Molecular Phylogenetics and Evolution 94 (2016) 718–729

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Utility of combining morphological characters, nuclear and mitochondrial genes: An attempt to resolve the conflicts of species identification for ciliated protists ☆

Yan Zhao [a,b,c], Zhenzhen Yi [a,*], Eleni Gentekaki [d], Aibin Zhan [c], Saleh A. Al-Farraj [e], Weibo Song [b]

[a] Guangzhou Key Laboratory of Subtropical Biodiversity and Biomonitor, School of Life Science, South China Normal University, Guangzhou 510631, China
[b] Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao 266003, China
[c] Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China[1]
[d] School of Science, Mae Fah Luang University, Chiang Rai 57100, Thailand
[e] Zoology Department, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

ARTICLE INFO

Article history:
Received 5 June 2015
Revised 8 October 2015
Accepted 16 October 2015
Available online 2 November 2015

Keywords:
DNA-based taxonomy
Genetic distance
Morphological characters
Species identification
Ciliated protists
Frontonia

ABSTRACT

Ciliates comprise a highly diverse protozoan lineage inhabiting all biotopes and playing crucial roles in regulating microbial food webs. Nevertheless, subtle morphological differences and tiny sizes hinder proper species identification for many ciliates. Here, we use the species-rich taxon *Frontonia* and employ both nuclear and mitochondrial loci. We attempt to assess the level of genetic diversity and evaluate the potential of each marker in delineating species of *Frontonia*. Morphological features and ecological characteristics are also integrated into genetic results, in an attempt to resolve conflicts of species identification based on morphological and molecular methods. Our studies reveal: (1) the mitochondrial *cox1* gene, nuclear ITS1 and ITS2 as well as the hypervariable D2 region of LSU rDNA are promising candidates for species delineation; (2) the *cox1* gene provides the best resolution for analyses below the species level; (3) the V2 and V4 hypervariable regions of SSU rDNA, and D1 of LSU rDNA as well as the 5.8S rDNA gene do not show distinct barcoding gap due to overlap between intra- and inter-specific genetic divergences; (4) morphological character-based analysis shows promise for delimitation of *Frontonia* species; and (5) all gene markers and character-based analyses demonstrate that the genus *Frontonia* consists of three groups and monophyly of the genus *Frontonia* is questionable.

1. Introduction

Ciliates comprise one of the best-studied eukaryotic microbes contributing the structure and function of microbial food webs in different ecosystems (Azam and Malfatti, 2007; Caron et al., 2012; Lara et al., 2011). Though ciliates have been studied for over two centuries, more than 80% of their biodiversity remains undiscovered (Foissner et al., 2007). Small size, culturing difficulties and sometimes paucity of diagnostic phenotypic variation significantly hinder ciliated species identification (Lynn, 2008). Traditionally, ciliates have been recognized based solely on taxonomically relevant morphological features, an effort that requires substantial expertise (Fan et al., 2014; Lu et al., 2014; Pan et al., 2014; Shao et al., 2014). To that end, researchers have attempted to employ molecular taxonomy to complement or replace traditional taxonomy (Caron et al., 2009; Caron, 2013).

The small subunit ribosomal DNA (SSU rDNA) was the first molecular marker to be employed (Hillis and Dixon, 1991). There are numerous studies that assign SSU rDNA sequences to ciliates and examine phylogenetic positions and relationships among them (Agatha and Strüder-Kypke, 2014; Bachy et al., 2013; Gao et al., 2014a; Gao and Katz, 2014b; Zhang et al., 2010). Currently, SSU rDNA sequences comprise the largest assemblage of ciliate data. Consequently, this gene marker is widely chosen for species identification and estimation of molecular diversity in environmental samples (Dunthorn et al., 2012). In recent years, it has become clear that the SSU rDNA might be too conserved to uncover cryptic species of ciliates (Chantangsi et al., 2007; Strüder-Kypke and Lynn, 2010; Zhao et al., 2013). Hence, considerable effort has been dedicated in searching for alternative gene markers that will adequately reveal genetic diversity and provide better resolution especially at the tips of the ciliate tree.

