SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

February 2019 Vol.62 No.2: 203–214
https://doi.org/10.1007/s11427-018-9422-5

Further analyses of variation of ribosome DNA copy number and polymorphism in ciliates provide insights relevant to studies of both molecular ecology and phylogeny

Yurui Wang[1,2†], Chundi Wang[1,2†], Yaohan Jiang[1,2], Laura A. Katz[3], Feng Gao[1,2*] & Ying Yan[1,2,3*]

[1]*Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao 266003, China;*
[2]*Key Laboratory of Mariculture, Ocean University of China, Ministry of Education, Qingdao 266003, China;*
[3]*Department of Biological Sciences, Smith College, Northampton, MA 01063, USA*

Received September 8, 2018; accepted November 6, 2018; published online January 21, 2019

Sequence-based approaches, such as analyses of ribosome DNA (rDNA) clone libraries and high-throughput amplicon sequencing, have been used extensively to infer evolutionary relationships and elucidate the biodiversity in microbial communities. However, recent studies demonstrate both rDNA copy number variation and intra-individual (intra-genomic) sequence variation in many organisms, which challenges the application of the rDNA-based surveys. In ciliates, an ecologically important clade of microbial eukaryotes, rDNA copy number and sequence variation are rarely studied. In the present study, we estimate the intra-individual small subunit rDNA (SSU rDNA) copy number and sequence variation in a wide range of taxa covering nine classes and 18 orders of the phylum Ciliophora. Our studies reveal that: (i) intra-individual sequence variation of SSU rDNA is ubiquitous in all groups of ciliates detected and the polymorphic level varies among taxa; (ii) there is a most common version of SSU rDNA sequence in each cell that is highly predominant and may represent the germline micronuclear template; (iii) compared with the most common version, other variant sequences differ in only 1–3 nucleotides, likely generated during macronuclear (somatic) amplification; (iv) the intra-cell sequence variation is unlikely to impact phylogenetic analyses; (v) the rDNA copy number in ciliates is highly variable, ranging from 10^3 to 10^6, with the highest record in *Stentor roeselii*. Overall, these analyses indicate the need for careful consideration of SSU rDNA variation in analyses of the role of ciliates in ecosystems.

ciliates, SSU rDNA, sequence variation, phylogenetic analyses, rDNA copy number, ecological significance

Citation: Wang, Y., Wang, C., Jiang, Y., Katz, L.A., Gao, F., and Yan, Y. (2019). Further analyses of variation of ribosome DNA copy number and polymorphism in ciliates provide insights relevant to studies of both molecular ecology and phylogeny. Sci China Life Sci 62, 203–214. https://doi.org/10.1007/s11427-018-9422-5

INTRODUCTION

Culture-independent molecular techniques have been used in ecological research on microbial eukaryotes in a wide range of ecosystems, including terrestrial, aquatic, anoxic and deep sea habitats (Bass and Cavalier-Smith, 2004; Stock et al., 2013; Stoeck et al., 2010; Zhao F et al., 2017). These techniques rely on either PCR-based (e.g. PCR, clone library and Sanger sequencing, Stoeck et al., 2010) or high-throughput sequencing (HTS) analyses (Zhao F et al., 2017). The resulting data are interpreted by defining microbial operational taxonomic units (OTUs) and subsequently estimating the community composition and structure (Grattepanche et al., 2016a; Grattepanche et al., 2016b; Jones et al., 2017; Mahé et al., 2017; Tucker et al., 2017).

The variation in evolutionary rates between highly conserved and variable regions and the concerted evolution

†Contributed equally to this work
*Corresponding authors (Feng Gao, email: gaof@ouc.edu.cn; Ying Yan, email: qilin_238@163.com)

© Science China Press and Springer-Verlag GmbH Germany, part of Springer Nature 2019

among different copies make ribosomal DNA (rDNA) an important molecular marker for ecological research, phylogenetic inference, and species identification (Gao et al., 2016; Wylezich et al., 2010; Zhao et al., 2016). However, both intraspecific and intra-individual (intra-genomic) variation of rDNA have been reported in different organisms, including some ciliates (Alverson and Kolnick, 2005; Ganley and Kobayashi, 2007; Gong et al., 2013; Gribble and Anderson, 2005; Pillet et al., 2012; Simon and Weiss, 2008; Wang C et al., 2017). Considering the importance of rDNA, it is essential to assess how intraspecific and intra-individual polymorphisms influence studies of ecology and evolution. Moreover, the copy number of the rDNA locus varies by lineages, and has been estimated in microalga, diatoms, dinoflagellates, fungi, animals, plants and ciliates (Galluzzi et al., 2004; Godhe et al., 2008; Gong et al., 2013; Heyse et al., 2010; Kapler, 1993; Prescott, 1994; Prokopowich et al., 2003; Simon et al., 2005; Taniguchi et al., 2011; Wang C et al., 2017; Yao et al., 1974; Zhu et al., 2005). The highly varied rDNA copy numbers among different organisms would introduce bias in research about community composition and species abundance (Fu and Gong, 2017).

Ciliates are a large group of predominantly-unicellular eukaryotes that have been used as model organisms in a wide range of biological studies (Chen et al., 2016; Gao et al., 2017; Gao et al., 2013; Huang et al., 2016; Luo et al., 2017; Wancura et al., 2018; Wang P et al., 2017; Wang et al., 2017a; Wang et al., 2017b; Wang et al., 2017c; Yan et al., 2017; Zhao X et al., 2017), mainly due to their unique biological features (e.g., nuclear dimorphism and extensively genome rearrangements). Ciliates are very diverse and play central roles in nutrient cycling among different trophic levels in microbial food loop (Kathol et al., 2009; Xu et al., 2014). More and more papers are emerging every year using ciliates as models to investigate issue of modern ecology and biogeography (e.g., Petz et al., 2007; Stoeck et al., 2010). However, interpreting patterns from molecular data from this clade is complicated as ciliates have extremely high rDNA copy number (highest record is ~570,000 per cell) and both intraspecific and intra-individual (intra-genomic) variability in rDNA sequences (Gong et al., 2013; Wang C et al., 2017). The question arises whether the copy number and sequence variation will impact species identification, phylogenetic analyses, and ultimately the estimation of microbial biodiversity.

Previous studies about the polymorphisms of ciliates focused on only a few taxa and/or tended to use the error-prone *Taq* DNA polymerase to amplify rDNA sequences (Gong et al., 2013; Wang C et al., 2017). In order to investigate the single nucleotide polymorphism of SSU rDNA in most ciliate classes, we used the high-fidelity polymerase *PfuTurbo* DNA polymerase to amplify the SSU rDNA sequence from 20 species belonging to nine classes and 18 orders of phylum Ciliophora (Figure 1). Then we analyzed whether the SSU rDNA sequence variation influences the topology of phylogenetic trees. We also measured the rDNA copy numbers of these species using single-cell quantitative PCR approach. The impact of sequence polymorphisms and copy number variation on molecular ecological studies were also discussed.

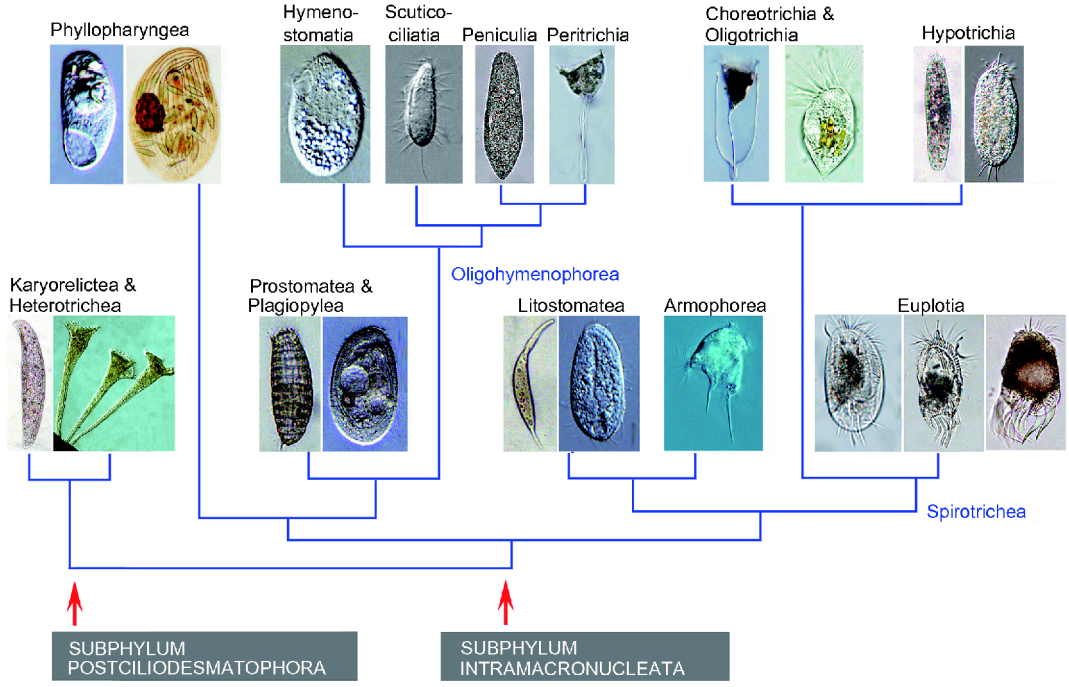

Figure 1 (Color online) Relationships among the 20 species analyzed in the present study. Topology based on Gao et al. (2016).

