The genus *Bertiella*, which has 29 known tapeworm species, belongs to the subfamily *Anoplocephalinae* of the *Anoplocephalidae* family (1). These tapeworms are common parasites in the small intestine of primates (2). Of these species, only *B. studeri*, *B. mucronata*, and *B. satyri* (3), which was recently redescribed as a different species (4), can infect humans (4,5). Children acquire this infection usually by eating contaminated fruits or by ingesting contaminated soil. The earliest identified cases of human bertiellosis in Sri Lanka occurred in 1975; these cases and 1 further case were reported in 1976. Six cases were reported in the literature from 1988–2006 (6). The most recent report was in 2006 from Rathnapura, Sabaragamuwa Province, Sri Lanka (7).

The morphologic, taxonomic, and molecular analysis of several species classified in the family *Anoplocephalidae* are not well documented (6). A recent study has identified an unexpected genetic diversity that suggests the existence of several *Bertiella* species in primates and humans (6,8). Multiple species of *Bertiella* tapeworms may infect humans in the New World and the Old World. It is not certain whether the Old World and New World *Bertiella* infections, previously all identified as *B. studeri* or *B. mucronata*, actually represent multiple different species; the true taxonomic distinction and geographic distribution of these 2 species are not entirely clear (6). Furthermore, diagnosis entirely based on egg morphology, size, and geographic distribution is insufficient to discriminate *B. studeri* tapeworms from other *Bertiella* spp infecting humans (9).

We provide a detailed molecular and phylogenetic description of *Bertiella studeri* tapeworms infecting children in Sri Lanka. Our findings can be used to identify multiple species of *Bertiella* tapeworms that can infect human hosts in the Old World.

This study provides the molecular analysis of the *B. studeri* tapeworms infecting children in Sri Lanka and describes phylogenetic relationships for this species. The Ethics Review Committee in the Faculty of Medicine, University of Peradeniya, Sri Lanka approved this study (protocol no. 2019/EC/03).

The Study

We conducted a retrospective study using tapeworm proglottids (Appendix Figure 1, https://wwwnc.cdc.gov/EID/article/26/8/20-0324-App1.pdf) from 24 pediatric patients referred to the Department of Parasitology, Faculty of Medicine, University of Peradeniya, Peradeniya, Sri Lanka, during 2007–2017. Patients were all <10 years of age (range 3.5–9 years). No other epidemiologic data were available.

We extracted genomic DNA separately using a commercial DNA extraction kit (PureLink; Invitrogen, https://www.thermofisher.com). We amplified 2 mitochondrial markers, nicotinamide adenine dinucleotide hydrogenase subunit 1 gene (NAD1) and cytochrome c oxidase subunit 1 gene (COX1), and 3 nuclear ribosomal markers, the second internal transcribed spacer region (ITS2), 28S large subunit ribosomal region (28S), and 18S rRNA gene (18S), using the specified primers and PCR conditions (Appendix Table). We subjected the PCR products to Sanger sequencing; only the ethanol-preserved samples provided a sufficient amount of DNA for sequencing (Appendix). We inferred molecular phylogenetic analysis and evolutionary history using maximum-likelihood and Bayesian methods.

Phylogenetic analysis identified a monophyletic group of *Bertiella* species in all 5 maximum-likelihood trees. The NAD1 region revealed several clades within the *Bertiella* monophyletic group (Figure 1, panel A); sequence similarity search identified 90.19% match with *Bertiella* species (GenBank accession no. JQ771111). The COX1 sequence similarity search identified 95.10% match with *Bertiella* species (GenBank accession no. JQ771106); COX1 analysis identified 2 clades in the *Bertiella* monophyletic
group (Figure 1, panel B). Bertiella species from human hosts, acquired in Equatorial Guinea and Argentina, and B. mucronata (New World) from Callcebus oenanthe monkeys were separated from the Sri Lanka clade (Figure 1, panel B).

The ITS2 sequences showed 99.35% similarity with B. studeri (GenBank accession no. AB586129) and 100% similarity with Bertiella species (GenBank accession no. JQ771096). All the B. studeri sequences from Asia are in 1 clade. The second clade included 3 sequences from Pan troglodytes chimpanzee in Kenya, 1 from a human infection acquired in Equatorial Guinea, and 1 from a human host in Brazil (Figure 2, panel A). The 28S rRNA gene analysis revealed 2 clades for Bertiella species. The sequence similarity search revealed 94.66% similarity with Bertiella species (GenBank accession no. KJ888951). Furthermore, we identified a single-nucleotide polymorphism in 28S rRNA region (T to C) between the samples from Sri Lanka that suggest genetic diversity (Appendix Figure 2, panel A). In the ML tree for 18S rRNA region, Sri Lanka samples and B. studeri obtained from Macaca fascicularis macaque formed a single clade (Figure 2, panel C). The sequence similarity search for 18S rRNA region identified 99.84% match with B. studeri (GenBank accession no. GU323706).

Furthermore, 18S rRNA region in the Sri Lanka samples have a single-nucleotide polymorphism (T to C) with the Bertiella sequence from M. fascicularis (Appendix Figure 2, panel B).

