A bioinformatics-based update on microRNAs and their targets in rainbow trout (Oncorhynchus mykiss)

Liandong Yang a,b, Shunping He a,*

a The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei, 430072, PR China
b University of Chinese Academy of Sciences, Beijing 100049, PR China

Abstract

MicroRNAs (miRNAs) participate in various vitally biological processes via controlling target genes activity and thousands of miRNAs have been identified in many species to date, including 18,698 known animal miRNA in miRBase. However, there are only limited studies reported in rainbow trout (Oncorhynchus mykiss) especially via the computational-based approaches. In present study, we systematically investigated the miRNAs in rainbow trout using a well-developed comparative genome-based homologue search. A total of 196 potential miRNAs, belonging to 124 miRNA families, were identified, most of which were firstly reported in rainbow trout. The length of miRNAs ranged from 17 to 24 nt with an average of 20 nt while the length of their precursors varied from 47 to 152 nt with an average of 85 nt. The identified miRNAs were not evenly distributed in each miRNA family, with only one member per family for a majority, and multiple members were also identified for several families. Nucleotide U was dominant in the pre-miRNAs with a percentage of 30.04%. The rainbow trout pre-miRNAs had relatively high negative minimal folding free energy (MFE) and adjusted MFE (AMFE). Not only the mature miRNAs but their precursor sequences are conserved among the living organisms. About 2466 O. mykiss genes were predicted as potential targets for 189 miRNAs. Gene Ontology (GO) analysis showed that nearly 2093, 2107, and 2081 target genes are involved in cellular component, molecular function, and biological processes respectively. KEGG pathway enrichment analysis illuminated that these miRNAs targets might regulate 105 metabolic pathways, including those of purine metabolism, nitrogen metabolism, and oxidative phosphorylation. This study has provided an update on rainbow trout miRNAs and their targets, which represents a foundation for future studies.

Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.

1. Introduction

MicroRNAs (miRNAs) are a new class of endogenously small non-coding single-stranded RNAs, which are involved in various developmental and physiological processes by regulating target genes expression by translational repression or transcript destabilization (Bartel, 2009; Carthew and Sontheimer, 2009; Kim, 2005). MiRNAs make up one of the largest classes of gene regulatory elements in animal genome (Roux et al., 2012), and large quantity of miRNAs reside in the genomes of animals, plants, viruses, and algae pervasively (Avesson et al., 2012; Berezikov et al., 2005; Colaiacovo et al., 2010; Lagos-Quintana et al., 2003; Reese et al., 2010). Therefore, since the first miRNA (lin-4) was discovered in Caenorhabditis elegans, a large number of miRNAs have been identified in the past few years through direct cloning, sequencing and northern blot, high-throughput sequencing, and even homologue search. MiRNAs are primarily produced from a longer primary miRNAs (pri-miRNAs), which are then processed to generate the characteristic stem-loop RNA precursors (pre-miRNAs) by the ribonuclease III (RNase III) enzyme Drosha within the nucleus. The pre-miRNAs are then exported to the cytoplasm and further processed to form the mature miRNAs by the RNase III Dicer (Lee et al., 2002, 2004). Finally, the mature miRNAs are assembled into the miRNA-induced silencing complexes (miRISCs) to exert their functions and induce their target miRNAs translational repression or mRNA cleavage in association with Argonaute (AGO) proteins (Krol et al., 2010). However, although the biogenesis and functions of miRNAs have been focused on for the past two decades, the high-throughput identification of miRNAs has only been performed recently with the development of high-throughput sequencing technology and the bioinformatic methods.