RESULTS

Single nucleotide polymorphisms (SNPs) of rDNA

We choose 20 species from nine classes and 18 orders of ciliates to investigate intraspecific and intra-individual SNPs in SSU rDNA. For each species, we amplify the SSU rDNA of three individuals using *PfuTurbo* DNA polymerase and sequenced 10 clones per individual. We find that there is a common sequence that occurred at the highest frequency within individual and species (Table 1, Table S1 in Supporting Information). The number of common sequences differs from 5–10 of the 10 clones among individuals within a species (Table S1 in Supporting Information) and the total number of common sequences within species varies from 18–30 of 30 clones (Table 1). As shown in Figure 2A and Table 1, the lower portion (<70%) of common sequences is detected in *Deviata* sp., *Favella ehrenbergii*, *Plagiopyla* sp. and *Tetrahymena thermophila*. The common sequences among the three individuals per species generally are the same, except *Deviata* sp. and *Favella ehrenbergii*. In *Deviata* sp. and *F. ehrenbergii*, the common sequences of the first two individuals are the same, but have one base pair divergence with that of the third one.

Among the 10 clones within an individual, 1–3 polymorphic sites are found when compared with the most common sequence (Figure 2B, Table S2 in Supporting Information). The highest total number of intra-individual polymorphic sites is 10 (0.65%), existing in *Loxodes* sp.-2 (Figure 3, Table 1). Total number of polymorphic sites within a species ranges from 0 to 13 among the 20 species with the lowest and the highest record in *Vorticella* sp. and *Loxodes* sp., respectively. Individual nucleotide diversity (π) varies from 0 to 0.130% with the highest number being in *Loxodes* sp. Besides, the levels of SSU rDNA polymorphism are also varied among the three individuals within a species (Table 1, Table S2 in Supporting Information). For instance, the haplotype diversity (Hd) of *Loxodes* sp. varies from 0.200% to 0.778%, and the nucleotide diversity (π) varies from 0.013% to 0.130% (Table S2 in Supporting Information).

The polymorphic sites of each species distribute randomly throughout the whole region of SSU rDNA (Figure 2B). The highest intraspecific pairwise distance of V2 region is detected in *Loxodes* sp. with 1.71% (2 bp). The numbers of OTUs for V2 region of the detected 20 species are 24, 20 and 20 when we choose 99%, 98% and 97% as cutoff respectively. The intraspecific pairwise distance of V4 region in *Uronychia binucleata* is the highest record in all detected species with 0.88% (2 bp). For the detected 20 species, the OTUs number of V4 region stabilizes at 20 when the cutoff value is not higher than 99%. The highest intraspecific pairwise distance of V9 region (2.33% in *E. vannus*, 2 bp) is remarkably higher than that in V2 and V4 region. 34 OTUs for V9 region of the 20 species are found when using 99% cutoff as threshold. This number declines to 21 and 20 when

Table 1 SSU rDNA polymorphisms and rDNA copy numbers of the 20 species

Species	SNP sites of each individual	Haplotypes of each individual	Common sequence of each species	Copy numbers of each species
Oxytricha trifallax	1	2	27/30	$1.1\times10^5\pm5.7\times10^4$
Deviata sp.	1–2	2–3	18/30	$4.6\times10^4\pm2.5\times10^4$
Favella ehrenbergii	0–2	1–3	19/30	$4.9\times10^5\pm2.7\times10^5$
Strombidium stylifer	1	2	27/30	$3.6\times10^4\pm1.4\times10^3$
Uronychia binucleata	0–2	1–3	26/30	$6.2\times10^4\pm1.9\times10^4$
Diophrys scutum	0–1	1–2	29/30	$4.7\times10^4\pm1.6\times10^4$
Euplotes vannus	1	2	27/30	$1.0\times10^5\pm7.2\times10^4$
Litonotus sp.	1–3	2–4	24/30	$2.3\times10^4\pm9.7\times10^3$
Apodileptus visscheri	2–4	2–3	22/30	$4.1\times10^4\pm3.7\times10^4$
Chilodontopsis depressa	1	2	27/30	$1.8\times10^4\pm8.4\times10^3$
Trithigmostoma sp.	0–1	1–2	29/30	$1.7\times10^4\pm8.3\times10^3$
Pseudocohnilembus persalinus	0–3	1–3	26/30	$1.1\times10^4\pm1.0\times10^3$
Vorticella sp.	0	1	30/30	$4.6\times10^4\pm2.9\times10^4$
Paramecium caudatum	1–3	2–4	25/30	$1.7\times10^5\pm1.3\times10^5$
Tetrahymena thermophila	2–5	3–6	20/30	$1.5\times10^4\pm7.1\times10^3$
Coleps sp.	0–2	1–3	27/30	$5.8\times10^3\pm2.1\times10^3$
Plagiopyla sp.	3	4	13/20	$8.4\times10^4\pm1.5\times10^3$
Caenomorpha medusula	0–1	1–2	28/30	$2.5\times10^4\pm1.0\times10^4$
Stentor roeselii	1–2	2–3	25/30	$3.5\times10^6\pm1.1\times10^5$
Loxodes sp.	1–10	2–6	22/30	$7.3\times10^3\pm6.5\times10^2$

