ABSTRACT

Order Cyclophyllidea (of platyhelminth cestode) has a rich diversity of parasites and includes many families and species that are known to cause serious medical condition in humans and domestic and wild animals. Given the severity of medical implications inflicted by the species members of the families in this order, specific identification of the aetiological agent becomes imperative. To serve the purpose, molecular characterization of some food-borne cyclophyllidean cestodes of zoonotic potential from Northeast India was carried out in the present study, using PCR-based molecular techniques and sequence analysis.

- To commence the study, tapeworm parasites (adult or metacestode) were collected from naturally infected hosts pig (Sus scrofa), cattle (Bos indicus), goat (Capra hircus), and black rat (Rattus rattus) from different regions of Northeast India and also from laboratory maintained albino rat (Rattus norvegicus); the collection sites included various localities in states of Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura.

- The collected cestode parasites included metacestode (larval stage) from muscle tissue of pigs, peritoneal cavity of goats, and from parenteric sites in cow and goat. Adult tapeworms were also collected from small intestine of black rat, albino rat and goat.

- Using light microscopy, the morphological features of the collected parasites were determined, following which they were identified as Cysticercus cellulosae (the bladder worm metacestode of Taenia solium), C. tenuicollis (the bladder worm of T. hydatigena), hydatid cyst (metacestode of Echinococcus granulosus) and adult
tapeworms—*Hymenolepis nana*, *H. diminuta* and *Moniezia expansa*. The identifications of the parasites under study were confirmed by comparing with the voucher specimens deposited in helminthological collection at Department of Zoology, North-Eastern Hill University, Shillong.

**Molecular characterization** of all the six tapeworm taxa recovered was carried out using DNA-based PCR technology. For the purpose, the nuclear ribosomal (rDNA) markers including internal transcribed spacer 2 (ITS2), smaller subunit of ribosomal (SSR) or 18S and larger subunit of ribosomal (LSR) or 28S and the mitochondrial cytochrome c oxidase subunit 1 (mtCO1) gene regions were chosen.

- For this perspective, DNA was isolated from individual parasite specimen using the standard phenol-chloroform method. ITS2, 18S and 28S rDNA and mtCO1 gene regions were PCR-amplified following standard procedures.
- The PCR-amplicons were purified and sequenced in both directions using the forward and reverse primers on automated sequencer; the sequences were submitted to GenBank and their accession numbers obtained. A total of 27 sequences were submitted to GenBank, which include nine ITS2, seven 18S, two 28S and nine mtCO1.

In order to accomplish the aim of the present study, following approaches were used for each of the cestode taxa under study:

- **Sequence analysis**
  - To validate all the sequenced gene regions, BLAST similarity search was performed for all cestode taxa under study and on the basis of similarity search
results, sequences of the same species (representing geographical/host isolates) and also of related species from the same family, were mined from the NCBI database available in public domain for analysis.

- Sequence analysis was carried out using various bioinformatics tools such as ‘Clustal W’ and ‘MUSCLE’ incorporated in MEGA6 software for multiple sequence alignment, and BioEdit for generating sequence identity matrix, with an aim to see intra- and inter-species genetic variations among the various geographical/host isolates and related taxa, respectively.

- The results of intra-specific genetic variation revealed ITS2 to be highly variable among the taeniids (Taenia and Echinococcus) and least among Hymenolepis and Moniezia species. Conversely, mtCO1 gene showed significantly higher rates of variation among the Hymenolepis and Moniezia species and least among the taeniids. Among the coding regions of rDNA gene array, 18S displayed considerably lower degree of genetic variations in most of the taxa studied and 28S gene, studied only for the species of Hymenolepis, showed little or negligible variation.

- With regard to inter-specific genetic variation, it was seen that the species of E. granulosus and M. expansa lacked the sequence data for ITS2 gene. Nonetheless, the results obtained disclosed ITS2 to be the most variable genetic marker with variation going upto 75.8% as in case of species of Taenia and 38.1% in case of Hymenolepis spp. The rDNA genes 18S and 28S as well as the mtCO1 gene revealed considerably lower rates of genetic variability in comparison to ITS2.
Secondary structure analysis

- Structural model-based analyses of DNA sequence data have become increasingly important for comparative analysis. In the present study, ITS2 sequences were used in the secondary structure analysis. ITS2 raw sequences were annotated using ITS2 database and folded by minimum free energy method using Mfold and RNAfold web servers. For all the species under study, the predicted secondary structure confirmed the presence of a typical four-helix structure common to almost all eukaryotic taxa, showing Helix-III to be the longest with the presence of UGGU and other variant motifs in it and a pyrimidine-pyrimidine mismatch in Helix-II. Considerable structural variations in the secondary structure were also noticed among and even within the tapeworm species studied.

