Computational Biology and Chemistry 68 (2017) 92–100

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Research Article

Genome-wide identification and characterization of conserved and novel microRNAs in grass carp (*Ctenopharyngodon idella*) by deep sequencing

Wangbao Gong[a], Yong Huang[b,*], Jun Xie[a], Guangjun Wang[a], Deguang Yu[a], Xihong Sun[b]

[a] Key Laboratory of Tropical & Subtropical Fishery Resource Application & Cultivation, Ministry of Agriculture, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou, 510380, China
[b] College of Animal Science and Technology, Henan University of Science and Technology, Luoyang, 471003, China

ARTICLE INFO

Article history:
Received 10 October 2016
Received in revised form 14 January 2017
Accepted 26 February 2017
Available online 2 March 2017

Keywords:
Grass carp
miRNA
Deep sequencing
Targets

ABSTRACT

MicroRNAs (miRNAs) are post-transcriptional regulators which bind to target to regulate protein expression by repressing translation or promoting degradation of the target mRNA. Studies have shown that deep sequencing is a powerful tool for the identification of miRNAs, and it is believed that may more miRNAs remain to be discovered in grass carp. In the present study, a pool of equal amounts of RNA obtained from 8 different adult grass carp tissues (spleen, liver, muscle, kidney, skin, testis, gut and heart) was sequenced using deep sequencing technology. A total of 16.579.334 raw reads were yielded from the pooled small RNA library. Using bioinformatics analysis, we identified 160 conserved miRNAs and 18 novel miRNAs in grass carp. Randomly selected 6 conserved and 2 novel miRNAs were confirmed their expression by stem-loop qRT-PCR assay. Furthermore, the 1212 potential targets of these miRNAs were predicted using miRNA target prediction tool. GO and KEGG pathway enrichment analyses indicated relevant biological processes. Our study provides the first genome-wide investigation of miRNAs from 8 mixed tissues of grass carp, and the data obtained expand the known grass carp miRNA inventory and provide a basis for further understanding functions of grass carp miRNAs.

1. Introduction

MicroRNAs (miRNAs) are endogenous, small, non-coding RNAs, comprising approximately ~22 nucleotides (nts) (Bartel 2004; He and Hannon 2004). It modulates the expression of target genes through transcriptional or posttranscriptional regulation by binding target mRNAs for either mRNA cleavage or the inhibition of mRNA translation (Ambros 2004; Budak and Akpinar, 2015; Filipowicz et al., 2005). MiRNAs are derived from transcripts that are processed in the nucleus to form miRNA precursors (pre-miRNAs) with hairpins of approximately 60–80 nts (Cullen 2004; Han et al., 2006). The pre-miRNAs translocate to the cytoplasm, where the pre-miRNAs are further cleaved by the RNAse III enzyme Dicer into duplexes of about 22 nts. One of the RNA strands is loaded by Dicer into an RNA-induced silencing complex (RISC) (Lee et al., 2003). And then mature miRNAs or miRNA star (named miRNA*) guide the binding of RISC to mRNA targets, forcing mRNA degradation and/or translational inhibition (Brennecke et al., 2005; Budak et al., 2016). Moreover, miRNAs can also activate the translation of target mRNAs (Vasudevan et al., 2007). Since the identification and characterization of the first miRNAs, lin-4 and let-7, as regulators of developmental timing in *Caenorhabditis elegans* (Lee et al., 1993; Reinhart et al., 2000), miRNAs have been identified in diverse organisms including animals, plants, and unicellular eukaryotes. It is estimated that miRNA genes represent approximately ~1% of the known eukaryotic genomes and around 10–30% of genes might be regulated by miRNAs (Bartel 2004; Cui et al., 2006). Accumulating evidence showed that miRNAs are implicated in a wide variety of cellular processes, including proliferation, organ development, cell differentiation, apoptosis, autophagy, immune response, tumorigenesis and metastasis (Budak et al., 2015; Carrington and Ambros, 2003; Clark et al., 2014; Huang et al., 2011; Williams 2008).

Due to the importance of miRNA in regulation of gene expression, extensive investigation aiming at the discovery of new miRNAs in several fish species has been carried out and the number of known miRNAs is increasing with over 3300 fish miRNA deposited sequences in the miRBase database release 21.0 in June

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

http://dx.doi.org/10.1016/j.compbiolchem.2017.02.010
1476-9271/

2014 (http://www.mirbase.org/). The identification and characterization of miRNA have become an essential firstly step for understanding functions of miRNAs. Traditionally, miRNA is mainly identified by two kinds of approaches: (i) experimental methods such as direct cloning sequencing from a small RNA library; (ii) computational identification based on RNA secondary structure predictions and sequence conservation across species. However, this methodology has limited application in discovering low abundant miRNAs. Currently, deep sequencing technologies are robust tools for the identification of poorly expressed, species-specific, non-conserved and novel miRNAs in plants or animals lacking whole genome sequence information (Hong et al., 2016; Kurtoglu et al., 2014; Lucas and Budak, 2012; Lv et al., 2012; Ouyang et al., 2015; Xia et al., 2011). Grass carp is one of the most important cultivated species in freshwater aquaculture and its farming results in the largest yield for a single species worldwide in China. Considering the importance of the fish species for protein foods and health products, understanding the regulation of growth and development is very important.