During the past few years, miRNAs have been identified by direct cloning, sequencing and northern blot analyses (Allen et al., 2004; Yu et al., 2008). However, the capability of these methods to detect miRNAs is too low-throughput and random. Since most miRNAs are evolutionarily highly-conserved among animals and plants (Altvia et al., 2005;
Liu et al., 2008; Reinhart et al., 2002), it is a powerful approach to predict new miRNA homologues by comparative genomics based on previously known miRNA sequences (Weber, 2005; Zhang et al., 2005). And large numbers of miRNAs have been successfully identified through comparative genomic computational strategies in a wide range of animals and plants (Bonnet et al., 2004; Lim et al., 2003). However, as for rainbow trout, regarded as the “aquatic lab-rat” and the most-widely cultivated cold freshwater fish species in the world (Palti et al., 2011), only limited researches have been performed experimentally (Ma et al., 2012; Ramachandra et al., 2008; Salem et al., 2010) and no studies about miRNAs have been reported via the computational approaches considering the large numbers of EST (expression sequence tags) available in the public database.

In this study, we employed a well-defined comparative genomic-based homologue search to systematically investigate miRNAs and their target genes in rainbow trout using the publicly available EST in the NCBI Genbank database and the previously known animal miRNAs in the miRBase. A total of 196 potential miRNAs were identified, which belong to 124 miRNA families and most of which were firstly reported in rainbow trout. We then extensively characterized a lot of features of these pre-miRNA sequences, including families, original locations in the gene, length, nucleotide composition, MFE, and AMFE. And the conservation of miRNAs and their precursors were also investigated. Finally, we predicted these miRNAs’ targets and made a deep analysis of targets functions through GO and the KEGG pathway database. These results will no doubt provide useful resource for the miRNAs research in rainbow trout in the future.

2. Materials and methods

2.1. Data collection

All currently available animal microRNAs were defined as a reference miRNA sequences for identifying potential conserved miRNAs in rainbow trout and were downloaded from the miRbase database (http://www.mirbase.org/). Release 19: August 2012) (Kozomara and Griffiths-Jones, 2011). There are a total of 18,698 known animal miRNA in miRBase that were derived from 97 animal species.

In order to identify more potential miRNAs, the rainbow trout expressed sequence tags (ESTs), CDNs, and miRNAs were downloaded from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/NCBI) and were put together as source sequences of potential pre-miRNAs. Additionally, all the EST sequences were used for predicting rainbow trout miRNA targets. To date, a total of 287,967 ESTs can be obtained for rainbow trout from the NCBI EST database (dbEST release 120,701, July 1, 2012; http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) (Boguski MS and Tolstoshev, 1993).

The nonredundant (nr) protein sequence database was downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/) and was used to remove the potential protein coding sequences. The GO and the KEGG databases which were used for analyzing the functions classification and pathway enrichment of their target were downloaded from the Gene Ontology web-site (http://www.geneontology.org/GO.downloads.shtml) (Ashburner et al., 2000) and the KEGG web-site (http://www.kegg.jp/kegg/download/) (Kanehisa and Goto, 2000).

2.2. Computational prediction of conserved microRNA genes in rainbow trout

In this study, we used a method of comparative genome-based homologous search to identify rainbow trout conserved microRNA genes (Fig. 1). We combined the BLASTN (Altschul et al., 1990) and water (Smith and Waterman, 1981) to identify the homologous sequences between the known animal miRNAs and the sequences in rainbow trout. The mature sequences of all currently known animal miRNAs were subjected to a BLASTN search to all of the rainbow trout EST sequences.

Adjusted BLASTN parameter settings were as follows: 1) the E-value was set at 10; 2) –F F; 3) the default word-match size was set at seven; and 4) –m 8. Although BLASTN is very fast, it is difficult to detect deletions, insertions and gaps within so short miRNA sequences. Therefore, following the initial search using blastn, we collected the homolog sequences and then employed water to detect the insertions, deletions, and gaps. The initial candidate rainbow trout miRNA genes were retained with no more than 3 nt substitutions between known miRNAs and the homolog sequences. Then we extracted the 100 nts sequences upstream and downstream from the matched boundaries from the rainbow trout EST and removed the protein-coding genes using BLASTX with default parameter (Altschul et al., 1997) against the nonredundant (nr) protein databases from NCBI. Finally, the hairpin structures for the remaining candidate’s homologous sequences were generated using a publicly available program, RNAfold (http://www.tbi.univie.ac.at/~ivo/RNA/index.html) (Hofacker, 2003; Zuker and Stiegler, 1981). The secondary structures, which have the lowest free energy (MFE, ΔG) ≤ – 18 kcal/mol, were selected for manual inspection.

