

# Species of *Tetrahymena* Identical by Small Subunit rRNA Gene Sequences are Discriminated by Mitochondrial Cytochrome *c* Oxidase I Gene Sequences

DENIS H. LYNN and MICHAELA C. STRÜDER-KYPKE

Department of Integrative Biology, University of Guelph, Guelph, ON, Canada N1G 2W1

**ABSTRACT.** The mitochondrial cytochrome *c* oxidase I (COI) genes of two isolates of each of the seven mating types of *Tetrahymena thermophila* were sequenced and found to differ by < 1% in nucleotide sequence and to be identical by putative protein sequence. As this gene was highly conserved in this species, the COI gene sequence was determined for four pairs of *Tetrahymena* species identical in their small subunit rRNA gene sequences. The following pairs of species showed from 1% to 12% divergence at the nucleotide level, enabling discrimination of all these species: (1) *Tetrahymena pyriformis* strain T and *Tetrahymena setosa* strain HZ-1; (2) *Tetrahymena canadensis* strain UM1215 and *Tetrahymena rostrata* strain ID-3; (3) *Tetrahymena pigmentosa* strain UM1285 and *Tetrahymena hyperangularis* strain EN112; and (4) *Tetrahymena tropicalis* strain TC-105 and *Tetrahymena mobilis*. However, because of the synonymous nature of the majority of substitutions, the pairs of species were identical based on the putative protein sequence.

**Key Words.** Barcode, COI PCR, species identification, SSrRNA.

**I**DENTIFICATION of animal species based on the sequence of a particular fragment of DNA has been referred to as DNA barcoding (Hebert et al. 2003). A DNA barcode, a fragment of the cytochrome *c* oxidase I (COI) gene of mitochondria, can be effectively used to identify many animal species (e.g. Folmer et al. 1994; Hajibabaei et al. 2006; Hebert et al. 2003, 2004). To our knowledge, there has been only one report applying the COI barcode to ciliates. Barth et al. (2006) demonstrated that this gene could be effectively used to separate out several *Paramecium* species, with interspecific nucleotide sequence divergences ranging from 12% to 27%. Although intrahaplogroup diversity was <2% for clades within *Paramecium caudatum* and *Paramecium multimicronucleatum*, it was as high as 6%–7% between haplogroups within these species, suggesting by this criterion the presence of putative biological species (Barth et al. 2006). Nevertheless, these data clearly show that the COI gene will be very useful in identifying both animal and protist species.

*Tetrahymena* is a ciliate genus in which molecular phylogenetic research has been used to not only recognize new species (Lynn et al. 2000), but also to outline the major trends of evolution within the genus (Strüder-Kypke et al. 2001). Given the success of the COI gene in identifying species of animals and *Paramecium*, we set out to sequence an ~ 980 base-pair fragment of the COI gene of *Tetrahymena* species. This may provide another marker to identify several species that are known to be identical along the length of the small subunit rRNA (Strüder-Kypke et al. 2001) and a portion of the large subunit rRNA gene (Nanney et al. 1998), but which can be identified by using histone genes (Sadler and Brunk 1992). We first sequenced 14 strains of *Tetrahymena thermophila*, represented by two strains of each of the seven mating types, to determine intraspecific variation of the COI gene. As this was extremely low, we applied the COI barcode and could discriminate between pairs of *Tetrahymena* species that had proven identical using the sequences of the rRNA genes. This research was first presented at the FASEB Conference on Ciliate Molecular Biology held at the Vermont Academy, Saxtons River, in July 2003.

## MATERIALS AND METHODS

**Strains and cultivation.** Fourteen strains of *Tetrahymena thermophila* were obtained from Dr. Paul Doerder, Cleveland

State University. These were distributed equally among the seven mating types of *T. thermophila* as follows: Mating type (MT) I – ANF689-2, isolated from the Allegheny National Forest by P. Doerder, and INBRED F\*; MT 2 – Strain A, B1298-3c and INBRED X; MT 3 – A\*III and Strain C, C\*; MT 4 – Strain B, B3140-8a and INBRED B DERIVED 2092-1-2a; MT 5 – Strain B, B1968-1b and Strain A, 1298-7a; MT 6 – Strain B, B1968-1c and INBRED B DERIVED rseC1 1275-16a; and MT 7 – Strain B3, B2493-4c1 and ANF 1359-1, the latter also isolated from the Allegheny National Forest by P. Doerder.

