Genomic DNA Sequence of *Leidynema appendiculata* from Meerut, U.P., India

Anshu Chaudhary, Sangeeta Pal, Malti and Hridaya Shanker Singh
Department of Zoology, Molecular Taxonomy Lab., Ch. C.S. University, Meerut (U.P.), 250004, India

Corresponding Author: Hridaya Shanker Singh, Department of Zoology, Molecular Taxonomy Lab., Ch. C.S. University, Meerut (U.P.), 250004, India

**ABSTRACT**

The present communication deals with taxonomy of *Leidynema appendiculata* from the hind gut of cockroach, *Periplaneta americana* L. An attempt has been made to summarize morphological, molecular data and taxonomical status of *Leidynema appendiculata*. The sequence of segment of LSU rDNA were obtained and analyzed. Besides this, based on molecular data, an attempt to find the phylogenetic position of this nematode has also been made. The analysis of morphological and molecular-taxonomic characters supported the independent species status of *Leidynema appendiculata*. An attempt to analyze the secondary structure of the 28S ribosomal (rRNA) encoding gene region has also been made.

**Key words:** *Leidynema appendiculata*, large subunit, ribosomal DNA, secondary structure, ribosome

**INTRODUCTION**

The super family Thelastomatoida is one of the two super families of the order Oxyurida. They are the nematodes parasitic or commensal in gut of terrestrial arthropods and usually feed upon the bacterial microfauna found there (Jex *et al.*, 2005). Adamson and Van Waerebeke (1992a-c) divided the superfamily into five families viz., Thelastomatidae, Protrelloididae, Hystrignathidae, Travassoscinematidae and Pseudonymidae which is followed by subsequent workers. The family Thelastomatidae is by far the largest family of the superfamily Thelastomatoida reported that the family Thelastomatidae is represented by 28 genera. Recently, three more genera to the family were added and hence the family now represented by 31 genera. So far most of these genera are characterized on morphological basis which often creates taxonomic confusions.

Recently, molecular markers have often been used for taxonomic identification and phylogenetic analyses in different species groups (Abdul Majeed *et al.*, 2001; Li *et al.*, 2011). Molecular methods have been used to assess the relationships of the major species and to examine intraspecific variations. It has been demonstrated that the 28S rRNA gene and nuclear ribosomal DNA containing the first internal-transcribed spacer (ITS1), 5.8S rRNA and second internal transcribed spacer (ITS2) regions are informative for molecular diagnostics of nematodes and in the phylogenetic relationship analyses between nematodes from population to class level (Adams *et al.*, 1998; Kampfer *et al.*, 1998; Iwahori *et al.*, 1998, Beckenbach *et al.*, 1999; De Giorgi *et al.*, 2002; Wang *et al.*, 2005; Nadler *et al.*, 2006).

During the survey of insect parasitic nematodes from Meerut, U.P., India, investigator came across, a known species of the family Thelastomatidae, the *Leidynema appendiculata*
(Leidy, 1850; Chitwood, 1932) having minor variations harbouring almentary canal of common cockroach. In this study, the 28S rRNA gene sequences were used to investigate the identification, validity and phylogenetic affinities with different geographical isolates in order to provide molecular evidence of this nematode in India along with detailed morphological and morphometric analysis as original description lacks some morphological and morphometric details. Reason behind selecting the 28S ribosomal RNA gene has been made because it represents a well-conserved gene that evolves relatively slowly. Moreover, sequences of this gene have recently been used in the studies resolving the phylogenetic relationships between nematode and other animals and this gene was shown to be a suitable marker for barcoding of nematodes (Barthelemy et al., 2006). Besides this, RNA secondary structure has also been used for phylogenetic reconstruction.

