Thaparocleidus devraji (Gusev, 1976) Lim, 1996, Infesting Ompok bimaculatus (Bloch, 1794) (Siluriformes: Siluridae): Morphological and Molecular Study

Saroj Rajvanshi¹, Nirupama Agrawal¹, Manoj Kumar Upadhyay²,*

¹Department of Zoology, University of Lucknow, Lucknow, 226007, Uttar Pradesh, India
²Scientist C, Biotech Park, Lucknow
*Corresponding Author: drmkupadhyay@gmail.com

Abstract Thaparocleidus devraji [8, 12] from Ompok malabaricus [27] was established by Gusev [8]. However, its description was based only on the hard parts. On a careful examination of its live specimens from an additional host and a new locality i.e. river Gomti of Lucknow, structural details of the parasite were added, including its egg. Furthermore, phylogenetic relationship was also established with other Indian species of this genus using Urocleidus similis and Ancyrocephalus paradoxus included in an out group. Sequence analysis of partial 28S rDNA had been done using Minimum evolution (ME), Neighbour-joining (NJ) and UPGMA method of MEGA 5 (Molecular Evolutionary Genetics Analysis-5, Tamura et al. [20]). Furthermore, secondary RNA structure, centroid structure, ss- count, energy dot plot and motif identification had been done, along with GC calculation (for conservedness). All the species of this genus under study, had been found morphologically distinct (reproductive and haptoral armature). The morphological distinctions had been supported by genetical evidences. The results of phylogeny had shown that all the species (understudy) are genetically distinct and T.devraji had monophyletic origin.

Keywords Thaparocleidus, Ompok bimaculatus, Phylogenetic Analysis

1. Introduction

Ompok bimaculatus (Bloch 1794), variously known as butter fish, two spot glass catfish or pabda catfish. It inhabits in freshwater body of plains and sub montane regions of India, Pakistan, Sri Lanka, Bangladesh and Myanmar. Till now only three species of the genus Thaparocleidus had been reported from this host viz., Thaparocleidus malabaricus [8, 12] (earlier described from Ompok bimaculatus, type locality not given), Thaparocleidus octotylus [10, 12] (from Ompok Pabda at Hyderabad) and Thaparocleidus devraji (from Ompok malabaricus from Bhavanisagar reservoir). In the present work, only Thaparocleidus devraji had been included whose description from Ompok bimaculatus, was based only on hard parts. On careful examination of its live specimens, from an additional host, in a new locality river Gomti of Lucknow district, structural details of the parasites were observed. Differentiation of a species genetically as well as morphological was one of the most important aims of modern biology, which effectively determines inter as well as intraspecies variations. Molecular tools, using molecular data of conserved domain was being used for rapid and accurate assessment of species. In monogenes, sequences of ribosomal subunits were widely used to infer phylogenetic relationships at the level of families and subfamilies [13, 14, 15] and also to investigate the evolutionary association between the parasites and their hosts [8, 16]. The structural parameters of the 28S secondary structure (geometrical parameters, bond energy and base composition) were used to study phylogenetic relationships among the species [1]. The topology of secondary structure (type/ number/position of loops and single/double stranded portion) was also considered important diagnostic genetic characters for species identification [9, 16 and 24]. Chaudhary and Singh [3] gave secondary structures and phylogenetic utility of the ribosomal subunit (28S) in Thaparocleidus parvulus [8, 12], parasitizing gill filaments of Mystus vittatus in India. In the present work, being a distinct valid species, Thaparocleidus devraji is described in detail because the original description lacked whole mount drawings etc. Its phylogeny on molecular basis is also established along with other Indian species of the genus Thaparocleidus (Thaparocleidus parvulus, Thaparocleidus sp. 1 HS, Thaparocleidus sp. 2 HS, Thaparocleidus sp. HSS, retrieved from Gen Bank). Furthermore, motif identification, secondary RNA structures, centroid structure, ss- count, and energy dot plot were generated along with GC percentage, using a GC calculator for species differentiation (earlier work of Chaudhary and Singh [3, 4], Chairy et al., [5] was for genera differentiation).
2. Materials & Methods

Fishes were collected from River Gomti at Lucknow (26.84°N 80.94°E), Uttar Pradesh; India. Live hosts were also bought from fish markets and maintained in glass aquaria. Hosts were identified by Fishbase [28]. 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). The methods for staining, mounting and illustrating the dactylogyrids were as described by Kritsky et al. [11] and numbering of hooks was that of Kulweic. Measurements were taken in µm, using a calibrated micrometer, following the procedure and terminology of Gusev [8]; means were followed by the range and the number (n) of specimens measured in parentheses. Unstained glycerine mounts, sealed with selant, were used for measurement of soft as well as hard parts. Measurements and illustrations were made with the help of images taken with camera (Olympus-Photometrics Coolsnap), attached with microscope, using Image-ProExpress 6.0 (for Image analysis).

