Gene 654 (2018) 57–63

Contents lists available at ScienceDirect

Gene

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

Research paper

Discovery and characterization of conserved and novel microRNAs from blunt snout bream (*Megalobrama amblycephala*) by deep sequencing

Yong Huang[a,*], Wangbao Gong[b], Jianli Xiong[a], Xiao Chan Gao[a], Hong Tao Ren[a]

[a] College of Animal Science and Technology, Henan University of Science and Technology, Luoyang, China
[b] Key Laboratory of Tropical and Subtropical Fishery Resource Application and Cultivation, Ministry of Agriculture, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, China

ARTICLE INFO

Keywords:
MicroRNA
Megalobrama amblycephala
Deep sequencing
qRT-PCR
Target gene

ABSTRACT

MicroRNAs (miRNAs) are short, single stranded RNA molecules with approximately 22 nts in length, which regulate the stability and translation of messenger RNAs in several organisms. To increase the repertoire of miRNAs characterized in M. amblycephala, we used the deep sequencing technology to sequence a small RNA library using pooled RNA sample isolated from the 4 different tissues of *M. amblycephala*. A total of 309 conserved miRNAs that originated from 131 miRNA families were detected. 15 novel candidates miRNA were identified. Randomly selected 6 miRNAs were analyzed by stem-loop qRT-PCR and differential expression patterns were observed in 6 different tissues of *M. amblycephala*. Furthermore, the potential targets were predicted. GO analysis showed that most of the targets were involved in a broad range of physiological functions including fish growth, development, metabolism, stress responses and so on. Overall, our results significantly increased the number of novel miRNAs in M. amblycephala, which should be useful for further investigation into the role of miRNAs in regulating diverse biological processes.

1. Introduction

MicroRNAs (miRNAs) are small, non-coding RNAs, 19–25 nts with an average of 22 nts in length, they regulate gene expression at the post-transcriptional level through binding mostly to the 3′UTRs of mRNAs, which leads to the inhibition of translation or the degradation of mRNAs (John et al., 2004; Lee et al., 2009). miRNAs are transcribed as pri-miRNA by RNA polymerase II in the nucleus, and then pri-miRNA are cleaved by Drosha to produce a short hairpin pre-miRNAs (Lund et al., 2004; Zeng et al., 2005). Finally, pre-miRNAs is further cleaved to form a miRNA:miRNA* duplex and incorporated into the RNA-induced silencing complex (RISC) which then regulates gene expression (Cambronne et al., 2012; Gregory et al., 2005). An increasing amount of evidences demonstrated that miRNAs play an important role in almost all biological processes, including developmental timing, organ differentiation, pathogenesis of diseases, proliferation, apoptosis, oncogenesis, immunity and in response to environmental stress (Carthew et al., 2017; Gao et al., 2017; Huang et al., 2011; Wienholds and Plasterk, 2005; Zhang et al., 2015).

Blunt snout bream (*Megalobrama amblycephala*) is a fish species of high economic value and accounts for an high proportion of Chinese aquaculture yields, which is farmed in freshwater polyculture systems since 1960s (Ji et al., 2014; Li et al., 2012). Although it is an important economic trait, the growth rate of *M. amblycephala* has not been well defined by molecular regulation mechanisms due to the limited genomic information of *M. amblycephala*. Thus, the identification of miRNAs information contributes to the gene expression profile analysis and miRNA functional annotation. Recently, deep sequencing technologies have been widely utilized to precisely identify conserved and novel miRNAs in several species (Andreassen et al., 2013; Castellano and Stebbing, 2013; Fu et al., 2013; Liu et al., 2017; Yuan et al., 2014; Zhang et al., 2013; Zhao et al., 2016).

