Genomics Data 8 (2016) 81–84

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

Identification of myogenic regulatory genes in the muscle transcriptome of beltfish (*Trichiurus lepturus*): A major commercial marine fish species with robust swimming ability

Hui Zhang [a,b,1], Chung-Ming Chang [c,1], Kang-Ning Shen [d], Weiwei Xian [a,b,*], Chung-Der Hsiao [e,**]

[a] Key Laboratory of Marine Ecology and Environment Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China
[b] Laboratory of Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China
[c] Department of Medical Biotechnology and Laboratory Science, College of Medicine, Chang Gung University, Tao-Yuan 33302, Taiwan
[d] Center of Excellence for the Oceans, National Taiwan Ocean University, 20224 Keelung, Taiwan
[e] Department of Bioscience Technology, Chung Yuan Christian University, 32023 Chung-Li, Taiwan

ARTICLE INFO

Article history:
Received 11 April 2016
Accepted 12 April 2016
Available online 17 April 2016

Keywords:
Muscle
Transcriptome
Beltfish
Myogenic regulatory gene

ABSTRACT

The beltfish (*Trichiurus lepturus*) is considered as one of the most economically important marine fish in East Asia. It is a top predator with a robust swimming ability that is a good model to study muscle physiology in fish. In the present study, we used Illumina sequencing technology (NextSeq500) to sequence, assemble and annotate the muscle transcriptome of juvenile beltfish. A total of 57,509,280 clean reads (deposited in NCBI SRA database with accession number of SRX1674471) were obtained from RNA sequencing and 26,811 unigenes (with N50 of 1033 bp) were obtained after de novo assembling with Trinity software. BLASTX against NR, GO, KEGG and eggNOG databases show 100%, 49%, 31% and 96% annotation rate, respectively. By mining beltfish muscle transcriptome, several key genes which play essential role on regulating myogenesis, including *pax3*, *pax7*, *myf5*, *myoD*, *mrf4*/*myf6*, *myogenin* and *myostatin* were identified with a low expression level. The muscle transcriptome of beltfish can provide some insight into the understanding of genome-wide transcriptome profile of teleost muscle tissue and give useful information to study myogenesis in juvenile/adult fish.

Specifications [standardized info for the reader]	
Organism/cell line/tissue	*Trichiurus lepturus*/muscle
Sex	N/A
Sequencer or array type	Illumina NextSeq500
Data format	Raw and processed
Experimental factors	Transcriptome profiling of muscle at adult stage
Experimental features	Muscle tissues were dissected from the juvenile beltfish and total RNAs were extracted by using TRIZOL reagent. Prepared cDNA libraries were paired-end sequenced by NextSeq500 platform. The obtained data was subjected for de novo transcriptome assembly using Trinity. The assembled unigene was later functionally annotated by searching NR, GO, KEGG, eggNOR and Swissprot databases.
Consent	N/A
Sample and location	One juvenile fish of beltfish was captured at Yangtze Estuary on Nov 4th, 2015

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/sra/SRX1674471 for muscle transcriptome.

2. Introduction

The muscle development process is spatially and temporally orchestrated by multiple myogenic regulatory factors. At cell fate determination stage of embryo, the external signals trigger mesodermal muscle progenitors transforming into myoblast. The interplay of *pax3*/*pax7*, *myoD*, *myf5* and *mrf4*/*myf6* promotes myoblast differentiating into multinucleate myotube (by a cell fusion process) which highly expressed *myogenin* and *mrf4*/*myf6*. The *myostatin* (also known as GDF-8), on the contrary, plays an inhibitory role on muscle differentiation. Later, the multinucleate myotube further assembles into muscle fiber and expresses variety of muscle specific proteins (MSPs) to build up the major architecture of muscle fiber [1,10,12,14,17]. Basically, the muscle development process is highly conserved between fish and high vertebrates. However, unlike their high vertebrate counterparts, fish continue their hyperplastic and hypertrophic muscle growth through adult stage [9,13]. Study on how muscle can have continuous cell growth

* Correspondence to: W. Xian, Key Laboratory of Marine Ecology and Environment Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China.
** Corresponding author.
E-mail addresses: wwxian@qdio.ac.cn (W. Xian), cdhsiao@cycu.edu.tw (C.-D. Hsiao).
[1] Equal contribution authors.

