International Journal of Systematic and Evolutionary Microbiology (2007), **57**, 2412–2425 **DOI** 10.1099/ijs.0.64865-0

Barcoding ciliates: a comprehensive study of 75 isolates of the genus *Tetrahymena*

Chitchai Chantangsi,[1]† Denis H. Lynn,[1] Maria T. Brandl,[2] Jeffrey C. Cole,[3] Neil Hetrick[3] and Pranvera Ikonomi[4]

Correspondence
Denis H. Lynn
ddr@uoguelph.ca

[1]Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

[2]United States Department of Agriculture, Agriculture Research Service, Western Regional Research Center, Food Safety and Health Unit, 800 Buchanan St, Albany, CA 94710, USA

[3]Protistology Department, American Type Culture Collection, 10801 University Blvd, Manassas, VA 20110-2209, USA

[4]Molecular Authentication Resource Center, American Type Culture Collection, 10801 University Blvd, Manassas, VA 20110-2209, USA

The mitochondrial cytochrome-*c* oxidase subunit 1 (*cox1*) gene has been proposed as a DNA barcode to identify animal species. To test the applicability of the *cox1* gene in identifying ciliates, 75 isolates of the genus *Tetrahymena* and three non-*Tetrahymena* ciliates that are close relatives of *Tetrahymena*, *Colpidium campylum*, *Colpidium colpoda* and *Glaucoma chattoni*, were selected. All tetrahymenines of unproblematic species could be identified to the species level using 689 bp of the *cox1* sequence, with about 11 % interspecific sequence divergence. Intraspecific isolates of *Tetrahymena borealis*, *Tetrahymena lwoffi*, *Tetrahymena patula* and *Tetrahymena thermophila* could be identified by their *cox1* sequences, showing <0.65 % intraspecific sequence divergence. In addition, isolates of these species were clustered together on a *cox1* neighbour-joining (NJ) tree. However, strains identified as *Tetrahymena pyriformis* and *Tetrahymena tropicalis* showed high intraspecific sequence divergence values of 5.01 and 9.07 %, respectively, and did not cluster together on a *cox1* NJ tree. This may indicate the presence of cryptic species. The mean interspecific sequence divergence of *Tetrahymena* was about 11 times greater than the mean intraspecific sequence divergence, and this increased to 58 times when all isolates of species with high intraspecific sequence divergence were excluded. This result is similar to DNA barcoding studies on animals, indicating that congeneric sequence divergences are an order of magnitude greater than conspecific sequence divergences. Our analysis also demonstrated low sequence divergences of <1.0 % between some isolates of *T. pyriformis* and *Tetrahymena setosa* on the one hand and some isolates of *Tetrahymena furgasoni* and *T. lwoffi* on the other, suggesting that the latter species in each pair is a junior synonym of the former. Overall, our study demonstrates the feasibility of using the mitochondrial *cox1* gene as a taxonomic marker for 'barcoding' and identifying *Tetrahymena* species and some other ciliated protists.

†**Present address:** Department of Zoology, University of British Columbia, Biological Sciences Bldg, 6270 University Blvd, Vancouver, British Columbia V6T 1Z4, Canada.

Abbreviations: CVP, contractile vacuole pore; K2P, Kimura two-parameter; LSU, large subunit; NJ, neighbour-joining; PBG, polar basal granule; SSU, small subunit.

The GenBank/EMBL/DDBJ accession numbers for the *cox1* and SSU rDNA sequences determined in this study are EF070242–EF070328, as detailed in Supplementary Table S1.

Details of the strains examined in this study, including sequence accession numbers, details of the nucleotide compositions of the *cox1* and SSU rDNA sequences, values of overall, within-genus and between-genera divergence of datasets of the *cox1* sequences and alignments of *cox1* and SSU rDNA sequences are available as supplementary material with the online version of this paper.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14
64865 © 2007 IUMS *Printed in Great Britain*

INTRODUCTION

DNA-based identification has been proposed to serve as an alternative taxonomic approach for identifying the immense diversity of living organisms (Dawkins, 1998; Tautz *et al.*, 2002, 2003). Hebert *et al.* (2003a, b) suggested a 650 bp 5′ fragment of the mitochondrial cytochrome-*c* oxidase subunit 1 (*cox1*) gene as a universal marker or 'DNA barcode' for global biological identification of animal species. Use of this mitochondrial gene has a number of advantages. Firstly, among mitochondrial genes, the *cox1* gene is one of only two protein-encoding genes [the other encoding cytochrome *b* (*cytb*)] that are present in all eukaryotes. In addition, it functions homologously in a wide range of eukaryotes, enabling its sequence to be compared across a diverse array of organisms. Secondly, cytochrome-encoding genes such as *cox1* and *cytb* are more variable than other commonly studied mitochondrial genes, such as 12S RNA and tRNA genes (Janczewski *et al.*, 1995; Kumazawa & Nishida, 1993). Therefore, these protein-encoding genes can be used to discriminate closely related species (Hebert *et al.*, 2003b). Thirdly, the *cox1* gene of various animal phyla is easily amplified using universal primers designed from conserved regions of the gene (Folmer *et al.*, 1994). Fourthly, because mitochondria reproduce by binary fission and without sexual recombination, their genes are less subject to insertions, deletions or other large-scale rearrangements that introduce more ambiguous variation into the sequence. Finally, it has long been recognized that the mitochondrial genome evolves at a faster rate than the nuclear genome and therefore mitochondrial genomic sequences at a particular region will be more informative in differentiating or distinguishing closely related species (Hebert *et al.*, 2003b). Preliminary studies have proven this gene to be a good taxonomic marker for discriminating species of animals such as gastropods (Remigio & Hebert, 2003), collembolans (Hogg & Hebert, 2004), mayflies (Ball *et al.*, 2005), flies (Scheffer *et al.*, 2006), moths (Brown *et al.*, 2003), butterflies (Hajibabaei *et al.*, 2006; Hebert *et al.*, 2004a), beetles (Monaghan *et al.*, 2005), ants (Smith *et al.*, 2005), spiders (Barrett & Hebert, 2005), fishes (Ward *et al.*, 2005) and birds (Hebert *et al.*, 2004b).

Among living organisms, protists have long been recognized as an assemblage of organisms of complex forms and with polymorphic life histories. A majority of them are microscopic, and specific staining procedures and electron microscopy are often required in order to reveal key features for taxonomic identification (Corliss & Daggett, 1983). So far, a DNA barcoding approach using *cox1* gene sequences has been applied to identify only a few groups of protists, such as red algae and some ciliate genera (Barth *et al.*, 2006; Lynn & Strüder-Kypke, 2006; Saunders, 2005). However, among those ciliates examined, only a few species within two genera, *Paramecium* and *Tetrahymena*, were investigated.

To examine the usefulness of the *cox1* barcode for ciliate species identification, species of the genus *Tetrahymena* were investigated. *Tetrahymena* includes a number of closely related species, both sexual outbreeders and asexual forms. Previous identification of *Tetrahymena* species has been based on several approaches such as morphology, a combination of ecology and life histories, mating tests, isozyme mobilities and PCR-RFLP (Corliss, 1973; Czapik, 1968; Holz & Corliss, 1956; Jerome & Lynn, 1996; Meyer & Nanney, 1987; Nanney & McCoy, 1976; Nyberg, 1981). However, because of the close relatedness of some members within the genus, species can often not be discriminated using morphological features, even at the ultrastructural level (Corliss & Daggett, 1983). Furthermore, several *Tetrahymena* isolates have long been known that share morphological similarities but are genetically isolated from each other and were initially named as different syngens, and they have proven to be difficult to discriminate without mating tests (Gruchy, 1955). In addition, some *Tetrahymena* species show phenotypic plasticity in response to different environmental conditions during their polymorphic life cycles, making them even more difficult to identify morphologically and ecologically (Corliss, 1973; Strüder-Kypke *et al.*, 2001).

Among identification approaches, mating reactivity remains the 'gold standard' to determine conspecificity and to discover new species of *Tetrahymena*. Using this method, a large number of cryptic species were discovered within the *Tetrahymena pyriformis* species complex (Elliott, 1970; Gruchy, 1955; Nanney & McCoy, 1976). However, a complete set of living reference strains is not available and the approach is impossible for amicronucleate strains, which have been known to be common in *Tetrahymena* (Nanney & McCoy, 1976). With the advent of molecular techniques, several approaches, such as isozyme mobilities and RFLP, were used to discriminate *Tetrahymena* species without requiring mating tests. Borden *et al.* (1977) reported that the most closely related syngens shared an isozyme similarity coefficient of 67 %, while the distantly related ones had a coefficient of 0. This criterion was used to support the establishment of several new *Tetrahymena* species (Nanney & McCoy, 1976). However, these identification approaches are now known to have their own limitations: RFLP patterns are identical for several species pairs (Jerome & Lynn, 1996), while similar polymorphisms among species are shown for isozyme mobilities (Meyer & Nanney, 1987; Nanney *et al.*, 1998).

Thus, the *cox1* DNA barcoding approach, which has proven to be useful for identifying species of animals, was chosen to identify *Tetrahymena* species. Since few DNA barcoding studies have been done on protists so far, it is uncertain whether the 650 bp 5′ region of the *cox1* gene, which works well taxonomically for the discrimination of animal species, would be suitable for differentiating ciliated protists. Therefore, in this study, we amplified and sequenced almost the entire length of the *cox1* gene of representatives of 36 *Tetrahymena* species and six wild isolates of undescribed *Tetrahymena* species as well as three non-*Tetrahymena* ciliates that are close allies of

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

Tetrahymena, *Colpidium campylum*, *Colpidium colpoda* and *Glaucoma chattoni*. This sampling covered all valid species of this ciliate genus that have been described so far and whose cultures are available. This enabled us to evaluate which part of the gene was the most appropriate diagnostic barcoding region for identifying species of ciliates. In addition, 33 strains of six different *Tetrahymena* species isolated from different geographical localities were examined to demonstrate that levels of genetic variation within species are sufficiently low to ensure unambiguous identification with the *cox1* barcode.