**MATERIALS AND METHODS**

Hosts, *P. americana* were caught from Meerut, U.P., India, alive with the help of nets and brought to the laboratory in live condition in insect cages and were kept in these cages till surgical procedures. On dissection, nematodes were removed from the digestive tract, placed in 0.6% saline and fixed in hot 70% alcohol solution. For light microscopical examination, the nematodes were cleared gradually in glycerin. A light microscope equipped with differential interference contrast, digital image analysis system (Motic digital microscope for Windows) was used for morphological and morphometric analysis and drawings were made by camera lucida.

Initially methods to extract genomic DNA were following from the work of Oboh et al. (2009) and Al-Saghir (2009). Genomic DNA was extracted from ethanol-preserved parasite using the DNeasy Tissue Kit (Qiagen). 28S rDNA was amplified using the Master Cycler Personal (Eppendorf) in a final volume of 25 μL PCR reaction. Each amplification reaction contained 10X PCR buffer, 0.4 mM dNTP, 1 U Taq polymerase (Biotools) and 10 pM of each primer-specifically designed forward (5'-TTGGCGTCTCAAGTGAAAG-3') and the reverse primer (5'-TTCACATCATCTTGCGGTCTC-3'). PCR was carried out with the following steps: an initial denaturation at 94°C for 3 min, 35 cycles of 94°C for 30 sec, 56°C for 45 sec and 72°C for 1 min, and a final extension at 72°C for 10 min. PCR products were separated on 1.5% agarose gel. The products were then purified by Chromous PCR cleanup kit (# PCR 10), according to manufacturer’s instructions. Both DNA strands were sequenced using a Big Dye Terminator v3.1 cycle sequencing kit in an ABI 3130 Genetic Analyzer using same primers. Sequences were analyzed using the MEGA 5.01 (Tamura et al., 2011; MEGA: Molecular Evolutionary Genetics Analysis), using neighbor-joining, minimum evolution, maximum likelihood and maximum parsimony methods. Bootstrap confidence values, reported as percentages, were calculated based on 1,000 replicates (Felsenstein, 1985). Distances (base substitutions per site) were computed using the Kimura 2-parameter method (Kimura, 1980). The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (Nei and Kumar, 2000).

RNA secondary structure for 28S rRNA was predicted by using the Sfold program (Ding et al., 2004) in Sfold (Statistical and Rational Design of Nucleic Acids). Sfold predicts RNA structure by identifying suboptimal structure using the free energy optimization methodology at a default temperature 37°C. The dynamic programming algorithm used in Sfold is based on the work of Zuker and Stiegler (1981). GC percentage was determined using CG calculator (http://www.genomicsplace.com/gc_calc.html).
RESULTS AND DISCUSSION

Leidynema appendiculata (Leidy, 1850; Chitwood, 1932):

- Host Type: *Periplaneta americana* L.
- Locality Type: Meerut (29° 01' N, 77° 45' E), U.P., India
- Site of infection: Hind gut
- Material Type: The holotype and paratype slides have been deposited in the museum of Department of Zoology (Voucher number Nem/2009/01), Ch. C.S. University, Meerut, U.P., India.

Fig. 1: *L. appendiculata* female, (a) Whole mount, (b) Anterior end of female, (c) Posterior end of female and (d) Eggs
The worms are small. Males are with evenly tapering anterior end and very short tail. The caudal appendage in female is elongated. Mouth is surrounded by very large sub median labial papillae bearing amphids or lateral organs in the form of circular protuberances. Body of the worm in both sexes is provided with well developed lateral alae. The cuticular covering of body bears transverse striations which are coarse in the anterior region in either sex. The excretory pore is marked in post oesophageal part of body in both the sex (Fig. 1, 2).

