PROCEEDINGS B

Research

Authors for correspondence:
Yan Zhao
e-mail: yanzhao@rcees.ac.cn
Zhenzhen Yi
e-mail: zyi@scnu.edu.cn

Species delimitation for the molecular taxonomy and ecology of the widely distributed microbial eukaryote genus *Euplotes* (Alveolata, Ciliophora)

Yan Zhao[1,2], Zhenzhen Yi[3], Alan Warren[4] and Weibo B. Song[2,5]

[1]Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, People's Republic of China
[2]Laboratory of Protozoology, Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao 266003, People's Republic of China
[3]Guangzhou Key Laboratory of Subtropical Biodiversity and Biomonitoring, School of Life Science, South China Normal University, Guangzhou 510631, People's Republic of China
[4]Department of Life Sciences, Natural History Museum, London SW7 5BD, UK
[5]Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266003, People's Republic of China

YZ, 0000-0002-3697-0282; ZY, 0000-0002-0693-9989

Recent advances in high-throughput sequencing and metabarcoding technologies are revolutionizing our understanding of the diversity and ecology of microbial eukaryotes (protists). The interpretation of protist diversity and the elucidation of their ecosystem function are, however, impeded by problems with species delimitation, especially as it applies to molecular taxonomy. Here, using the ciliate *Euplotes* as an example, we describe approaches for species delimitation based on integrative taxonomy by using evolutionary and ecological perspectives and selecting the most appropriate metabarcoding gene markers as proxies for species units. Our analyses show that: *Euplotes* (*sensu lato*) comprises six distinct clades, mainly as result of ecological speciation; the validity of the genera *Euplotes* (*sensu stricto*), *Euplotoides*, *Euplotopsis* and *Moneuplotes* are not supported; the *vannus*-type group, which includes species without distinct morphological differences, seems to be undergoing incipient speciation and contains cryptic species; the hypervariable V4 region of the small subunit rDNA and D1–D2 region of the large subunit rDNA are the promising candidates for general species delimitation in *Euplotes*.

1. Introduction

Unicellular eukaryotes (protists) are valuable biological models that have caused much of the recent upsurge of interest in microbial ecology because of their enormous diversity [1,2]. *Euplotes* Ehrenberg, 1830 (Alveolata, Ciliophora) is a highly differentiated and speciose genus that plays key ecological roles in a wide range of environments; marine, freshwater, soil and hypersaline [1]. Nowadays, there is growing interest in using *Euplotes* in various research fields including: as biological indicators of environmental quality [3]; as model organisms for investigating symbiotic relationships [4]; and in investigations of adaptive mechanisms in strictly psychrophilic species [5]. In common with most protists, however, ecological studies of *Euplotes* are constrained by lack of understanding of species delimitation.

Species delimitation, which lies at the core of taxonomy and ecology, is related to the boundary between micro- and macro-evolution and determines the boundaries between, and number of, species units [6,7]. For unicellular eukaryotes, speciation is a result of adaptation to different environments and is often related to adaptive phenotypic plasticity, diverse ecological strategies and genetic differentiation [7]. Thus, species delimitation should be assessed with integrative approaches and the properties of molecular species

THE ROYAL SOCIETY
PUBLISHING

2

delimitation should be verified within this framework. This is necessary because incorrect interpretation of molecular species delimitation would, for example, limit our ability to interpret species diversity based on high-throughput sequencing (HTS) data, a technique which has revolutionized our understanding of the diversity and ecology of protists in general [8,9].

Traditionally, the morphospecies concept has dominated protist diversity studies [10,11]. A morphospecies is defined as a collection of forms that all fit into a defined range of morphological variation—forms that, so far as we can tell, all occupy the same ecological niche. Ciliates in particular are rich in morphological characters that are taxonomically informative. Some morphological traits (body size, shape, etc.) used for species identification are, however, variable even within clones, and are not suitable for separating species that have subtle morphological differences or high phenotypic plasticity [12,13]. In the case of *Euplotes*, body shape and size vary considerably under different nutritional conditions [13,14], or in the presence of certain predators [15,16]. Gates [17] found that certain patterns of the fronto-ventral cirri (FVC) may be variable within both *E. harpa* and *E. minuta*, after being cultured in different conditions of temperature and salinity. Hufnagel & Torch [18] observed intra-clonal dimorphism of caudal cirri in *E. vannus*. Finally, the dargyrome, an important character in the systematics of *Euplotes*, has been shown to exhibit an extensive range of variability, even within clonal cultures [19,20].