Records we examined showed patients had white, flat, motile worm segments in stools, and some patients had reported abdominal disturbances and intermittent diarrhea. Previous studies reported recurrent abdominal pain and continuous perianal itching, anorexia, weight loss, and intermittent diarrhea in infected patients (6); however, these symptoms are not unique to Bertiella infection, and so the correct diagnosis of bertiellosis is important. Treatment failure for B. studeri worms using niclosamide was reported in a 30-month-old patient in Sri Lanka in 2004 (10) and in a 5-year-old patient in Sabaragamuwa Province, Sri Lanka (7).

Conclusions
Our results suggest an intraspecific diversity of Bertiella tapeworms. Such diversity may occur according to the host and the geographic location. A previous study conducted by Doležalová et al. (8,11) has suggested a broad genetic diversity among the Bertiella species in primates and humans; further studies are required to support this suggestion. According to the available demographic data, most of the patients resided in Central
Province, Sri Lanka; the most likely reason that they comprised most patients is the Bertiella tapeworm reservoir hosts, particularly Ceylon torque monkey (Macaca sinica) and gray langur (Presbytis entellus), that inhabit this region (12,13). Over time, these monkey populations have lost their habitats due to deforestation and rapid urbanization in Sri Lanka; they are now regular visitors in suburban and urban areas scavenging for food near human settlements, which has increased human exposure to B. studeri infection (14,15).

Figure 2. Molecular phylogeny of the nuclear ribosomal markers in study of Bertiella tapeworms in children in Sri Lanka. Bold text indicates Bertiella studeri samples from Sri Lanka. A) Maximum-likelihood tree containing 17 taxa, constructed by the analysis of partial ITS2 sequence alignment. B) Maximum likelihood tree containing 24 taxa, constructed by the analysis of partial 28S sequence alignment. C) Maximum-likelihood tree containing 13 taxa, constructed by the analysis of partial 18S sequence alignment. Numbers above the nodes indicate the percentages of 1,000 nonparametric bootstrap pseudoreplicates (>70) and below the nodes the percentages of 1,000 Bayesian posterior probabilities (>70). GenBank accession numbers are provided for reference sequences. Scale bars represent nucleotide divergence.
Unavailability of molecular data for B. studeri 28S, COX1, and NAD1 markers in GenBank was a constraint that we encountered during phylogenetic analysis. In our study, we generated molecular data for 2 mitochondrial markers (NAD1, COX1), and 3 nuclear ribosomal markers (28S, 18S, ITS2) and submitted them to GenBank (Appendix). The molecular data obtained can be used for further analysis in diagnostics, to discern phylogenetic relationships and evolutionary correlations, and to understand the transmission dynamics of B. studeri tapeworms. Our data may also be used to assist in elucidating if multiple species of Bertiella sp. tapeworms infect human hosts in the Old World.

This study was funded by the University Research Grant (no. URG/2018/31/M), University of Peradeniya, Sri Lanka.

A.A. carried out the laboratory work, sequence analysis, and wrote the manuscript with input from all authors. T.H.L. carried out DNA sequencing. S.W. performed sequence analysis, manuscript writing, and finalized the manuscript. S.W. had the final responsibility for the decision to submit for publication. All authors reviewed the draft and approved the decision to submit for publication.

About the Author
Ms. Amarasinghe is a research assistant in the Department of Parasitology, Faculty of Medicine, University of Peradeniya, Sri Lanka. She is particularly interested in parasitology research.

References

Address for correspondence: Susiji Wickramasinghe, Department of Parasitology, Faculty of Medicine, University of Peradeniya, Peradeniya, 20400, Sri Lanka; email: susiji@pdn.ac.lk or susijjip@yahoo.co.jp
**Bertiella studeri** Infection in Children, Sri Lanka

**Appendix**

**PCR Conditions and Concentrations**

We analyzed nicotinamide adenine dinucleotide hydrogenase subunit 1 gene (NAD1), cytochrome c oxidase subunit 1 gene (COX1), and 3 nuclear ribosomal markers, the second internal transcribed spacer region (ITS2), 28S large subunit ribosomal region (28S), and 18S rRNA gene (18S). For all 5 markers, we conducted PCR in a final volume of 25 μL containing 4 μL 25mM MgCl2, 2 μL of 2.5 mM dNTPs, 2.5 μL of 10X PCR buffer, 0.5 μL of 5U/μL Taq DNA polymerase, 10 pmol μL each primer, and 5 μL of template DNA. For NAD1, the PCR conditions were used as in Littlewood et al., 2008 (1). For COX1, ITS2, 28S, and 18S, the PCR conditions were as follows: initial denaturation at 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 1 min, and the final extension at 72°C for 5 min. Then, PCR products were examined by running 5 μL of each product in a 1.5% agarose gel and subsequently stained with Diamond Nucleic Acid Dye (Promega, https://www.promega.com).