METHODS

Source of samples. Seventy-eight isolates of tetrahymenine ciliates were examined in this study (see Supplementary Table S1 in IJSEM Online). Of 42 taxonomically valid species of *Tetrahymena* (Fenchel & Finlay, 2004), 36 are available in culture and at least one isolate of each of these species was selected as a species representative. These species were *Tetrahymena americanis*, *T. asiatica*, *T. australis*, *T. bergeri*, *T. borealis*, *T. canadensis*, *T. capricornis*, *T. caudata*, *T. corlissi*, *T. cosmopolitanis*, *T. elliotti*, *T. empidokyrea*, *T. farleyi*, *T. furgasoni*, *T. hegewischi*, *T. hyperangularis*, *T. leucophrys*, *T. limacis*, *T. lwoffi*, *T. malaccensis*, *T. mimbres*, *T. mobilis*, *T. nanneyi*, *T. nipissingi*, *T. paravorax*, *T. patula*, *T. pigmentosa*, *T. pyriformis*, *T. rostrata*, *T. setosa*, *T. shanghaiensis*, *T. silvana*, *T. sonneborni*, *T. thermophila*, *T. tropicalis* and *T. vorax*. The remaining six valid *Tetrahymena* species, *Tetrahymena chironomi*, *T. dimorpha*, *T. edaphoni*, *T. rotunda*, *T. sialidos* and *T. stegomyiae*, were not examined because of the unavailability of their cultures. In addition, six isolates of undescribed *Tetrahymena* species and three other tetrahymenine species, *C. campylum*, *C. colpoda* and *G. chattoni*, were also investigated. Furthermore, among the *Tetrahymena* species, additional isolates of *T. borealis* (4), *T. lwoffi* (1), *T. patula* (1), *T. pyriformis* (4), *T. thermophila* (3) and *T. tropicalis* (4) were examined to assess intraspecific sequence divergence. Moreover, the 16 *cox1* sequences available from GenBank for one *T. pyriformis* and 15 *T. thermophila* isolates were also included. Detailed information is provided in Supplementary Table S1 for all isolates, including catalogue numbers, designations, localities and GenBank accession numbers.

Culture methods and maintenance. All tetrahymenine ciliates except for the two species of *Colpidium* were cultured axenically in 10 ml sterile proteose peptone yeast extract (PPYE) medium (0.5 g glucose, 2.0 g proteose peptone, 2.0 g yeast extract and 400 ml distilled water). The two species of *Colpidium* were cultured in 10 ml bacterized dried cereal grass leaves (Cerophyl) medium, prepared as follows. Cerophyl (20 g) was added to 400 ml distilled water and boiled for 10 min. Distilled water was added to compensate for evaporation and the medium was filtered through Whatman no. 1 filter paper. An aliquot of 10 ml of this concentrated medium was transferred into culture test tubes and autoclaved for 20 min at 121 °C. Ten millilitres of the concentrated Cerophyl medium was diluted with 200 ml distilled water in a sterile flask. Using sterile technique, *Enterobacter aerogenes* was inoculated into the diluted Cerophyl medium. The bacterized Cerophyl medium was left at room temperature for 24 h prior to inoculation with *Colpidium* spp. All ciliate cultures were transferred biweekly and maintained at room temperature using sterile technique.

DNA extraction, amplification and sequencing. In contrast to dead cells and cell debris, which settle on the bottom of the culture tube, healthy cells swim actively just below the surface of the medium. Thus, about 1 ml of this portion of the culture medium containing ciliates was placed in a 1.5 ml microcentrifuge tube. The cells were pelleted by centrifugation at 11 000 ***g*** for 5 min. Two DNA extraction protocols, using Chelex beads as described by Walsh *et al.* (1991) and total nucleic acid purification as specified by EPICENTRE, were performed to yield total genomic DNA of the ciliates. An appropriate amount of DNA template was used in amplification of the *cox1* gene and the small-subunit (SSU) rDNA using puReTaq Ready-To-Go PCR beads (GE Healthcare).

***cox1* gene.** Approximately 2000 bp of the mitochondrial *cox1* gene were amplified by PCR using the forward primer 5′-ATGTGAGTTG-ATTTTATAGAGCAGA-3′ and the reverse primer 5′-GGDATA-CCRTTCATTTT-3′, which were newly designed in this study. The thermal cycler was programmed as follows: hold at 94 °C for 4 min; 5 cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 1 min and extension at 72 °C for 105 s; 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 1 min and extension at 72 °C for 105 s; and hold at 72 °C for 10 min. PCR products corresponding to the expected size were separated by agarose gel electrophoresis, purified using the GENECLEAN kit (Qbiogene) and sequenced in both the forward and reverse directions with an ABI 3730 DNA Analyser using the standard BigDye Terminator version 3.1 cycle-sequencing kit using amplification and internal primers. Almost the entire length, about 1821 bp, of the *cox1* gene of 45 species representatives was sequenced, while at least 689 bp of the 5′ region of the *cox1* gene of 17 additional intraspecific isolates of six *Tetrahymena* species mentioned previously were sequenced to assess the diagnostic barcoding region.

SSU rDNA. Approximately 2900 bp of an rDNA fragment were amplified using the forward primer A (5′-CAACCTGGTTGATC-CTGCCAGT-3′) and the reverse primer C (5′-TTGGTCCGTGTTT-CAAGACG-3′) (Jerome & Lynn, 1996; Medlin *et al.*, 1988). The PCR product included the SSU rDNA, internal transcribed spacer (ITS) 1, the 5.8S rDNA, ITS2 and a portion of the large-subunit (LSU) rDNA. The thermal cycler was programmed as follows: hold at 94 °C for 4 min; 35 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 90 s and extension at 72 °C for 3 min; and hold at 72 °C for 10 min. PCR products corresponding to the expected size were processed as described above. However, only a 1800 bp region of the SSU rDNA was sequenced and included in further analyses.

The SSU rDNA was amplified and sequenced for 19 tetrahymenine species that were chosen as species representatives and examined in this study but which have not yet been investigated in any previous studies. These species were *T. americanis*, *T. asiatica*, *T. caudata*, *T. cosmopolitanis*, *T. elliotti*, *T. furgasoni*, *T. leucophrys*, *T. limacis*, *T. lwoffi*, *T. mimbres*, *T. nipissingi*, *T. paravorax*, *T. shanghaiensis*, *T. silvana*, *T. sonneborni*, *Tetrahymena* sp. 1 (Foissner), *Tetrahymena* sp. 3 (RA9), *Tetrahymena* sp. 5 (NI) and *C. colpoda*. In addition, the SSU rDNA of two isolates of *T. pyriformis* and four isolates of *T. tropicalis* were also amplified and sequenced. Detailed information on these species and isolates is given in Supplementary Table S1.

Sequence analyses. The *cox1* and SSU rDNA sequences were first aligned automatically by CLUSTAL W (Thompson *et al.*, 1994) using the MEGA program version 3.1 (Kumar *et al.*, 2004) and then further refined by eye. The alignments for *cox1* and SSU rDNA sequences of the examined tetrahymenine ciliates are available as supplementary material in IJSEM Online.

The *cox1* gene sequences obtained from this study were 1821 bp in length except for that of *G. chattoni*, which was 1785 bp long. The 1821 nucleotides span positions 52–1872 with reference to the complete *cox1* genes of *T. pyriformis* and *T. thermophila* published in GenBank (Brunk *et al.*, 2003; Burger *et al.*, 2000).

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

These *cox1* sequences were then divided into the following five datasets.

Dataset 1. This dataset comprised the 1821 bp sequences of the *cox1* genes from a representative of each species, for a total of 45 sequences. Thirty-six positions (1786–1821) at the 3′-end of the *cox1* gene sequence of *G. chattoni* were treated as missing data.

Dataset 2. This dataset comprised only the 5′ half of the 1821 bp sequences of the *cox1* genes from a representative of each species, for a total of 45 sequences. This region is 912 bp in length.

Dataset 3. This dataset comprised only the 3′ half of the 1821 bp sequences of the *cox1* genes from a representative of each species, for a total of 45 sequences. This region is 909 bp in length. Thirty-six positions (874–909) at the 3′-end of the *cox1* gene sequence of *G. chattoni* were treated as missing data.

Dataset 4. This dataset comprised the 689 bp sequences of the *cox1* genes from a representative of each species, for a total of 45 sequences. This region starts at position 169 with reference to the 1821 sites of our *cox1* gene sequences and at position 220 with reference to the 2067 sites of the complete *cox1* gene sequence. Compared to animal barcoding regions, the starting position of this 689 bp sequence is in close proximity to the beginning site of the diagnostic barcoding region used for identifying animal species.

Dataset 5. This dataset comprised the 689 bp sequences of the *cox1* genes from all 78 isolates that were included in this study. This is the same region as dataset 4.

The SSU rDNA of 19 tetrahymenine species and of six isolates was newly sequenced in this study. In addition, another 24 species whose SSU rDNA sequences were available at GenBank were also included in this study. These species are *T. australis*, *T. bergeri*, *T. borealis*, *T. canadensis*, *T. capricornis*, *T. corlissi*, *T. empidokyrea*, *T. farleyi*, *T. hegewischi*, *T. hyperangularis*, *T. malaccensis*, *T. mobilis*, *T. nanneyi*, *T. patula*, *T. pigmentosa*, *T. pyriformis*, *T. rostrata*, *T. setosa*, *T. thermophila*, *T. tropicalis*, *T. vorax*, *Tetrahymena* sp. 6 (Brandl), *C. campylum* and *G. chattoni*. Moreover, full sequences of the SSU rDNA of *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 4 (SIN), kindly provided by Dr Michaela C. Strüder-Kypke (Department of Integrative Biology, University of Guelph, Canada), were included in our study. Detailed information on these species and isolates is provided in Supplementary Table S1. Aligned SSU rDNA sequences of all 45 species representatives were constructed as a dataset for sequence divergence calculation. This nucleotide dataset was 1650 positions in length and included a few gaps and some ambiguous characters. Exclusion of these gaps and ambiguous positions gave a total of 1639 sites, which were used in sequence divergence analyses. In addition, SSU rDNA sequences of two *T. pyriformis* isolates and four *T. tropicalis* isolates were added to the 45-sequence tetrahymenine SSU rDNA alignment to illustrate phylogenetic relationships among these organisms.