2.1. DNA Isolation

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

2.2. Polymerase Chain Reaction (PCR)

Partial 28S rDNA region of *Thaparocleidus devraji* was amplified in an Eppendorf Master Cycler Personal (PCR machine: Polymerase chain reaction machine) using forward (5’- ACCCGCTGAATTTAAAGCAT-3’) and reverse (5’- TTCTTCAGAGTACTTTTTA-3’) primers [2]. The reaction volume was 25µl, containing 2µl 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 30 Sec. (extension), 72°C for 7 min. (Final extension), hold at 72°C for 4 min. PCR products were checked on 1.5% agarose gels in TAE buffers stained with ethidium bromide (EtBr) and visualized under UV light. Amplicons were sequenced with the same primers using automated sequencer (Inst Model/Name: 3730xl/ -SYNGENE-373XL/-1403624-004 of Applied Biosystems).

2.3. Data Analysis

Sequencing products were subjected to BLAST (Basic Local Alignment Search Tool) for homology search. Multiple sequence alignment was performed using Clastal W [19]. The sequence of query species (*Thaparocleidus devraji*) was compared with retrieved sequences (Table 2) to infer phylogenetic relationship among them. Sequence data (obtained/retrieved) were analyzed using minimum evolution and neighbour-joining methods of MEGA 5 for generating phylogenetic trees. The robustness of the inferred phylogeny were assessed using bootstrap values at 1000 replications. Genetic relatedness among the analyzed monogenes was because of conserved as well as identical regions. Secondary structure of RNA was reconstructed from obtained sequence of partial 28S ribosomal region. Thus the inferred secondary structure and (folding pattern) significantly improved the phylogenetic studies. Further energy dot plot for searching all possible folding (with energy values) and ss-count were used to predict propensity of each base respectively. Percentage of Guanocine (G) and Cytocine (C) was calculated using GC calculator [26]. Secondary RNA structure, ss-count, and energy dot plot was generated using Mfold [24]. Secondary structure and centroid structure of each species was also generated for species discrimination [17, 21, 22]. Motifs were predicted by MEME [18]. Sequence (partial 28S rDNA) of *Thaparocleidus devraji* was submitted to Genbank under accession number KC962229.

2.5. Redescription

Body 657 (398-823; n=10) long, maximum width 126 (78-150; n=10) at mid length (Figure 1). Cephalic region well developed; cephalic lobes 2 pairs, having 2 pairs of eye spots, posterior pair larger, accessory granules present; pharynx spherical 41 (24-55; n=10) in diameter; oesophagus short to non-existent.

Testis 98 (65-136; n=10) long, maximum width at mid length of testis, 55 (35-75; n=10); vas deferens loops left intestinal caecum; seminal vesicle sigmoid dilation of vas deferens. One pair prostatic reservoirs, opening at the base of copulatory complex. Copulatory complex of a sclerotised copulatory tube, having about one and half clockwise coil, 116 (93-133; n=10) long and proximally articulated massive accessory piece (Figure 1.3) in two pieces; measuring 15 (10-18; n=10) and 21 (16-25; n=10) in length. Ovary round to oval, inter-caecal, 99 (70-130; n=10) long, maximum width at mid length, 58 (50-62; n=10). Vagina lightly sclerotised (Figure 1.7). Seminal receptacle round to oval. Vitellaria dense, throughout trunk, except in the regions of reproductive organs.