Currently, some miRNAs in *M. amblycephala* has been identified. Wan et al. (2015) investigated the miRNAs from Intermuscular bone and connective tissue of 6-month-old *M. amblycephala*. A total of 218 conserved miRNA genes (belonging to 97 families) were identified and expressed, of these miRNAs, 44 conserved miRNA sequences exhibited significant expression differences between the two tissues. Yi et al. (2013) constructed two small RNA libraries and sequenced with the mixed pools of tested tissues (brain, pituitary, liver and muscle) from the big-size group and small-size group with relatively high and low growth rates of *M. amblycephala*. 347 conserved miRNAs (belonging to

Abbreviations: nt, Nucleotides; UTRs, 3′ untranslated regions; RISC, RNA-induced silencing complex; pri-miRNAs, Primary miRNAs; pre-miRNAs, precursor miRNAs; LPS, Lipopolysaccharide; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; NCBI, National Center for Biotechnology Information; MFE, Minimum folding free energies

* Corresponding author.
E-mail address: huangyong1979111@163.com (Y. Huang).

https://doi.org/10.1016/j.gene.2018.02.045
Received 8 September 2017; Received in revised form 27 January 2018; Accepted 15 February 2018
Available online 18 February 2018
0378-1119/

123 families) and 22 novel miRNAs were identified. Zhang et al. (2014) performed the deep sequencing to identify the in hepatic miRNA from blunt snout bream fed normal-fat and high-fat diets, and 202 (193 conserved and 9 novel) miRNAs were found, of which 12 were differentially expressed between the normal-fat and high-fat diet groups. Sun et al. (2016) identified some miRNA from the gill samples of control or ammonia-exposed *M. amblycephala*, which related to the immune regulation. Recently, Yuhong et al. (2016) identified immune-related miRNAs form two small RNA libraries which were from immune tissues with or without LPS stimulation in *M. amblycephala*. A total of 324 miRNAs, including 218 conserved miRNAs and 106 putative novel miRNAs were identified. Herein, we employed Illumina 2500 sequencing platform to extend the repertoire of *M. amblycephala* miRNAs for future functional research.

In present study, we first constructed a small RNA library using pooled RNA sample isolated from four types of tissues, and find more miRNAs than one tissue in *M. amblycephala.* The identification and characterization of miRNAs can initiate further study on miRNA regulatory mechanisms, and GO and KEGG pathway analysis help toward a better understanding of the important roles of miRNAs in physiological processes in *M. amblycephala*.

2. Materials and methods

2.1. Animal materials

M. amblycephala fishes from three different individuals were approximately two years old (300–500 g) were purchased from Taihu fishery company, Jingzhou, Hubei province. The four tissues (kidney, spleen, muscle and liver) were sample from freshly fish, respectively. These tissues were frozen immediately in liquid nitrogen and stored in −80 °C freezers until use. All handling of fish was conducted under the guidance of the Care and Use of Laboratory Animals in China. This research was approved by the Committee on the Ethics of Animal Experiments of Henan University of Science and Technology, Luoyang, China.

2.2. Small RNA library construction and sequencing

The four tissues were collected in triplicate and pooled as one sample for RNA extraction with Trizol reagent (Invitrogen, USA) following the manufacturer's instructions. Any potential genomic DNA contamination was removed with RNase-free DNase I (Takara, China). *M. amblycephala* small RNA library was constructed according to the previously described methods (Huang et al., 2017). In brief, small RNA fractions with the length of 18–30 nt were purified by electrophoretic separation on a 15% denaturing polyacrylamide gel. The small RNAs were ligated with 5′- and 3′-RNA adapters using T4 RNA ligase (Takara, China). The small RNAs with adapter were subsequently reverse-transcribed into cDNA following the Illumina protocol. The small RNA library was sequenced with Illumina HiSeq 2500 platform serviced by the Beijing Biomarker Technologies Co., Ltd. All sequencing data has been deposited in the NCBI with accession number SRP083094.