Sequence divergences were calculated for each dataset using the Kimura two-parameter (K2P) distance model (Kimura, 1980). A neighbour-joining (NJ) phylogenetic tree was inferred from genetic distances of *cox1* sequences calculated by DNADIST with the K2P model of sequence evolution using PHYLIP version 3.65 (Felsenstein, 2004; Saitou & Nei, 1987). Using *G. chattoni* as the outgroup species, this tree-building approach was performed on the dataset of 689 bp *cox1* gene sequences of 78 tetrahymenine ciliates and the dataset of 1639 bp SSU rDNA sequences of 51 tetrahymenine ciliates to provide an illustration of the patterning of divergence between and within species. In addition, 1000 bootstrap resamplings were carried out using SEQBOOT to determine confidence levels of deduced relationships. CONSENSE within the PHYLIP package was used to construct a consensus tree.

RESULTS AND DISCUSSION

With tetrahymenine *cox1*-specific primers, the *cox1* gene was amplified and recovered for all tetrahymenine ciliates that were examined in this study. However, because of considerable divergence in the *cox1* gene of *G. chattoni*, as indicated by its high percentage sequence divergence, our internal primers that were used to sequence toward the 3′ region of the gene could not sequence that area, leaving 36 bp unavailable. Therefore, only 1785 bp of the *cox1* gene was obtained from *G. chattoni*. Alignment of the entire length of 1821 bp of *cox1* sequences of all 45 species representatives was straightforward as there were no insertions and no deletions, corroborating research on other organisms (Hebert *et al.*, 2003b; Mardulyn & Whitfield, 1999; Ward *et al.*, 2005).

Nucleotide composition

Mean frequencies of thymine (T), cytosine (C), adenine (A) and guanine (G) and molar percentage G+C content for the SSU rDNA of examined tetrahymenines are in the same range as in other ciliates (Elwood *et al.*, 1985; Sogin & Elwood, 1986; Schlegel *et al.*, 1991). Mean base frequencies and G+C and A+T contents of the *cox1* gene for each genus were similar in five different datasets of the gene, as listed in Supplementary Table S2 in IJSEM Online. Relative nucleotide frequencies for the *cox1* genes of the examined tetrahymenine species were 38.4–43.1 mol% for T, 11.0–13.6 mol% for C, 30.5–33.6 mol% for A and 13.9–15.6 mol% for G, yielding mean G+C and A+T contents of 25.2–29.2 and 70.8–74.8 mol%, respectively. The G+C content of the 1821 bp *cox1* gene sequences was similar among the three tetrahymenine genera, with a mean of 26.7 mol%, and lower than that of the *cox1* gene of *Paramecium aurelia*, which is 41.8 mol% (Burger *et al.*, 2000). The G+C content of *cox1* genes of tetrahymenine ciliates is considerably different from that of animal *cox1* genes: the G+C content of several orders of insects is about 35 mol% (Hebert *et al.*, 2003b), while that of 207 species of Australian fish ranges from 42 to 47 mol% (Ward *et al.*, 2005). This G+C content is also characteristic of the nucleotide composition of the entire mitochondrial genomes of *T. pyriformis* and *T. thermophila*, which show low G+C contents, of 21.3 and 20.7 mol%, respectively (Brunk *et al.*, 2003; Burger *et al.*, 2000), considerably lower than the 41.2 mol% G+C of the *P. aurelia* mitochondrial genome (Cummings, 1992). A search was also carried out for stop codons of the ciliate mitochondrial genomes, UAA and UAG, which encode glutamine in the ciliate nuclear genomes; UGA was excluded as it encodes tryptophan in ciliate mitochondrial genomes (Brunk *et al.*, 2003). These stop codons were not observed in any of the amplified *cox1* gene sequences, demonstrating that these were fully functional mitochondrial *cox1* gene sequences.

Sequence analyses for *cox1* and SSU rDNA

Sequence analyses for the *cox1* gene. Mean sequence divergences of all 45 species representatives analysed from

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

Table 1. Overall, within-genus and between-genera sequence divergence based on the K2P model

Results from comparisons of the 689 bp *cox1* barcoding dataset and the 1639 bp SSU rDNA dataset of 45 tetrahymenine species representatives examined in this study are shown. Genus T, 42 species of *Tetrahymena*; genus C, two species of *Colpidium*; genus G, one species of *Glaucoma*. For the SSU rDNA, sequence divergence was calculated with pairwise or complete deletion in effect, as shown. NC, Not calculated.

Gene	Nucleotide positions (*n*)	Sequence divergence (%) (mean ± SEM)						
		Overall	Within-genus			Between-genera		
			Genus T	Genus C	Genus G	Genera T and C	Genera T and G	Genera C and G
cox1	689	11.13 ± 0.78	10.47 ± 0.74	12.82 ± 1.46	NC	13.13 ± 0.98	20.31 ± 1.59	21.19 ± 1.80
SSU rDNA	1639 (pairwise)	1.71 ± 0.16	1.39 ± 0.15	3.41 ± 0.45	NC	4.05 ± 0.42	3.48 ± 0.42	3.85 ± 0.39
	1639 (complete)	1.56 ± 0.16	1.24 ± 0.16	3.26 ± 0.44	NC	3.87 ± 0.42	3.41 ± 0.37	3.76 ± 0.42

three different *cox1* datasets, including 1821 bp, the 912 bp 5′ half and the 909 bp 3′ half, were closely similar, ranging from 10.59 ± 0.61 to 11.14 ± 0.52 % (mean ± SEM), as shown in Supplementary Table S3. Starting at position 169 with reference to 1821 sites, about 689 bp in the 5′ half of the *cox1* gene were selected as the diagnostic barcoding region. The first portion of about 240 bp and the last portion of about 90 bp in this 689 bp region are comparable to the beginning part and the middle part of 650 bp animal barcodes, respectively. However, the middle portion of about 360 bp of the 689 bp tetrahymenine barcode is an insert unique to the ciliate mitochondrial *cox1* gene (Ziaie & Suyama, 1987). The mean sequence divergence value calculated from the 689 bp barcoding region between species pairs for the 45 species representatives was 11.13 ± 0.78 % (Table 1). This divergence value was very similar to those of the three other regions, indicating that any parts of the *cox1* gene can be used as a diagnostic barcode to discriminate between the tetrahymenine species examined. Furthermore, 983 of 990 interspecific pairwise comparisons based on the barcoding region showed more than 2 % sequence divergence (Fig. 1). Among these pairwise comparisons, the following seven species pairs showed apparently low values, ranging from 0 to 1.17 %: *T. canadensis* ATCC® 30368™ and *T. rostrata* ATCC® 30770™ (0.73 %), *T. furgasoni* ATCC® 30006™ and *T. lwoffi* (1630/1G) (0 %), *T. nanneyi* ATCC® 50071™ and *T. nipissingi* ATCC® 30837™ (0.29 %), *T. pyriformis* ATCC® 30005™ and *T. setosa* ATCC® 30782™ (0 %), *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 3 (RA9) (0.15 %), *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 5 (NI) (1.17 %) and *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI) (1.02 %) (Fig. 1). Analyses of percentage sequence divergences of 1821 bp, the 912 bp 5′ half and the 909 bp 3′ half for these species pairs showed similar results. However, there were no differences in sequence divergence values in the 5′ and 3′ halves between *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 3 (RA9) (both 0.22 ± 0.16 %) and between *T. furgasoni* ATCC® 30006™ and *T. lwoffi* (1630/1G) (both 0 %).

All 42 species of *Tetrahymena*, two species of *Colpidium* and *G. chattoni* were grouped by genus. Within-genus sequence divergence values for *Tetrahymena* and *Colpidium* were approximately 10 and 13 %, respectively, in all four different *cox1* datasets. Calculation of between-genera sequence divergence based on *cox1* gene sequences showed that *Tetrahymena* is more genetically related to *Colpidium* than to *Glaucoma* (Table 1 and Supplementary Table S3).

Calculated from the 689 bp barcoding region, mean values (± SEM) for sequence divergences among five isolates of *T. borealis*, two isolates of *T. lwoffi*, two isolates of *T. patula*, six isolates of *T. pyriformis*, 19 isolates of *T. thermophila*

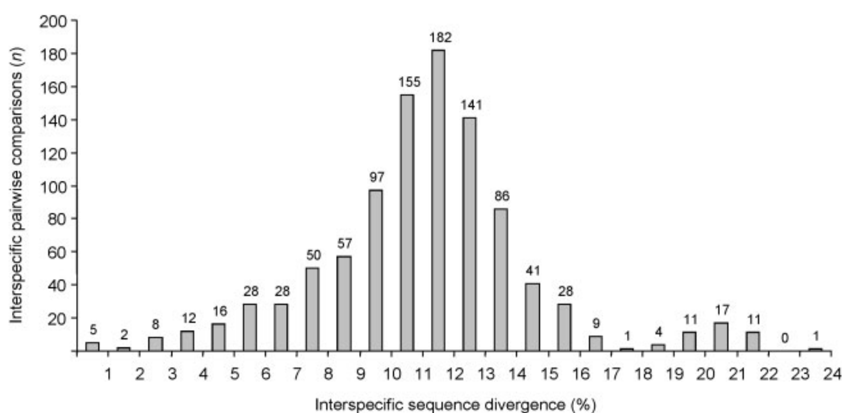

Fig. 1. Distribution of interspecific pairwise sequence divergence based on the 689 bp *cox1* barcoding region of 45 tetrahymenine species representatives, showing that a majority of pairwise comparisons have more than 2 % sequence divergence.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

and five isolates of *T. tropicalis* ranged from 0 to 9.07 ± 0.83 % (Fig. 2). Among these six species, conspecific isolates of *T. borealis*, *T. lwoffi*, *T. patula* and *T. thermophila* showed unproblematic results in that they showed <1 % intraspecific sequence divergence, whereas some conspecific isolates of *T. pyriformis* and *T. tropicalis* showed considerably higher sequence divergences, leading to suspicion of the validity of cultures and/or the status of the isolates of these species.

The mean sequence divergence of individuals within species was 0.95 %, compared with 10.47 % between species within the genus *Tetrahymena* (Table 1). Therefore, there is about 11-fold more variation among congeneric species than among conspecific isolates. Exclusion of all isolates of species with high intraspecific sequence divergences (i.e. *T. pyriformis* and *T. tropicalis*) reduced this mean intraspecific sequence divergence to 0.18 %. This increased the value to about 58.2-fold more variation among congeneric isolates than among conspecific ones. In previous 'barcode of life' studies, a difference in relative sequence divergence of about an order of magnitude is typical between intra- and interspecific comparisons. For example, congeneric sequence divergences are 18 and 25 times greater than conspecific sequence divergences for North American birds (Hebert *et al.*, 2004b) and Australian fishes (Ward *et al.*, 2005), respectively.