☆ This paper was edited by the Associate Editor J.B. Dacks.

* Corresponding author.
E-mail address: zyi@scnu.edu.cn (Z. Yi).

[1] The author's present address.

http://dx.doi.org/10.1016/j.ympev.2015.10.017
1055-7903/

The mitochondrial cytochrome *c* oxidase subunit I (*cox1*) was the first marker to be tested in a broad scale. In their seminal study, Strüder-Kypke and Lynn (2010) revealed that *cox1* is highly divergent among the ciliate classes and that it provided fine scale resolution at below the genus level. Since then researchers have undertaken several studies on specific ciliate taxa, whereby various markers are tested in an effort to identify the ideal barcoding candidate. The nuclear ITS regions, histone 4 and the mitochondrial cytochrome *b* genes have all been tested as candidate barcoding markers with varying degrees of success (Barth et al., 2008; Gentekaki and Lynn, 2009; Greczek-Stachura et al., 2012; Yi et al., 2008). Studies focusing on select oligohymenophorean genera suggested that *cox1*-based barcoding can delineate closely related species and uncover cryptic ones (Chantangsi et al., 2007; Greczek-Stachura et al., 2012; Kher et al., 2011; Przyboś et al., 2009; Tarcz et al., 2013, 2014; Zhao et al., 2013). Recently, Santoferrara et al. (2013) and Stoeck et al. (2013) proposed the D1-D2 region of the LSU rDNA for ciliate species identification, because it shows clear barcoding gaps within Tintinnida and among congeners of *Paramecium*.

Nonetheless, there is still no universal gene marker for species discrimination of ciliates. And the usefulness of candidate nuclear and mitochondrial gene markers is rarely tested and compared on a taxon containing congeneric and conspecific sequences. In the present investigation, we used the species-rich genus *Frontonia* Ehrenberg, 1838 (for nomenclature, see Aescht, 2001), a peniculine ciliate that occupies a variety of habitats and is part of the micro-zooplankton (Lynn, 2008), to evaluate and validate candidate barcoding genes. The morphology of the genus has been studied extensively with a large amount of data, and their available morphological features can provide a useful template against which to test the accuracy of DNA-based taxonomy (Bullington, 1939; Fan et al., 2011; Fokin, 2008; Long et al., 2008; Pan et al., 2012, 2013). We assessed the resolving power of nuclear ITS1, ITS2, 5.8S rDNA, the hypervariable region V2 and V4 of the SSU rDNA and D1 and D2 of the LSU rDNA, as well as mitochondrial *cox1* gene both at the congeneric and conspecific levels. Furthermore, morphological features, habitat salinity information and genetic divergence were also integrated into a character-based barcode approach to further examine species delineation in this species-rich taxon. We then reconstructed the phylogenetic relationships of *Frontonia*, using the selected candidate barcoding genes. Subsequently, we examined whether the resulting relationships are congruent with the broadly accepted SSU rDNA phylogeny of the genus (Fan et al., 2013; Gao et al., 2008; Pan et al., 2013).

2. Materials and methods

2.1. Sampling information and species identification

Specimens of *Frontonia* spp. and *Stokesia* sp. were collected from brackish and fresh waters in South and North of China (Table 1, Fig. 1). We used a total of 34 samples from 16 species representing the major lineages of *Frontonia* and one species of its presumably sister genus *Stokesia* to evaluate the candidate barcoding loci. Isolation of cells was carried out within a week after sampling. Cells were isolated and observed *in vivo* using bright field and Nomarski differential interference contrast microscopy. Protargol and Chatton-Lwoff staining methods were used to reveal the ciliary and silverline pattern, respectively. Identification of the different morphospecies was according to morphological features described in detail in previous studies (Bullington, 1939; Fan et al., 2011, 2013; Fokin, 2008; Long et al., 2008; Pan et al., 2012, 2013).