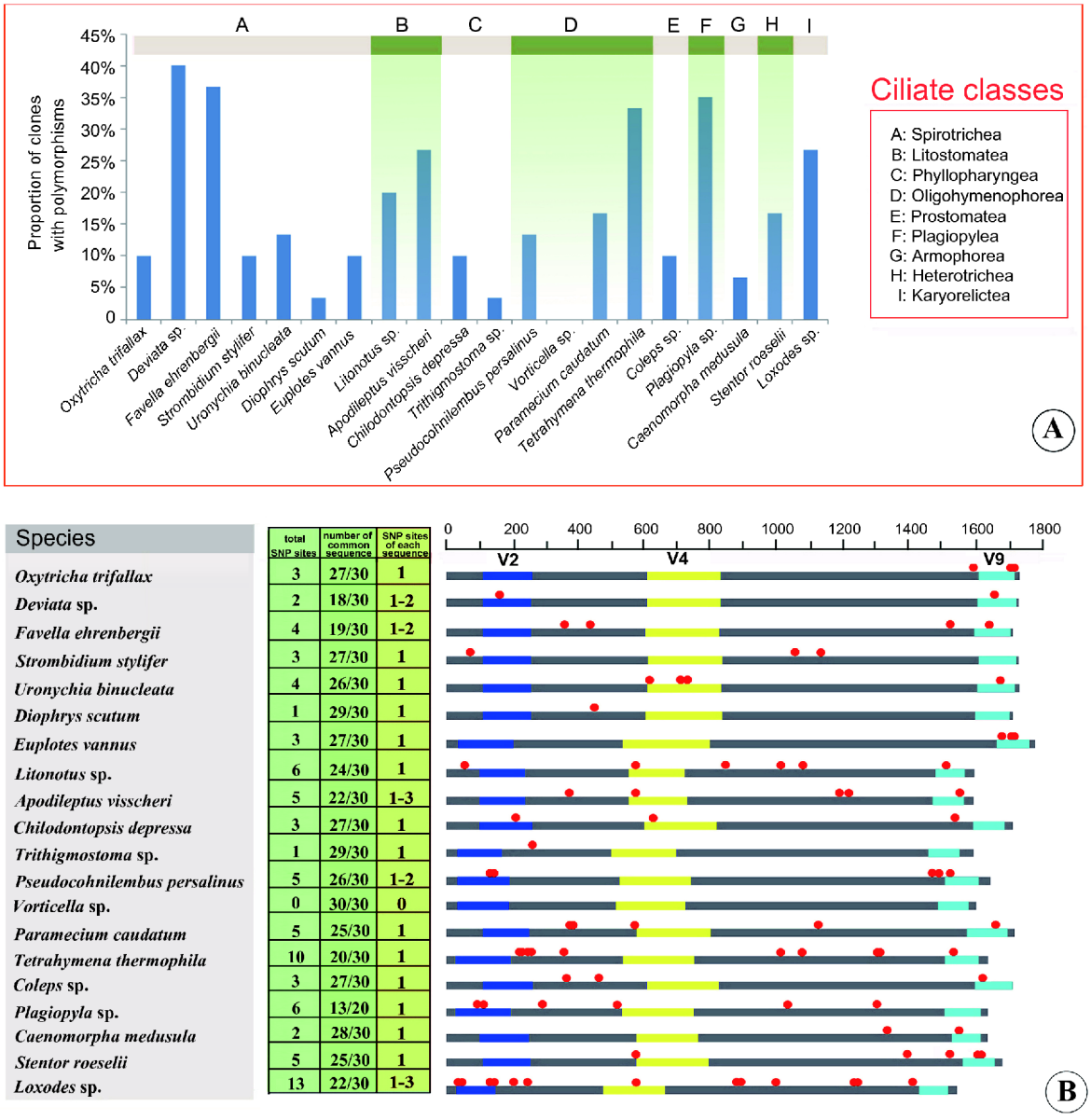

Figure 2 (Color online) The proportion of clones with single nucleotide sites and the distribution of single nucleotide polymorphisms (SNPs) in small subunit rDNA in the 20 species. A, The 20 species cover nine classes in Ciliphora, which are shown in A–I in figure respectively. B, Polymorphic sites are indicated by red solid circles. The dark blue, yellow and light blue segments indicate the variable regions 2, 4 and 9 respectively. Lengths of sequences are drawn to scale.

the cutoff falling to 98% and 97%, respectively (Table 2).

Table 2 The number of defined OTUs for V2, V4 and V9 regions of the 20 species when 97% to 99% cutoffs are used

Cutoff	V2	V4	V9
99%	24	20	34
98%	20	20	21
97%	20	20	20

Phylogenetic analyses based on SSU rDNA sequences

To investigate whether the polymorphism of SSU rDNA would influence the topology of phylogenetic tree in ciliates, we combined all haplotypes (sequences with the same nucleotides are classified as one haplotype) of the 20 species with 161 sequences downloaded from the National Center for Biotechnology Information (NCBI) to build maximum likelihood (ML) and bayesian inference (BI) trees (accession numbers are shown in Figure 4). The topologies of the ML and BI trees inferred from SSU rDNA sequences are congruent though with variable support values; therefore, only the topology of ML tree (with support values from both methods) is shown (Figure 4). The phylogenetic analyses show that except *F. ehrenbergii*, all haplotypes within the same species form monophyletic groups. The 5 haplotypes of present *F. ehrenbergii* and previous population of *F. ehren-*

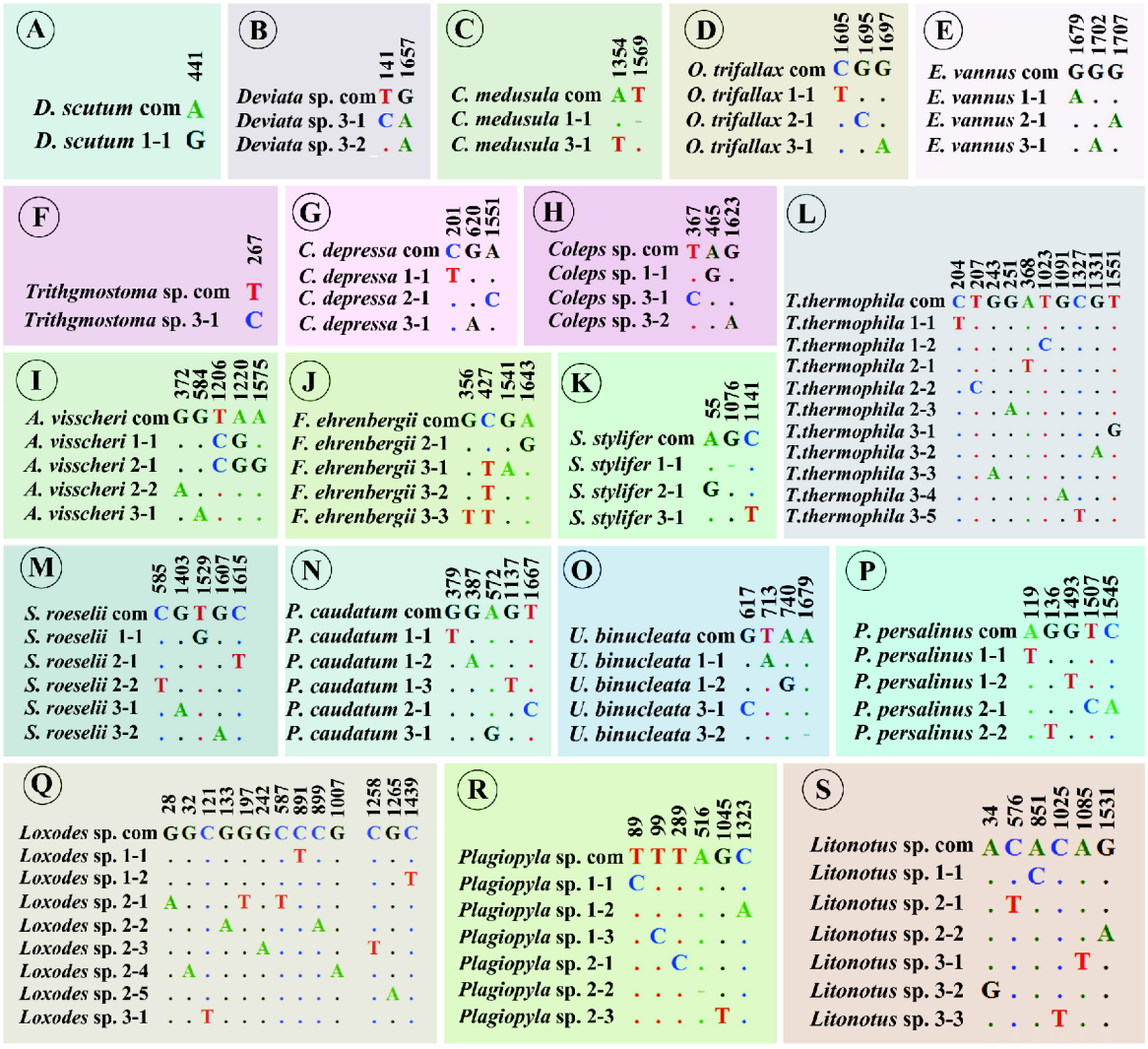

Figure 3 (Color online) The position of polymorphic sites compared with common sequence (com). For the numbers *x*-*y* behind the species name, *x* means the individual number while *y* means the clone number of this individual. A, *Diophrys scutum*; B, *Deviata* sp.; C, *Caenomorpha medusula*; D, *Oxytricha trifallx*; E, *Euplotes vannus*; F, *Trithigmostoma* sp.; G, *Chilodontopsis depressa*; H, *Coleps* sp.; I, *Apodileptus visscheri*; J, *Favella ehrenbergii*; K, *Strombidium stylifer*; L, *Tetrahymena thermophila*; M, *Stentor roeselii*; N, *Paramecium caudatum*; O, *Uronychia binucleata*; P, *Pseudocohnilembus persalinus*; Q, *Loxodes* sp.; R, *Plagiopyla* sp.; S, *Litonotus* sp.

bergii (GU574769) do not form monophyletic groups but are interdigitated with *F. panamensis* (AY143572) and *F. markusovszkyi* (JN871725) with high support (90%ML, 1.00BI), which may due to the existence of cryptic species or junior synonym (Bachvaroff et al., 2012; Kim et al., 2010).

rDNA copy numbers of each species

The cycle threshold (C_t) and the logarithms of rDNA copy number of standards in all the qPCR reactions form linear relationship, indicating the amplification efficiency is from 90% to 110% and R^2 is greater than 0.99 (Figure S1 in Supporting Information). The rDNA copy number of each cell was determined based on its own C_t value and corresponding standard curve (Table S1 in Supporting Information).