http://dx.doi.org/10.1016/j.gdata.2016.04.005

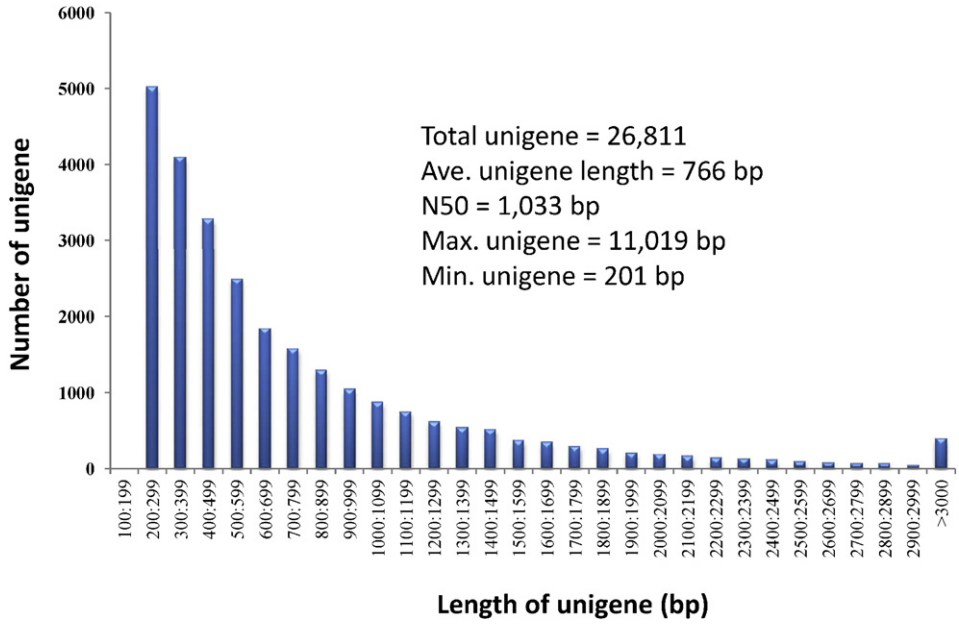

Fig. 1. Length distribution of the assembled unigene of beltfish muscle transcriptome.

and proliferation in fish can benefit better understanding on boosting the growth/repair of muscle tissue at juvenile/adult stage in high vertebrates.

Beltfish (*Trichiurus lepturus*) is a member of the cutlassfish family (Trichiuridae) and is a major commercial marine fish species with robust swimming ability. It is a long, slender fish found throughout the tropical and temperate oceans of the world. Beltfish is a major commercial species in Northwest Pacific, especially in China, South Korea and Japan. Juvenile beltfish participate in the diet vertical migration, rising to feed on krill and small fish during the night and returning to the sea bed in the day. The movement pattern is reversed by large adults, which mainly feed on fish, squid and shrimp. Adults are highly carnivorous and will cannibalize younger individuals regularly [2]. Since Beltfish play as a top predator in the marine ecosystem, they provide a good model to monitor the bioaccumulation of heavy metals (like mercury and selenium) in the muscle tissues for a long time [4,15,16]. In this study, we took Beltfish as a marine fish model to explore muscle gene regulation at a molecular level by RNAseq approach. The establishment of muscle transcriptome provides not only useful information for evaluating the biological impact of heavy metal bioaccumulation on muscle, but also provides fundamental information for myogenic regulatory gene expression at juvenile/adult stages.

3. Experimental design, materials and methods

3.1. RNA extraction

The muscle tissue was dissected from one wild captured juvenile beltfish (body length around 15 cm) and immediately stored in RNAlater (Qiagen, Hilden, Germany) and then stored at −80 °C prior to RNA extraction. Total RNAs were extracted by using the TRIZOL Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Total RNA samples were then digested by DNase I to remove potential genomic DNA contamination. Integrity and size distribution were checked with Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA).

3.2. RNA isolation, library construction and Illumina sequencing

Initially, about 2.5 μg of starting total RNAs were used to synthesize cDNA libraries by following the standard protocols of the Illumina TruSeq RNA Sample Preparation Kit (Illumina). The final library had an insert size about 200–300 bp. After qPCR quantitation and dilution, the library was sequenced on an Illumina NextSeq500 with 150 bp paired-end reads. A total of 57,851,358 raw paired-end reads were generated. Adaptor sequences were trimmed; and reads with low quality were removed by cutadapt software [8]. After the removal of ambiguous nucleotides, duplicates and low-quality sequences (Phred quality scores < 20), a total of 57,509,280 cleaned reads (99.4%) were obtained. The raw transcriptome sequences in the present study were deposited in the NCBI SRA database (SRX1674471).

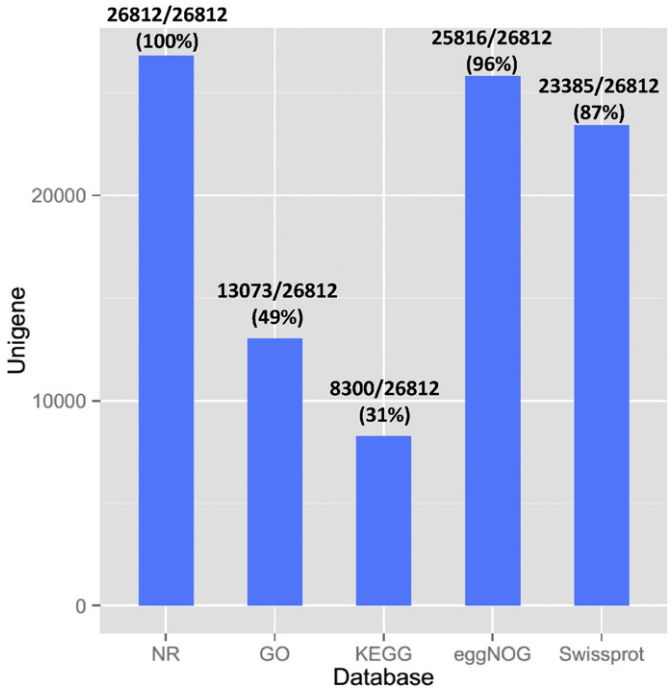

Fig. 2. Comparison of the gene annotation rate of unigene against NR, GO, KEGG, eggNOR and Swissprot databases.