Research paper

Bichir microRNA repertoire suggests a ray-finned fish affinity of Polypteriforme

Liandong Yang a,b, Zhaolei Zhang c,d, 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, People's Republic of China
b University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China
c Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
d Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

A R T I C L E   I N F O
Article history:
Received 23 December 2014
Received in revised form 2 April 2015
Accepted 21 April 2015
Available online 24 April 2015

Keywords:
Phylogeny
Polypteriformes
MicroRNAs
Bichir

A B S T R A C T
The phylogenetic position of Polypteriforme (bichirs) remains elusive, despite extensive research both on morphological and molecular datasets. Unfortunately, morphological cladistic analyses and molecular phylogenetic analyses had reached conflicting conclusions, as Polypteriformes were either grouped with lobe-finned fishes (Sarcopterygii) or ray-finned fishes (Actinopterygii), or even classified as their own group, the Brachiopterygii. In this study, we applied a third independent source of datasets, the presence versus absence of microRNAs, to re-investigate the phylogenetic relationship of bichirs. Through deep sequencing of small RNA library, we showed that bichirs should be grouped into ray-finned fishes rather than lobe-finned fishes. Phylogenetic analysis confirmed that bichirs were placed as the most basal member of the ray-finned fishes. Bichirs shared five unique microRNA families with teleosteans, which were not found in any other species investigated to date. Bichirs have also retained three of five microRNAs that were previously deemed to be lost exclusively in teleosteans. Furthermore, we report more than one hundred novel microRNAs that are unique to bichir. The identification of microRNAs in bichir provides overwhelming evidence for their affinity of ray-finned fishes. The sarcopterygian-like characteristics such as lobed fins, paired ventral lungs, and external gills in juveniles should not be considered as derived traits shared with those of sarcopterygians.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The phylogenetic position of polypteriform fish species (bichirs and reedfish) has received intensive attention over the past 150 years. Both morphological and molecular approaches had been applied to resolve this; however these two approaches had reached conflicting conclusions. The earliest morphological analyses conducted in the 1860s concluded that bichirs were sarcopterygians and they were thus grouped into the suborder Crossopterygii together with lungfish, coelacanth and fossil rhipidistians (Tree 1; Fig. 1A) (Huxley, 1861; Cope, 1871). Later morphological analyses classified bichirs as basal actinopterygians (Tree 3; Fig. 1C) (Goodrich, 1928; Gardiner, 1973; Schaeffer, 1973; Patterson, 1982; Lauder and Liem, 1983; Gardiner and Schaeffer, 1989; Nelson, 1994) or even a distinct subclass of Osteichthyes, the Brachiopterygii (Tree 2; Fig. 1B) (Jessen, 1973; Nelson, 1973; Jarvis, 1981; Bjerring, 1985). On the other hand, although the majority of molecular phylogenetic studies have favored the view that bichirs are the most basal member of the ray-finned fishes (Normark et al., 1991; Le et al., 1993; Noack et al., 1996; Venkatesh et al., 2001; Inoue et al., 2003), there are also studies that reported alternative phylogenetic trees (Rasmussen and Arnason, 1999; Rocco et al., 2004), albeit with low confidence. The difficulties to assess the phylogenetic position of bichirs are mainly that their fossil evidence can only extend back to the Eocene (Greenwood, 1974; Carroll, 1988) and that these fishes exhibit a lot of primitive (e.g. cartilaginous skeleton, ganoid scales, and the intestine with a spiral valve) and sarcopterygian-like characteristics. For example, their jaw structure is more similar to that of the tetrapods than that of the teleosts, and their pectoral fins with lobed base covered with scales are extremely similar to those of lobe-finned fishes. Bichirs also possess highly vascularized paired lungs and a pair of external gills in juveniles, which are not seen in any other ray-finned fishes, except in some species of lungfish (Robinson, 2011). These conflicting characteristics are so interesting that some researchers are still arguing that Polypteriformes are morphologically too distinct from actinopterygians to be classified as member of this.
group, and should be placed into sarcopterygians or at least more closely related to sarcopterygians than to actinopterygians (Jessen, 1973; Nelson, 1973; Jarvik, 1981; Bjerring, 1985).