MiRNA homolog sequences in rainbow trout were considered to be potential miRNAs when they fulfill all of the criteria below: (1) there were no more than three nucleotide substitutions between the known mature miRNA sequence and the homolog sequence from ESTs; (2) the minimum length of the pre-miRNA was 45 nt; (3) the pre-miRNA could fold into an appropriate stem-loop hairpin secondary structure; (4) there were no more than six mismatches in the miRNA/ miRNA∗ duplex; (5) there were no loops or breaks in the miRNA/ miRNA∗ sequences; and (6) the mature miRNA should be localized in one arm of the stem-loop structure.

2.3. Conservation analysis

A total of 18,698 known mature miRNAs derived from 97 metazoa animals was get from miRBase (release 19) to create a local blast database. All identified 196 rainbow trout miRNAs were used to search to the local database using BLASTN to find the high homology miRNAs with an e-value cutoff of 10⁻⁵. We counted the distribution of conserved rainbow trout miRNAs across some representative species. Then we also analyzed the rainbow trout pre-miRNAs (omny-mir-126) conservation with their orthologues in human, mouse, rat and zebrafish by the publically available WebLogo: a sequence logo generator (Crooks et al., 2004).

2.4. Prediction of miRNA target genes via computational analysis

To predict the potential targets of miRNA in rainbow trout, all the ESTs were obtained from NCBI database. Because the 3’UTRs are the primary target regions of miRNAs (Betel et al., 2008), we first identified the approximate 3’UTRs for all the rainbow trout transcripts. We used BLASTX to identify the homologous genes between gene models of rainbow trout and protein sequences of zebrafish. All the 42,157 zebrafish protein sequences were obtained from Ensembl (http://www.ensembl.org/) and rainbow trout assembled unigenes of 117,120 sequences were downloaded from NCBI. Then the unigenes were searched against the protein sequences by BLASTX with an e-value of 10⁻⁵. If there were more than one gene matched, we only picked the one with maximal match as the homolog in rainbow trout. Then we further filtered the bad matches with cutoff of at least 60% to the shorter sequence of either query or subject. We used the method from Ma et al. (2012) to get the 5’UTRs, coding regions and 3’UTRs for all the matched genes in rainbow trout unigenes. Finally, we fetched 10,568 5’UTRs, 12,012 coding regions and 10,568 3’UTRs for all matched genes.

Both the 3’UTRs, 5’UTRs and CDS regions were used to predict targets by two widely used miRNA target prediction algorithms, miRanda (Enright et al., 2004) (http://www.microrna.org/microrna/home.do) and PITA (Kertesz et al., 2007) (http://genie.weizmann.ac.il/pubs/
The criteria of miRanda for potential target sites were as follows: 1) $S$, the sum of single-residue-pair match scores over the alignment, $\geq 140$ and 2) $\Delta G$, the free energy of duplex formation, $\leq -17$ kcal/mol. The other parameters for miRanda and parameters for PITA were all default (Betel et al., 2010). And in order to reduce the total number of false positive target results, only the genes predicted by both programs were regarded as potential target genes.