The within-species groupings showed a low degree of genetic distance, as represented by short branch length in the 689 bp *cox1* NJ tree, and were generally supported by 100 % bootstrap values (Fig. 3). However, conspecific isolates of *T. pyriformis* and *T. tropicalis* showed high sequence divergence values, explaining the interspersion of isolates of these two species on the NJ tree (Fig. 3). Four of six isolates of *T. pyriformis* were clustered together with 70.6 % bootstrap support, but separately from the other two isolates (Fig. 3). In addition, *T. setosa* ATCC® 30782™ and *T. tropicalis* ATCC® 205060™ were also clustered with this *T. pyriformis* clade with 100 % bootstrap support in the NJ analyses (Fig. 3). The remaining two *T. pyriformis* isolates, *T. pyriformis* ATCC® 205038™ and *T. pyriformis* ATCC® 205062™, were grouped with low bootstrap support with the main *T. pyriformis* clade, but on a separate well-supported clade with 79.8 % bootstrap support (Fig. 3). Moreover, the terminal branches of these two *T. pyriformis* isolates appeared to be long, showing larger genetic distances and suggesting the possibilities of their status as novel species.

Interspersion of the other four *T. tropicalis* isolates was nearly complete: *T. tropicalis* ATCC® 205083™ and *T. tropicalis* ATCC® 205097™ were clustered with *Tetrahymena* sp. 2 (CO), *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI) with 99.4 % bootstrap support; *T. tropicalis* ATCC® 30276™ was grouped with *T. mobilis* ATCC® PRA-174™ with 87.6 % bootstrap support; and, finally, *T. tropicalis* ATCC® 205156™ was placed close to *T. empidokyrea* ATCC® 50595™ and near *G. chattoni* (Fig. 3).

Sequence divergence analyses of different portions of the 1821 bp region showed that the 689 bp 5′-region could be used effectively as a DNA barcode for identifying *Tetrahymena* spp., as supported by very similar values of percentage sequence divergence calculated from the four regions of the *cox1* gene. A comparable region of the *cox1* gene has been used extensively as a diagnostic barcoding region in identifying species for several animal groups, as noted above.

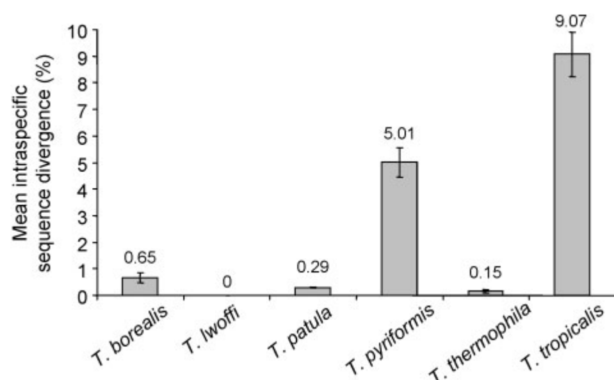

Fig. 2. Intraspecific sequence divergence (%, mean ± SEM) of *T. borealis*, *T. lwoffi*, *T. patula*, *T. pyriformis*, *T. thermophila* and *T. tropicalis* calculated from the 689 bp *cox1* barcoding region. Intraspecific isolates of *T. pyriformis* and *T. tropicalis* showed considerably higher sequence divergence values when compared with those of the remaining four species, suggesting the presence of cryptic species or misidentification or contamination of cultures.

Sequence analyses for SSU rDNA. The 1639 bp of SSU rDNA were included in the sequence divergence analysis. Calculation of 1599 positions of the gene, excluding gaps and missing data, for 45 species representatives showed a mean sequence divergence of 1.56 ± 0.16 %. However, the mean value of sequence divergence for the 1639 sites increased slightly to 1.71 ± 0.16 % when gaps and missing data were deleted in a pairwise manner (Table 1). The sequence divergence values calculated from the SSU rDNA were considerably lower than those obtained from the *cox1* gene in every analysis (Table 1). A majority of interspecific pairwise comparisons based on SSU rDNA showed 0–2 % sequence divergence (Fig. 4). Except for the two species pairs *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI), which both showed 0.06 % sequence divergence in SSU rDNA, the other five species pairs that showed a low degree of sequence divergence based on *cox1* gene sequences had 0 % sequence divergence between their SSU rDNAs.

Certain other species pairs could not be discriminated using the SSU rDNA sequences, showing 0 % sequence divergence (Fig. 4). However, these species pairs showed some degree of sequence divergence based on *cox1*

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

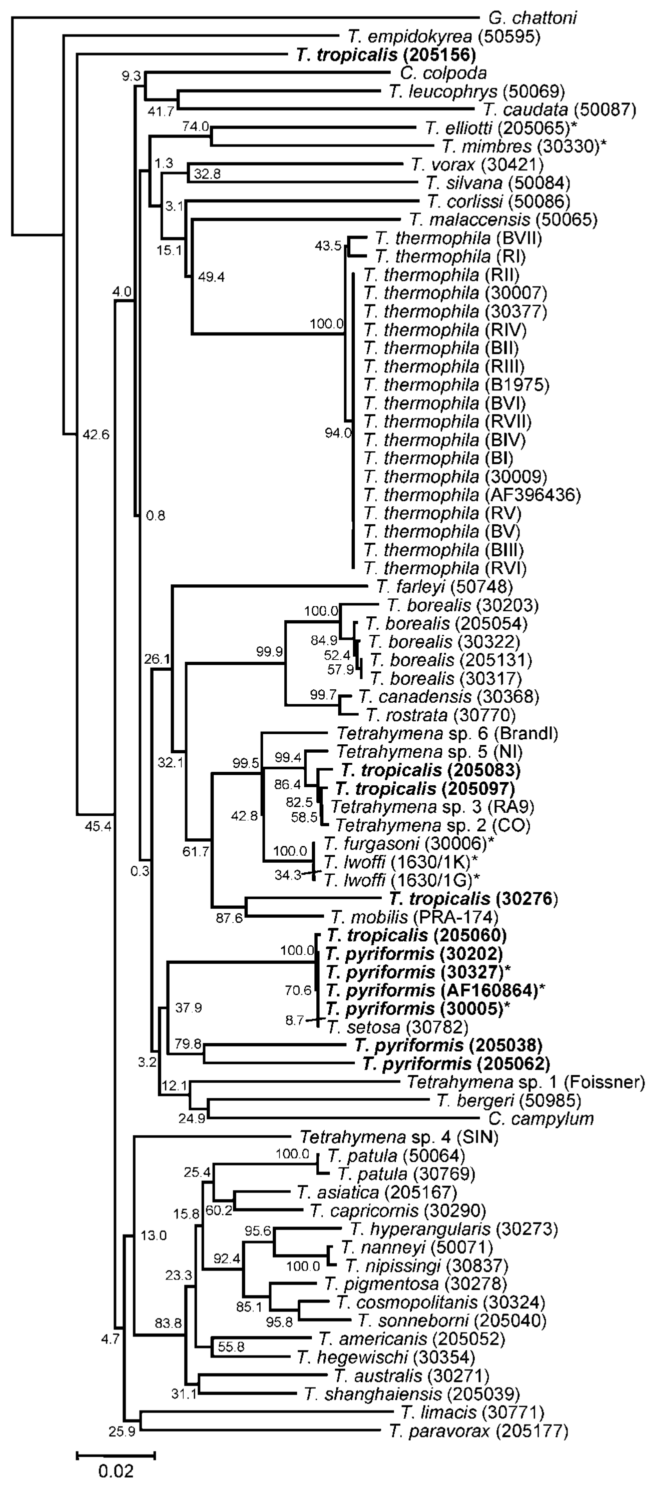

Fig. 3. NJ tree inferred from 689 bp of the diagnostic barcoding region of the *cox1* gene for all 78 tetrahymenine species isolates. Genetic distance calculation was based on the K2P model of sequence evolution. Branch lengths separating taxa represent genetic distances. Bar, 0.02 nucleotide substitutions per site. Percentage bootstrap support for taxa within the *T. pyriformis* and *T. thermophila* clades, which are not shown in the figure, were typically lower than 20 and 5 %, respectively. Values are percentages of bootstrap support from 1000 resamplings for the NJ analysis. Because of the interspersion of intraspecific isolates of *T. pyriformis* and *T. tropicalis*, those strains are labelled in bold. Amicronucleate species and strains (if known) are labelled with asterisks (*).

sequences, such as 5.30 % between *T. americanis* ATCC® 205052™ and *T. hegewischi* ATCC® 30823™, 4.70 % between *T. asiatica* ATCC® 205167™ and *T. capricornis* ATCC® 30291™, 4.01 % between *T. borealis* ATCC® 205012™ and *T. canadensis* ATCC® 30368™ and 4.43 % between *T. cosmopolitanis* ATCC® 30324™ and *T. hyperangularis* ATCC® 30273™. This result indicated that SSU rDNA sequences are not sufficiently variable to differentiate tetrahymenine ciliates, especially in closely related species, and also confirmed the usefulness of the *cox1* gene in identifying *Tetrahymena* species.