Here, we attempt to reassess the problematic classification of bichirs through analysis of bichir’s microRNAs (miRNAs) repertoire, an independent source of molecular dataset. miRNA sequences, which have been successfully applied to many metazoan phylogenetic studies (Sperling et al., 2009; Heimberg et al., 2010; Sperling et al., 2010; Campbell et al., 2011; Philippe et al., 2011; Rota-Stabelli et al., 2011; Sperling et al., 2011; Wiegmann et al., 2011; Helm et al., 2012; Lyson et al., 2012; Pisani et al., 2012; Fromm et al., 2013; Peterson et al., 2013; Tarver et al., 2013), have five properties making them among the most reliable phylogenetic markers: (i) identification of novel miRNAs does not necessarily require fully sequenced genome sequence, (ii) new lineage-specific miRNA families are continually added to metazoan genomes through time which makes them ideal as phylogenetic markers, (iii) low levels of secondary miRNA loss, (iv) rarity of substitutions to the mature miRNA sequence, and (v) almost impossible scenario of convergent evolution of miRNAs (Wheeler et al., 2009; Tarver et al., 2013). Therefore, miRNAs should provide an alternative and valuable evidence for the evolutionary position of Polypteriformes. In the present study, we constructed a small RNA library from five organs (heart, brain, liver, kidney, and spleen) of bichir (Polypterus senegalus) and conducted high-depth next generation sequencing. We found that bichir shares five unique miRNA families with teleostei and none with sarcopterygians; this provides the strongest support for the Actinopterygii affinity for Polypteriformes.

2. Materials and methods

2.1. Small RNA isolation and deep sequencing

All animal experiments were performed in accordance with the ethics committee of Institute of Hydrobiology, Chinese Academy of Sciences. One adult Senegal bichir (P. senegalus) was purchased and dissected. To obtain the whole miRNA transcriptomes, we isolated total RNAs from five pooled organs (heart, brain, liver, kidney, and spleen) using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer’s protocol. The quality of RNA was examined by electrophoresis and the RNA 6000 Nano Assay (Agilent) following the manufacturer’s protocol.

After RNA quality control, small RNAs of 16–30 nt in length were isolated from total RNA by size fractionation with 15% TBE urea polyacrylamide gel (TBU) and then adaptors were ligated to the 5’ and 3’ terminals of these small RNAs. Afterwards, reverse transcription followed by PCR (RT-PCR) was used to create cDNA constructs, which were then purified on 6% Novex TBE PAGE gel and sequenced at the Beijing Biomarker Technologies Co., Ltd. using Illumina HiSeq (TM) 2000 platform.

2.2. Identification of known and novel miRNAs

Initial reads obtained from Illumina sequencing were first processed to remove lower quality reads, adaptor sequences and reads shorter than 16 nucleotides. The remaining clean reads were searched against the RFam database (ftp://ftp.sanger.ac.uk/pub/databases/Rfam/) and the NCBI GenBank noncoding RNA database (http://blast.ncbi.nlm.nih.gov/) to annotate rRNA, tRNA, snRNA, and snoRNA.

Our group had initiated sequencing of the bichir genome and has assembled a draft genome at 90x coverage (data not published). The processed clean small RNA reads were then analyzed by miRDeep2 software package (mapper and miRDeep2 analysis modules) (Friedlander et al., 2008, 2012) to predict microRNAs in bichir. We used all default parameters except that parameter g was set to −1 to allow all precursor sequences from automatic excision gearing to be analyzed. All mature sequences, star sequences, and precursor sequences above the threshold score produced by miRDeep2 were retained and regarded as putative miRNAs and used for further analysis to identify known and novel miRNAs.

Known miRNAs in bichir were annotated by identifying their homologous miRNAs in miRBase database (http://www.mirbase.org/, Release 20: June 2013) (Kozomara and Griffiths-Jones, 2011) using the following criteria: (1) seed region, nucleotides 2–7 must be identical; and (2) the remainder of the sequence alignment must contain no more than three mismatches. The putative miRNA produced by miRDeep2 analysis was regarded as a true bichir miRNA when it met both criteria and annotated as the bichir ortholog of the known miRNA gene in miRBase.