**Sequence Annotation**

Of the 24 samples, 22 gave positive results for at least 1 primer set. In this study, only the ethanol-preserved samples gave a sufficient amount of DNA to carry out the sequencing. Previous studies have shown that formalin preservation makes the tapeworm tissues unsuitable for DNA extraction (6,7). Sequence editing was carried out manually using BioEdit version 7.0.5.3 (8). The sequence similarity search was done using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignments were obtained using ClustalW version 2.0 (https://www.genome.jp/tools-bin/clustalw).

**Sequence Analysis**

We analyzed NAD1 sequence alignment and constructed a maximum-likelihood tree of the mitochondrial markers (Figure 1). The nucleotide sequence obtained for NAD1 was 891 bp.
long. The translated NAD sequence contained 296 aa and was submitted to the GenBank (accession no. MN 982427). COX1 sequences were submitted to GenBank under accession nos. MN982420, MN982421, MN982422, MN982423, MN982424, MN982425, MN982426). Lengths of these sequences were 387 bp, 393 bp, 386 bp, 386 bp, 394 bp, 393 bp, and 394 bp respectively.

Translated sequences in ExPASy (https://web.expasy.org/translate) contained 124 aa and all 7 sequences were similar to each other. Analysis of both mitochondrial markers was conducted by ML method in MEGA version 7.0.26 (https://www.megasoftware.net). Bayesian analysis was implemented in MrBayes (https://nbisweden.github.io/MrBayes), with the model GTR. Four Markov Chain Monte Carlo chains were run for 1 million generations. They were applied as 3 heated chains and 1 cold chain. The four chains reached burn-in time by 200,000 generations. The frequency of clades in trees was sampled for every 100 generations. Both ML and Bayesian trees had the same topologies.

We conducted molecular phylogeny of the nuclear ribosomal markers (Figure 2). In analysis of partial ITS2 sequence alignment, the sequences obtained did not show any nucleotide variance. We submitted sequences to GenBank (accession nos. MN982881, MN982882, MN982883, and MN982884); lengths were 240 bp, 221 bp, 213 bp, and 219 bp respectively.

Analysis of partial 28S sequence alignment yielded 2 sequences, the lengths of which were 789 bp and 791 bp. Sequences were deposited into GenBank (accession nos. MN982722–MN982723. C) Analysis of partial 18S sequence alignment yielded 4 sequences, the lengths of which were 642 bp, 642 bp, 642 bp, and 620 bp; sequences were deposited into GenBank (accession nos. MN982715, MN982716, MN982717, MN982718, MN982719, and MN982720.

We analyzed 3 nuclear ribosomal markers by maximum-likelihood using MEGA version 7.0.26 (https://www.megasoftware.net). Furthermore, we conducted Bayesian analysis in MrBayes (https://nbisweden.github.io/MrBayes/index.html), with the model general time reversible plus gamma 4 plus invariate sites. Four Markov Chain Monte Carlo chains were run for 1 million generations, applied as 3 heated chains and 1 cold chain. The 4 chains reached burn-in time by 250,000 generations. The frequency of clades in trees was sampled for every 100 generations. The maximum-likelihood and Bayesian trees had the same topologies.
References


Appendix Table. Amplified regions and the primers used for the PCR amplification in analysis of Bertiella sp. tapeworms in children, Sri Lanka

<table>
<thead>
<tr>
<th>Amplified region</th>
<th>Primer name*</th>
<th>Sequence (5’-3’)</th>
<th>Size of the amplicon, bp</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAD1</td>
<td>Cyclo-Nad1F (f)</td>
<td>GGN TAT TST CAR TNT CGT AAG GG TTC YTG AAG TTA ACA GCA TCA</td>
<td>730–900</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>Cyclo-tmNR (r)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX1</td>
<td>JB3 (f)</td>
<td>TTT TTT GGG CAT CCT GAG GTT TAT</td>
<td>393</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>JB4.5 (r)</td>
<td>TAA AGA AAC ATA ATG AAA ATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28S</td>
<td>LSU5 (f)</td>
<td>ACC CGC TGA AT1 TAA GCA</td>
<td>794</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>LSU5 (r)</td>
<td>TCC TGA GGG AAA CTT CGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18S</td>
<td>BF (f)</td>
<td>GGA CAC TAT GAG GAT TGA CAG A</td>
<td>600</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>BR (r)</td>
<td>CCT TTC GGG GCA CCA AGA TGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITS-2</td>
<td>3S (f)</td>
<td>CGG TGG ATC ACT CGG CTC GT</td>
<td>598 and 664</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>28A (r)</td>
<td>CCT GGT TAG TTT CTT TCT CTC CGC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* f, forward; r, reverse.

Appendix Figure 1. Morphological characteristics observed during the study of Bertiella spp. in children, Sri Lanka. A) Bertiella studeri proglottids. B) B. studeri eggs liberated from the proglottids.
**Appendix Figure 2.** Nucleotide sequence alignments showing SNPs in the 28S and 18S regions. A) alignment of the 2 sequences obtained for the 28S region. B) Alignment of the 6 sequences of 18S region Sri Lankan (SL) samples with *B. studeri* reference sequence (GenBank accession no. GU323706). Highlighting indicates SNPs in each alignment. Dots indicate identity with the base in the top line.