Journal of Biodiversity and Environmental Sciences (JBES)
ISSN: 2220-6663 (Print) 2222-3045 (Online)
Vol. 4, No. 1, p. 116-121, 2014
http://www.innspub.net

RESEARCH PAPER OPEN ACCESS

Comparative studies of Indian and Malaysian *Thapaocleidus siamensis* using molecular marker

Saroj Rajvanshi, Nirupama Agrawal*

Department of Zoology, University of Lucknow, Lucknow, 226007, Uttar Pradesh, India

Article published on January 16, 2014

Key words: *Thparocleidus*, comparative study, motif identification, secondary structure.

Abstract

In the present communication the sequence of an exotic species of genus *Thaparocleidus* i.e *T. siamensis* (Lim, 1990) Lim, 1996 (collected in India) has been compared with that of Malaysian, on in silico basis using 28S rDNA. Since, both are morphologically more or less similar and effect of environment on *T. siamensis* cannot access structurally. Thus to evaluate the environmental effects on 28S rDNA region of *T. siamensis*, a comparative study has been done for Indian (obtained) and Malaysian (retrieved) *T. siamensis* using MEME Suite, MARNA and ExpaRNA (online softwares)). The results obtained showing a shift in genetic material because of climatic changes in India.

***Corresponding Author:** Nirupama Agrawal ✉ dr_neeru_1954@yahoo.co.in

Introduction

Pangasianodon hypophthalamus, (Sauvage, 1878) belongs to family Pangasiidae, is an exotic fish in India. We have collected two Malaysian species of the genus *Thaparocleidus* namely *T. siamensis* and *T. caecus* (Mizelle and Kritsky, 1969) Lim, 1996 from gills of *P. hypophthalamus*. Morphologically, Indian and Malaysian specimens of *T. siamensis* are significantly similar. We have compared them genetically, using 28S ribosomal DNA, to evaluate the effects of environmental changes on *T. siamensis* in India. Thus, sequence of 28S rDNA, has been obtained and submitted to GenBank, from *T. siamensis* in India and compared with that of retrieved sequence from Malaysia.

The utility of 28S rDNA being a highly conserved region throughout the evolution has been mentioned by earlier workers (Billoud *et al* ., 2000, Zweib *et al* ., 1981 Schultz *et al* ., 2005 Grajales *et al* ., 2007 and Chaudhary and Singh, 2012, Rajvanshi and Agrawal, 2013) and the structural parameters of this region has been used in systematic and species differentiation. However, motifs have been used as diagnostic biomarkers for species discrimination (Moritz and Cicero, 2005, Chaudhary and Singh, 2012). Thus, we are also using 28S rDNA for the comparative study of *T. siamensis* of Indian and Malaysian sub-continents.

Being same species, they must have common regions (motifs) in the sequence. These are conserved domains within the sequence. The number of motifs, their frequency and position must be similar in a species. Since the morphological distinction cannot be observed, an attempt has been made to evaluate, weather environment has exerted any impact on genetic material of *T. siamensis* or it is the same as in Malaysia. For this, the motifs and their regular expressions have been predicted with the help of MEME software (Timothy *et al* ., 1994). The structural alignment of both species (by MARNA) shows the consensus regions. This result is further tested and verified by the common secondary structure which shows same conservation profile. Secondary structure for each sequence has also been generated separately by ExpaRNA to compare the conservedness. It presents that the conserved sequences are present in both but at distinct positions. These genetical variations in Indian sequence of *T. siamensis* has been incorporated because of environmental changes. These changes are responsible for the gene shift in Indian species.

Materials and methods

Fishes were collected from ornamental fish aquarium at Lucknow and fish farms of Barabanki, Uttar Pradesh; India and maintained in glass aquaria. Hosts were identified by Fishbase (Forese and Pauly, 2012). Gills of freshly dead hosts were examined fresh as well as fixed (3% formaline diluted with lukewarm water). Parasites were dislodged with micro needles in glass petri-dishes and studied under a phase contrast microscope (Olympus BX 51). Unstained glycerine mounts, sealed with sealant, were used for study.

DNA isolation

Single parasite was collected in absolute ethanol for DNA extraction. Total DNA was extracted from the collected parasite using Qiagen's Dneasy Blood and Tissue Kit (Cat. No. 69504) by following protocol as per DNA extraction kit with slight modifications.