The results indicate that: (ⅰ) rDNA copy numbers vary among ciliate species and classes. The highest rDNA copy number is found in the class Heterotrichea (*Stentor roeselii*) with the value $3.5\times10^6\pm1.1\times10^5$ while the lowest one is detected in the class Prostomatea (*Coleps* sp., $5.8\times10^2\pm2.1\times10^2$). The class Spirotrichea, with the highest number of species included, has an average of 1.3×10^5 rDNA copy number per cell, which is the second highest record among the nine classes. (ⅱ) Even in the same class, rDNA copy number varies between different species. These classes, Spirotrichea, Litostomatea, Phyllopharyngea, and Oligohymenophorea, include more than one species in the present study. Spirotrichea has highest standard deviation (1.5×10^5) among these four classes. The second and third ones are 6.3×10^4 and 8.6×10^3, detected in Oligohymenophorea and Litostomatea, respectively. The rDNA copy numbers in two species of the class Phyllopharyngea, *Chilodontopsis de-*

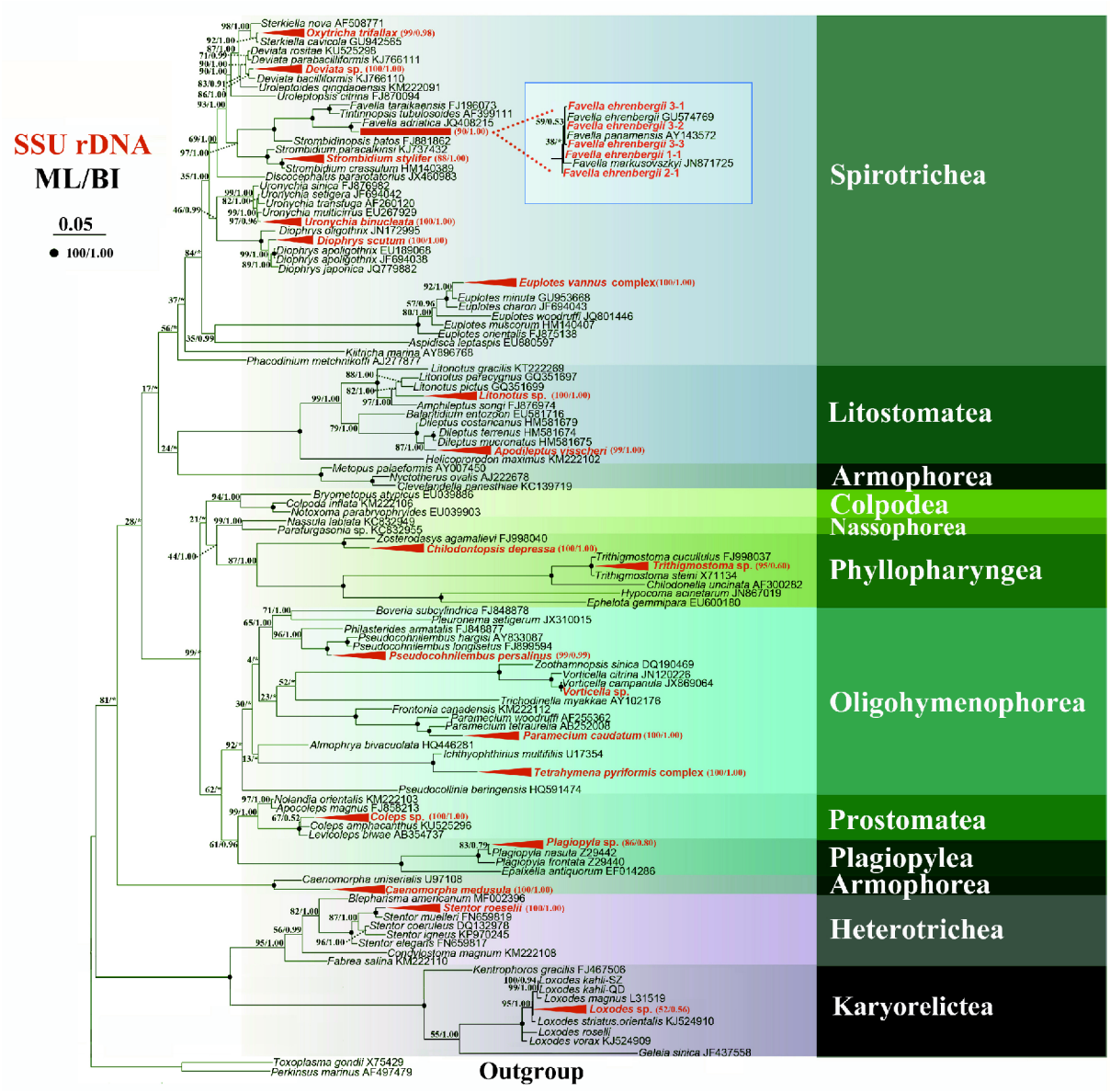

Figure 4 (Color online) The ML tree based on the small subunit ribosomal DNA (SSU rDNA) sequences showing phylogenetic positions of all haplotypes of the newly sequenced 20 species (in red and bold). Numbers near the branches denote bootstrap values for ML and posterior probability value for BI. Asterisk indicates the disagreement in topology between ML and BI trees. Fully supported (100%/1.00) nodes are represented by solid circles. Red triangles represent all haplotypes of the species can form a monophyly while red rectangle shows that all haplotypes of this species group with other species. All branches are drawn to scale. Scale bar corresponds to 5 substitutions per 100 nucleotide positions.

pressa and *Trithigmostoma* sp., are nearly the same ($1.8\times10^4\pm8.4\times10^3$ and $1.7\times10^4\pm8.3\times10^3$). (iii) The intraspecific copy number variation is taxon-dependent. For example, the rDNA copy numbers in the three individuals of *Strombidium stylifer* are nearly the same, whereas the copies in the three individuals of *Paramecium caudatum* differ over 40-fold.

DISCUSSION

Single nucleotide polymorphisms in ciliates

Single nucleotide polymorphisms (SNPs) of rDNA have been detected in a wide range of organisms (Alverson and Kolnick, 2005; Ganley and Kobayashi, 2007; Gong et al., 2013; Gribble and Anderson, 2005; Pillet et al., 2012; Simon and Weiss, 2008; Wang C et al., 2017). For example, the proportion of the intraspecific polymorphic sites in the LSU region of some heterotrophic dinoflagellates can be 0.80% (*Dinophysis* species), 1.00% (*Preperidinium meunerii*), 3.50% (*Diplopsalis lenticular*) or as high as 22.00% (*Protoperidinium depressum*) (Gribble and Anderson, 2005; Rehnstam-Holm et al., 2002). Intraspecific SSU rDNA polymorphism has been reported in a few ciliates (Gong et al., 2013; Wang C et al., 2017), but only three were verified

using high-fidelity DNA polymerase, with the highest proportion 1.68% in *Halteria grandinella* (Wang C et al., 2017). In this study, we expand the taxa sampling to 20 species, covering nine classes and 18 orders of the phylum Ciliophora. Our results indicate that SNPs exist in the SSU rDNA sequences in most ciliates, with varied levels in different species (Figures 2 and 3, Table 1). The highest proportion of the intraspecific polymorphic sites in this study is 0.84% (*Loxodes* sp.), which is likely an underestimate as only a small number of copies are sequenced (Table 1).

Additionally, not only the intraspecific but also the intra-individual (intra-genomic) polymorphic sites in the rDNA are detected (Figure 3, Table 1). In the present study, the highest proportion of intra-individual polymorphic sites of SSU rDNA is found in *Loxodes* sp.-2 (10, 0.65%, Figure 3, Table S2 in Supporting Information). It is also reported in other organisms. For example, about 1% intra-individual nucleotide polymorphism in the SSU rDNA sequence is observed in the diatom genus *Skeletonema* (Alverson and Kolnick, 2005). These data are a little higher in four plant pathogenic fungi, which is 2.00%–3.70% in the SSU rDNA region, 2.20%–3.60% in ITS region and 1.70%–3.60% in the LSU region, respectively (Simon and Weiss, 2008).

Despite the sequence variation in rDNA, one most common version of the SSU rDNA sequence exists among the 10 clones of each individual, with frequencies varying from 60%–100%. The common sequences are usually the same among different individuals in the same species. Compared with the most common sequence, usually less than three polymorphic sites are detected in other sequences. Considering that ciliates contain both the diploid micronucleus and the polyploid macronucleus within each cell, and the macronucleus rDNA are amplified from the zygotic template after conjugation (e.g., Prescott, 1994). The most common version may represent the micronuclear template SSU rDNA. The polymorphic sites found in other sequences could be generated through DNA amplification during macronuclear development, amitosis of macronuclei (e.g., Prescott, 1994) and/or represent sequencing error. A few shared polymorphisms within species can be found, suggesting these might represent allelic variation.