### 2.5. Function analyses of target genes by GO and KEGG pathway

In order to further understand the miRNA targets function and classification as well as their metabolic regulatory networks, function analysis of miRNA targets based on the GO database and the KEGG pathway database were performed using the Blast2GO suit (Conesa et al., 2005; Gotz et al., 2008). First, the miRNA targets sequences were searched against the NCBI nonredundant (nr) protein database and the UniProtKB/Swiss-prot database using BLASTX (Altschul et al., 1997) with an $E$-value cutoff of $1 \times 10^{-5}$. Target gene names were assigned to each sequence according to their best BLAST hit. Then we used the Blast2GO suit to annotate the function category of target genes using the function for the mapping of Gene Ontology (GO) terms with following parameters: with an $E$-value of $1 \times 10^{-5}$, annotation cut-off $>55$ and a GO weight $>5$. And finally, the metabolic pathway analysis by KEGG was also performed.

### 3. Results and discussion

#### 3.1. Identifying potential miRNAs in rainbow trout

Although lots of miRNAs have been identified in some species using experiment approaches, it is believed that computational prediction is a highly rapid and effective way to explore the potential miRNAs in species whose complete genomic sequence is not available (Zhang et al., 2005). In this study, through comparative genomics search and bioinformatic tools the rainbow trout pre-miRNA sequences were predicted. As described in Materials and methods section, a total of 196 potential $O.$ mykiss miRNAs in 191 pre-miRNAs were predicted which belong to 124 families (Fig. 2); (omy-let-7, mir-7, 17, 21, 22, 26, 27, 33, 81, 92, 126, 128, 133, 137, 138, 139, 142, 143, 153, 159, 190, 234, 281, 301, 308, 345, 354, 371, 378, 383, 483, 499, 541, 574, 632, 669, 725, 750, 932, 961, 969, 988, 1010, 1178, 1224, 1234, 1281, 1287, 1434, 1584, 1599, 1603, 1638, 1643, 1645, 1692, 1695, 1701, 1770, 1777, 1814, 2127, 2189, 2230, 2284, 2293, 2304, 2325, 2345, 2351, 2361, 2390, 2391, 2420, 2429, 2437, 2441, 2488, 2491, 2493, 2709, 2790, 2881, 2888, 2982, 3138, 3141, 3529, 3600, 3737, 3792, 3901, 4003, 4039, 4042, 4099, 4099, 4102, 4114, 4111, 4114, 4128, 4150, 4154, 4166, 4170, 4184, 4186, 4194, 4198, 4209, 4520, 4805, 4868, 4931, 4951, 4961, 5393, 5594, 6006, 6056, 6528 and 6631). Detail information on predicted $O.$ mykiss miRNAs, including families, names, mature sequences, lengths, original locations in the gene, and other aspects, was listed in supplementary files (Table S1 and Fig. S1).

MiRNAs can be identified by direct cloning (Ramachandra et al., 2008; Sunkar et al., 2005), high-throughput sequencing (Chi et al., 2011; Soares et al., 2009) and computational homology-based methods (Jones-Rhoades and Bartel, 2004; Li et al., 2010). And recently, a lot of miRNAs have been identified by high-throughput sequencing in various animals and plants for its high performance, including rainbow trout (Ma et al., 2012). However, miRNAs were expressed limiting to specific developmental stages and tissues, so the discovery of miRNAs obtained in few tissues will certainly miss lots of miRNAs expressed in other tissues or developmental stages. As a result, computational-based approaches through the homology search at the EST level were used to identity additional miRNAs, which were often missed by methods such as direct-cloning and sequencing (Brameier, 2010; Weber, 2005; Zhang et al., 2005). Here in this work, amongst the 196 identified miRNAs in rainbow trout, only 29 miRNAs have been reported in previous researches (Table S1) (Ma et al., 2012; Ramachandra et al., 2008; Salem...
et al., 2010), and the remaining 167 miRNAs were newly identified rainbow trout miRNAs. There may be two reasons for this. The first one is that previous reports have focused on limited tissues and developmental stages. Because most miRNAs are tissue- and developmental stage-specific expressed, there will be many miRNAs to be missed out in limited sampling. The second reason is that integration of EST data will significantly facilitate novel miRNAs discovery (Krzyzanowski et al., 2011).