Polymerase chain reaction (PCR)

Partial 28S rDNA region of *T. siamensis* was amplified in an Eppendorf Master Cycler Personal (PCR machine: Polymerase chain reaction machine) using forward (5'- ACCCGCTGAATTTAAGCAT-3') and reverse (5'-CTCTTCAGAGTACTTTTCAAC-3') primers (Tandon, 2007). The reaction volume was 25µl, containing 2µl polymerase chain reaction (PCR) buffer (10X), 0.5µl dNTPs (10mM), 0.5µl forward primer (19.6 nMol.), 0.5µl reverse primer (31.9 nMol.), 0.5µl Taq polymerase (5 Units), 1µl MgCl2 (25 mM), 5µl genomic DNA and 15µl miliQ water. PCR conditions were 95° C for 4 min (initial denaturation), followed by 35 cycles of 95° C for 1

min (denaturation), 55° C for 45 sec (annealing), 72 °C for 1 min (extension) and 72° C for 10 min (final extension). PCR products were checked on 1.5 % agarose gels in TAE buffer stained with ethidium bromide (EtBr) and visualized under UV light. Amplicons were sequenced with the same primers using automated sequencer (3730/ABI-3730XL-1409-023), Xcelris Labs Limited, India.

Data analysis

Sequence products were subjected to BLAST (Basic Local Alignment Search Tool) for homology search. Multiple sequence alignment was performed using Clastal W (Thompson *et al* ., 1994). The sequence of query species was compared with retrieved sequences (Malaysian sequence of *T. siamensis*-AF218123.1) to infer genetical distinction between them. Motifs were identified, using MEME (Tamura *et al* ., 2004, 2011). A conservation profile was also generated, using MARNA (Will *et al* ., 2007, 2012). Conservedness in loops was predicted with help of ExpaRNA (Simth *et al* ., 2010). Percentage of Guanocine (G) and Cytocine (C) was calculated using GC calculator (http://www.genomicsplace.com/gc_calc.html) and found 48.9% and 47.8% for Mayasian and Indian sequence of *T. siamensis* respectively. 28S rDNA region of *T. siamensis*, from India was submitted to GenBank under accession numbers JX947852.

Observations

Motif indentification

Nucleotide sequences of 28S r DNA region of *T. siamensis* (Indian and Malaysian) have been tested for the motif identification using MEME software. The results of MEME have predicted three different kinds of motifs in 28S rDNA sequences of *T. siamensis*. Minimum motif width is of 41 bases and maximum, of 50 bases. Base length of motif one and two is 50 bases. However, motif three is of 41 bases. Motif one is coded as sky blue, two as deep blue and three as red colour. Motif one (Fig. 2.) is repeated three times, motif two and three (Fig. 3., 4.) only once in Malaysian sequence of *T. siamensis*. Same results have been obtained for motifs in sequence of *T. siamensis* (Indian sequence). The position of motifs in Indian sequence is clearly distinct from Malaysian. The order of motif is found to be same in both, but showing definite shifting in motif localisation. The shifting encountered in India is of 48 bases. Probably this change has been influenced by environmental changes. The p-value of the motifs has been found to be different within each sequence. The combined P-value of all the three motifs of *T. siamensis* is 1.82e-62 and 4.28e-63 for Indian and Malaysian sequence respectively.

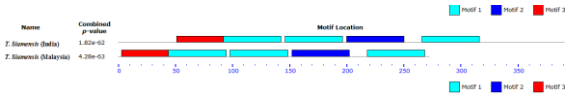

Fig. 1. Cambined bar diagram showing motifs in sequence of 28S region of T. *siamensis*.

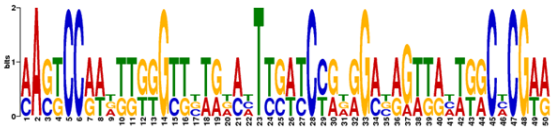

Fig. 2. Motif one.

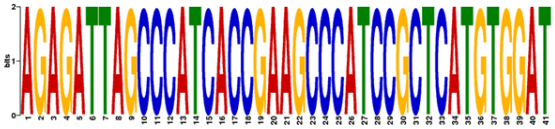

Fig. 3 Motif two.