**Female:** Body of female is elongated, measuring 2.70-2.75 mm in length. Maximum width 0.25-0.28 mm recorded in pre-gonadal region of the body. Mouth leads into well developed, rectangular buccal cavity measuring 0.012-0.014 × 0.015-0.017 mm. Oesophagus is long measuring

---

Fig. 2: *L. appendiculata* male, (a) Whole mount, (b) Anterior end of male, (c) Posterior end of male and (d) Spicule
0.35-0.42 mm in total length and is divisible into three parts: corpus, isthmus and oesophageal bulb. Anterior most corpus measures 0.25-0.28 × 0.03-0.04 mm. Middle oesophageal part, isthmus is smaller in comparison to corpus and is highly muscularized, measuring 0.021-0.024 × 0.028-0.032 mm. Posterior rounded valvular oesophageal bulb measuring 0.08-0.12 mm in diameter. Nerve ring is located at 0.12-0.14 mm from anterior end of the body. Excretory pore is post-oesophageal in location, at 0.55-0.58 mm from anterior end of the body. Intestine is well developed and equipped with prominent cardia and a diverticulum at its anterior end but posteriorly it is cylindrical and simple. Anus is located slightly anterior to posterior most part of the alae at 0.65-0.68 mm. Ovaries are two, uteri are amphid nextlic. Vulva is more or less equatorial, located at 1.50-1.80 mm away from anterior end of the body. Vagina is non-muscular and straight. Uteri filled with large number of eggs which are oval in outline, elongated and slightly flattened on one side measuring 0.065-0.068 × 0.038-0.034 mm. Tail is filiform, elongated measuring 0.14-0.16 mm in length (Table 1).

**Male:** Males are much smaller than females measuring 0.82-0.85 mm in length. Maximum width, 0.09-0.12 mm recorded in the post oesophageal region of the body. Mouth leads into rectangular
Fig. 3: Phylogenetic tree of the relationships between the *L. appendiculata* and related species

buccal cavity measuring 0.013-0.015×0.012-0.014 mm. Oesophagus is elongated measuring 0.08-0.094 mm in length and comprises of well a developed corpus, isthmus and oesophageal bulb. In the original description measurement of oesophagus is possibly miscalculated for the reasons best known to the authors. Length of oesophageal corpus is 0.05-0.06 mm and its width is 0.015-0.017 mm. Isthmus measures 0.005-0.007×0.012-0.016 mm. Oesophageal bulb is oval in outline measuring 0.025-0.027×0.024-0.028 mm in diameter. Nerve ring is located at 0.062-0.065 mm from the anterior end of the body. Excretory pore is located at 0.11-0.13 mm form anterior end of the body. Intestine is simple, without any diverticulum and it is provided with a well developed cardia. Diameter of intestine is more or less double as compared to the oesophageal bulb. Anus is located at 0.021-0.025 mm from the posterior end of body. Testis elongated, occupying two third posterior part of the body. Tail is pointed, very small and measures 0.009-0.012 mm in length. Spicule is single, slightly curved, measuring 0.049-0.052 mm in length. Papillae are five pairs (one pair pre-anal, one pair ad-anal and three pairs post-anal). Pre-anal papillae are protruded whereas; ad-anal and first pair of post-anal papillae is sessile (Table 1).

The fragment of 28S rDNA is 220 bp in length (sequence of *L. appendiculata* was deposited in GenBank database accession number GQ925910). Significant alignment was produced in BLAST with sequences of different nematodes species. Close similarity between the sequence of *L. appendiculata* and other nematodes sequences included in alignment, ranged between 90 and 99% with the proportion of gaps ranging from 0 to 5% (Fig. 3). On the basis of base pair sequences, BLAST clustered present worm together with *L. appendiculata* from Russia (accession number EU355330) and showed that worms at the disposal of authors is 99% related. Phylogenetic tree of the relationships between the *L. appendiculata* and related species with high bootstrap values in a minimum evolution tree. Similar results were obtained using neighbor-joining, maximum likelihood and maximum parsimony maximum parsimony.

