into 14,029 contigs of which 100% displayed homology to the Atlantic cod transcriptome. Despite a high similarity between transcripts, evidence for significant differences between Baltic and Atlantic cod was found.

1. Introduction

The Atlantic cod (*Gadus morhua* L.) is a key species in the North Atlantic Ocean, with significant ecological and commercial value. It lives at low-salinity (7–12‰) in the Baltic Sea, but requires salinity over 14‰ for successful spawning (Nissling and Westin, 1997). Areas suitable for reproduction of cod are accessible in a few zones only in the Baltic Sea (Bornholm, Belt Sea) thanks to rare intrusions of high salinity and well oxygenated water from the North Sea. The Baltic cod produces eggs with thinner chorion compared to fish from outside the Baltic as a result of adaptation to low salinity (Nissling et al., 1994). The Baltic and Atlantic populations differ genetically (Nielsen et al., 2003; O'Leary et al., 2007; Poćwierz-Kotus et al., 2015), but current understanding of Baltic cod genetics is still limited. Several Atlantic cod transcriptomes obtained from eggs and embryos have been characterized with micro-arrays (Rise et al., 2014; Škugor et al., 2014). Furthermore, Roche GS-FLX 454 pyrosequencing based transcriptomes from different tissues (brain, gonad, head kidney, hindgut, liver, spleen) and eggs have been published (Star et al., 2011; Lanes et al., 2013). However, no transcriptome data is currently available for Atlantic cod from the Baltic Sea. The aim of this study was to assemble, annotate and analyze Baltic cod transcriptome from gills and compare it to the Atlantic cod available data from the Cod Genome Project.

2. Data description

2.1. RNA preparation and sequencing

Atlantic cod were sampled in November 2011 from the Schleifjord, Kiel Bight (KIEL) and Gdańsk Bay (GDA). Live fish were transported to the Marine Station of the University of Gdańsk in Hel and settled in tanks (3500 l). They were kept at 10 °C in water which simulated the natural salinities of the geographic source region of the cod (KIEL — 18‰ and GDA — 8‰). The fish were maintained at natural photoperiod and acclimated to laboratory conditions for a minimum of 14 days until they start feeding and displayed typical behaviour. All experiments complied with EC Directive 2010/63/EU for animal experiments and with the permission no. 60/2012 of the Local Ethics Committee on Animal Experimentation No. 3.

Samples of gills were collected and stored in RNAlater, following the manufacturer's instruction (Quiagen, Hilden, Germany). Total RNA from 36 individuals was extracted using the ISOLATE II RNA Mini Kit (Bioline, London, UK) and stored at −20 °C. Concentration of extracted RNA was determined at 260 nm on microplate using the Epoch Microplate Spectrophotometer (BioTek Instruments, Inc., Winooski, USA). Samples were processed using a 454 pyrosequencing by CD Genomics (USA).

2.2. Transcriptome assembly

After trimming adapters and low quality bases, a total of 962,516 reads with mean length 300–400 bp (Fig. 1) were used for de novo

* Corresponding author. Tel.: +48 58 731 17 63; fax: +48 58 551 21 30.
E-mail address: rwenne@iopan.gda.pl (R. Wenne).

http://dx.doi.org/10.1016/j.margen.2015.04.005

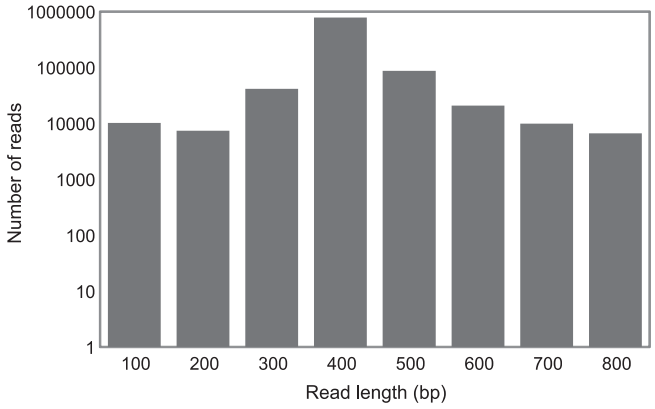

Fig. 1. Frequency distribution of raw reads lengths on a logarithmic scale.

transcriptome assembly using the CLC Genomics Workbench ver. 7.5.1, CLC Bio, Aarhus, Denmark (see Supplementary methods for details).