The following pairs of species, which are identical in small subunit rRNA (SSrRNA) gene sequence (Strüder-Kypke et al. 2001), were obtained from the American Type Culture Collection (ATCC) as a single strain of each species: (1) *Tetrahymena pyriformis* strain T (ATCC Accession No. 30202) and *Tetrahymena setosa* strain HZ-1 (ATCC 30782); (2) *Tetrahymena canadensis* strain UM1215 (ATCC 30368) and *Tetrahymena rostrata* strain ID-3 (ATCC 30770); (3) *Tetrahymena pigmentosa* strain UM1285 (ATCC 30278) and *Tetrahymena hyperangularis* strain EN112 (ATCC 30273). (4) *Tetrahymena tropicalis* strain Tc-105 (ATCC 30276) was obtained from the ATCC while *Tetrahymena mobilis*, which has an identical SSrRNA gene sequence, was obtained from Prof. Dr. W. Foissner, University of Salzburg.

All strains were cultured on 1% (w/v) proteose peptone and 1% (w/v) yeast extract with 0.2% (w/v) dextrose and transferred every other week.

**DNA extraction, amplification, and sequencing.** All strains were harvested by centrifugation and washed in TE buffer (10 mM Tris base, 1 mM EDTA, pH 8.0). Chelex DNA extraction followed the method of Walsh, Metzger, and Higuchi (1991) as described by Strüder-Kypke et al. (2001). In subsequent PCR reactions from 1–19 µl of the Chelex supernatant solution were used.

The PCR amplification of the 986 nucleotides long COI gene fragment, corresponding to the region 38427 to 39412 of the COI gene of *T. thermophila* (GenBank Accession No. AF396436), was performed in a Perkin-Elmer GeneAmp 2400 Thermal Cycler (PE Applied Biosystems, Mississauga, ON, Canada). The forward primer 288 5'-TCAGGTGCTGCACTAGC-3' and the reverse primer FolB 5'-TAAACTTCAGGGTGACCAAAAAATCA-3' (Folmer et al. 1994) were used with the following reaction conditions: initial denaturation at 94 °C, 4 min; five cycles with 94 °C—1 min, 45 °C—1.50 min, 72 °C—2 min; 35 cycles with 94 °C—1 min, 50 °C—1.50 min, 72 °C—2 min; final extension at 72 °C for 7 min. Polymerase chain reaction products were purified using the GeneClean kit (BIO/CAN, Mississauga, ON, Canada). They were sequenced in both directions using an ABI Prism 377

Corresponding Author: D. H. Lynn, Department of Integrative Biology, University of Guelph, Guelph, ON, N1G 2W1 Canada—Telephone number: 519-824-4120, ext. 52746; FAX number: 519-767-1656; e-mail: ddr@uoguelph.ca



Automated DNA Sequencer (Applied Biosystems Inc., Foster City, CA), using dye terminator and Taq FS with the forward and reverse PCR primers.

## RESULTS AND DISCUSSION

Using a fragment of the CO1 gene as a DNA barcode has proven effective as a method of identifying a variety of animal species (Folmer et al. 1994; Hajibabaei et al. 2006; Hebert et al. 2003, 2004). Thus, this region of the CO1 gene of 14 isolates of *T. thermophila* was amplified to assess how much nucleotide sequence divergence occurs within this species. This region of the CO1 gene was identical in nucleotide sequence in all isolates of *Tetrahymena thermophila* with the exception of the two isolates from the Allegheny National Forest, ANF689-2 and ANF 1359-1 for which there were six and seven substitutions, respectively, amounting to <1% sequence divergence between them and the other isolates (Fig. 1) (GenBank Accession Numbers DQ411866-DQ411879). This is not surprising as the 12 inbred strains of *T. thermophila* sequenced here, all derived from a single geographical isolate from Woods Hole, MA (Allen and Gibson 1973), with only the isolates from Allegheny National Forest showing differences. However, as these differences were all synonymous substitutions, there were no differences among the putative protein sequences of all 14 isolates of *T. thermophila* (data not shown).

Within species divergences using the CO1 gene of animals are typically less than 1% for butterflies (Hajibabaei et al. 2006; Hebert et al. 2003), birds (Hebert et al. 2004), and cypraeid gastropods (Meyer and Paulay 2005). These low levels of divergence were also found within haplogroups of *Paramecium* species (Barth et al. 2006). This contrasts sharply with between species divergences in animals that range from 4% to 7% for butterflies (Hajibabaei et al. 2006; Hebert et al. 2003), to ~8% for birds (Hebert et al. 2004), and from 12% to 27% for *Paramecium* species (Barth et al. 2006). Thus, given our results with *T. thermophila*, we conclude that these low levels of divergence within species in diverse groups of organisms should also enable both identification of known species and the recognition of novel taxa of ciliates.