Comparison of intraspecific sequence divergence in *Tetrahymena* and other organisms

The low intraspecific sequence divergence values for all four unproblematic *Tetrahymena* species examined (*T. borealis*, *T. lwoffi*, *T. patula* and *T. thermophila*) are similar to those calculated for the barcoding region of various groups of animals. For example, mean values of within-species sequence divergence are 0.39 % in Australian fishes (Ward *et al.*, 2005), 0.43 % in North American birds (Hebert *et al.*, 2004b), 1 % in mayflies (Ball *et al.*, 2005) and <1 % in springtails (Hogg & Hebert, 2004), lepidopterans (Hajibabaei *et al.*, 2006; Hebert *et al.*, 2003b) and several arachnid species (Barrett & Hebert, 2005). These results are consistent with intraspecific divergences of mitochondrial genes hardly greater than 2 % and most less than 1 % (Avise, 2000). However, several species of animals, as well as some protists, have been reported to have high degrees of intraspecific divergence. For example, values of 5 and

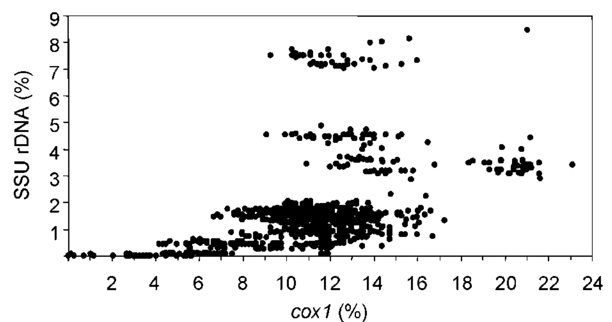

Fig. 4. Relationship between interspecific sequence divergences of the 689 bp *cox1* barcoding region and 1639 bp SSU rDNA for 45 tetrahymenine species representatives. Note that the *cox1* gene provides better discrimination among these tetrahymenine species.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

13 % within-species divergence were reported in the collembolans *Sminthurides malmgreni* and *Folsomia quadrioculata*, respectively (Hogg & Hebert, 2004). Two specimens of the shortnose chimaera *Hydrolagus novaezelandiae* showed 14.08 % intraspecific distance (Ward *et al.*, 2005). A single individual of the mayfly *Maccaffertium modestum* showed 13.7 % intraspecific divergence from the other specimens examined, implying the presence of a cryptic species complex as suggested by some behavioural evidence (Ball *et al.*, 2005). In addition, intraspecific divergence values of the ciliates *Paramecium caudatum* and *Paramecium multimicronucleatum* were 7 and 9.5 %, respectively (Barth *et al.*, 2006). However, Barth *et al.* (2006) suspected that their intraspecific isolates of these nominate *Paramecium* species might belong to more than one syngen or biological species and suggested future investigation of these isolates by mating reactivities to determine whether they belong to known or novel syngens.

Intraspecific sequence divergence analyses for *T. pyriformis*. When Nanney & McCoy (1976) established nominal species for the syngens of *Tetrahymena*, they did not at the same time establish holotype strains for each species. Therefore, we suggest that the *T. pyriformis* strain phenoset A ATCC® 30327™ be established as the holotype strain based on its original isolation in 1922 by Professor Dr André Lwoff (ATCC catalogue). With this isolate as the type for the species, four *T. pyriformis* isolates (an unnamed isolate represented by GenBank accession no. AF160864, ATCC® 30327™, ATCC® 30005™ and ATCC® 30202™), *T. setosa* HZ-1 (=ATCC® 30782™) and *T. tropicalis* G1-R (=ATCC® 205060™) clustered together with 100 % bootstrap support in the *cox1* NJ tree: these six strains differed from each other by less than 0.2 % in their *cox1* gene sequences. The relatedness between *T. pyriformis* and *T. setosa* has been suggested by several previous studies. For example, there are no nucleotide differences in the 5.8S, SSU rDNA or partial LSU rDNA sequences between *T. pyriformis* and *T. setosa* (Nanney *et al.*, 1989; Preparata *et al.*, 1989; Strüder-Kypke *et al.*, 2001). However, this species pair showed an isozyme similarity coefficient of only 62 %, indicating their independent species status based on the isozyme species criterion (Meyer & Nanney, 1987; Nanney *et al.*, 1989).

In contrast to our study, previous studies have not shown a close relationship between *T. tropicalis* G1-R (=ATCC® 205060™) and strains of *T. pyriformis*. Referred to as strain G1-R, *T. tropicalis* ATCC® 205060™ showed isozyme mobilities distinct from those of other strains of *Tetrahymena*, including strains of *T. pyriformis* (Borden *et al.*, 1973a). In addition, based on sequences of the D2 domain of the LSU rDNA, *T. tropicalis* ATCC® 205060™ was grouped together with other *T. tropicalis* isolates, but separately from the '*T. pyriformis*' cluster (Nanney *et al.*, 1998). However, our SSU rDNA sequence corroborated the result from the *cox1* data, confirming the clustering of *T. tropicalis* ATCC® 205060™ with *T. pyriformis* (GenBank accession no. X56171): there was 0 % sequence divergence in the SSU rDNA (data not shown).

Based on *cox1* gene sequences, deep intraspecific divergence was found between *T. pyriformis* ATCC® 205038™ and *T. pyriformis* ATCC® 205062™, as represented by 8.46 % genetic distance between these two isolates and the remaining four *T. pyriformis* isolates. However, analysis of sequences of 190 bases of the LSU rDNA showed a close relationship between these two *T. pyriformis* isolates and other *T. pyriformis* isolates (Nanney *et al.*, 1998). Given a mean intraspecific sequence divergence value of <1.0 %, these two *T. pyriformis* isolates showed more than eight times this conspecific divergence value. The incongruence between our study and previous ones may indicate either the presence of cryptic species within *T. pyriformis* or misidentification or contamination of the cultures.

Based on the success of the *cox1* gene in identifying animal and other *Tetrahymena* species, our work implies that *T. setosa* HZ-1 (=ATCC® 30782™) and *T. tropicalis* G1-R (=ATCC® 205060™) should be assigned to *T. pyriformis* (see below). Analyses of the SSU rDNA to assess the status of *T. pyriformis* ATCC® 205038™ and *T. pyriformis* ATCC® 205062™ confirmed their placement, based on *cox1* sequences, close to *T. pyriformis* (GenBank accession no. X56171), with a mean of 0.04 % sequence divergence in the SSU rDNA for these three strains (data not shown).

Intraspecific sequence divergence analyses for *T. tropicalis*. As indicated by nearly complete interspersion of its conspecific isolates in the NJ tree, *T. tropicalis* showed a high degree of intraspecific sequence divergence. Again, since Nanney & McCoy (1976) did not designate a holotype strain for *T. tropicalis*, we suggest that strain TC-105 of *T. tropicalis* ATCC® 30276™ be chosen as the holotype strain on the basis of its origin from one of the type localities of the species, which are the American tropics and the Pacific islands (Nanney & McCoy, 1976), and its original isolation from Rio Martin Sanchez, Panama, by A. M. Elliott (Nanney *et al.*, 1998). Given this isolate as the holotype strain, the other four isolates did not show close intraspecific relationships to it based on the *cox1* gene sequence. This result is consistent with our analyses of SSU rDNA sequences but contradictory to that based on analysis of sequences of 190 bases of the LSU rDNA, which showed a grouping together of all four isolates, *T. tropicalis* ATCC® 30276™, *T. tropicalis* ATCC® 205097™, *T. tropicalis* ATCC® 205060™ and *T. tropicalis* ATCC® 205083™ (Nanney *et al.*, 1998). In addition, *T. tropicalis* ATCC® 205156™ was placed distantly from these other *T. tropicalis* isolates in the 689 bp *cox1* NJ tree and confirmed by the SSU rDNA tree, which placed this isolate as a sister taxon of *T. empidokyrea* ATCC® 50595™ (Fig. 5). These inconsistencies between our study and previous ones may indicate either the presence of cryptic species within *T. tropicalis* or mislabelling or contamination of the cultures. Amplification and sequencing of the *cox1* gene of

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

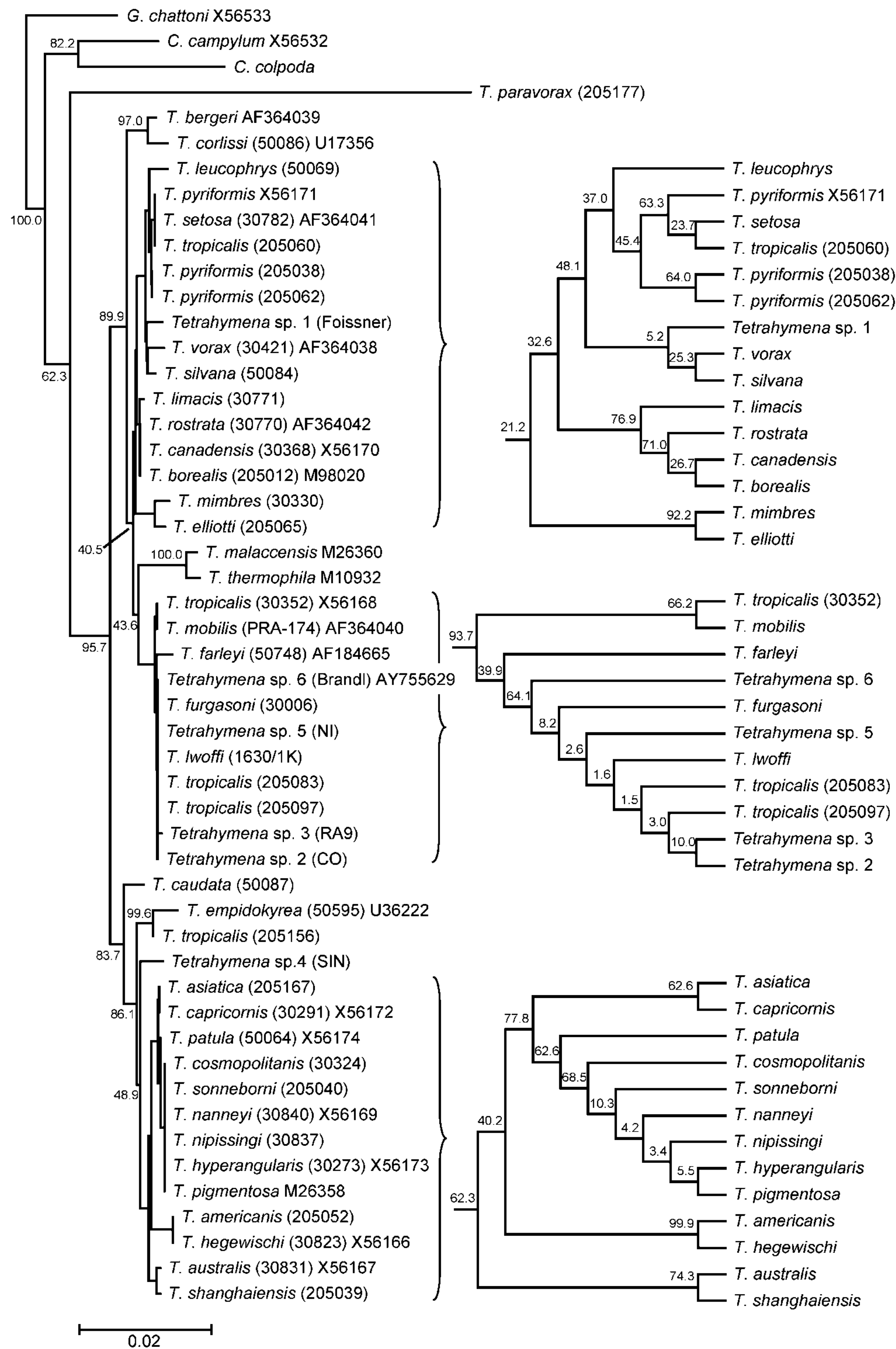

Fig. 5. NJ tree of 51 tetrahymenine species, including 45 species representatives, two additional intraspecific isolates of *T. pyriformis* and four additional intraspecific isolates of *T. tropicalis*, inferred from 1639 bp of the SSU rDNA. Percentage bootstrap values at nodes are based on 1000 resamplings of the data. Magnified clades show the topology more clearly, but percentage bootstrap values at the nodes are typically extremely low. Genetic distance calculation was based on the K2P model of sequence evolution. Branch lengths separating taxa represent genetic distance. Bar, 0.02 nucleotide substitutions per site.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

additional *T. tropicalis* isolates should be performed to assess its species status.