### 3.2. Characterization of miRNAs in rainbow trout

The 196 identified miRNAs belong to 124 families with an average of about 1.6 miRNA members per family. However, the identified miRNAs were not evenly distributed in each miRNA family. For a majority of the miRNA families identified, there was only one member. Whereas, for some miRNA families (let-7, mir 133, 1287, 1603, 1777, 2391, and 2881) that each contained 3 members. Interestingly, based on dre-miR-2189, we have identified 11 mir 2189 members in rainbow trout. However, there is currently only one mir-2189 from zebrafish in the miRBase, which has been detected by deep sequencing (Soares et al., 2009). It suggests that *O. mykiss* has more dependence on this family.

Among the 196 identified potential miRNAs, 106 (54.1%) were located within the 3′ arm of the secondary hairpin stem-loop structure of pre-miRNAs and the other 90 miRNAs (45.9%) were located within the 5′ arm. Interestingly, however, the number of miRNAs obtained from the plus strand was the same as that identified from the minus strand. The length of these 196 mature miRNAs ranged from 17 nt to 24 nt long with an average of 20.6 ± 0.12 nucleotides. The majority of miRNAs have lengths of approximately 20 and 22 nt, which account for 30.6% and 19.4% of the total population, respectively, followed by 21 nt (17.3%), 19 nt (11.7%), 18 nt (8.7%), 24 nt (7.7%), 17 nt (2.6%) and 23 nt (2%) (Fig. 3A). The length of their precursor sequences was also varied from miRNA to miRNA, which ranged from 47 nt to 152 nt in length with an average of 84.8 ± 1.6 nt (Fig. 3B). However, a significantly large percentage of pre-miRNAs (94.9%) have lengths between 50 and 120 nt in length. There are only 10 miRNAs (5.1%) with less than 50 nt and more than 120 nt in length.

The percentage composition of the four nucleotides (A, U, G, and C) in the precursor sequences was not evenly distributed (Table 1). According to previous research, the nucleotide uracil (U) is the most presented in mature and precursor sequences both in plant and animal...
Fig. 3. Characterization of miRNAs in rainbow trout. The length distribution of mature miRNA (A) and pre-miRNA sequences (B). The distribution of the four nucleotides (C, D), MFE (E) and AMFE (F) of pre-miRNA sequences.
miRNAs (Wang et al., 2012). In this study, our results showed that the average of U content was also dominant nucleotide with a mean of 30.04 ± 7.44%, which is significantly higher than the content of other nucleotides (t-tests, \( P < 1 \times 10^{-5} \)) (Fig. 3C). A significant majority (79.1%) of precursor miRNAs contained more than 25% of the nucleotide U. As a result of forming three hydrogen bonds with each other between G and C, nucleotides G and C make a contribution to the formation and stabilization of the hairpin structure of pre-miRNAs. Commonly, the more GC content a sequence contains, the more stable its secondary structure will be (Wang et al., 2012). However, within the identified potential pre-miRNAs, the GC content (45.38 ± 10.67%) was significantly lower than the AU content (54.55 ± 10.79%) (t-tests, \( P < 1 \times 10^{-5} \)) (Fig. 3D), consistent with previous findings in Asiatic cotton (Wang et al., 2012). Meanwhile, the A/U and C/G ratios were also calculated, and the ratio of A/U and C/G were 0.87 and 0.81, respectively, which suggests that more of the nucleotide U and G existed in the precursor sequences.

Minimal folding free energy (MFE) is an important criterion for measuring the stability of the secondary hairpin structure of pre-miRNAs. Generally speaking, the lower the MFE value, the more stable the secondary structure of a RNA sequence (Zhang et al., 2006). In the present study, our results showed that the value of MFIs range from −18.1 kcal/mol to −74.7 kcal/mol with average value of −27.02 ± 8.19 kcal/mol (Fig. 3E). It has an issue for determining the stability of the secondary structure using MFE because different pre-miRNAs contains different number of nucleotides. Therefore, the adjusted minimal folding free energy (AMFE), which is the MFE of a pre-miRNA sequence that is 100 nt in length, was adopted. The AMFE of the 196 identified rainbow trout pre-miRNAs ranged from −15.47 to −78.63 kcal/mol with an average of −33.31 ± 10.48 kcal/mol (Fig. 3F).