As this level of variation within *T. thermophila* was similar to that observed in animal species and *Paramecium*, we proceeded to sequence the CO1 gene of the pairs of *Tetrahymena* species that have identical SSrRNA gene sequences. The %GC of this region of the CO1 gene ranged from a low of 25% in *T. pigmentosa* (GenBank Accession Number DQ411882) and *T. mobilis* (DQ411881), to 26% in *T. hyperangularis* (DQ411880), *T. canadensis* (DQ411885), and *T. rostrata* (DQ411883), to a high of 27% in the remaining three species (DQ411884, DQ411886, DQ411887). The percentage of nucleotide sequence divergences between species ranged from 1% to 12.4%. In all cases, those species identical by SSrRNA gene sequence showed some nucleotide divergence in this region of the CO1 gene (Fig. 1).

At least nine positions in this portion of the CO1 gene varied in amino acids among these species. At the level of the protein sequence, however, all substitutions were synonymous within pairs of *Tetrahymena* species identical by the SSrRNA gene sequences (data not shown).

While there is apparently no transfer of mitochondria between conjugating partners of *Tetrahymena* (Roberts and Orias 1973), there is evidence that mitochondrial transfer does occur in *Paramecium* species (Adoutte and Beisson 1970). However, the frequency of mitochondrial transfer among individuals in natural populations of *Paramecium* must be quite low as both nuclear and mitochondrial gene trees demonstrated the same terminal clades (Barth et al. 2006). Nevertheless, even if the nuclear and mitochondrial gene trees differed, the important point here is that

this CO1 gene fragment can unambiguously identify species of *Paramecium* and *Tetrahymena*.

In conclusion, this fragment of the CO1 gene shows low variability at least within the species *T. thermophila*, indicating that this species can be identified to a high degree of precision based on this gene. While further sampling will be needed to definitively demonstrate that this variability is typical of all species of *Tetrahymena*, our results indicate that species pairs of *Tetrahymena*, identical by the SSrRNA gene sequences, can be discriminated using this mitochondrial marker. These results, together with those of Barth et al. (2006), demonstrate that it will be profitable to extend this approach more broadly, both within the species complexes of *Paramecium* and *Tetrahymena* and to other ciliates and protists.

## ACKNOWLEDGMENTS

We thank Dr. Paul Doerder, Cleveland State University, for providing us with the 14 strains of *Tetrahymena thermophila*. We also thank Christina Hay whose initial probings of the CO1 barcode demonstrated that it could be successfully applied to *Tetrahymena* species. This research was supported by Natural Sciences and Engineering Research Council Discovery Grant awarded to D.H.L.

## LITERATURE CITED

- Adoutte, A. & Beisson, J. 1970. Cytoplasmic inheritance of erythromycin resistant mutations in *Paramecium aurelia*. *Mol. Gen. Genet.*, **108**: 70–77.
- Allen, S. L. & Gibson, I. 1973. Genetics of *Tetrahymena*. In: Elliott, A. M. (ed.), *Biology of Tetrahymena*. Dowden, Hutchinson & Ross Inc., Stroudsburg, PA. Ch. 11, p. 307–373.
- 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.
- 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.*, **33**:294–299.
- 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. USA*, **103**:968–971.
- Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. London, Biol. Sci., Ser. B*, **270**:313–321.
- Hebert, P. D. N., Stoeckle, M. Y., Zemplak, T. S. & Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLOS Biol.*, **2**:e312 (DOI: 10.1371/journal.pbio.0020312).
- Lynn, D. H., Gransden, S. G., Wright, A.-D. G. & Josephson, G. 2000. Characterization of a new species of the ciliate *Tetrahymena* (Ciliophora: Oligohymenophorea) isolated from the urine of a dog: first report of *Tetrahymena* from a mammal. *Acta Protozool.*, **39**:289–294.
- Meyer, C. P. & Paulay, G. 2005. DNA barcoding: error rates based on comprehensive sampling. *PLOS Biol.*, **3**:e422 (DOI: 10.1371/journal.pbio.0030422).
- Nanney, D. L., Park, C., Preparata, R. & Simon, E. M. (1998). Comparison of sequence differences in a variable 23S domain among sets of cryptic species of ciliated protozoa. *J. Eukaryot. Microbiol.*, **45**:91–100.
- Roberts, C. T., Jr. & Orias, E. 1973. Cytoplasmic inheritance of chloramphenicol resistance in *Tetrahymena*. *Genetics*, **73**:259–272.
- 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.
- 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.
- Walsh, P. S., Metzger, D. A. & Higuchi, R. 1991. Chelex<sup>®</sup> 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, **10**:506–513.