Comparison of interspecific sequence divergence in *Tetrahymena* and other organisms

In general, interspecific *cox1* sequence divergence values of *Tetrahymena* spp. are about 10 %. This result is similar to the interspecific sequence divergence values calculated from the barcoding region of various groups of animals: mean values of between-species sequence divergence range from 4.41 to 6.02 % in three families of tropical lepidopterans (Hajibabaei *et al.*, 2006), 7.93 % in North American birds (Hebert *et al.*, 2004b), >8 % in springtails (Hogg & Hebert, 2004), 9.93 % in Australian fishes (Ward *et al.*, 2005), 16.4 % in arachnid species (Barrett & Hebert, 2005) and 18 % in mayflies (Ball *et al.*, 2005). Moreover, between-species sequence divergence of 11 animal phyla ranged from 6.6 to 15.7 % with the exception of some cnidarians, which showed only 1 % interspecific sequence divergence (Hebert *et al.*, 2003a). Although a majority of interspecific sequence divergence values were normally many times greater than intraspecific ones, several congeneric pairs have been found to exhibit a low degree of interspecific divergence, sometimes lower than 2 %. For example, 13 species in four bird genera, *Larus*, *Haematopus*, *Corvus* and *Anas*, showed interspecific genetic distances lower than 1.25 % (Hebert *et al.*, 2004b).

In our study, there were several *Tetrahymena* species pairs that showed <1 % interspecific sequence divergence calculated from the 689 bp barcoding region. These species pairs are *T. canadensis* ATCC® 30368™ and *T. rostrata* ATCC® 30770™, *T. furgasoni* ATCC® 30006™ and *T. lwoffi* (1630/1G), *T. nanneyi* ATCC® 50071™ and *T. nipissingi* ATCC® 30837™, *T. pyriformis* ATCC® 30005™ and *T. setosa* ATCC® 30782™, *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 3 (RA9), *Tetrahymena* sp. 2 (CO) and *Tetrahymena* sp. 5 (NI) and *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI).

Taxonomic relationship of *T. canadensis* and *T. rostrata*. Formerly known as syngen 7 of the *T. pyriformis* species complex, *T. canadensis* is morphologically indistinguishable but reproductively isolated from other members of the *T. pyriformis* species complex and, based on isozyme data, the organism was established as a named species by Nanney & McCoy (1976). *T. rostrata* was first described as *Paraglaucoma rostrata* by Kahl (1926). Corliss (1952) transferred this species to the genus *Tetrahymena* as *T. rostrata*. On the basis of its capacity to form both reproductive and resting cysts, to exhibit edaphic and parasitic forms and to possess histophagous or parasitic habits, Corliss (1952) established it as the 'type' of the *rostrata* complex.

The low value of interspecific sequence divergence between *T. canadensis* ATCC® 30368™ and *T. rostrata* ATCC® 30770™ implies either convergent molecular evolution of these two species, phenotypic morphological variations within a single genetic species, contamination or misidentification of the cultures. Interspecific sequence divergence values inferred from 1821 bp and 689 bp of the *cox1* gene were 1 % (18 nucleotide differences) and 0.73 % (5 nucleotide differences), respectively. In addition, the 1639 sites of the SSU rDNA sequences were identical. Close relationships between *T. canadensis* and *T. rostrata* were also revealed by several previous molecular studies. For example, based on analysis of 180 sites of 23S rRNA gene sequences, no differences were observed (Nanney *et al.*, 1989). In addition, these two species were grouped within riboset A, since identical sequences were also found for 120 bp of the LSU rDNA and 154 bp of the 5.8S rDNA (Preparata *et al.*, 1989). Furthermore, only 3 of 579 bp of the histone H3II/H4II region were different between these two species (Sadler & Brunk, 1992). Although analysis of isozyme mobility between *T. canadensis* and *T. rostrata* showed a similarity coefficient of 66 %, which suggested their discrete but very closely related species status (Borden *et al.*, 1977; Meyer & Nanney, 1987), all of the other molecular evidence apparently indicates their identity as a single species.

Since we observed neither reproductive nor resting cysts in *T. rostrata* ATCC® 30770™, misidentification of this culture is likely. New field isolates of *T. rostrata* demonstrating species-specific characters must be collected to confirm the taxonomic validity and barcode for this species. This will also help us determine whether *T. canadensis* is just a pyriform stage of *T. rostrata*, and the former should then be regarded as a junior synonym of the latter.

Taxonomic relationship of *T. furgasoni* and *T. lwoffi*. *T. furgasoni* and *T. lwoffi* as well as *T. elliotti* are all amicronucleate and distinct from *T. pyriformis* and from each other based on heterogeneous isozyme patterns (Borden *et al.*, 1973a, b). Nanney & McCoy (1976) established these three isozymically distinct groups as new species and named them after three pioneer tetrahymenologists, Waldo H. Furgason, André Lwoff and Alfred M. Elliott.

The failure to discriminate between *T. furgasoni* ATCC® 30006™ and *T. lwoffi* (1630/1G) was due to the 0 % sequence divergence in both the *cox1* and SSU rDNA sequences. These two species were assigned to the *pyriformis* complex and, before formally receiving Latinized binomial names, *T. furgasoni* was known as two strains of *T. pyriformis*, GL-5 and GL-10, whereas *T. lwoffi* was recognized as four strains of *T. pyriformis*, PP, CH-S, GP and H-1 (Borden *et al.*, 1973a; Meyer & Nanney, 1987). These two groups are isozymically different and were therefore later established as the new species *T. furgasoni* and *T. lwoffi* by Nanney & McCoy (1976). Re-examination of isozyme mobilities of *T. furgasoni* and *T. lwoffi* by Meyer & Nanney (1987) showed different results from the previous study by Borden *et al.* (1973a): these two species then had indistinguishable isozyme patterns. Meyer &

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

Nanney (1987) suggested that they should be synonymized and suppressed the species name *T. lwoffi*. In addition, a high similarity in cytoskeleton proteins between *T. furgasoni* and *T. lwoffi* was observed by Williams *et al.* (1984). Furthermore, since both species lack a micronucleus, a mating test and the production of fertile F1 could not be performed. Thus, DNA-based identification using *cox1* gene sequences has supported the synonymous status of these two species. Other isolates of *T. furgasoni* and *T. lwoffi* (if any) are required in order to support their conspecificity further.

Taxonomic relationship of *T. pyriformis* and *T. setosa*. *T. setosa*, formerly known as *T. setifera*, was described by Holz & Corliss (1956) as a new species on the basis of having a caudal cilium at its posterior end. Possession of a caudal cilium by *T. setosa* gives darkly stained granules surrounded by a delicate fibril at the posterior pole of the cell, the so-called polar basal granule (PBG) complex. The species was also distinguished based on its contractile vacuole pores (CVP) compared with those of *T. pyriformis*: the two CVPs of *T. setosa* are placed parallel to each other, whereas those of *T. pyriformis* are positioned obliquely to each other.

One nucleotide difference in the *cox1* gene sequences between *T. pyriformis* ATCC® 30005™ and *T. setosa* ATCC® 30782™ led to sequence divergence values of only 0.11 and 0 % calculated from the 3′-region and the barcoding region, respectively. In addition, the 1639 sites of their SSU rDNA sequences were identical. This gave rise to the suspicion of two possible scenarios, as discussed below. Firstly, the cultures may have been either mislabelled or contaminated. Secondly, these two tetrahymenines are actually the same genetic species. Observation and staining of starved live cells of *T. setosa* showed that none had a caudal cilium (data not shown). Cells from the *T. setosa* culture were stained using the Chatton–Lwoff procedure (Corliss, 1953). Fifty silver-stained cells were examined to determine the position of CVPs and the existence of a PBG complex. The parallel CVP type was found in 14 cells ($n=50$) of the *T. setosa* culture. The remaining 36 cells had an oblique pattern of CVPs (i.e. *T. pyriformis*-type CVPs). In addition, when the posterior ends of cells were observed, only two cells ($n=12$) had both a PBG complex and parallel CVPs. The remaining 10 cells had no PBG complex and oblique CVPs (i.e. *T. pyriformis*-specific characteristics). Although these microscopic observations suggest either that *T. setosa* is polymorphic for these traits or that the culture of *T. setosa* had been contaminated with *T. pyriformis* either prior to submission to the ATCC or at some time during its history, the possibility of them being the same genetic species cannot be ruled out.

As previously mentioned, *T. setosa* was described on the basis of possessing a caudal cilium, arrangement of CVPs and other morphological features. However, a study by Nelsen & Debault (1978) showed that a caudal cilium could be induced in *T. pyriformis* syngen 1, later named *T. thermophila*. Therefore, the use of this feature for taxonomic diagnosis seems uncertain.

Single-cell isolation should be performed to establish a pure *T. setosa* culture. Observation and staining of live cells should be then carried out to ensure that they truly manifest the *T. setosa* taxonomic characters described by Holz & Corliss (1956). Once the pure culture is established, both *cox1* and SSU rDNA should be reamplified and sequenced to compare with the results presented here. If this yields results identical to ours, *T. setosa* should be considered a junior synonym of *T. pyriformis*, confirming that *T. setosa* might be only a strain of *T. pyriformis*, as stated once by Corliss (1972).