2.3. Deep sequencing analysis and miRNA identification

The sequencing reads were processed to remove adaptors, and cleaned by Q30 value. Reads with > 20% bases < Q30, and N base > 10% were filtered. The remaining reads were further filtered if read length was > 30 bp (base pairs) or < 18 bp. large amounts of clean small RNAs were analyzed by BlastX against the zebrafish genome. The matched sequences were blasted against the Rfam and NCBI GenBank non-coding RNA database to filter out rRNAs, tRNAs, snRNAs, snoRNAs and other ncRNA sequences by Bowtie (Langmead et al., 2009) for annotation. The unannotated sequences were analyzed through blasting the miRbase 21.0 database to identify the conserved miRNAs (with no more than 2 mismatches). Then the novel miRNAs were predicted using miRDeep2 (Friedlander et al., 2012) from remaining small RNAs. The miRDeep2 output figures were manually checked by the existence of miRNA star sequence and the precision of editing of mature sequences was as described previously. Furthermore, their hairpin structures were then analyzed using RNAfold software. Finally, the base bias composition on the first position of identified miRNAs with certain length and on each position of all miRNAs was analyzed, respectively.

2.4. Validation of miRNAs and expression analysis using stem-loop qRT-PCR

Randomly selected six miRNAs (four conserved and two novel miRNAs) were used to validate their expression by stem-loop qRT-PCR. Total RNA from heart, liver, kidney, muscle, spleen and intestine of *M. amblycephala* was isolated using Trizol reagent following the recommendations of the manufacturer. *M. amblycephala* cDNA was synthesized from respective total RNA with specific stem-loop primers designed according to previous study (Gong et al., 2017; Yang et al., 2014). All stem-loop RT-PCR primers are listed in **Table S1**. Briefly, 1.5 μg of total RNA were reversely transcribed to cDNA using M-MLV reverse transcriptase with stem-loop RT primers (Takara, Dalian, China). qPCR was carried out using the SYBR Green PCR Core Reagents Kit (Applied Biosystems, USA) on an ABI Prism 7900 Sequence Detection System (Applied Biosystems, USA) according to the manufacturer's instructions. Relative expression levels of the novel miRNAs were measured in terms of threshold cycle value (Ct) and were normalized to 5S rRNA using the equation $2^{-\triangle\triangle Ct}$, in which $\Delta Ct = Ct_{miRNA} - Ct_{5S}$. All qPCR was performed in triplicates.

2.5. Prediction of potential miRNA target genes and functional analysis

To better explore the potential functions of the miRNAs, target genes were predicted with miRnada software (http://www.microrna.org/microrna/home.do). The parameters of miRanda were set as score > 50 and free energy threshold ≤ 20 kcal/mol. As the genome references of *M. amblycephala* was not available, the sequences of zebrafish genome sequences were selected to predict miRNA target genes. Furthermore, for obtaining a better miRNA target biological functions analysis, GO analysis of targeted mRNA transcripts were generated using the database for annotation, and the gene numbers of each GO term were calculated. Subsequently, the main pathways of biochemical and signal transduction significantly associated with the predicted target genes of the miRNAs were determined. The analysis of GO categories and pathway were based primarily on the DAVID Bioinformatics Resources 6.7 (NIAID/NIH).

3. Results and discussion

3.1. Sequencing and analysis of small RNAs in M. amblycephala

To increase the coverage of *M. amblycephala* miRNAs, a small RNA library was constructed from the pooled RNA samples isolated from four different tissues and sequenced. A total of 14,337,875 raw reads were generated by deep sequencing technology. After the removal of the 5′ and 3′ adaptors, adaptor–adaptor ligations, low quality tags containing 'N' reads, length < 18 or > 30 nt reads, a total of 12,526,813 high quality reads (clean reads) were obtained. To assess the reads quality, the length distribution based on both total abundance and distinct reads were analyzed (Fig. 1). The majority of reads were in the range between 20 nt and 23 nt in length, with 21 nt and 22 nt as the two major size groups, which is the typical size range for Dicer-derived product. The length distribution trend is also consistent with the length distribution trend of miRNAs in fishes, and the result was same as that of *Gobiocypris rarus* (Hong et al., 2016), *Misgurnus anguillicaudatus* (Huang et al., 2016), *Ctenopharyngodon idella* (Xu et al., 2014), *Cyprinus*