Haptor 118 (90-170; n=10) long, 97 (74-130; n=10) wide. Dorsal anchor: outer length 58 (53-70; n=10), inner length 68 (60-72; n=10), recurved point 25 (20-28; n=10) (Figure 1.1); dorsal patch 19 (18-22; n=10) (Figure 1.1). Ventral anchor: without roots, outer length 20 (18-26; n=10), inner length 19 (17-24; n=10), recurved point 6 (6-7; n=10) (Figure 1.2). Dorsal bar 19 (18-22; n=10) (Figure 1.5). Ventral bar 31 (21-38; n=10) long (Figure 1.4). Seven pairs similar hooks 12 (12-13; n=10) long (Figure 1.6). Egg round to oval, unipolar, 57 (56-58; n=5) long, 39 (37-40; n=5) wide; single polar filament 9 (9-11; n=5) long (Figure 1.8).
Table 2. Genbank reference sequences used and their respective information

| S. no. | Parasite                     | Host              | Accession no. | Country    |
|-------|------------------------------|-------------------|---------------|------------|
| 1     | *Thaparocleidus parvulus*    | -                 | GU014844.1 (R)| India      |
| 2     | *Thaparocleidus* sp. 2 HS    | *Silonia silondia*| GU980973.1(R) | India      |
| 3     | *Thaparocleidus* sp. 1 HS    | *Wallago attu*    | GU980972.1(R) | India      |
| 4     | *Thaparocleidus* sp. HSS     | *Wallago attu*    | JN020351.1(R) | India      |
| 5     | *Thaparocleidus devraji*     | *Ompok bimaculatus*| KC962229 (Q) | India      |
| 6     | *Urocleidus similis*         | -                 | AJ969938.1(R) | Czech Republic |
| 7     | *Ancyrocephalus paradoxus*   | -                 | AJ969952.1(R) | Slovakia   |

Syn: *Silurodiscoides devraji* [8]; *Parancyrodiscoides devraji* [6, 8]
Type host: *Ompok malabaricus* [27]
Additional host: *Ompok bimaculatus* [27]
Infection site: Gills
Type locality: Bhavanisagar reservoir, India
Present record and locality: River Gomti, Lucknow
No. of Fish hosts examined: 15
No. of hosts found infected: 6
No. of parasites studied: 25
Prevalence and intensity: Prevalence of *Thaparocleidus devraji* was 20% and intensity reached up to 20 to 30 parasites per fish host
Specimen studied: 5 Vouchers (w9330/1-w9335/1) were submitted in Helminthological collection of ZSI (Zoological survey of India)
Gene sequences: Sequence (partial) of 28S rDNA was submitted in NCBI under accession no. KC962229

Figure 1. *Thaparocleidus Devraji* (Gusev, 1976) Lim, 1996

Figure 1.1. Dorsal anchors, 1.2- Ventral anchors, 1.3- Copulatory complex, 1.4- Ventral bar, 1.5- Dorsal bar, 1.6- hook, 1.7- Vaginal armature, 1.8- Egg
3. Results AND Discussion

*Thaparocleidus devraji* was established by Gusev, [8] as *Silurodiscoides devraji* from gills of *Ompok malabaricus* at Bhavanisagar reservoir, India. Dubey et al. [6] placed it under *Parancylodiscoides devraji*. Lim [12] considered *Thaparocleidus*, a senior synonym of *Silurodiscoides*, transferring this species in *Thaparocleidus*. However, the original description was based only on hard parts and vaginal armature and egg were also not observed, which we could add. We could also observe the detailed soft parts. Of all the known Indian species, the species under study closely resembles with *Thaparocleidus octotylus* in the comparative morphology of hooks, dorsal patch, copulatory tube and ventral bar, but differs in the structure of dorsal anchor (smaller in *Thaparocleidus octotylus*), ventral anchor (roots significant in *Thaparocleidus octotylus*), dorsal bar (longer in *Thaparocleidus octotylus*), accessory piece of copulatory complex (pincer shaped in *T. octotylus* but massive in *Thaparocleidus devraji*). It also differs significantly from *Thaparocleidus malabaricus* in the structure of copulatory complex, vaginal armature, dorsal anchor, ventral anchor, dorsal patch, dorsal bar and ventral bar. We are of the opinion that it is a distinct and valid species of the genus.