Those miRNAs produced by miRDeep2 analysis that did not meet above criteria were considered as novel miRNAs in bichir. To remove the fragment of tandem repeats or interspersed repeats, we searched these putative novel miRNA precursors against the draft genome assembly of bichir to exclude the multiple loci (>5), defined as E-value ≤ 1 × 10−6 (Andreasen et al., 2013). Finally, all remaining sequences were blasted against the nonredundant (nr) database from NCBI (http://blast.ncbi.nlm.nih.gov/) and the functional small RNA database ver. 3.4 (http://www.ncrna.org/frnadb/) to exclude other kinds of small RNA. The remaining putative novel miRNAs were confirmed manually using the following criteria: the reads should match both mature and star sequences perfectly and should show a consistent

![Fig. 1. The three possible phylogenetic relationships between bichirs, lobe-finned fishes, and ray-finned fishes.](image)
processing of the 5′ end of the mature sequences. The miRNAs meeting all these criteria were considered as novel miRNAs.

2.3. Phylogenetic analysis

A total of 190 miRNA families were coded as absence/presence for 15 chordate taxa with data generated during this study (P. senegalus), and taken from the dataset used in Heimberg et al. (Heimberg et al., 2010) (Branchiostoma, Ciona, Myxine, Petromyzon, Lampetra, shark, Danio, Tetraodon, Anolis, Taeniopterygia, Gallus, Ornithorhynchus, Monodelphis, Homo), which covered both embryos and different adult organs (Supp. File 3). All miRNAs were treated as Dollo type characters and assigned equal weight (Sperling and Peterson, 2009). Phylogenetic analyses were performed using branch and bound Maximum Parsimony searches (Farris, 1977) implemented in PAUP* v.4.0b10 (Swofford, 2002). Bayesian analysis was performed using BEAST 1.8 (Drummond et al., 2006, 2012) with parameter settings following Field et al. (2014).

3. Results

Small RNA transcriptomes from the heart, brain, liver, kidney, and spleen of one individual male were pooled and analyzed through high-throughput sequencing. In total, 20,111,426 raw reads were generated and 16,695,943 clean reads were retained after quality filtering, adaptor trimming, and length selection, which were then clustered into 777,856 unique sequences for subsequent analysis. The length distribution analysis showed that very high percentage (71.8%) of clean reads was 21–23 nucleotides (nt) long (Supp. File 1). After searching against the Rfam 11.0 database and NCBI noncoding RNA database (nr) for small RNA sequences, 2,904,214 reads of rRNA, tRNA, snRNA, snoRNA, and other non-coding RNAs were annotated and removed (Supp. File 2) and the remaining reads were retained for miRNA annotation.

We used the software miRDeep2 (Friedlander et al., 2008, 2012) to map the retained sequence reads to the bichir draft genome and to identify candidate bichir miRNAs. The detailed procedures are described in the Materials and methods section; in the end we derived a total of 598 high-confidence candidate miRNAs. We next conducted sequence homology search of these candidate miRNAs against known mature miRNAs in miRBase (Kozomara and Griffiths-Jones, 2011). Any miRNA that met the sequence homology criteria (see Materials and methods) was considered a true bichir ortholog of an evolutionarily conserved miRNA gene. A bichir miRNA is deemed as “evolutionarily conserved” if it has at least one ortholog in the miRBase. Through this analysis, we identified a total of 186 distinct evolutionarily conserved mature and star miRNA sequences, corresponding to 225 distinct precursor sequences, and belonging to 120 miRNA families. Almost all of these miRNA precursors had only one copy found in the preliminary assembly of the bichir genome except for one precursor (pse-mir-208a) that had two identical copies and another precursor (pse-mir-122b) that had three identical copies. To summarize our findings, we identified a total of 228 putative precursors in the bichir genome, which corresponded to the 186 conserved mature miRNAs. Detailed information of all evolutionarily conserved miRNA sequences identified as well as their corresponding mature and star sequences is given in Supp. File 4.

Those miRNA sequences, which met the threshold of miRDeep analysis but did not have any known homologous miRNA gene families in miRBase (370 putative novel miRNAs), were further analyzed to predict species-specific novel miRNAs in bichir. Through BLAST searches against the draft bichir genome assembly, putative novel precursor sequences that had multiple significant hits in the bichir genome (more than 5 hits with an E-value of $1 \times e^{-6}$) were discarded in order to
remove interspersed repeats or tandem repeats. Then, the remaining putative precursors were searched against the nonredundant (nr) database in GenBank and other small RNA databases to remove mRNAs or other kinds of small non-coding RNAs. The remaining miRNAs, which met the consensus criteria used to identify novel miRNAs, were considered to be true novel miRNAs in bichir (see Materials and methods). A total of 107 distinct novel mature miRNAs and their corresponding star sequences and 111 different putative precursors (two of them are present as two copies) were discovered in bichir (Supp. File 5).