Details of the secondary RNA structure prediction are: G+C content for the 28S region of rDNA of *L. appendiculata* is 51.4%. The minimum free energy is estimated by summing individual energy contributions from base pair stacking, hairpins, bulges, internal loops and multi-branch loops. RNA secondary structure consists of stems and loops (Fig. 4). From the data we were able to draw the rRNA secondary structure where each residue is identified by a base pair, the backbone and the hydrogen bonds are represented as dots between the base pair. Mainly five types of loops are present in RNA secondary structure viz., interior, hairpin, exterior, multi and bulge (Fig. 5). Each residue is represented on the abscissa and semi-elliptical lines connect bases that pair with each other. In a centroid diagram, bases are positioned along a circle, in a clockwise orientation. An arc
Fig. 4: Predicted 28S RNA secondary structure of *L. appendiculata*

connecting two bases across the circle indicates pairing between the bases. The lack of pseudoknots in the secondary structure is reflected by the absence of intersecting lines in the centroid structure (Fig. 5). From the two-dimensional histogram (2Dhist), the patterns of base pair frequencies are nearly identical for the sample (Fig. 7). It contains base pair frequencies for constructing 2D histogram. A two-dimensional histogram (2Dhist) displays base pair probabilities computed from a statistical sample of structures. In the 2Dhist, base pair probabilities are shown by solid squares in the upper left triangle, with the nucleotide positions on both axes. The areas of the solid squares are proportional to the frequencies of the base pairs in the sampled structures.
Fig. 5: Histogram showing types of loops in 28S region of *L. appendiculata*

Fig. 6: Centroid structure of predicted 28S RNA structure of *L. appendiculata*

The probability profile displays predicted accessible sites on the target RNA. For prediction of target accessibility, a complete probability profile of single-stranded regions is generated for the entire target RNA. Sites with high probabilities of being single-stranded are predicted to be accessible. The probability profiles ($W = 4$) for the sample are also computed and in a plot, by plotting against nucleotide position (Fig. 8). On a profile for fragment width $W$, the probability that $W$ consecutive bases are all unpaired is plotted against the first base of the segment. This approach has proved to make substantially better predictions than the MFE structure. The significance of assigning probability as a measure of confidence in prediction is highlighted (Fig. 8). A single-stranded region predicted by both the MFE structure and the ss-count has low probabilities on the
Fig. 7: The two-dimensional histogram of base pairs in the *L. appendiculata*

Fig. 8: Line plot representing secondary structure of 28S RNA of *L. appendiculata*

probability profile (*W* = 4 bases, as described in Ding and Lawrence, 2001). The ss-count statistic gives the propensity of a base to be unpaired, as measured by the frequency with which it is unpaired in a group of the optimal and suboptimal foldings within a specified increment of the MFE. At nucleotide position *i*, the probability that nucleotide *i*, *i* + 1, *i* + 2, *i* + 3 (i.e., fragment width *W* = 4) are all single stranded is plotted against *i*. This probability is computed by MFE structure and by ss-count from mfold for the nucleotides 1-220 of *L. appendiculata*. A 3D energy landscape plot of the sampled ensemble and representative structures (Fig. 9). The optimal number of clusters determined by our software is 2. Structures belonging to the clusters are marked as solid dots of two different colors. The MFE structure and the ensemble centroid are both in the largest cluster (light blue color), with a probability of 0.767. The coordinates for a structure are (MDS axis 1, MDS axis 2, free energy), where the horizontal axes are from MDS and the vertical axis is the free energy of a secondary structure. The coordinates are for the MFE structure (-14.68, 1.03, -84.90), for the ensemble centroid (3.49, -1.99, -85.10), for the centroid of cluster 1, (3.49, -1.99, -85.10) and for the centroid of cluster 2, (-14.60, 0.95, -84.60).
Fig. 9: The energy landscape of the sampled ensemble and representative structures for *L. appendiculata* of 220 nt

Clusters are sorted in descending order of cluster size. The cluster containing the MFE structure is marked using a red asterisk. The size of a cluster is the sampling estimate for the probability of the cluster i.e., the sample frequency of the cluster. The MFE structure is excluded from the calculation of cluster sizes and the sizes of all clusters sum to one. Multi-Dimensional Scaling (MDS) plot of the sampled ensemble and representative structures MDS is a technique for representing high-dimensional objects in typically two dimensions (Kruskal and Wish, 1977). For RNA secondary structures, base-pair distances are used as an input to MDS. Members of all other clusters are plotted as small circles.