Phylogenetic analysis among *Thaparocleidus* species (understudy) had been established by Minimum evolution (ME), Neighbour joining (NJ) and UPGMA methods of MEGA 5. The evolutionary trends were generated using sequences of 28S rDNA region of five Indian species of genus *Thaparocleidus* namely *Thaparocleidus parvulus*, *Thaparocleidus* sp. 2 HS, *Thaparocleidus* sp. 1 HS, *Thaparocleidus* HSS and *Urocleidus similis* and *Ancyrocephalus paradoxus* (treated as out group). *Thaparocleidus* belongs to the sub-family Ancylodiscoidinae and Urocleidina similiis and Ancyrocephalus paradoxos to Ancyrocephalinae of the family Dactylogyridae. All the phylogenetic methods (ME, NJ and UPGMA) showed *Thaparocleidus* to be of monophyletic origin [14, 15, 23]. These five species showed similar evolutionary pattern by all the three phylogentic methods. Query sequence (*Thaparocleidus devraji*) evolved from the same node as other species, showing closeness. All the methods formed two clusters for *Thaparocleidus*. Cluster one further formed two sub clusters. Sub cluster one included *Thaparocleidus parvulus* and *Thaparocleidus* sp. 1 HS and sub cluster two had *Thaparocleidus* sp. 2 HS. The second cluster included *Thaparocleidus* sp. HSS and *Thaparocleidus devraji*. The ME tree (Figure 2) of phylogeny showed 41% similarity between species of sub cluster one, 26% between species of sub cluster one and two and 55% between species of the cluster two. However, NJ (Figure 3) and UPGMA (Figure 4) tree had 40%, 27%, 59% and 42%, 41% 59% similarity respectively. The similarity between cluster one and two was 39% for ML tree. Estimate of average evolutionary divergence of all the five sequence pair was 0.73. Estimate of pair wise mean lied between 0.65-0.77 (Table 1). Molecular study therefore supported the morphological observations. *Thaparocleidus devraji* (query sp.) was close to *Thaparocleidus* sp. HSS. Correct identification and designation of *T*. sp. 1 HS, *T*. sp. 2 HS and *T*. sp. HSS was required for comparison of Indian species.

| Species                  | Pair wise mean (0.65-0.77) |
|--------------------------|----------------------------|
| *Thaparocleidus parvulus*| 0.66                       |
| *Thaparocleidus* sp. 2 HS| 0.64 0.66                  |
| *Thaparocleidus* sp. 1 HS| 0.71 0.74 0.73             |
| *Thaparocleidus* HSS     | 0.73 0.68 0.77 0.65        |
| *Urocleidus similis*     | 0.79 0.67 0.79 0.77 0.77   |
| *Ancyrocephalus paradoxus*| 0.76 0.8 0.73 0.73 0.83 0.74 |

Reconstructed secondary structure of RNA (Figure 9) of partial 28S ribosomal region of *Thaparocleidus devraji* had highest negative free energy $\Delta G$ -97.60 kcal/mol. It was -101.10, -226.90, -227.30 and -269.10 in *Thaparocleidus parvulus* (Figure 5), *Thaparocleidus* sp. 1 HS (Figure 6), *Thaparocleidus* sp. 2 HS (Figure 7) and *Thaparocleidus* sp. HSS (Figure 8) respectively. The highest negative free energy of hairpin, bulge, multi, interior and external loops for *Thaparocleidus parvulus* was 31.6, 1.80, 7.5, 11.9 and -11.50 respectively. *Thaparocleidus* sp. 1 HS had 62.8, 20.5, 9.8, 37.4, -18.90, *Thaparocleidus* sp. 2 HS had 58.5, 24.1, 9.8, 39.4, -14.70, *Thaparocleidus* sp. HSS had 54.4, 22.8, 26.6, 48.4, -9.00 and *Thaparocleidus devraji* had 29.7, 3.10, 5.10, 21.4 and -8.20 respectively. Remaining energy was contributed by stacking energy (Table 3).

![Figure 2. Minimum Evolution method](image-url)
Figure 3. Neighbour-joining method

Figure 4. UPGMA method of phylogenetic tree (partial 28S rDNA region of 5 Indian sp. of genus *Thaparoleiodus* with *Urocleidus similis* and *Ancyrocephalus paradoxus* (as outgroup) using MEGA 5).
Figure 5. Schematic representation of secondary structure of partial 28S rRNA for *T. parvulus*
Figure 6. Schematic representation of secondary structure of partial 28S rRNA for T. sp. 1 HS

dG = -226.90
Figure 7. Schematic representation of secondary structure of partial 28S rRNA for *T*. sp. 2 HS

dG = -227.30
Figure 8. Schematic representation of secondary structure of partial 28S rRNA for *T*. sp. HSS