Phylogenetic analysis

- Phylogenetic trees were constructed by using Maximum likelihood (ML) and Bayesian Inference (BI) methods through RaxML and MrBayes softwares, respectively for each of the mentioned gene regions of the parasites under study. Both ML and BI phylogenetic trees were well resolved; however, the BI approach showed much better results with high Bayesian posterior probability (Bpp) values as supports for the nodes. Of the genetic markers, ITS2 and mtCO1 genes showed robust trees resolving all the taxa with significant bootstrap and Bpp values. However, the ITS2 gene showed poor taxon coverage and only a limited number of representative species. The 18S phylogeny also showed well-resolved trees for most of the taxa studied except for the *Echinococcus* species where all the taxa could not be well resolved. The phylogeny of *Hymenolepis* inferred through 28S gene was unable to resolve the phylogenetic tree.
Thus, for phylogenetic reconstruction of the cyclophyllidean cestodes species under study, the mtCO1 gene emerged as the reliable marker region of choice.

Phylogeny of Cyclophyllidea

The cestode Order Cyclophyllidea has a rich diversity of parasites and includes many families and species that are known to cause serious medical condition in humans and domestic and wild animals. Despite various attempts to resolve phylogenetic relationships at the inter-family level, uncertainty remains. In order to add resolution to the existing phylogeny of the order, the generated partial mtCO1 sequences were combined with those available from GenBank and phylogeny was inferred taking a total of 83 representative species spanning 8 key families using Bayesian analysis.

- The phylogenetic tree revealed Dilepididae as the most basal taxon and showed early divergence in the phylogenetic tree. Paruterinidae, Taeniidae and Anoplocephalidae showed non-monophyletic assemblage; the result suggested that the family Paruterinidae may represent a polyphyletic group.

- The diverse family Taeniidae appeared in two separate clades; while one of them included all the members of the genus *Echinococcus* and also *Versteria*; the representatives of the genera *Taenia* and *Hydatigera* clubbed in the other clade. A close affinity of Dipylidiidae with *Taenia* and *Hydatigera* was seen, whereas existence of a close relationship between Mesocestoididae and *Echinococcus* (of Taeniidae) was also demonstrated.
The crown group comprised the families Anoplocephalidae, Davaineidae, Hymenolepididae and Mesocestoididae, and also all species of the genus *Echinococcus* and *Versteria mustelae*. 

Monophyly of these families (excepting Anoplocephalidae) and the genus *Echinococcus* as well as its sister-taxon relation with *V. mustelae* is also confirmed.

Furthermore, non-monophyly of Anoplocephalidae is suggested to be correlated with divergence in the host selection.

**Molecular dating of zoonotic cyclophyllidean cestodes**

Being soft-bodied organisms, cestodes as such lack any fossil records, because of which molecular dating studies on this group of parasites are scanty or rare. In such cases, the only way to trace back the evolutionary history of tapeworms is to infer phylogeny using genetic markers. In an attempt to resolve evolutionary history of the zoonotic cestodes in the Order Cyclophyllidea and for tracing back their divergence, emergence and diversification times, the generated partial mtCO1 data were combined with data from GenBank and based on a Bayesian statistical platform, a time-calibrated phylogenetic analysis was carried out.

The results revealed that the zoonotic cyclophyllidean cestodes known today arose of a common ancestor possibly during the mid-Miocene epoch (18.4 Mya). The time estimates for the emergence of members of the genus *Hymenolepis* and lineages of *Echinococcus vogeli* and *E. oligarthus* were shown to be in accordance with the diversification of modern murine rodents that served as definitive hosts to these parasites.
Likewise, the results also corroborate co-appearance of the cestodes parasitizing canids (such as *E. granulosus*, *E. multilocularis* and *Dipylidium caninum*) and their definitive hosts.

Furthermore, the results showed that while diversification of *Taenia* species occurred earlier (14.3 Mya) in geological times, the association of *Taenia* with humans is only a recent event (4.9-1.3 Mya).