2.3. Functional annotation

Of the 14,029 contigs, 100% homologous transcripts (13,585 genes) were identified using BLASTn with an E-value threshold of 10^{-10}. All contigs were divided into biotypes (Fig. 2A). Gene ontology (GO) terms were assigned to each contig using the Blast2GO tool (www.blast2go.com, Götz et al., 2008). Among 13,585 genes 75.39% had an ontology definition and were classified into three main GO categories (biological process — BP, molecular function — MF and cellular component — CC) and 45 subcategories (Fig. 2B). Among these genes, 48.34% were assigned to at least one GO term in the molecular function category, 33.24% in biological processes, and 18.42% in cellular component (see Supplementary methods for details).

In addition, we compared Baltic cod contigs with the Atlantic Ocean cod reference transcriptome (http://www.ensembl.org/Gadus_morhua) using BLASTn. The identified Baltic cod transcripts covered 58% of the Atlantic cod transcriptome (Fig. 3A). While 13,827 (98.56%) Baltic transcripts aligned perfectly and unambiguously, 202 (1.44%; 170 genes) demonstrated significant distinctiveness from Atlantic cod (Fig. 3B). Among those 170 genes 100% were protein coding sequences; 68.82%

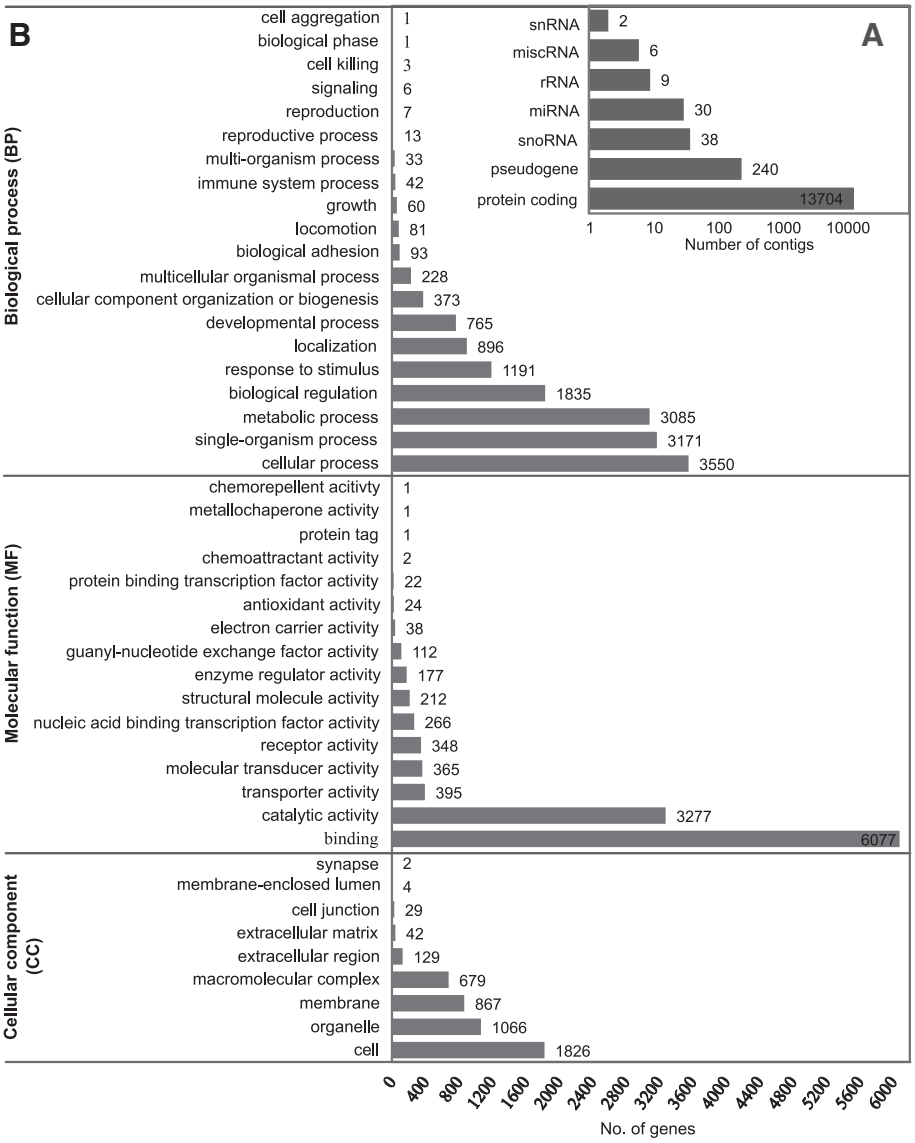

Fig. 2. Summary of Baltic cod transcriptome assembly. (A) Among 14,029 contigs 13,704 were coding proteins, 240 were annotated as pseudogenes, and 85 as small noncoding RNAs (logarithmic scale). (B) A total of 10,242 genes were assigned to at least one GO term and were grouped into three main categories and 45 subcategories.