Taxonomic relationship of *T. nanneyi* and *T. nipissingi*. *T. nanneyi* and *T. nipissingi* were described as new biological species on the basis of mating reactivities and isozyme mobilities by Simon *et al.* (1985) and Nyberg (1981), respectively. Isozyme similarity coefficient values between them are >62 % (Meyer & Nanney, 1987; Simon *et al.*, 1985). In addition, Nyberg (1981) found a correlation between temperature tolerance and *Tetrahymena* species and used this species-specific correlation as a supporting criterion to describe new *Tetrahymena* species such as *T. hegewischi*, *T. nipissingi* and *T. sonneborni*. D. Nyberg (personal communication) has suggested that the following strains be established as holotypes: *T. hegewischi* KP7 (=ATCC® 30832™), *T. nanneyi* LB2 (=ATCC® 50071™), *T. nipissingi* X2-AM (=ATCC® 30837™) and *T. sonneborni* EA2 (=ATCC® 30834™). In this study, holotype cultures of *T. nanneyi* and *T. nipissingi* were examined.

Interspecific sequence divergences lower than 1 % for the *cox1* gene have rarely been found. This phenomenon was observed in our study of *T. nanneyi* and *T. nipissingi*. The genetic distance between these two species is <0.33 % based on *cox1* gene sequences, demonstrating their close relationship. However, nuclear gene sequences of *T. nanneyi* and *T. nipissingi* are identical for several genes, for example a partial sequence of the 23S rRNA gene (Preparata *et al.*, 1989) and the entire SSU rDNA. Although molecular data showed great similarities between these species, reproductive isolation from each other and other species of *Tetrahymena* confirms their status as true biological species (Nyberg, 1981; Simon *et al.*, 1985). However, this close relationship between *T. nanneyi* and *T. nipissingi* remains suspicious in light of several lines of evidence. When Nyberg (1981) described *T. nipissingi* as a new biological species, he also found two other *Tetrahymena* isolates of a new mating group numbered 15: WX0 and XQ5. In his mating reactivity experiment, Nyberg (1981) refrained from establishment of these two isolates as a new species since crossing between them failed to produce viable progeny. In addition, Nyberg (1981) reported the observation of mating reactivity between strains of group 15 and *T. nipissingi* by E. Simon, although mating reactivity between them had not been repeatable.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

Subsequently, viable immature progeny were produced by crossing *Tetrahymena* strain LB2 and isolates of Nyberg's group 15, providing corroborative evidence of a new biological species. This led Simon *et al.* (1985) to establish a new species, *T. nanneyi*. Given the close molecular relatedness between *T. nanneyi* and *T. nipissingi* from previous studies and our own and the evidence mentioned above, the close relationships between these two species may indicate either a recent divergence between *T. nanneyi* and *T. nipissingi* or that these two ciliates belong to the same species.

In order to validate the biological species status of *T. nanneyi* and *T. nipissingi*, additional strains from culture collections and nature should be thoroughly examined using the *cox1* gene and SSU rDNA. Since the isolate of *T. nipissingi* X2-AM (=ATCC® 30837™) that was examined in this study was only one representative of the species, other isolates, such as *T. nipissingi* XK 2 (=ATCC® 205160™), X13-R1 (=ATCC® 30838™), X32H3 (=ATCC® 205029™) and X7E3 (=ATCC® 205091™), should particularly be examined. In addition, other isolates of *T. nanneyi* [i.e. the isolate strains XQ5 (=ATCC® 50067™) and WX0 (=ATCC® 205090™)] that were originally used in describing this species should also be investigated.

Taxonomic relationships among undescribed *Tetrahymena* species. In our study, six undescribed species, *Tetrahymena* sp. 1 (Foissner), *Tetrahymena* sp. 2 (CO), *Tetrahymena* sp. 3 (RA9), *Tetrahymena* sp. 4 (SIN), *Tetrahymena* sp. 5 (NI) and *Tetrahymena* sp. 6 (Brandl), were examined with the aim of identifying and assigning them to a nominal species using the DNA-based identification criteria of the 689 bp *cox1* barcoding region.

Based on the 689 bp *cox1* gene sequences, *Tetrahymena* sp. 1 (Foissner), *Tetrahymena* sp. 4 (SIN), and *Tetrahymena* sp. 6 (Brandl) showed considerable genetic distances of 11.47, 9.71 and 9.16 %, respectively, from other *Tetrahymena* species, while, based on SSU rDNA sequences, *Tetrahymena* sp. 1 (Foissner), *Tetrahymena* sp. 4 (SIN) and *Tetrahymena* sp. 6 (Brandl) showed a lower evolutionary distance, 1.18, 1.35 and 1.11 %, respectively, from other *Tetrahymena* species. Furthermore, these three undescribed *Tetrahymena* showed considerable *cox1* sequence divergence from each other: 11.60 % between *Tetrahymena* sp. 1 (Foissner) and *Tetrahymena* sp. 4 (SIN), 9.76 % between *Tetrahymena* sp. 1 (Foissner) and *Tetrahymena* sp. 6 (Brandl) and 8.77 % between *Tetrahymena* sp. 4 (SIN) and *Tetrahymena* sp. 6 (Brandl). In addition, these three species did not cluster together in the *cox1* NJ tree. *Tetrahymena* sp. 1 (Foissner) and *Tetrahymena* sp. 6 (Brandl) were found free-living in natural habitats, whereas *Tetrahymena* sp. 4 (SIN) is a parasite of the guppy *Poecilia reticulata*. Wilhelm Foissner (personal communication) suggested that *Tetrahymena* sp. 1 (Foissner) could be a new species because of its capacity to form a macrostome. A study by Brandl *et al.* (2005) reported a close relationship of *Tetrahymena* sp. 6 (Brandl) to *T. mobilis* and *T. tropicalis* based on SSU rDNA sequences. Similar relationships were found in our study, in which *T. mobilis* and some *T. tropicalis* isolates grouped with this isolate using both the *cox1* gene and SSU rDNA sequences. Although *Tetrahymena* sp. 4 (SIN) as well as *Tetrahymena* sp. 2 (CO), *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI) were found as parasites of the guppy *Poecilia reticulata*, *Tetrahymena* sp. 4 (SIN) was not closely related to the other three isolates, showing a considerable degree of genetic distance, 7.79 % based on the *cox1* sequence. Thus, *Tetrahymena* sp. 1 (Foissner), *Tetrahymena* sp. 4 (SIN) and *Tetrahymena* sp. 6 (Brandl) likely represent new species of *Tetrahymena* on the basis of *cox1* gene sequences.

Interspecific sequence divergence among *Tetrahymena* sp. 2 (CO), *Tetrahymena* sp. 3 (RA9) and *Tetrahymena* sp. 5 (NI) ranged from 0.22 to 1.17 % based on *cox1* gene sequences. Furthermore, these three new isolates showed a sequence divergence of 9.08 % from other *Tetrahymena* species. Given that the range of intraspecific divergence among animals and *Tetrahymena* species is generally around 1–2 %, these three isolates, together with two *T. tropicalis* strains (ATCC® 205083™ and ATCC® 205097™), likely belong to the same new species, as indicated by less than 0.73 % *cox1* sequence divergence among them. Additionally, this is supported by the ecological fact that the three new isolates are all apparently parasites of the guppy *Poecilia reticulata*, although a parasitic nature has not yet been demonstrated for *T. tropicalis*.

CONCLUSIONS

Our study illustrates the potential utility of *cox1* gene sequences in identifying the closely allied species of the ciliate genus *Tetrahymena* as well as other tetrahymenine species examined. This study extends the spectrum of applicability of universal DNA barcodes for biological identification to ciliophoran protists. In addition, the study showed that the nucleotide length of 689 bp of the *cox1* gene is sufficient to provide a diagnostic signal for the identification of tetrahymenine ciliates. Furthermore, this 689 bp *cox1* barcoding region gives a unique diagnostic sequence for each taxonomic species, with a few exceptions that may be explained by cultures being misidentified or contaminated. Two noteworthy exceptions, *T. pyriformis* and *T. tropicalis*, will require additional taxonomic re-examination, as they may represent new cryptic species complexes.

Our results indicated that the *cox1*-based identification system should be applied as a tool by culture collections and research laboratories to authenticate and verify their collections, since unintentional mix-ups or losses of cell lines can happen. On one hand, our study indicated that the *cox1* barcode will be a powerful alternative in the biological identification of ciliates, especially those whose SSU rDNAs, which have long been used as putative

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

'barcodes', are identical or very similar. This will speed up the discovery of cryptic ciliate species and also provide new insights into ciliate biodiversity. On the other hand, several scientific communities, such as ecologists, population biologists, parasitologists and evolutionary biologists, whose work requires the identification of the organisms, will be able to use our *cox1*-based identification approach without needing taxonomic expertise on a particular group.

ACKNOWLEDGEMENTS

We are grateful to the ATCC for kindly providing cultures of *Tetrahymena*. In addition, our sincere gratitude is extended to all scientists who generously provided us with wild isolated cultures: Professor Dr Wilhelm Foissner, Dr Dina Zilberg and Miss Marcia P. Leibowitz. We wish to express our gratitude to Professor Clifford F. Brunk for kindly providing sequences of the *cox1* gene for *Tetrahymena paravorax* and other *Tetrahymena* species to facilitate our primer design and to Dr Michaela C. Strüder-Kypke for kindly providing sequences of the *cox1* gene for some *Tetrahymena* species for our analyses. This research was supported through funding to the Canadian Barcode of Life Network from Genome Canada through the Genomics Institute, a NSERC of Canada Discovery Grant awarded to D. H. L. and other sponsors (listed at http://www.BOLNET.ca). C. C. was supported by a national scholarship awarded by the Cooperative Research Network (CRN), Government of Thailand. Finally, we are thankful to three anonymous reviewers for their constructive and helpful comments on earlier drafts of the manuscript.

REFERENCES

Avise, J. C. (2000). *Phylogeography: the History and Formation of Species*. Cambridge, MA: Harvard University Press.

Ball, S. L., Hebert, P. D. N., Burian, S. K. & Webb, J. M. (2005). Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *J N Am Benthol Soc* **24**, 508–524.

Barrett, R. D. H. & Hebert, P. D. N. (2005). Identifying spiders through DNA barcodes. *Can J Zool* **83**, 481–491.

Barth, D., Krenek, S., Fokin, S. I. & Berendonk, T. U. (2006). Intraspecific genetic variation in *Paramecium* revealed by mitochondrial cytochrome *c* oxidase I sequences. *J Eukaryot Microbiol* **53**, 20–25.