### 3.3. Conservation analysis

The miRNAs always showed highly conserved nature among organisms (Wienholds and Plasterk, 2005). In this study, our results also showed that many rainbow trout miRNAs were conserved across different species (Fig. 4A). Totally, there were twenty-nine rainbow trout miRNAs conserved in other species (BLASTN e-value \( ≤ 10^{-5} \)), in which fifteen miRNAs were highly conserved across more than five different species. Notably, omy-miR-133a, omy-miR-7a and omy-miR-92a displayed highly conservation across Bilateria (Supplementary Fig. S2).

Besides the mature miRNA sequences, the rainbow trout pre-miRNA (omy-mir-126) conservation investigation across human (Homo sapiens), mouse (Mus musculus), rat (Rattus norvegicus) and zebrafish (Danio rerio) orthologues was also performed, which showed highly conserved nature for precursor sequences (Fig. 4B). These results demonstrated that not

---

**Table 1**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimal</th>
<th>Maximal</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFE (ΔG, −kcal/mol)</td>
<td>27.02</td>
<td>8.19</td>
<td>74.7</td>
<td>18.1</td>
</tr>
<tr>
<td>AMFE (−kcal/mol)</td>
<td>33.31</td>
<td>10.48</td>
<td>15.47</td>
<td>78.63</td>
</tr>
<tr>
<td>Length (nt)</td>
<td>84.83</td>
<td>22.33</td>
<td>47</td>
<td>152</td>
</tr>
<tr>
<td>(G+C)%</td>
<td>45.38</td>
<td>10.67</td>
<td>26.67</td>
<td>94.74</td>
</tr>
<tr>
<td>(U+A)%</td>
<td>54.55</td>
<td>10.79</td>
<td>1.05</td>
<td>73.33</td>
</tr>
<tr>
<td>A%</td>
<td>24.51</td>
<td>7.01</td>
<td>0.00</td>
<td>36.75</td>
</tr>
<tr>
<td>C%</td>
<td>19.46</td>
<td>7.01</td>
<td>1.28</td>
<td>59.09</td>
</tr>
<tr>
<td>G%</td>
<td>25.91</td>
<td>7.55</td>
<td>12.22</td>
<td>58.82</td>
</tr>
<tr>
<td>U%</td>
<td>30.04</td>
<td>7.44</td>
<td>1.05</td>
<td>48.72</td>
</tr>
<tr>
<td>A/U ratio</td>
<td>0.87</td>
<td>0.38</td>
<td>0.00</td>
<td>3.36</td>
</tr>
<tr>
<td>C/G ratio</td>
<td>0.81</td>
<td>0.36</td>
<td>0.05</td>
<td>2.60</td>
</tr>
</tbody>
</table>

---

**Fig. 4.** Conservation of Oncorhynchus mykiss miRNAs (A). The highly conserved miRNAs were showed in the X-axis, and the numbers of the conserved species were shown in the Y-axis. Conservation investigation of O. mykiss pre-miRNA (B). Alignment of pre-miRNAs (mir-126) with human, mouse, rat and zebrafish was generated by the WebLogo: a sequence logo generator. The mature miRNA sequences were marked in rectangular boxes.
3.4. Target prediction of miRNAs

Regulation exhibited by miRNAs is mainly resulted from the interaction between miRNAs and their target genes via perfect or near-perfect complementation (Brown and Sanseau, 2005; Sethupathy et al., 2006). Therefore, it is of prime importance to identify the miRNA target genes to understand the miRNA regulatory function. However, there is no certain accepted standard to identify the miRNA targets experimentally nowadays, so bioinformatics based approaches were effective and rapid strategy to predict miRNA targets (Alexiou et al., 2009; John et al., 2004; Lewis et al., 2005; Ma et al., 2012) and only the genes predicted by both programs were regarded as potential target genes.