The sequence variation is unlikely to impact phylogenetic analyses as revealed from the phylogenetic analyses that all the haplotypes in the same species can group in monophyletic clade. This is because one is most likely to obtain the most common version of the SSU rDNA sequence when performing PCR amplification and direct sequencing, and also when cloning before sequencing. Even if the minor versions of the sequence are picked, they differ by less than three nucleotides compared to the common one, which will not alter the topology. However, using high-fidelity DNA polymerase is essential to avoid experimental errors (Wang C et al., 2017).

rDNA copy numbers in ciliates

Ciliates are reported to have extremely high rDNA copy numbers per cell. Before this study, the rDNA copy number of ciliates is ranging from 3,300 (*Zoothamnium* sp. 2) to 570,000 (*Halteria grandinella*), estimated from 15 species covering four classes, including Oligohymenophorea, Spirotrichea, Prostomatea and Heterotrichea (Gong et al., 2013; Heyse et al., 2010; Kapler, 1993; Prescott, 1994; Taniguchi et al., 2011; Wang C et al., 2017; Yao et al., 1974). We quantified the rDNA copy number of another 20 species in this study, covering nine classes and 18 orders of ciliates. Combined with previous studies, the rDNA copy number in ciliates varies even more greatly among species (Figure 5, Gong et al., 2013; Huang and Katz, 2014;Wang C et al., 2017). The rDNA copy number in *Stentor roeselii* (3.5×10^6 $\pm1.1\times10^5$) is about 1,000 fold of that in the lowest record of *Zoothamnium* sp. 2 ($3.4\times10^3\pm3.9\times10^2$) (Gong et al., 2013). Even among the three individuals of the same species, the difference is sometimes significant. For example, it differs over 40-fold in *Paramecium caudatum*-1 and *P. caudatum*-2.

The copy number variation among individuals within the same species may be explained by that: (ⅰ) they may be in different stages of macronucleus endoreplication after conjugation, during which the rDNA would replicate via extrachromosomal amplification (Torres-Machorro et al., 2010); (ⅱ) during asexual propagation, the progeny macronuclei undergo random distribution, which may result in the copy number variation in different cells (e.g., Huang et al., 2017; Orias and Flacks, 1975); (ⅲ) they may be under different nutritional conditions (Engberg and Pearlman, 1972), as most of the cells were isolated directly from the environmental samples; (ⅳ) the copy number of each macronuclear chromosome may be individually regulated in order to up- or down-regulate gene expressions (Xu et al., 2012).

Based on analyses of other lineages, there are additional explanations for the rDNA copy number variation among different species. In diatoms and dinoflagellates, the rDNA copy number has a positive correlation with biovolume (Godhe et al., 2008; Zhu et al., 2005). Prokopowich et al. (2003) revealed that a strong positive relationship between rDNA copy number and genome size (referred to *C*-value) using the data from 162 species of plants and animals. The rDNA copy number observed in some plants and protists are related to environmental effects (e.g., temperature, fire-related stress, Bobola et al., 1992; Fu and Gong, 2017; Govindaraju and Cullis, 1992; Strauss and Tsai, 1988). Some studies found that rDNA copy number is associated with genome stability in yeast (Ide et al., 2010). Studies in *Drosophila* indicated that rDNA copy number is related to genome-wide gene expression and subtle changes to rDNA copy number between individuals may contribute to biologically relevant phenotypic variation (Paredes et al., 2011).

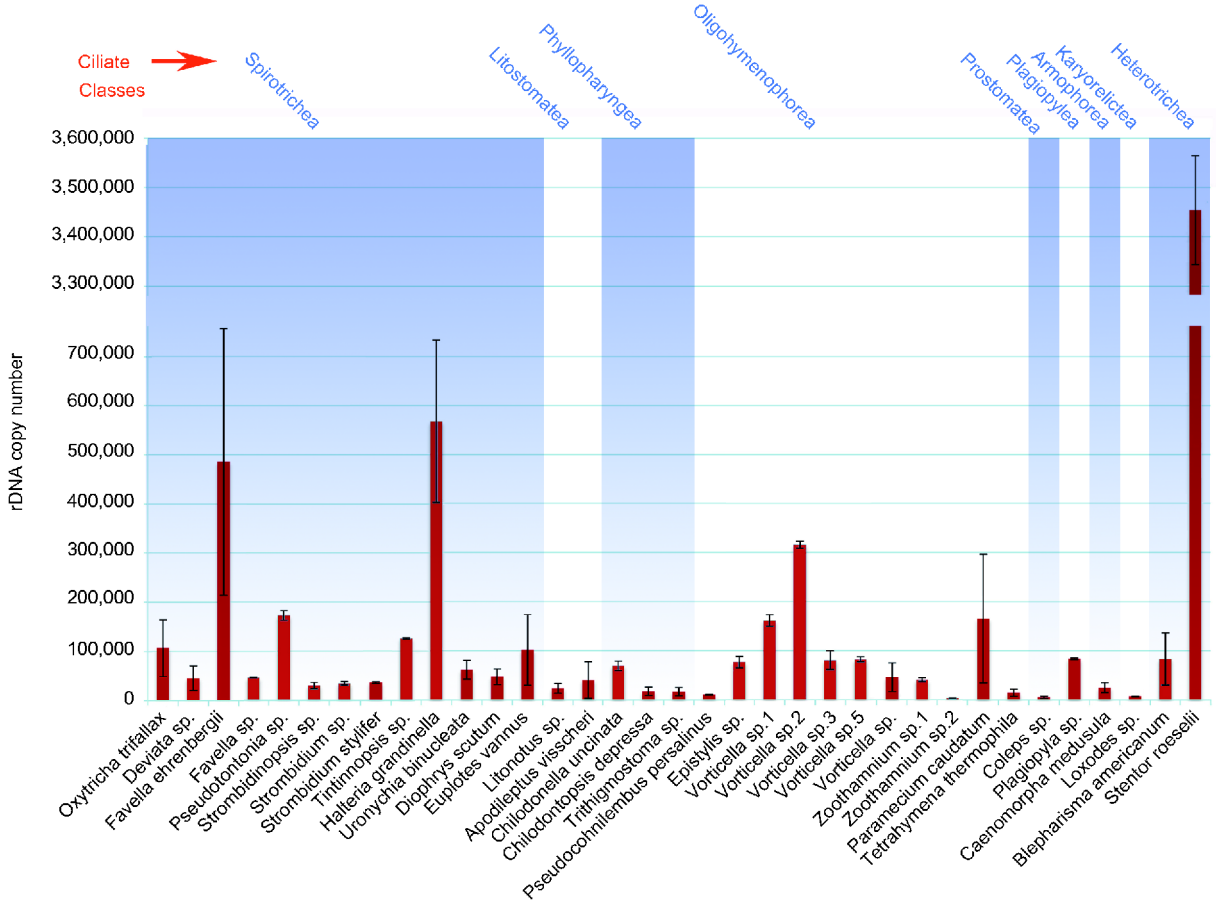

Figure 5 (Color online) Estimated rDNA copy numbers in different ciliates according to present and previous studies (Gong et al., 2013; Huang and Katz, 2014; Wang C et al., 2017).

Studies in human also reveal rDNA copy number is coupled with gene expression variation but negatively associated with mtDNA abundance (Gibbons et al., 2014).

Ecological significance and suggested cutoffs

As an important marker gene, rDNA is widely used to assess the environmental biodiversity and species abundance, especially for microorganisms (Godhe et al., 2008; Liu and Gong, 2012; Taniguchi et al., 2011; Zhu et al., 2005). Previous studies proposed that rDNA sequences, including SSU rDNA, LSU rDNA and ITS-5.8S rDNA, are candidate marker genes for ciliated species delimitation, while the most suitable marker genes are group-dependent (Dunthorn et al., 2012; Santoferrara et al., 2013; Stoeck et al., 2014; Zhao et al., 2016; Zhao et al., 2018). For instance, Dunthorn et al. (2012) indicated that, compared to V9 region of SSU rDNA, V4 region is more phylogenetically informative and more suitable for accessing ciliate environmental diversity. The D1-D2 region of LSU rDNA was proved to be the most useful potential barcoding tool for *Paramecium* species and tintinnids (Santoferrara et al., 2013; Stoeck et al., 2014) and Zhao et al. (2018) found that the V4 region of SSU rDNA and the D1-D2 region of LSU rDNA are the promising candidates to distinguish the *Euplotes* species. Besides, the mitochondrial cytochrome c oxidase subunit I (*cox1*) could discriminate among and below species level in particular genera, however, it is difficult to design universal *cox1* primers for all groups of ciliates (Zhao et al., 2016).