According to the reports of Tarver et al. (2013), the origin of almost every clade investigated so far is accompanied by the gain of at least one miRNA family, that is to say, every lineage contains some lineage-specific miRNAs. This makes miRNAs ideal markers for phylogenetic analysis. For example, the Bilateria, Vertebrate, and Eutheria share 31, 46, and 93 common miRNA families, respectively. Therefore, here we first focused on the 9 miRNA families that had been previously identified to be specific to teleostei, i.e. they had not been found in any other animal species (Tarver et al., 2013). Through secondary structure and sequence comparative analyses, we identified five of these nine teleostei-specific miRNAs among the candidate bichir miRNA repertoire: miR-462, miR-722, miR-724, miR-725, and miR-2187 (Fig. 2) and all of these miRNAs have both the 5’ processing data and a star sequence supported by many sequenced reads. As for all the other miRNA families specific to amniota or their sub-groups (refer to Sarcopterygii), we did not find orthologous miRNAs for any of them in bichir.

In addition to lineage specific gain of microRNAs, lineage specific loss of microRNAs can also be used to infer phylogenies. For example, it was reported that teleostei have lost five miRNA families (miR-32, miR-191, miR-551, miR-875, and miR-1329) that are present in vertebrata or gnathostomata (Tarver et al., 2013). Interestingly, bichirs have at least retained three of them (miR-551, miR-875, and miR-1329) (Supp. File 3). Therefore, these three miRNA families must have been lost in teleosts exclusively, but not in bichirs. Taken together, these evidences indicate that bichir may be a member of ray-finned fishes.

We next adopted the same approach as in Heimberg et al. (2010), added the new bichir miRNAs into the evolution matrix previously generated by these researchers to re-calculate the phylogenetic tree. The finalized matrix contains absence or presence of data for 190 miRNA families in 15 chordate taxa (Supp. File 6). The phylogenetic analysis using both parsimony method and Bayesian analysis shows that bichirs are branched in the basal position of ray-finned fishes with high posterior probability (PP = 1). All the phylogenetic positions of other species we used were consistent with previous studies, which further confirmed the correct position of bichir (Fig. 3).

4. Discussion

The phylogenetic position of bichirs has always been controversial. With sarcopterygian-like characteristics, such as fleshy pectoral fins, paired ventral lungs, and a pair of external gills in juveniles, bichirs were firstly placed into the sarcopterygians (Noack et al., 1996). Even as of today, some researchers still consider them as a member of the sarcopterygians or at least being more closely related to sarcopterygians than to the actinopterygians (Nelson, 2006). This position is supported by morphological analyses and by molecular analyses using selected genes such as Cytochrome b, 16S, and 12S genes (Noack et al., 1996; Rocco et al., 2004). However, the prevailing opinion in the field, which is supported by molecular analysis using mitochondrial genomes and multiple nuclear genes (Noack et al., 1996; Inoue et al., 2003; Near et al., 2012), places polypterids as basal actinopterygians. It is noted that sometimes these genes are not able to resolve the phylogenetic placement of bichirs in relation to the sarcopterygians (Inoue et al., 2003; Near et al., 2012).
2003). Indeed, bichirs show a mixture of derived and ancestral morphological characters and share many actinopterygian-like and sarcopterygian-like characters, which make it extremely difficult to determine the exactly phylogenetic position using morphological or even molecular characters. On the other hand, because ancient lineages including bichirs, lungfishes, and coelacanths, originated within a small (20–30 million years) time window approximately 400 million years ago (Mayer, 1995), it was too short to accumulate lineage-specific molecular changes. Therefore, molecular phylogenetic analyses based on slowly evolving genes have to some degree been unsuccessful (Mayer, 1995).

miRNAs offer an alternative source of data to resolve the phylogenetic position of bichirs. As elaborated above, miRNAs possess unique properties such as continuous acquisition and low level of loss, which make them ideal candidates to resolve difficult phylogenetic problems. In order to be reliable phylogenetic markers, miRNA validation should make them ideal candidates to resolve difficult phylogenetic problems. As elaborated above, miRNAs possess unique molecular changes. Therefore, molecular phylogenetic analyses based on slowly evolving genes have to some degree been unsuccessful (Mayer, 1995).

References