dG = -269.10
Figure 9. Schematic representation of secondary structure of partial 28S rRNA for *T. devraji*
The order of loop preference (descending order) in constructed. RNA secondary structure of *Thaparocleidus devraji* had interior loop (13), hairpin loop (9) followed by bulge loop (1) multi loop (1) and external loops (1). In *Thaparocleidus parvulus*, it was hairpin, interior, multi, bulge and external. In *Thaparocleidus* sp. 1 HS interior, hairpin, bulge, multi and external loop. However, in *Thaparocleidus* sp. 2 HS and *Thaparocleidus* sp. HSS, it was similar i.e. interior, hairpin, bulge, multi and external (Figure 10). The centroid structures for *Thaparocleidus parvulus* and *Thaparocleidus* sp. 1 HS, *Thaparocleidus* sp. 2 HS, *Thaparocleidus* sp. HSS and *Thaparocleidus devraji* showed bases along the circle. The topology of RNA secondary structures (geometry of base pairing) of each species showed distinct highest negative free energy (sum of the free energies of its all loops), free energy of a loop is independent of all other loops. Number, position and types of loop were also different. Thus, RNA secondary structures are also significant in the identification of species. Since, the highest negative free energy and topology of RNA secondary structures is species specific [3, 4, 5].

**Figure 10.** Comparative bar chart showing the number and type of loops in *Thaparocleidus* species

The base pairing (helices) were shown in the form of arcs. The arc pattern was quite different in all the five species (Figure 11, 12,13, 14, 15). The centroid structures of each species of *Thaparocleidus* showed different helix pattern (base pairing). Different species never showed similar topology of RNA structure and arc pattern and could additionally be utilized in species identification. Similar interpretations were given by Chaudhary and Singh [3, 4, 5].

The GC content within the sequence is responsible for the conservedness as reported by earlier workers Chaudhary and Singh [3, 4, 5]. Percentage of GC lies between 45% to 51.2% in species understudy. *Thaparocleidus devraji* had 182 GC (48.9%) and 190 AT, out of 372 nucleobases. *Thaparocleidus parvulus* had 162 GC (49.2%) and 167 AT/329, *Thaparocleidus* sp. 1 HS had 341 GC (45%) and 417 AT/758, *Thaparocleidus* sp. 2 HS had 339 GC (45.1%) and 412 AT/751 and *Thaparocleidus* sp. HSS had 427 GC (49.3%) and 439 AT/866.

The energy dot plot (Figure 16) represented superposition of all possible folding in *Thaparocleidus devraji*. Different colours were, therefore, used to indicate varying levels of sub optimality. Chairy et al., [5] have used energy dot plot for species discrimination of *Dactylogyroides loongicirrus*. *D. longicirrus* had different optimal energy from *T. devraji* [4].

The ss-count of *Thaparocleidus devraji* (Figure 17) showed propensity of each base of 372 nucleobases to be single stranded, within 24 times folding. However, *Dactylogyrosis longicirrus* had 738 nucleobases [4].

The sequences of all the five species of *Thaparocleidus* were aligned to predict the consensus region, represented by brackets (Figure 18). However, consensus profile was grey in colour (given as bottom line in each alignment row).
Motif one (Figure 19) was repeated twice in all the five species. Motif two (Figure 20) four times in *Thaparocleidus* sp. 1 and *Thaparocleidus* sp. 2, three times in *Thaparocleidus* sp. HSS. However, this was absent in *Thaparocleidus pravulus* and *Thaparocleidus devraji*. Motif three (Figure 21) was repeated four times again in *Thaparocleidus* sp. 1 and *Thaparocleidus* sp. 2, twice in *Thaparocleidus pravulus* and *Thaparocleidus devraji* and once in *Thaparocleidus* sp. HSS. Regular expression of motif one was

\[
[CAT][TA][CG][GA][TA][CT][TC][CG][GA][TA][AG][CA][AG][AT][GT][TC][AG][AG][AT][AG][GA][TC][AG][AG][AT][AG][GA][TA]
\]

Three motifs were found common in these five species (Figure 25). The width of motif one, two and three was 50, 50 and 40 nucleotide bases respectively. Motif one, two and three were coded as green, deep blue and red respectively. Frequency of each motif varied per species (Figure 26).