However, the interspecific sequence variation of rDNA would cause overestimation of the biodiversity (Medinger et al., 2010; Zhu et al., 2005). As high-throughput sequencing is being widely used in studies of microbial biodiversity, it is critical to consider rDNA variability when determining operational taxonomic units (OTUs) (Caron et al., 2009). Different cutoff values (e.g., 95%–100%) have been used in different lineages or even in the same lineages (e.g., Grattepanche et al., 2016a; Izhaki et al., 2013; Santoferrara et al., 2016; Stackebrandt and Ebers, 2006). A cutoff value of 97% has been widely used in bacteria (e.g., Izhaki et al., 2013) but some researchers suggested that cutoff value should be elevated to about 99% (Stackebrandt and Ebers, 2006). For microbial eukaryotes, it is even more difficult to determine the cutoff value. A cutoff value between 97% and 99% or even 100% was used (Santoferrara et al., 2016). Gong et al. (2013) and Wang C et al., (2017) both proposed that 99%

cutoff for full length of SSU rDNA can exclude the influence of interspecific sequence variation, even though sometimes it might be too stringent for some species.

Appropriate cutoff values will also depend on the region sequenced (e.g., V2 or V4 region). For example, a cutoff value of 98% for V4 region whereas 97% cutoff for V9 region could eliminate the interspecific sequence variation in the maximum extent (Wang C et al., 2017). Based on the data collected here, we use 97% to 99% cutoff to define ciliate OTUs for V2, V4 and V9 regions respectively (Table 2). 99% cutoff defines the real number for V4 region but overestimate the OTUs for V2 and V9 regions while 98% and 97% cutoffs are suitable for V2 region and V9 region, respectively. As the cutoff value could be taxon-dependent, researchers should explore the impact of changing cutoff values on inferences.

The great variation of rDNA copy number within and among species also increases the difficulty in estimating the species abundance in ecological research (Terrado et al., 2011). As we discussed above, the rDNA copy number can vary over 1000-fold among different species. Even within the same ciliate species, rDNA copy number is dynamic during life cycles and under varying environmental conditions. Therefore, caution is needed when interpreting data derived from rDNA-based surveys to infer abundance in microbial ecology.

MATERIAL AND METHODS

Specimens collection and identification

We analyzed 20 ciliate species chosen to capture diversity across the ciliate tree of life. Detailed information, including its systematic classification (based on Gao et al., 2016), body size and sampling sites, is shown in Table S1 in Supporting Information. We isolated cells from populations for either molecular work or identification by detailed observation *in vivo* and protargol impregnation (Wilbert, 1975).

DNA extraction, PCR amplification and sequencing

Ciliates were washed five times with filtered and sterilized water to remove potential contamination. Then each cell was transferred to a 1.5 mL microfuge tube with a minimum volume of water (about 0.5 μL). Single-cell genomic DNA was extracted using REDExtract-N-Amp Tissue PCR Kit (Sigma, St. Louis, USA) according to the modified manufacturer's protocol (1/10 of suggested volume for each solution). Then we amplified the full length of SSU rDNA (1,546–1,764 bp) of three individuals (isolated from the same sample) for each species with universal primers 18SF (5′-AACCTGGTTGATCCTGCCAGT-3′) or 82F (5′-GAAACTGCGAATGGCTC-3′) and 18SR (5′-TGATCCTTCTGCAGGTTCACCTAC-3′) (Jerome et al., 1996; Medlin et al., 1988) using *PfuTurbo* DNA polymerase (Cat. #600250, Agilent Technologies, USA), which has high fidelity (Cline et al., 1996; Wang C et al., 2017). The PCR products were purified by EasyPure PCR Purification Kit (TransGen Biotech, China) and then cloned by pClone007 Blunt Simple Vector Kit (Tsingke, China). For each individual, 10 clones were selected randomly and sequenced in Shanghai Personal Biotechnology Company (Qingdao, China).

Intra-individual DNA polymorphism and nucleotide diversity

The contigs were assembled with SeqMan v.7.1.0 (DNAStar, Anson and Myers, 1997). Then the 10 sequences of each individual were aligned using BioEdit v.7.0.1 (Hall, 1999). We excluded the primers and counted the numbers of polymorphic sites manually. Gaps were considered as polymorphic sites. Pairwise genetic distance was calculated with MEGA v.6.06 (Tamura et al., 2007) and the haplotype numbers and nucleotide diversity were calculated using DnaSP v.5.10 (Librado and Rozas, 2009). The OTUs were assigned using the QIIME v.1.9.0 (Caporaso et al., 2010), with threshold from 97% to 99% pair-wise nucleotide sequence identity.

Phylogenetic analyses

All haplotypes of SSU rDNA sequences of 20 species and other 161 sequences downloaded from NCBI database were aligned using the GUIDANCE2 Server with default parameters (Sela et al., 2015). Resulting alignments were modified manually using BioEdit v.7.0.1 (Hall, 1999). The ML tree, with 1,000 bootstrap replicates, was constructed using RAxML-HPC2 on XSEDE v.8.2.10 (Stamatakis, 2014) in CIPRES Science Gateway with GTR+I+G model. Bayesian inference (BI) analysis was performed using MrBayes on XSEDE v.3.2.6 in CIPRES Science Gateway with GTR+I+G model which was selected by MrModeltest v.2.0 (Ronquist et al., 2012) and PAUP (Nylander, 2004). Markov chain Monte Carlo (MCMC) simulations were run for 10,000,000 generations with a frequency of 100 generations and a burn-in of 10,000 trees. A majority rule consensus tree with posterior probabilities (PP) was constructed by all remaining trees. The clades with values higher than 80% ML and 0.90 BI were viewed as well supported. Tree topologies were visualized with MEGA v.6.06 (Tamura et al., 2007).

Quantitative real-time PCR assays

We extracted the whole DNA of each cell to measure its rDNA copy number. Before carrying out quantitative real-time PCR (qPCR), we amplified the partial sequence of SSU

rDNA of every individual with designed qPCR primers 1474F (5′-GTTGGTGGAGTGATTTGTCTGG-3′) and 1633R (5′-AGACCTGTTATTGCCTTAAACTTCC-3′) to test the applicability of the primers. The primers are suitable for most species, except *Coleps* sp., so we designed a pair of specific primers 705F (5′-GTCGGTATGGGAATCAGG-3′) and 857R (5′-CCCAACTGTCCCTGTTAA-3′) for it. We constructed a plasmid containing the SSU rDNA of *Coleps* as standard for its own qPCR while a plasmid containing the SSU rDNA of *Stentor roeselii* was used for the others as standard. Subsequently, we measured the concentrations of plasmids by Qubit 3.0 (Invitrogen, USA) and calculated their copy numbers online (http://cels.uri.edu/gsc/cndna.html). Plasmids were serially diluted ten-fold (10^{-1}–10^{-8}) to obtain standard curves. The quantitative real-time PCR reactions were performed using EvaGreen qPCR MasterMix-Low Rox Kit (Applied Biological Materials Inc., Canada) with an ABI 7500 Fast Real-Time PCR System (Applied Biosystems). The qPCR was performed according to Wang C et al. (2017).

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Natural Science Foundation of China (31772428), the National Science Foundation of the USA (1541511), Young Elite Scientists Sponsorship Program by CAST, and Fundamental Research Funds for the Central Universities (201841013 and 201762017). We thank Prof. Weibo Song, Ocean University of China (OUC), for the helpful suggestions in drafting this manuscript. Many thanks are also due to Wen Song, Chunyu Lian, Mingjian Liu, Borong Lu, Song Li, Rui Wang, Lun Wang and Yang Bai, graduate students in OUC, for their help in species identification.*

References

Alverson, A.J., and Kolnick, L. (2005). Intragenomic nucleotide polymorphism among small subunit (18s) rDNA paralogs in the diatom genus *Skeletonema* (bacillariophyta). J Phycol 41, 1248–1257.

Anson, E.L., and Myers, E.W. (1997). ReAligner: a program for refining DNA sequence multi-alignments. J Comput Biol 4, 369–383.

Bachvaroff, T.R., Kim, S., Guillou, L., Delwiche, C.F., and Coats, D.W. (2012). Molecular diversity of the syndinean Genus *Euduboscquella* based on single-cell PCR analysis. Appl Environ Microbiol 78, 334–345.

Bass, D., and Cavalier-Smith, T. (2004). Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of *Cercozoa* (Protozoa). Int J Syst Evol Microbiol 54, 2393–2404.

Bobola, M.S., Eckert, R.T., and Klein, A.S. (1992). Restriction fragment variation in the nuclear ribosomal DNA repeat unit within and between *Picearubens* and *Picea mariana*. Can J For Res 22, 255–263.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7, 335–336.