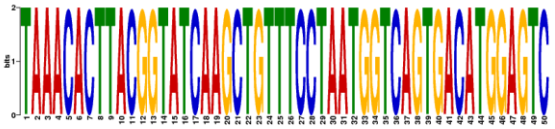

Fig. 4. Motif three.

Secondary structure predictions

Sequences of 28S rDNA region of Indian and Malaysian sequence of *T. siamensis* have been aligned using MARNA, to predict the conserved region, showing colour coded plots of alignment (primary results). Consensus sequences are deeply coloured. The hueness of the colour increases with base difference. The conservation profile for aligned sequences is given at the base line of each row of alignment in grey colour (Fig. 5.). A common RNA secondary structure has also constructed with the

assist of primary results of sequence alignment of both sequence (Fig. 6.). The RNA secondary structure further confirms and refines the primary results of alignment and showing conserved region throughout sequence along with various types of loops (hairpin, interior, multi, bulge and exterior loop), constructed by non-matched bases of sequence. A separate RNA secondary structure for each, Indian and Malaysian sequence of *T. siamensis* have been generated, using ExpaRNA. It is showing species specific, topological differences in secondary structure of Indian as well in Malaysian sequence. Besides topology, these structures infer the consensus sequences using five colour code patterns (light green, violet, yellow, red and blue). These colour coded bases are exactly similar in a particular loop but at different position in sequence.

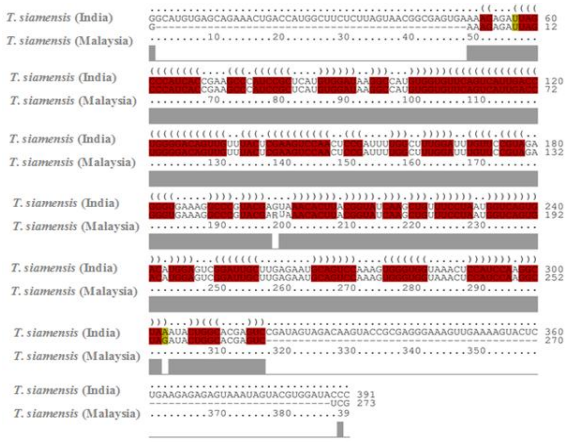

Fig. 5. Sequence alignment of *T. parvulus* and *T. siamensis* showing consensus regions with conservation profile (grey bar).

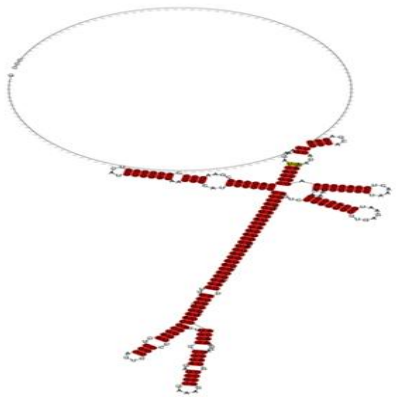

Fig. 6. Secondary structure of *T. parvulus* and *T. siamensis* showing consensus regions (dark colored), hueness of colour increases with the differences.

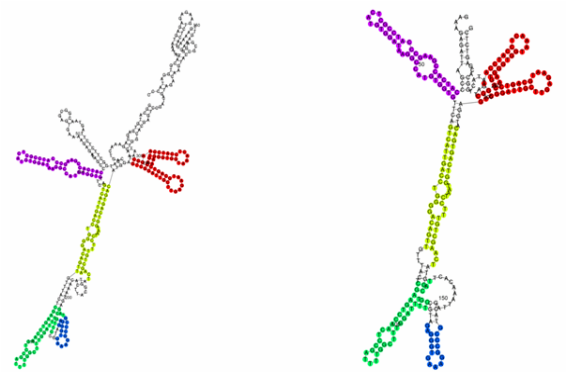

Fig. 7. A & B Secondary structures of Indian and Malaysian *T. siamensis,* showing conserved regions in different loops.

Result and Discussion

T. siamensis (Lim, 1990) Lim, 1996, collected from India has been compared morphologically as well as molecularly, using 28S rDNA with Malaysian *T. siamensis*. The morphological differences between Indian and Malaysian specimens of *T. siamensis* are not significant, having higher degree of similarity in sclerotised part. Thus, it was difficult to asscess the environmental effects. The sequences of 28S rDNA regions have been therefore used to answer if there is any genetical differences between Indian and Malaysian *T. siamensis*or not? Primary and secondary results predicted for *T. siamensis,* using 28S rDNA region clearly shows that there is high degree of similarity in genetic material of Indian and Malaysian sequences but still the variations have been incorporated in sequence of *T. siamensis* in Indian region.