Because the 3′UTRs of target genes are the mainly base-pairing region of miRNAs (Betel et al., 2008; Lewis et al., 2005; Linsley et al., 2007), the 3′UTRs, 5′UTRs and CDS of all unigenes of rainbow trout were first identified (see Materials and methods) and then screened for binding with the rainbow trout miRNAs respectively. Under the strict criteria described in Materials and methods section, a total of 2466 potential target genes were identified for the 196 miRNAs via searching against rainbow trout EST sequences with an average of 68.4 target genes per miRNA (Table 2). For most miRNAs (186), more than one potential target genes were predicted and there were only 7 miRNAs that did not have target genes (Supplementary Table S2.1). These results further suggested that one miRNA could have many target genes (Bentwich et al., 2005; Lewis et al., 2005). Several genes are targeted both in the 3′UTRs and the coding regions (CDS) (Supplementary Table S2.5) or the 5′UTRs and the coding regions (Supplementary Table S2.7) while no genes are targeted in both the 3′UTRs and 5′UTRs (Supplementary Table S2.6), which was consistent with the previous report on the egg-predominant miRNAs in rainbow trout (Ma et al., 2012).

Previous reports have shown that many miRNAs coordinate regulated multiple target genes (Grun et al., 2005; Ivanovska et al., 2008). In consistence with the well known modes of the interaction between miRNAs and their target genes, the diversified relationship include: 1) one single miRNA targets a single gene, 2) one single miRNA targets multiple genes, 3) multiple miRNAs target one single gene, and 4) multiple miRNAs target multiple genes, which suggested that the complicated regulatory networks of miRNAs should also exist in rainbow trout and opened the possibility of substantial functional redundancy among miRNAs.

3.5. GO function classification and KEGG pathway analysis of miRNA targets

Both Gene Ontology (GO) and KEGG pathway analyses are promising approaches to uncover the miRNA-targets regulatory network at a whole level (Ashburner and Bergman, 2005). In this present study, the 2466 rainbow trout miRNA target genes were annotated using Gene Ontology (GO) terms that were categorized into three classes including “cellular component” (C), “molecular function” (F), and “biological process” (P) (Ashburner et al., 2000). A total of 2093 target genes are involved in 736 different cellular component functions, 2107 target genes taken part in 1248 different molecular functions, and 2081 target genes participate in 3661 biological processes. The distribution of assigned GO terms was shown in Fig. 5. Approximately 86% of miRNA targets were located in cell and cell part, followed by those located in organelle (62%) and organelle part (42%) according to the GO terms. Strikingly and interestingly, the miRNA target genes involved in binding and catalytic were the most and second represented compared with other

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Maximal</th>
</tr>
</thead>
<tbody>
<tr>
<td>3′UTRs/5′UTRs/CDS</td>
<td>68.96</td>
<td>68.57</td>
<td>279</td>
</tr>
<tr>
<td>3′UTRs</td>
<td>2.20</td>
<td>3.18</td>
<td>17</td>
</tr>
<tr>
<td>5′UTRs</td>
<td>19.16</td>
<td>21.66</td>
<td>82</td>
</tr>
<tr>
<td>CDS</td>
<td>47.31</td>
<td>47.88</td>
<td>203</td>
</tr>
<tr>
<td>3′UTRs + CDS</td>
<td>0.18</td>
<td>0.56</td>
<td>4</td>
</tr>
<tr>
<td>3′UTRs + 5′UTRs</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5′UTRs + CDS</td>
<td>0.13</td>
<td>0.38</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 5. Gene Ontology (GO) categories and distribution of miRNA target genes in rainbow trout. The results are classified in three main categories: cellular component, molecular function, and biological process. The y-axis on the left indicates the percent of genes in a category. The y-axis on the right indicates the number of genes in a specific category.
relatively low represented molecular function. More than half overall target genes were participated in biological process including biological regulation, cellular process, and metabolic process (Supplementary Table S3). In addition, approximately 200 target genes are involved in immune system process. Immune response is the basic process of organisms, which can protect the organisms from pathogen invading. And accumulating evidences have suggested that regulation of the immune related genes against the invading pathogens by miRNAs is conserved among most target genes, which may be a common phenomenon of all the organisms (Bennasser et al., 2004; Hagen and Lai, 2008; Yu et al., 2006; Zhang et al., 2009).