Caron, D.A., Countway, P.D., Savai, P., Gast, R.J., Schnetzer, A., Moorthi, S.D., Dennett, M.R., Moran, D.M., and Jones, A.C. (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. Appl Environ Microbiol 75, 5797–5808.

Chen, X., Gao, S., Liu, Y., Wang, Y., Wang, Y., and Song, W. (2016). Enzymatic and chemical mapping of nucleosome distribution in purified micro- and macronuclei of the ciliated model organism, *Tetrahymena thermophila*. Sci China Life Sci 59, 909–919.

Cline, J., Braman, J.C., and Hogrefe, H.H. (1996). PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. Nucl Acids Res 24, 3546–3551.

Dunthorn, M., Klier, J., Bunge, J., and Stoeck, T. (2012). Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. J Eukaryot Microbiol 59, 185–187.

Engberg, J., and Pearlman, R.E. (1972). The amount of ribosomal RNA genes in *Tetrahymena pyriformis* in different physiological states. Eur J Biochem 26, 393–400.

Fu, R., and Gong, J. (2017). Single cell analysis linking ribosomal (r)DNA and rRNA copy numbers to cell size and growth rate provides insights into molecular protistan ecology. J Eukaryot Microbiol 64, 885–896.

Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garces, E., and Magnani, M. (2004). Development of a real-time PCR assay for rapid detection and quantification of *Alexandrium minutum* (a dinoflagellate). Appl Environ Microbiol 70, 1199–1206.

Ganley, A.R.D., and Kobayashi, T. (2007). Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res 17, 184–191.

Gao, F., Warren, A., Zhang, Q., Gong, J., Miao, M., Sun, P., Xu, D., Huang, J., Yi, Z., and Song, W. (2016). The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). Sci Rep 6, 24874.

Gao, F., Huang, J., Zhao, Y., Li, L., Liu, W., Miao, M., Zhang, Q., Li, J., Yi, Z., El-Serehy, H.A., et al. (2017). Systematic studies on ciliates (Alveolata, Ciliophora) in China: progress and achievements based on molecular information. Eur J Protistol 61, 409–423.

Gao, S., Xiong, J., Zhang, C., Berquist, B.R., Yang, R., Zhao, M., Molascon, A.J., Kwiatkowski, S.Y., Yuan, D., Qin, Z., et al. (2013). Impaired replication elongation in *Tetrahymena mutants* deficient in histone H3 Lys 27 monomethylation. Genes Dev 27, 1662–1679.

Gibbons, J.G., Branco, A.T., Yu, S., and Lemos, B. (2014). Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. Nat Commun 5, 4850.

Godhe, A., Asplund, M.E., Härnström, K., Saravanan, V., Tyagi, A., and Karunasagar, I. (2008). Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74, 7174–7182.

Gong, J., Dong, J., Liu, X., and Massana, R. (2013). Extremely high copy numbers and polymorphisms of the rDNA operon estimated from single cell analysis of oligotrich and peritrich ciliates. Protist 164, 369–379.

Govindaraju, D.R., and Cullis, C.A. (1992). Ribosomal DNA variation among populations of a *Pinus rigida* Mill. (pitch pine) ecosystem: I. Distribution of copy numbers. Heredity 69, 133–140.

Grattepanche, J.D., McManus, G.B., and Katz, L.A. (2016a). Patchiness of ciliate communities sampled at varying spatial scales along the New England shelf. PLoS ONE 11, e0167659.

Grattepanche, J.D., Santoferrara, L.F., McManus, G.B., and Katz, L.A. (2016b). Unexpected biodiversity of ciliates in marine samples from below the photic zone. Mol Ecol 25, 3987–4000.

Gribble, K.E., and Anderson, D.M. (2005). High intraindividual, intraspecific, and interspecific variability in large-subunit ribosomal DNA in the heterotrophic dinoflagellates *Protoperidinium*, *Diplopsalis*, and *Preperidinium* (Dinophyceae). Phycologia 46, 315–324.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41, 95–98.

Heyse, G., Jönsson, F., Chang, W.J., and Lipps, H.J. (2010). RNA-dependent control of gene amplification. Proc Natl Acad Sci USA 107, 22134–22139.

Huang, J., and Katz, L.A. (2014). Nanochromosome copy number does not correlate with RNA levels though patterns are conserved between strains of the ciliate morphospecies *Chilodonella uncinata*. Protist 165,

445–451.
Huang, J., Luo, X., Bourland, W.A., Gao, F., and Gao, S. (2016). Multigene-based phylogeny of the ciliate families Amphisiellidae and Trachelostylidae (Protozoa: Ciliophora: Hypotrichia). Mol Phylogenets Evol 101, 101–110.
Huang, L., Lu, X., Zhu, C., Lin, X., and Yi, Z. (2017). Macronuclear Actin copy number variations in single cells of different *Pseudokeronopsis* (Alveolata, Ciliophora) populations. Eur J Protistol 59, 75–81.
Ide, S., Miyazaki, T., Maki, H., and Kobayashi, T. (2010). Abundance of ribosomal RNA gene copies maintains genome integrity. Science 327, 693–696.
Izhaki, I., Fridman, S., Gerchman, Y., and Halpern, M. (2013). Variability of bacterial community composition on leaves between and within plant species. Curr Microbiol 66, 227–235.
Jerome, C.A., Lynn, D.H., and Simon, E.M. (1996). Description of *Tetrahymena empidokyrea* n. sp., a new species in the *Tetrahymena pyriformis* sibling species complex (Ciliophora, Oligohymenophorea), and an assessment of its phylogenetic position using small-subunit rRNA sequences. Can J Zool 74, 1898–1906.
Jones, A.C., Hambright, K.D., and Caron, D.A. (2017). Ecological patterns among bacteria and microbial eukaryotes derived from network analyses in a low-salinity lake. Microb Ecol 75, 917–929.
Kapler, G.M. (1993). Developmentally regulated processing and replication of the *Tetrahymena* rDNA minichromosome. Curr Opin Genets Dev 3, 730–735.
Kathol, M., Norf, H., Arndt, H., and Weitere, M. (2009). Effects of temperature increase on the grazing of planktonic bacteria by biofilm-dwelling consumers. Aquat Microb Ecol 55, 65–79.
Kim, S.Y., Yang, E.J., Gong, J., and Choi, J.K. (2010). Redescription of *Favella ehrenbergii* (Claparède and Lachmann, 1858) Jörgensen, 1924 (Ciliophora: Choreotrichia), with phylogenetic analyses based on small subunit rRNA gene sequences. J Eukaryotic Microbiol 57, 460–467.
Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Liu, X., and Gong, J. (2012). Revealing the diversity and quantity of peritrich ciliates in environmental samples using specific primer-based PCR and quantitative PCR. Microb Environ 27, 497–503.
Luo, X., Gao, F., Yi, Z., Pan, Y., Al-Farraj, S.A., and Warren, A. (2017). Taxonomy and molecular phylogeny of two new brackish hypotrichous ciliates, with the establishment of a new genus (Ciliophora, Spirotrichea). Zool J Linn Soc 179, 475–491.
Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., et al. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat Ecol Evol 1, 0091.
Medinger, R., Nolte, V., Pandey, R.V., Jost, S., Ottenwälder, B., Schlötterer, C., and Boenigk, J. (2010). Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. Mol Ecol 19, 32–40.
Medlin, L., Elwood, H.J., Stickel, S., and Sogin, M.L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71, 491–499.
Nylander, J. (2004). MrModeltest, a program to evaluate the fit of several models of evolution to a given data and unrooted tree (version 2.2). Program distributed by the author (Evolutionary Biology Centre, Uppsala University, Sweden).
Orias, E., and Flacks, M. (1975). Macronuclear genetics of *Tetrahymena*. I. Random distribution of macronuclear genecopies in *T. pyriformis*, syngen 1. Genetics 79, 187–206.
Paredes, S., Branco, A.T., Hartl, D.L., Maggert, K.A., and Lemos, B. (2011). Ribosomal DNA deletions modulate genome-wide gene expression: "rDNA-sensitive" genes and natural variation. PLoS Genet 7, e1001376.
Petz, W., Valbonesi, A., Schiftner, U., Quesada, A., and Cynan Ellis-Evans, J. (2007). Ciliate biogeography in Antarctic and Arctic freshwater ecosystems: endemism or global distribution of species? FEMS MicroBiol Ecol 59, 396–408.
Pillet, L., Fontaine, D., and Pawlowski, J. (2012). Intra-genomic ribosomal RNA polymorphism and morphological variation in *Elphidium macellum* suggests inter-specific hybridization in Foraminifera. PLoS ONE 7, e32373–264.
Prescott, D.M. (1994). The DNA of ciliated protozoa. Microbiol Rev 58, 233–267.
Prokopowich, C.D., Gregory, T.R., and Crease, T.J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. Genome 46, 48–50.
Rehnstam-Holm, A.S., Godhe, A., and Anderson, D.M. (2002). Molecular studies of *Dinophysis* (Dinophyceae) species from Sweden and North America. Phycologia 41, 348–357.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systatic Biol 61, 539–542.
Santoferrara, L.F., Grattepanche, J.D., Katz, L.A., and McManus, G.B. (2016). Patterns and processes in microbial biogeography: do molecules and morphologies give the same answers? ISME J 10, 1779–1790.
Santoferrara, L.F., McManus, G.B., and Alder, V.A. (2013). Utility of genetic markers and morphology for species discrimination within the order Tintinnida (Ciliophora, Spirotrichea). Protist 164, 24–36.
Sela, I., Ashkenazy, H., Katoh, K., and Pupko, T. (2015). GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. Nucl Acids Res 43, W7–W14.
Simon, D., Moline, J., Helms, G., Friedl, T., and Bhattacharya, D. (2005). Divergent histories of rDNA group I introns in the lichen family Physciaceae. J Mol Evol 60, 434–446.
Simon, U.K., and Weiss, M. (2008). Intragenomic variation of fungal ribosomal genes is higher than previously thought. Mol Biol Evol 25, 2251–2254.
Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 6, 152–155.
Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stock, A., Edgcomb, V., Orsi, W., Filker, S., Breiner, H.W., Yakimov, M. M., and Stoeck, T. (2013). Evidence for isolated evolution of deep-sea ciliate communities through geological separation and environmental selection. BMC Microbiol 13, 150.
Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H. W., and Richards, T.A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol Ecol 19, 21–31.
Stoeck, T., Przybos, E., and Dunthorn, M. (2014). The D1-D2 region of the large subunit ribosomal DNA as barcode for ciliates. Mol Ecol Resour 14, 458–468.
Strauss, S.H., and Tsai, C.H. (1988). Ribosomal gene number variability in Douglas-Fir. J Heredity 79, 453–458.
Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24, 1596–1599.
Taniguchi, A., Onishi, H., and Eguchi, M. (2011). Quantitative PCR assay for the detection of the parasitic ciliate *Cryptocaryon irritans*. Fish Sci 77, 607–613.
Terrado, R., Medrinal, E., Dasilva, C., Thaler, M., Vincent, W.F., and Lovejoy, C. (2011). Protist community composition during spring in an Arctic flaw lead polynya. Polar Biol 34, 1901–1914.
Torres-Machorro, A.L., Hernández, R., Cevallos, A.M., and López-Villaseñor, I. (2010). Ribosomal RNA genes in eukaryotic microorganisms: witnesses of phylogeny? FEMS Microbiol Rev 34, 59–86.
Tucker, S.J., McManus, G.B., Katz, L.A., and Grattepanche, J.D. (2017). Distribution of abundant and active planktonic ciliates in coastal and slope waters off New England. Front Microbiol 8, 2178.
Wancura, M.M., Yan, Y., Katz, L.A., and Maurer-Alcalá, X.X. (2018). Nuclear features of the heterotrich ciliate *Blepharisma americanum*: genomic amplification, life cycle, and nuclear inclusion. J Eukaryot