The present analysis showed that 28S ribosomal region exhibited conserved sequence patterns species specific key character (motifs). The number of motif, their frequency, order and position of occurrence within the sequence, were peculiar for them. Thus, it was important to predict species-level variability.
Figure 16. Energy dot plot for *T. devraji*

Figure 17. SS count plot of *T. devraji*
Figure 18. Sequence alignment of *Thaparocleidus* sp. (understudy) along with consensus profile (grey bars)
Figure 19. 18 Motif one

Figure 20. 19 Motif two

Figure 21. 20 Motif three

| Name   | Start | p-value | Sites                        |
|--------|-------|---------|------------------------------|
| 7. sp. HS | 97    | 4.01e-22 | TAAAATCTGG CACCAATCCGATATGACAAATGAAATCCTACCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. sp. 2 HS | 302   | 4.01e-22 | TAAAATCTGG CACCAATCCGATATGACAAATGAAATCCTACCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. sp. 1 HS | 270   | 4.01e-22 | TAAAATCTGG CACCAATCCGATATGACAAATGAAATCCTACCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. parva  | 250   | 4.01e-22 | TAAAATCTGG CACCAATCCGATATGACAAATGAAATCCTACCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. sp. 2 HS | 257   | 7.91e-21 | GTGTCATGAA AATGAGTGCATGATTCTGAGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. sp. 1 HS | 202   | 7.91e-21 | GTGTCATGAA AATGAGTGCATGATTCTGAGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. parva  | 212   | 1.35e-20 | GTGTCATGAA AATGAGTGCATGATTCTGAGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. drevigi | 231   | 1.89e-18 | TGAAATAGGTTAACCTAGGTATTTTCTGGAATGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. drevigi | 299   | 1.89e-18 | TGAAATAGGTTAACCTAGGTATTTTCTGGAATGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |
| 7. sp. HS  | 29    | 2.07e-18 | TTACCTCGCC AATGAGTGCATGATTCTGAGAAATGTGACATCTCAAGGAAATTTGAGCGTTTTCTTCGTAAC |

Figure 22. 21 Position of motif one in *Thaparocleidus* sp.
Thaparocleidus devraji (Gusev, 1976) Lim, 1996, infesting Ompok bimaculatus (Bloch, 1794) (Siluriformes: Siluridae): Morphological and Molecular Study

| Name         | Start | p-value |
|--------------|-------|---------|
| T. sp. 1 HS  | 544   | 6.51e-15 |
| T. sp. 2 HS  | 259   | 1.05e-15 |
| T. sp. 3 HS  | 425   | 7.45e-17 |
| T. sp. 4 HS  | 270   | 7.45e-17 |
| T. sp. 5 HS  | 359   | 2.56e-14 |
| T. sp. 6 HS  | 466   | 2.20e-13 |
| T. sp. 7 HS  | 501   | 2.59e-12 |
| T. sp. 8 HS  | 197   | 2.59e-12 |
| T. sp. 9 HS  | 114   | 2.59e-12 |
| T. sp. 10 HS | 79    | 2.59e-12 |
| T. sp. 11 HS | 422   | 7.71e-11 |

**Figure 23.** 22 Position of motif two in Thaparocleidus sp.

| Name         | Start | p-value |
|--------------|-------|---------|
| T. sp. 1 HS  | 70    | 5.64e-13 |
| T. sp. 2 HS  | 15    | 5.64e-13 |
| T. devraji   | 182   | 9.90e-13 |
| T. sp. 2 HS  | 165   | 5.71e-12 |
| T. sp. 3 HS  | 120   | 2.99e-11 |
| T. sp. 4 HS  | 145   | 2.99e-11 |
| T. sp. 5 HS  | 424   | 1.50e-10 |
| T. sp. 6 HS  | 500   | 2.52e-10 |
| T. sp. 7 HS  | 551   | 2.52e-10 |
| T. parvulus  | 143   | 7.75e-10 |
| T. devraji   | 75    | 2.58e-10 |
| T. sp. 9 HS  | 125   | 1.03e-09 |
| T. parvulus  | 15    | 1.03e-09 |

**Figure 24.** 23 Position of motif three in Thaparocleidus sp.

**Figure 25.** 24 Combined block diagram of motif in Thaparocleidus sp.
Figure 26. Comparative bar chart of motifs for *Thaparocleidus* sp.