Following GO analysis, KEGG was then used to do a pathway enrichment analysis of predicted miRNA target genes. A total of 735 miRNA target genes were found to be involved in 105 predicted KEGG metabolic pathways, which suggested that many pathways are involved in several miRNAs in rainbow trout. The number of sequences ranged from 1 to 60 (Supplementary Table S4). The top 10 pathways with the greatest number of involved target genes were shown in Table 3. The top 10 metabolic pathways were: purine metabolism (60), nitrogen metabolism (36), oxidative phosphorylation (32), pyrimidine metabolism (28), glutathione metabolism (20), alanine, aspartate and glutamate metabolism (18), phosphatidylinositol signaling system (15), pentose phosphate pathway (14).

4. Conclusions

In this work, we updated the information on rainbow trout miRNAs and their target genes using comparative genome-based homologous search. Totally, 196 candidate miRNAs were identified belonging to 124 miRNA families and most of which were firstly reported in rainbow trout. A total of 2466 potential targets were also predicted for the 189 miRNAs. These findings will be helpful for future studies in rainbow trout.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.09.060.

Conflict of interest

The authors declare that they have no conflict of interests.

Acknowledgments

This work was supported by the Major Research plan of the National Natural Science Foundation of China (Grant No. 91131014) and the State Key Program of National Natural Science of China (Grant No. 31130049).

Table 3: The top 10 pathways with highest sequence numbers.

<table>
<thead>
<tr>
<th>No.</th>
<th>Pathway</th>
<th>All target genes with pathway annotation</th>
<th>Pathway ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Purine metabolism</td>
<td>60 (8.16%)</td>
<td>map00230</td>
</tr>
<tr>
<td>2</td>
<td>Nitrogen metabolism</td>
<td>36 (4.90%)</td>
<td>map00910</td>
</tr>
<tr>
<td>3</td>
<td>Oxidative phosphorylation</td>
<td>32 (4.35%)</td>
<td>map00190</td>
</tr>
<tr>
<td>4</td>
<td>Pyrimidine metabolism</td>
<td>28 (3.81%)</td>
<td>map00240</td>
</tr>
<tr>
<td>5</td>
<td>Glutathione metabolism</td>
<td>30 (2.72%)</td>
<td>map00480</td>
</tr>
<tr>
<td>6</td>
<td>Alanine, aspartate and glutamate metabolism</td>
<td>18 (2.45%)</td>
<td>map00250</td>
</tr>
<tr>
<td>7</td>
<td>Glycolysis/gluconeogenesis</td>
<td>15 (2.04%)</td>
<td>map00010</td>
</tr>
<tr>
<td>8</td>
<td>Arachidonic acid metabolism</td>
<td>15 (2.04%)</td>
<td>map00590</td>
</tr>
<tr>
<td>9</td>
<td>Phosphatidylinositol signaling system</td>
<td>15 (2.04%)</td>
<td>map00407</td>
</tr>
<tr>
<td>10</td>
<td>Pentose phosphate pathway</td>
<td>14 (1.90%)</td>
<td>map00030</td>
</tr>
</tbody>
</table>

References