Microbiol 65, 4–11.
Wang, C., Zhang, T., Wang, Y., Katz, L.A., Gao, F., and Song, W. (2017). Disentangling sources of variation in SSU rDNA sequences from single cell analyses of ciliates: impact of copy number variation and experimental error. Proc R Soc B 284, 20170425.
Wang, P., Wang, Y., Wang, C., Zhang, T., Al-Farraj, S.A., and Gao, F. (2017). Further consideration on the phylogeny of the Ciliophora: analyses using both mitochondrial and nuclear data with focus on the extremely confused class Phyllopharyngea. Mol Phylogenets Evol 112, 96–106.
Wang, Y., Wang, Y., Sheng, Y., Huang, J., Chen, X., Al-Rasheid, K.A.S., and Gao, S. (2017a). A comparative study of genome organization and epigenetic mechanisms in model ciliates, with an emphasis on *Tetrahymena*, *Paramecium* and *Oxytricha*. Eur J Protistol 61, 376–387.
Wang, Y., Chen, X., Sheng, Y., Liu, Y., and Gao, S. (2017b). N6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. Nucl Acids Res 45, 11594–11606.
Wang, Y., Sheng, Y., Liu, Y., Pan, B., Huang, J., Warren, A., and Gao, S. (2017c). N 6 -methyladenine DNA modification in the unicellular eukaryotic organism *Tetrahymena thermophila*. Eur J Protistol 58, 94–102.
Wilbert, N. (1975). Eine verbesserte technik der protargolimprägnation für ciliaten. Mikrokosmos 64, 171–179.
Wylezich, C., Nies, G., Mylnikov, A.P., Tautz, D., and Arndt, H. (2010). An evaluation of the use of the LSU rRNA D1-D5 domain for DNA-based taxonomy of eukaryotic protists. Protist 161, 342–352.
Xu, H., Zhang, W., and Jiang, Y. (2014). Do early colonization patterns of periphytic ciliate fauna reveal environmental quality status in coastal waters? Environ Sci Pollut Res 21, 7097–7112.
Xu, K., Doak, T.G., Lipps, H.J., Wang, J., Swart, E.C., and Chang, W.J. (2012). Copy number variations of 11 macronuclear chromosomes and their gene expression in *Oxytricha trifallax*. Gene 505, 75–80.
Yan, Y., Rogers, A.J., Gao, F., and Katz, L.A. (2017). Unusual features of non-dividing somatic macronuclei in the ciliate class Karyorelictea. Eur J Protistol 61, 399–408.
Yao, M.C., Kimmel, A.R., and Gorovsky, M.A. (1974). A small number of cistrons for ribosomal RNA in the germinal nucleus of a eukaryote, *Tetrahymena pyriformis*. Proc Natl Acad Sci USA 71, 3082–3086.
Zhao, F., Filker, S., Stoeck, T., and Xu, K. (2017). Ciliate diversity and distribution patterns in the sediments of a seamount and adjacent abyssal plains in the tropical Western Pacific Ocean. BMC Microbiol 17, 192.
Zhao, X., Wang, Y., Wang, Y., Liu, Y., and Gao, S. (2017). Histone methyltransferase TXR1 is required for both H3 and H3.3 lysine 27 methylation in the well-known ciliated protist *Tetrahymena thermophila*. Sci China Life Sci 60, 264–270.
Zhao, Y., Yi, Z., Gentekaki, E., Zhan, A., Al-Farraj, S.A., and Song, W. (2016). Utility of combining morphological characters, nuclear and mitochondrial genes: an attempt to resolve the conflicts of species identification for ciliated protists. Mol Phylogenets Evol 94, 718–729.
Zhao, Y., Yi, Z., Warren, A., and Song, W.B. (2018). Species delimitation for the molecular taxonomy and ecology of the widely distributed microbial eukaryote genus *Euplotes* (Alveolata, Ciliophora). Proc R Soc B 285, 20172159.
Zhu, F., Massana, R., Not, F., Marie, D., and Vaulot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52, 79–92.

SUPPORTING INFORMATION

Figure S1 The liner relationships between the cycle threshold (Ct) and the logarithms of rDNA copy number of standards in all 11 times of qPCR reactions in this study. Correspondence between reactions and species: A, *Deviata* sp., *Paramecium caudatum*, *Caenomorpha medusula* and *Stentor roeselii*; B, *Strombidium stylifer*, *Diophrys scutum* and *Uronychia binucleate*; C, *Vorticella* sp.-1,2, *Apodileptus visscheri* and *Favella ehrenbergii*; D, *Oxytricha trifallax* and *Chilodontopsis depressa*; E, *Coleps* sp.; F, *Loxodes* sp.; G, *Vorticella* sp.-3, *Plagiopyla* sp. and *Euplotes vannus*; H, *Pseudocohnilembus persalinus* and *Tetrahymena thermophila*-1; I, *Tetrahymena thermophila*-2; J, *Trithigmostoma* sp.; K, *Tetrahymena thermophila*-3 and *Litonotus* sp.

Table S1 Portion of common sequence and rDNA copy numbers of the three individuals of the 20 species

Table S2 Genetic distances and polymorphic sites of SSU rDNA in the 20 species

The supporting information is available online at http://life.scichina.com and https://link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.