| Motif no. | Parasite | Start point | End point | P-value       | Lowest P-value | Combined p-value for each species |
|-----------|-----------|-------------|-----------|---------------|----------------|-----------------------------------|
|           |           |             |           |               |                |                                   |
| **Motif one** |          |             |           |               |                |                                   |
|           | *T. parvulus* | 212         | 262       | 1.35e-20      | 4.03e-22       | 1.97e-24                          |
|           | *T. sp.1 HS* | 202         | 252       | 7.91e-21      | 4.03e-22       | 1.32e-40                          |
|           | *T. sp. 2 HS* | 237         | 287       | 7.91e-21      | 4.03e-22       | 8.02e-41                          |
|           | *T. sp. HSS* | 29          | 79        | 2.07e-18      | 4.03e-22       | 4.03e-22                          |
|           | *T. devraji* | 299         | 349       | 1.29e-18      | 1.19e-18       | 6.36e-23                          |
| **Motif two** |          |             |           |               |                |                                   |
|           | *T. parvulus* | 79          | 129       | 3.53e-12      | 7.45e-17       | 1.05e-18                          |
|           | *T. sp.1 HS* | 661         | 711       | 2.86e-13      | 7.45e-17       |                                   |
|           | *T. sp. 2 HS* | 114         | 164       | 2.94e-12      | 6.53e-19       | 6.53e-19                          |
|           | *T. sp. HSS* | 197         | 247       | 35.9e-13      | 8.66e-14       | 8.66e-14                          |
|           | *T. devraji* | 49          | 89        | 1.86e-09      | 7.78e-11       |                                   |
| **Motif three** |          |             |           |               |                |                                   |
|           | *T. parvulus* | 35          | 75        | 1.68e-13      | 2.22e-10       | 1.68e-13                          |
|           | *T. sp.1 HS* | 130         | 170       | 2.99e-11      | 2.32e-10       |                                   |
|           | *T. sp. 2 HS* | 165         | 205       | 8.71e-12      | 2.32e-10       | 1.68e-13                          |
|           | *T. sp. HSS* | 253         | 293       | 1.03e-09      | 4.98e-13       |                                   |
|           | *T. devraji* | 73          | 113       | 5.85e-10      | 4.98e-13       |                                   |
The genus *Thaparocleidus* had a great species diversity in India and we could collect several of them, including a few new species. All the species described so far were primarily based on morphological basis. Some species are quite similar structurally. The use of these tools appeared very useful. Chaudhary and Singh [3, 4] have differentiated genera *Thaparocleidus* and *Bifurcohaptor*, using various tools. We tried to compare five species of one genus in the present work for the first time, using 28S ribosomal region. The phylogenetic trees, primary as well as secondary results (morphological identification).

**Acknowledgements**

We acknowledge the help received from our teacher Prof. K.C. Pandey. Facilities developed from UGC-SAP (DRS-I) programme under the thrust area “Helminth Taxonomy” of the Department of Zoology, University of Lucknow, utilized for the present work. We thank Head, Zoology for providing laboratory facilities. We also acknowledge UGC for the award of JRF (Endorsement no. 28943) and financial support to NA (F-4-10/2010 BSR).

**REFERENCES**

[1] B. Billoud, M.A. Guerrucci, M. Masselot, J.S. Deutsch. Cirripede phylogeny using a novel approach: molecular morphometrics, Molecular Biology and Evolution, Vol. 17, 1435–1445, 2000.

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

[3] A. Chaudhary, H.S. Singh, Secondary structure and phylogenetic utility of the ribosomal large subunit (28S) in monogeneans of the genus *Thaparocleidus* and *Bifurcohaptor* (Monogenea: Dactylogyridae), Journal of Parasitic Disease, Vol. 37, 74-83, 2012a.

[4] A. Chaudhary, H.S. Singh, Description of two new species of genus *Thaparocleidus Jain*, 1952 (Monogenea: Dactylogyridae) from freshwater fishes in India: genetic and morphological evidence with application of the secondary structure model of rRNA for phylogeny, Journal of Helminthology, Vol. 87, 160-170, 2012b.

[5] H.R. Chiary, A. Chaudary, H.S. Singh, Phylogenetic analysis of the *Dactylogyridae* longicirrus (Monogenea: Dactylogyridae) based on the 18S and ITS 1 ribosomal genes, Bioinformation, Vol. 9, No. 5, 250-254, 2013.