Borden, D., Whitt, G. S. & Nanney, D. L. (1973a). Electrophoretic characterization of classical *Tetrahymena pyriformis* strains. *J Protozool* **20**, 693–700.

Borden, D., Whitt, G. S. & Nanney, D. L. (1973b). Isozymic heterogeneity in *Tetrahymena* strains. *Science* **181**, 279–280.

Borden, D., Miller, E. T., Whitt, G. S. & Nanney, D. L. (1977). Electrophoretic analysis of evolutionary relationships in *Tetrahymena*. *Evolution* **31**, 91–102.

Brandl, M. T., Rosenthal, B. M., Haxo, A. F. & Berk, S. G. (2005). Enhanced survival of *Salmonella enterica* in vesicles released by a soilborne *Tetrahymena* species. *Appl Environ Microbiol* **71**, 1562–1569.

Brown, J. W., Miller, S. E. & Horak, M. (2003). Studies on New Guinea moths. 2. Description of a new species of *Xenothictis* Meyrick (Lepidoptera: Tortricidae: Archipini). *Proc Entomol Soc Wash* **105**, 1043–1050.

Brunk, C. F., Lee, C. L., Tran, A. B. & Li, J. (2003). Complete sequence of the mitochondrial genome of *Tetrahymena thermophila* and comparative methods for identifying highly divergent genes. *Nucleic Acids Res* **31**, 1673–1682.

Burger, G., Zhu, Y., Littlejohn, T. G., Greenwood, S. J., Schnare, M. N., Lang, B. F. & Gray, M. W. (2000). Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J Mol Biol* **297**, 365–380.

Corliss, J. O. (1952). Le cycle autogamique de *Tetrahymena rostrata*. *C R Acad Sci* **235**, 399–402 (in French).

Corliss, J. O. (1953). Silver impregnation of ciliated protozoa by the Chatton-Lwoff technic. *Stain Technol* **28**, 97–100.

Corliss, J. O. (1972). *Tetrahymena* and some thoughts on the evolutionary origin of endoparasitism. *Trans Am Microsc Soc* **91**, 566–573.

Corliss, J. O. (1973). History, taxonomy, ecology, and evolution of species of *Tetrahymena*. In *Biology of Tetrahymena*, pp. 1–55. Edited by A. M. Elliott. Stroudsburg, PA: Dowden, Hutchinson & Ross.

Corliss, J. O. & Daggett, P.-M. (1983). "*Paramecium aurelia*" and "*Tetrahymena pyriformis*": current status of the taxonomy and nomenclature of these popularly known and widely used ciliates. *Protistologica* **19**, 307–322.

Cummings, D. J. (1992). Mitochondrial genomes of the ciliates. *Int Rev Cytol* **141**, 1–64.

Czapik, A. (1968). La famille Tetrahymenidae et son importance dans la systématique et l'évolution des ciliés. *Acta Protozool* **5**, 315–357 (in French).

Dawkins, R. (1998). *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder*. Boston: Houghton Mifflin.

Elliott, A. M. (1970). The distribution of *Tetrahymena pyriformis*. *J Protozool* **17**, 162–168.

Elwood, H. J., Olsen, G. J. & Sogin, M. L. (1985). The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Mol Biol Evol* **2**, 399–410.

Felsenstein, J. (2004). PHYLIP (phylogeny inference package), version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, USA.

Fenchel, T. & Finlay, B. J. (2004). The ubiquity of small species: patterns of local and global diversity. *Bioscience* **54**, 777–784.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3**, 294–297.

Gruchy, D. F. (1955). The breeding system and distribution of *Tetrahymena pyriformis*. *J Protozool* **2**, 178–185.

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W. & Hebert, P. D. N. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proc Natl Acad Sci U S A* **103**, 968–971.

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. (2003a). Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc Biol Sci* **270** (Suppl. 1), S96–S99.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003b). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**, 313–321.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci U S A* **101**, 14812–14817.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. (2004b). Identification of birds through DNA barcodes. *PLoS Biol* **2**, e312.

Hogg, I. D. & Hebert, P. D. N. (2004). Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can J Zool* **82**, 749–754.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14

Holz, G. G. & Corliss, J. O. (1956). *Tetrahymena setifera* n. sp., a member of the genus *Tetrahymena* with a caudal cilium. *J Protozool* **3**, 112–118.

Janczewski, D. N., Modi, W. S., Stephens, J. C. & O'Brien, S. J. (1995). Molecular evolution of mitochondrial 12S RNA and cytochrome *b* sequences in the pantherine lineage of Felidae. *Mol Biol Evol* **12**, 690–707.

Jerome, C. A. & Lynn, D. H. (1996). Identifying and distinguishing sibling species in the *Tetrahymena pyriformis* complex (Ciliophora, Oligohymenophorea) using PCR/RFLP analysis of nuclear ribosomal DNA. *J Eukaryot Microbiol* **43**, 492–497.

Kahl, A. (1926). Neue und wenig bekannte Formen der holotrichen und heterotrichen Ciliaten. *Arch Protistenkd* **55**, 197–438 (in German).

Kimura, M. (1980). A simple method of estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

Kumazawa, Y. & Nishida, M. (1993). Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J Mol Evol* **37**, 380–398.

Lynn, D. H. & Strüder-Kypke, M. C. (2006). Species of *Tetrahymena* identical by small subunit rRNA gene sequences are discriminated by mitochondrial cytochrome *c* oxidase I gene sequences. *J Eukaryot Microbiol* **53**, 385–387.

Mardulyn, P. & Whitfield, J. B. (1999). Phylogenetic signal in the COI, 16S, and 28S genes for inferring relationships among genera of Microgastrinae (Hymenoptera; Braconidae): evidence of a high diversification rate in this group of parasitoids. *Mol Phylogenet Evol* **12**, 282–294.

Medlin, L., Elwood, H. J., Stickel, S. & Sogin, M. L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* **71**, 491–499.

Meyer, E. B. & Nanney, D. L. (1987). Isozymes in the ciliated protozoa. In *Isozymes* (*Current Topics in Biological and Medical Research*, vol. 13), pp. 61–101. Edited by M. C. Rattazzi. New York: Alan R. Liss.

Monaghan, M. T., Balke, M., Gregory, T. R. & Vogler, A. P. (2005). DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philos Trans R Soc Lond B Biol Sci* **360**, 1925–1933.

Nanney, D. L. & McCoy, J. W. (1976). Characterization of the species of the *Tetrahymena pyriformis* complex. *Trans Am Microsc Soc* **95**, 664–682.

Nanney, D. L., Meyer, E. B., Simon, E. M. & Preparata, R.-M. (1989). Comparison of ribosomal and isozymic phylogenies of tetrahymenine ciliates. *J Protozool* **36**, 1–8.

Nanney, D. L., Park, C., Preparata, R. & Simon, E. M. (1998). Comparison of sequence differences in a variable 23S rRNA domain among sets of cryptic species of ciliated protozoa. *J Eukaryot Microbiol* **45**, 91–100.

Nelsen, E. M. & Debault, L. E. (1978). Transformation in *Tetrahymena pyriformis*: description of an inducible phenotype. *J Protozool* **25**, 113–119.

Nyberg, D. (1981). Three new "biological" species of *Tetrahymena* (*T. hegewischi* n. sp., *T. sonneborni* n. sp., *T. nipissingi* n. sp.) and temperature tolerance of members of the "*pyriformis*" complex. *J Protozool* **28**, 65–69.

Preparata, R. M., Meyer, E. B., Preparata, F. P., Simon, E. M., Vossbrinck, C. R. & Nanney, D. L. (1989). Ciliate evolution: the ribosomal phylogenies of the tetrahymenine ciliates. *J Mol Evol* **28**, 427–441.

Remigio, E. A. & Hebert, P. D. N. (2003). Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. *Mol Phylogenet Evol* **29**, 641–647.

Sadler, L. A. & Brunk, C. F. (1992). Phylogenetic relationships and unusual diversity in histone H4 proteins within the *Tetrahymena pyriformis* complex. *Mol Biol Evol* **9**, 70–84.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Saunders, G. W. (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos Trans R Soc Lond B Biol Sci* **360**, 1879–1888.

Scheffer, S. J., Lewis, M. L. & Joshi, R. C. (2006). DNA barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. *Ann Entomol Soc Am* **99**, 204–210.

Schlegel, M., Elwood, H. J. & Sogin, M. L. (1991). Molecular evolution in hypotrichous ciliates: sequence of the small subunit RNA genes from *Onychodromus quadricornutus* and *Oxytricha granulifera* (Oxytrichidae, Hypotrichida, Ciliophora). *J Mol Evol* **32**, 64–69.

Simon, E. M., Meyer, E. B. & Preparata, R. M. (1985). New wild *Tetrahymena* from Southeast Asia, China, and North America, including *Tetrahymena malaccensis*, *Tetrahymena asiatica*, *Tetrahymena nanneyi*, *Tetrahymena caudata*, and *Tetrahymena silvana* n. spp. *J Protozool* **32**, 183–189.

Smith, M. A., Fisher, B. L. & Hebert, P. D. N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philos Trans R Soc Lond B Biol Sci* **360**, 1825–1834.

Sogin, M. L. & Elwood, H. J. (1986). Primary structure of the *Paramecium tetraurelia* small-subunit rRNA coding region: phylogenetic relationships within the Ciliophora. *J Mol Evol* **23**, 53–60.

Strüder-Kypke, M. C., Wright, A.-D. G., Jerome, C. A. & Lynn, D. H. (2001). Parallel evolution of histophagy in ciliates of the genus *Tetrahymena*. *BMC Evol Biol* **1**, 5.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. (2002). DNA points the way ahead in taxonomy. *Nature* **418**, 479.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. (2003). A plea for DNA taxonomy. *Trends Ecol Evol* **18**, 70–74.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Walsh, P. S., Metzger, D. A. & Higuchi, R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**, 506–513.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B Biol Sci* **360**, 1847–1857.

Williams, N. E., Buhse, H. E., Jr & Smith, M. G. (1984). Protein similarities in the genus *Tetrahymena* and a description of *Tetrahymena leucophrys* n. sp. *J Protozool* **31**, 313–321.

Ziaie, Z. & Suyama, Y. (1987). The cytochrome oxidase subunit I gene of *Tetrahymena*: a 57 amino acid NH_2-terminal extension and a 108 amino acid insert. *Curr Genet* **12**, 357–368.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Fri, 07 Dec 2018 09:08:14