Recently, several more attempts have been made to identify miRNAs from grass carp; for instance, 493 miRNAs have been identified from grass carp kidney tissue, 177 miRNAs from grass carp kidney tissue with *A. hydrophila* AH10 infection and 185 miRNAs were identified associate with motile aeromonad septicemia in grass carp (Gan et al., 2016; Xu et al., 2015; Xu et al., 2016). However, to our knowledge, no research has been found to identify miRNA genes from different grass carp tissues. In this study, using the deep sequencing approach, we performed the extensive miRNA sequencing from grass carp 8 mixed tissues (spleen, liver, muscle, kidney, skin, testis, gut and heart) to identify conserved and novel miRNAs. Our findings in this study will give fundamental information to aid understanding of miRNA populations in this species. In addition, our analysis also adds to the growing database of miRNAs, and advances the understanding of the important roles of miRNAs in Cyprinids, especially in the grass carp.

2. Materials and methods

2.1. Ethics statement

Grass carp (1.500–2.000 g body weight; 30–50 cm body length) were purchased from Taihu fishery company, Jingzhou, Hubei province. The experiments were performed in accordance with the International Guiding Principles for Biomedical Research Involving Animals as promulgated by the Society for the Study of Reproduction. The experimental procedures were also approved by the Institutional Animal Care and Use Committee (IACUC) of Henan University Science and Technology (Luoyang, China). All efforts were made to minimize suffering of the animals.

2.2. Fish tissue collections and RNA extraction

The tissue samples of adult grass carp were collected including spleen, liver, muscle, kidney, skin, testis, gut and heart, and then these tissues quickly frozen in liquid nitrogen and stored at −80 °C. The four tissues were then pooled as one sample for RNA extraction with Trizol (Takara, Japan) according to manufacturer instructions. The total RNA was quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and assessed for quality control using 2100 Bioanalyzer (Agilent, USA). All RNA samples had RNA integrity >7.

2.3. Construction of small RNA library and deep sequencing

The small RNA library construction and deep sequencing were performed in Biomarker Technologies (Beijing, China), using the TruSeq Small RNA Sample Preparation Kit and sequencing Kit (Illumina, USA) according to the manufacturer's protocol. Briefly, a small RNA fraction (from 18 to 30 nt) was isolated by polyacrylamide gel electrophoresis, and ligated with proprietary adaptors using T4 RNA ligase. The RNA was converted to single-stranded cDNA using Superscript II reverse transcriptase (Invitrogen, USA) and subsequently amplified in 15 cycles using Illumina's small RNA primer set and Phusion polymerase (New England Lab, USA). The PCR products were size-fractionated and recovered for sequencing with Illumina platform according to the manufacturer's instructions. The original image data obtained by the deep sequencing analyzer were automatically transformed into raw reads using base calling.

2.4. Bioinformatics analysis of grass carp small RNA sequences

The generated initial raw reads were processed by a data-cleaning pipeline to eliminate the reads of low quality, reads containing poly (A) stretches, reads less than 18 nt, 3′ adaptor nulls, 5′ adaptor contaminants and insert nulls. Then, the clean reads were mapped onto the zebrafish genome (NCBI assembly: GRCz10) to perform distribution analysis, and miRNAs were predicted by using SOAP software (Li et al., 2009). Remaining reads were then compared against the ncRNAs (rRNAs, tRNAs, snRNAs, and snoRNA) deposited in the NCBI GenBank database, the Ensembl and Rfam database using BLAST to annotate the small RNA sequences. Furthermore, reads overlapping with the exons of protein-coding genes were also excluded to avoid mRNAs contamination. Subsequently, the clean reads were aligned against the known miRNA precursors in the miRBase 21.0 database to identify the conserved miRNAs. The novel miRNAs were predicted by exploring the secondary structure, the Dicer cleavage site and the minimum free energy of the unannotated small RNA reads, which could be mapped to reference genome sequences by using miRDeep2 software (Friedlander et al., 2012). The criteria for selection of grass carp miRNA precursors were: (1) folding of miRNA precursor sequences into an appropriate secondary stem-loop structure; (2) one arm of a precursor stem portion contains mature miRNA whereas its opposite arm harbors miRNA strar sequence; (3) no loop or break and less than six nucleotide mismatches between mature miRNA and its opposite miRNA star sequence; (4) higher negative MFEI values of precursors than other types of RNAs; and (5) miRNA/miRNA star sequences form a duplex with two nucleotides 3′ overhang in hairpin secondary structure.

2.5. Validation and expression analysis of miRNAs by stem-loop qRT-PCR

To confirm the presence and the expression levels of the identified mature miRNAs, 8 miRNAs were chosen for stem-loop qRT-PCR analysis. This method is generally more sensitive than microarray and northern blot analysis and allows for the absolute quantification of transcript abundance. Total RNA was isolated using RNAiso Plus reagent (Takara, Japan) according to the manufacturer's protocol and then treated with RNase-free DNase I (Takara, Japan) to remove genomic DNA. RNA quality and quantity were determined using a NanoDropND-1000 spectrophotometer (NanoDrop Technologies, USA). Primers for stem-loop qRT-PCR were designed according to descriptions in prior studies (Table S1) (Huang et al., 2016). Briefly, 1 μg of total RNA of each sample was used to generate cDNA using specific stem–loop RT primers, and the product was stored at −20 °C for future use. The qPCR was performed using Power SYBR-Green PCR master mix kit (Takara, Japan) on the BIORAD iQ5 Real-time PCR Detection System. The 15 μL PCR included 1 μL RT product, 2 × SYBR-Green PCR master mix, 2 μM forward primers and 2 μM reverse primer. The