Additionally, the emergence of *T. solium*, *T. asiatica* and *T. saginata* lineages is revealed to be in agreement with the time of appearance of their respective host lineages.

➢ **Taxon/ species- specific primers**

To establish an easy and authentic tool for species discrimination and identification, the taxon/species-specific primers were designed using Primer3plus program, by targeting unique rDNA region spanning ITS2 region for four of the parasites under study.

The primers were designed to amplify the sequence bi-directionally, i.e., both in forward and reverse directions; these specific primers are TsF/TsR (*T. solium*); EgF/EgR (*E. granulosus*); HnF/HnR (*H. nana*) and HdF/HdR (*H. diminuta*).

The species-specific primers were tested against other parasitic species: Acanthocephala (*Pallisentis* sp.), cestodes (*Moniezia expansa*, *Taenia hydatigena*) and trematodes (*Fasciola gigantica*, *Clinostomum* sp.) that are prevalent in the region. The designed primers successfully amplified respective species DNA with negligible or no cross amplification, thus indicating their specificity.
• The result suggests that the designed species-specific primers can be used as a reliable diagnostic tool for species identification and discrimination in cases of multiple species infection.

➢ **DNA barcode**

• For platyhelminths, mitochondrial cytochrome oxidase c subunit 1 (mtCO1) is the gene of choice for barcoding. In the present study, DNA barcode could be generated for *T. solium, E. granulosus and H. diminuta* using the primer set Dice-1F and Dice-14R.

• The data was uploaded in the BOLD systems and the barcode index numbers acquired.

• The study is supported by citations of 257 references from the literature and includes 53 Tables and 60 figures.

❖ In conclusion, the study provides newer data on ITS2, which has been lacking for many of the cyclophyllidean cestodes (eg. *T. solium, M. expansa* and *E. granulosus*). Furthermore, the ITS2 secondary structure for all the six tapeworms has been described for the first time. Additionally, a comprehensive phylogenetic relationship among the various families in the Order Cyclophyllidea along with the molecular dating of the zoonotic tapeworms has also been discussed. The species-specific PCR primers designed and the barcodes generated under the present study provide additional tools for authentic diagnosis and taxon identification.
### Sequences submitted to GenBank

<table>
<thead>
<tr>
<th>Species</th>
<th>Locality</th>
<th>ITS2</th>
<th>18S</th>
<th>28S</th>
<th>CO1</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Taenia solium</em></td>
<td>Shillong, Meghalaya</td>
<td>KT899311</td>
<td>KX866813</td>
<td>_</td>
<td>KT213568</td>
</tr>
<tr>
<td><em>Taenia hydatigena</em></td>
<td>Shillong, Meghalaya</td>
<td>KX792106</td>
<td>KX388155</td>
<td>_</td>
<td>KR996173</td>
</tr>
<tr>
<td></td>
<td>Kohima, Nagaland</td>
<td>KX759100</td>
<td>KX827628</td>
<td>_</td>
<td>KT099323</td>
</tr>
<tr>
<td></td>
<td>Imphal, Manipur</td>
<td>KX781214</td>
<td>KX494910</td>
<td>_</td>
<td>KT099322</td>
</tr>
<tr>
<td><em>Echinococcus granulosus</em></td>
<td>Umdohlun, Meghalaya</td>
<td>KX013555</td>
<td>_</td>
<td>_</td>
<td>KR297266</td>
</tr>
<tr>
<td></td>
<td>Chirang, Assam</td>
<td>KX434757</td>
<td>KX388154</td>
<td>_</td>
<td>KR905550</td>
</tr>
<tr>
<td><em>Hymenolepis nana</em></td>
<td>Shillong, Meghalaya</td>
<td>KU748350</td>
<td>KX454312</td>
<td>KU748351</td>
<td>KU821727</td>
</tr>
<tr>
<td><em>Hymenolepis diminuta</em></td>
<td>Shillong, Meghalaya</td>
<td>KR349973</td>
<td>KX759099</td>
<td>KT026970</td>
<td>KT213567</td>
</tr>
<tr>
<td><em>Moniezia expansa</em></td>
<td>Agartala, Tripura</td>
<td>KX425620</td>
<td>_</td>
<td>_</td>
<td>KJ576906</td>
</tr>
</tbody>
</table>