[6] A. Dubey, A.K. Gupta, S.M. Agrawal, Studies on monogenean parasites in freshwater fishes at Raipur VIII. Validity of *Paracylindocotyle* Achmerow, 1964 (Siluridoids, Gusev, 1974) and two new species from Raipur, Indian Journal Helminthology, Vol. 44, 9-16, 1992.

[7] Y. Des devises, S. Morand, O. Jousson, P. Legendre, Co evolution between Lamellodiscus (Monogenea: Diplectanidae) and Sparidae (Teleostei): the study of a complex host-parasite system, Evolution, Vol. 56, 2459–2471, 2002.

[8] A.V. Gusev, Freshwater Indian Monogeneoidea. Principles of systematics, analysis of the world faunas and their evolution, Indian Journal Helminthology, Vol. 25-26, 1-241, 1976.

[9] A. Grajales, C. Aguilar, J. Sanchez, Phylogenetic reconstruction using secondary structures of internal transcribed spacer 2 (ITS2, rDNA): finding the molecular and morphological gap in Caribbean gorgonian corals, BMC Evolution Biology, Vol. 7, No. 90, 2007.

[10] T. Kulkarni, Studies on the monogenean trematodes of fishes found in Hyderabad, Andhra Pradesh (India) Part II, Rivista De Parasitologia, Vol. 30, No. 4, 263-281, 1969.

[11] D.C. Kritsky, V.E. Thatcher, W.A. Boeger, Neotropical Monogenea 8. Revision of Urocleidoides (Dactylogyridae, Ancoereophalinae), Proceedings of Helminthological Society of Washington, Vol. 53, No. 1, 1-37, 1986.

[12] L.H.S. Lim, Thaparocleidus Jain, 1952, the senior synonym of *Silurodiscoides* Gusev, 1976 (Monogenea: Ancylocleididae), Systematic Parasitology, Vol. 35, 207-215, 1996.

[13] L. Plaisance, D.T.J. Littlewood, P.D. Olson, S. Morand, Molecular phylogeny of gill monogeneans (Platyhelminthes, Monogenea, Dactylogyridae) and colonization of Indo-West pacific butterflyfish hosts (Perciformes, Chaetodontidae), Zoologica Scripta, Vol. 34, 425–436, 2005.

[14] A. Simkova, L. Plaisance, I. Matejusova, S. Morand, O. Verneau, Phylogenetic relationships of the Dactylogyridae Bychowsky, 1933 (Monogenea: Dactylogyridae): the need for the systematic revision of the Ancoereophalinae Bychowsky, 1937, Systematic Parasitology, Vol. 54, 1–11, 2003.

[15] A. Simkova, I. Matejusova, C.O. Cunningham, A molecular phylogeny of the Dactylogyridae sensu Kritsky and Boeger (1989) (Monogenea) based on the D1-D3 domains of large subunit rDNA, Parasitology, Vol. 133, 43-53, 2006.

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

[17] C. Smith, S. Heyne, A.S. Richter, S. Will, R. Backofen, Freiburg RNA Tools: a web server integrating IntaRNA, ExpRNA and LocARNA, Nucleic Acids Research, Vol. 38, W373-W377, 2010.

[18] L. Timothy, Bailey, C. Elkan, 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, 1994.

[19] J.D. Thompson, D.G. Higgins, T.J. Gibson. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions- specific gap penalties and weight matrix choice, Nucleic Acids Research, Vol. 22, 4673-4680, 1994.

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

[21] S. Will, K. Reiche, I.L. Hofacker, F.S. Peter, R. Backofen. Inferring non-coding RNA families and classes by means of genome-scale structure-based clustering. PLoS Computational Biology, Vol. 3, No. 4, e65, 2007.

[22] S. Will, I.T. Josh, I.L. Hofacker, P.F. Stadler, R. Backofen. LocARNA-P: Accurate boundary prediction and improved detection of structural RNAs, RNA, Vol. 18, No. 5, 900-914, 2012.

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

[24] C. Zwieb, C. Glotz, R. Brimacombe. Secondary structure comparisons between small subunit ribosomal RNA molecules from six different species, Nucleic Acids Research, Vol. 9, 3621–3640, 1981.

[25] M. Zuker. Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Research, Vol. 31, No. 13, 3406-3415, 2003.

[26] http://www.genomicsplace.com/gc_calc.html

[27] www.iucnredlist.org Downloaded on 11 March 2013

[28] www.fishbase.org, version (04/2013)