The genetic variations found between sequence of Indian and Malaysian *T. siamensis are* probably because of changes in climatic conditions of India. A significant gene shifting has been incorporated in *T. siamensis*. This gene shift is of 48 bases. The shifting pattern is clearly visible from combined block diagram of motifs, showing a forward shift. The order of motifs in Indian and Malaysian sequence is found to be similar, which is three, one (twice repetition), two and one respectively. The frequency of each motif is similar in both Indian and Malaysian

sequence of *T. siamensis* i.e motif one is repeated three times, motif two and motif three, only once in both sequences. This result is further supported by structural alignment, which is also showing high degree of similarity (represented by grey conservation bar). Similar interpretations have also been made in common secondary structure of RNA generated for primary structural alignment of the sequences. Colour coded conserved regions in secondary structure of each species further verifies that both have similar conserved regions in their genetic material but at different positions. It is significant to note that the same genetic material of both the species have some variations, incorporated in its position, to adapt to the new environmental conditions in India.

T. siamensis of Malaysia and India are more or less, morphologically and molecularly similar (conserved domains), but because of environmental changes, gene shifting has been incorporated in Indian species. These changes are currently not visible morphologically, but if it continues, can establish sibling/new species.

Acknowledgements

Facilities developed from UGC-SAP (DRS-I) programmed under the thrust area "Helminth Taxonomy" of the Department of Zoology were utilized for the present work. We acknowledge UGC for the award of JRF (Endorsement no. 28943) and financial support to NA (F-4-10/2010 BSR).

References

Billoud B, Guerrucci MA, Masselot M, Deutsch JS. 2000. Cirripede phylogeny using a novel approach: molecular morphometrics. Molecular Biology and Evolution **17,** 1435–1445.

Grajales A, Aguilar C, Sanchez J. 2007. Phylogenetic reconstruction using secondary structures of internal transcribed spacer 2 (ITS2, rDNA): finding the molecular and morphological gap in Caribbean gorgonian corals. BMC Evolutionary Biology **7,** 90.

Moritz C, Cicero C. 2005. DNA barcoding: promise and pitfalls. PLoS Biology **2,** 1529–1531.

Schultz J, Maisel S, Gerlach D, Mueller T, Wolf M. 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA **11,** 361–364.

Smith C, Heyne S, Richter AS, Will S, Backofen R. 2010. Freiburg RNA Tools: a web server integrating IntaRNA, ExpaRNA and LocARNA. Nucleic Acid Research **38,** W373-W377.

Rajvanshi S, Agrawal N. 2013. One known and unknown species of of the genus *Thaparocleidus*, Jain, 1952, infecting Sperata aor (Hamilton, 1822): Camparison with species from China, on molecular basis. Bioinformation **9 (11),** 577-582.

Timothy L, Bailey, Charles E. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology,* AAAI Press, Menlo Park, California. 28-36.

Thompson JD, Higgins DG, Gibson TJ. 2003. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions- specific gap penalties and weight matrix choice. Nucleic Acid Research **22,** 4673-80.

Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology,* AAAI Press, Menlo Park, California. **101,** 11030-11035.

Tandon V. 2007. Molecular taxonomy of trematodes parasites-protocols to follow. Manual of the workshop on fish-parasites-taxonomy capacity building. Andhra University, Visakhapatnam, India 79-86.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution **28**, 2731-2739.

Will S, Reiche K, Hofacker IL, Peter FS, Backofen R. 2007. Inferring non-coding RNA families and classes by means of genome-scale structure-based clustering. PLoS Computational Biology **3(4),** e65.

Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. 2012. LocARNA-P: Accurate boundary prediction and improved detection of structural RNAs. RNA **18(5),** 900-14.

Zwieb C, Glotz C, Brimacombe R. 1981 Secondary structure comparisons between small subunit ribosomal RNA molecules from six different species. Nucleic Acids Research **9,** 3621–3640.http://www.genomicsplace.com/gc_calc.html www.fishbase.org, version (04/2013).