The genus *Leidynema* was established by Schwenk (Travassos, 1926). The study of the worms at the disposal of authors places it under the genus *Leidynema* Schwenk (Travassos, 1929). Morphologically this worm was found closest to *Leidynema appendiculata* (Leidy, 1850; Chitwood, 1932) with minor variations in the measurement of various body parts (Table 1). Variations were noticed mainly in the measurements of various body parts. That could be due to the presence of parasite in different hosts and ecological niche.

In India, molecular phylogenetic studies of nematodes are still at an early stage. The position of *L. appendiculata* in the phylogenetic trees reconstructed in BLAST confirms its placement of the worm at our disposal within genus *Leidynema* Schwenk (Travassos, 1929). Moreover, it shows that the position in the molecular phylogenetic tree corresponds well to the morphological similarity with the representative of Oxyuroidea. Genetic relatedness between the studied species of *Leidynema* indicates closest similarity with *L. appendiculata* from Russia i.e., 99%. The reasons for differences in genetic similarity between same species from different continents are due to their geographical distribution or due to adaptation with their host in different continents. Present phylogenetic analysis based on the LSU rRNA gene sequence, *L. appendiculata* from India and Russia are same species. Therefore, we concluded that *L. appendiculata* may be the same species from different
region and different hosts. Genetic comparison with further specimens documented from other parts of the world may alter our taxonomical concept of this group and provide further clues to the understanding of the evolution of the species.

In the present study, we propose that L. appendiculata from India and Russia are the same species. The minor molecular biological differences might be due to the following reasons- (1) these two species are presently undergoing a speciation that precedes the molecular difference of 1%. (2) as these two morphological types are exclusively and respectively found on two different allopatric hosts, this molecular differentiation (i.e., 1%) could be host-induced. (3) This might also be due to the presence of this parasite in two different zoo-geographical regions which are separated way back by the mighty barriers.

Beside this, we also inferred the secondary structures of the long subunit (LSU) in L. appendiculata. The structure prediction method we propose presents a promising approach to reconstruct secondary structures of non-coding genes in taxa that have not been studied so far but has significant taxonomic and phylogenetic value. The consideration of taxon-specific secondary structure models helps to improve the inference. Different RNA folding algorithms also take into account the structural energy as the major determinant in furnishing RNA secondary structure models and conformation which will definitely add meaningful dimensions to our understanding of the relationships among the sequence features and structural parameters that come into play in determining the structural energy. This approach can be further fine-tuned in resolving ambiguities using differences at the RNA structural level for identification of sibling species complexes. Representative structures, including the MFE structure, ensemble centroid and centroids of all colored clusters, are drawn as large dots in the graph. The units on the axes of the MDS plot only serve the purpose of indicating the relative positions of the objects (Chan et al., 2005).

The present study demonstrates use of secondary structure on phylogenetic analyses of rRNA sequences. This effect was remarkable in sequence alignment and tree reconstruction of both simulated and empirical data. However, the favorable performance of structural alignments partly vanished in tree reconstructions, an aspect which clearly needs further investigations.

CONCLUSION

This study of the Leidynema appendiculata (Leidy, 1850; Chitwood, 1932) from P. americana represents the first study for phylogenetic purposes and concludes that sequences of large subunit are suitable to complement morphological data for the reconstruction of phylogenetic relationships. RNA secondary structure analysis could be a valuable tool for distinguishing species and completing L. appendiculata systematics, more so because 28S secondary structure contains more information than the usual primary sequence alignment. The analysis of morphological and molecular-taxonomic characters supported the independent species status of Leidynema appendiculata.

ACKNOWLEDGMENTS

The authors are thankful to Head, Department of Zoology, C.C.S. University, Meerut, for providing laboratory facilities. This work was funded by the Department of Science and Technology, (grant number SR/SO/A543/2005) awarded to HSS.

REFERENCES