2.2. DNA extraction and sequencing

Total genomic DNA was extracted from single cells with the DNeasy & Tissue Kit (Shanghai, QIAGEN China, China) according to modified manufacturer's protocol. Fragments of about 809 bp, 812 bp, 895 bp and 964 bp of the mitochondrial *cox1* gene were amplified using the forward primers CoxL11058 (5′-TGA TTA GAC TAG AGA TGGC-3′) or modified F388dT (5′-GGN KCB AAA GAT GTW GC-3′), and reverse primers CoxH10176 (5′-GAA GTT TGT CAG TGT CTA TCC-3′) or modified R1184dT (5′-TAD ACY TCA GGG TGA CCR AAA AAT CA-3′) designed by Barth et al. (2006) and Strüder-Kypke and Lynn (2010). Amplification cycle conditions were: 4 min at 94 °C followed by 35 cycles of 94 °C for 45 s, 55 °C for 60 s, 72 °C for 75 s, and a final extension at 72 °C for 10 min. A fragment of about 4200 bp containing SSU rDNA, ITS1, 5.8S rDNA, ITS2 and LSU rDNA was amplified using the forward universal primers Euk A (5′-AAY CTG GTT GAT YYT GCC AG-3′) (Medlin et al., 1988), and the reverse modified 28R (5′-AAC CTT GGA GAC CTG AT-3′) (Zhang et al., 2012). The amplification cycle conditions were 4 min at 94 °C followed by 35 cycles of 94 °C for 45 s, 60 °C for 75 s, 72 °C for 4:20 min, and a final extension at 72 °C for 10 min. The polymerase reactions were performed with TAKARA high fidelity *Ex Taq*™ polymerase (Takara, Shuzo, Japan). Following PCR amplification, the resulting amplicons of *cox1* gene were purified using the Spin Column PCR Product Purification Kit (Sangon Code: SK1132, Sangon Bio. Co., China) according to manufacturer's specifications. The purified PCR products were then inserted into the pMD®19-T simple vector (TaKaRa Code: D104A, Takara Bio Technology Co., Ltd., Dalian, China). The positive colonies of *cox1* gene were sequenced bi-directionally with the M13-47 forward (5′-CGC CAG GGT TTT CCC AGT CAC GAC-3′) and RV-M reverse (5′-GAG CGG ATA ACA ATT TCA CAC AGG-3′) primers. And the nuclear amplicons were sequenced directly with the amplification primers, combining with the designed walking primers in Life Technologies (Invigen™ sequencing facility in Guangzhou & Beijing, China). The new sequences have been deposited in the NCBI database (Table 1). In total, we generated 192 sequences: 24 for the SSU rDNA, 57 for the ITS1-5.8S-ITS2 fragment, 32 for the LSU rDNA and 79 for the *cox1* gene.

2.3. Genetic distance and phylogenetic analysis

Sequences were checked and assembled with Geneious ver. 5.4 (Drummond et al., 2011). Nuclear sequences were aligned using the MUSCLE software embedded in Geneious with default settings and then manually edited. The mitochondrial *cox1* gene sequences were aligned based on the predicted amino acid sequence with MEGA ver. 5.2 (using Protozoan mitochondrial genetic code) (Tamura et al., 2011). Using our newly generated data and sequences of peniculine species from GenBank, 13 datasets were constructed (Table 2). Datasets V2, V4, D1 and D2 were constructed according to previous data and secondary structure predictions (Wuyts et al., 2001, 2002; Zhang et al., 2010). Datasets rDNA and rDNA-*cox1*, covering as many taxa as possible, were concatenated from existing datasets SSU rDNA, ITS-5.8S, LSU rDNA and *cox1* using SeaView ver. 4 (Gouy et al., 2010).

Phylogenetic analyses were performed on the datasets *cox1*/SSU/ITS1-5.8S-ITS2/LSU, and the concatenated datasets rDNA and rDNA-cox1 (Table 2). Maximum likelihood (ML) and Bayesian Inference (BI) analyses were conducted with RAxML-HPC2 on XSEDE and MrBayes ver. 3.2.2 available on the Cipres Science Gateway (http://www.phylo.org/portal2/home.action) (Miller et al., 2010), respectively. ML bootstrap support was calculated from 1,000 bootstrap replicates. Bayesian posterior probabilities were computed by running four chains for 1,000,000 generations, with a sample frequency of each 100 generations and 2,500 trees were