As a consequence of these findings, species identification in *Euplotes* tends to be challengeable even for experienced protistologists. Therefore, although *Euplotes* has been studied for nearly two centuries resulting in a significant body of the literature [13], there remain numerous unresolved problems concerning the systematics of the genus, for instance, the identification and species separation of members of the *vannus*-type group, the validity controversy of *E. woodruffi*, *E. balteatus*, *E. charon*, etc., and the proposed division of *Euplotes sensu lato* into *Euplotes sensu stricto*, *Euplotoides*, *Euplotopsis* and *Moneuplotes* [21]. An integrative study of *Euplotes* taxonomy is therefore needed in order to resolve such problems. In a previous study, we attempted to identify *Euplotes* species by combining a single-gene marker (small subunit (SSU) rDNA) with dargyrome type and habitat preference [22]. Several gene markers of additional *Euplotes* species have since been sequenced providing the opportunity to study this genus with a more comprehensive dataset.

In this investigation, we focus on *Euplotes* species that have been identified by expert taxonomists. We assess species delimitation properties of gene fragments (SSU rDNA, large subunit (LSU) rDNA, internal transcribed spacer (ITS)-5.8S rDNA) in particular their hypervariable regions (ITS1, ITS2, D1–D2, V4, V9) that are commonly promoted as species proxies in molecular ecology. By combining DNA data, ecological niche properties and morphological characters, we outline assumable evolutionary trajectories of *Euplotes* and assess the feasibility of using DNA taxonomy for investigating *Euplotes* species diversity.

2. Material and methods

(a) Species collection and morphological identification

Most of the *Euplotes* species treated in this study were collected from northern and southern coastal regions of China, plus three species collected from the intertidal region of a Saudi Arabian sandy beach on the Arabian Gulf, and one species from Tibetan soil (see the electronic supplementary material, table S1 for details). *Euplotidium* sp. was collected from Zhanjiang, China (electronic supplementary material, table S1). Species were identified and described in detail by expert taxonomists [23–26]. Data for additional species and populations were obtained from the literature (electronic supplementary material, table S1). Terminology and systematic classification are according to Lynn [1] and Gao *et al.* [27,28].

The validity of each species, including those identified in the present study and in previous investigations, was carefully evaluated by comparing the data with the original (where available) and revised descriptions. It was found that some descriptions are extremely brief and lack sufficient detail for reliable species determination. Based on these findings, we scored the likely validity of each *Euplotes* species (electronic supplementary material, table S1) and used these scores to guide subsequent analyses.

(b) DNA extraction, gene amplification and sequencing

DNA extraction and gene amplification were performed according to Zhao *et al.* [29]. The purified nuclear amplicons were sequenced directly with the amplification primers combined with the designed walking primers in Life Technologies (Invitrogen™ sequencing facility, Guangzhou & Beijing, China). The new sequences have been deposited in GenBank (for accession numbers, see the electronic supplementary material, table S1).

(c) rDNA data retrieval and sequence alignment

We used python scripts to download from GenBank all available SSU rDNA, ITS-5.8S rDNA and LSU rDNA sequences for *Euplotes* and for four out of five outgroup species (*Certesia quadrinucleata*, *Aspidisca magna*, *Diophrys scutum* and *Uronychia setigera*), which include a total of 176 SSU rDNA, 62 ITS-5.8S rDNA and 41 LSU rDNA. Using another python script, 246 sequences of V4/V9-SSU rDNA, 132 sequences of the ITS1/ITS2 and 106 sequences of D1–D2-LSU rDNA were obtained for metabarcoding analysis. Sequence alignments were performed using the MUSCLE algorithm (https://www.ebi.ac.uk/Tools/msa/muscle/) with minor manual adjustments.

(d) Phylogenetic analyses

Phylogenetic relationships based on each gene marker were inferred separately using maximum likelihood (ML) and Bayesian inference (BI). ML analyses were conducted using randomized axelerated ML for high performance computing (RAxML HPC) at CIPRES (http://www.phylo.org) with the general time reversible (GTR) model and gamma distribution approximated by four rate categories based on 1000 bootstrap samples [30]. Models of nucleotide substitution were evaluated for the data using jMODELTEST 2 [31] based on Akaike information criterion (AIC) and Bayesian information criterion (BIC). BI analyses were conducted using MRBAYES v. 3.2 with the selected parameters (rates = invgamma, number of substitution types = 6, ncat = 4) corresponding to the model estimated (GTR + I + Γ) [32,33]. Posterior probability distributions were generated using Markov chain Monte Carlo methods. Bayesian posterior probabilities were computed by one cold and three heated chains running simultaneously for 1 000 000 generations, with trees being sampled every 100 generations and the first 2500 trees discarded as 'burn-in', as described previously [29].

In this study, we integrated information based on morphological and ecological properties (see details in the electronic supplementary material, table S1) with the phylogenetic relationships recovered using ML and built explicit phylogenetic networks using SPLITSTREE 4 [34]. Based on this, we provided hypothetical evolutionary trajectories of members of the genus *Euplotes*.
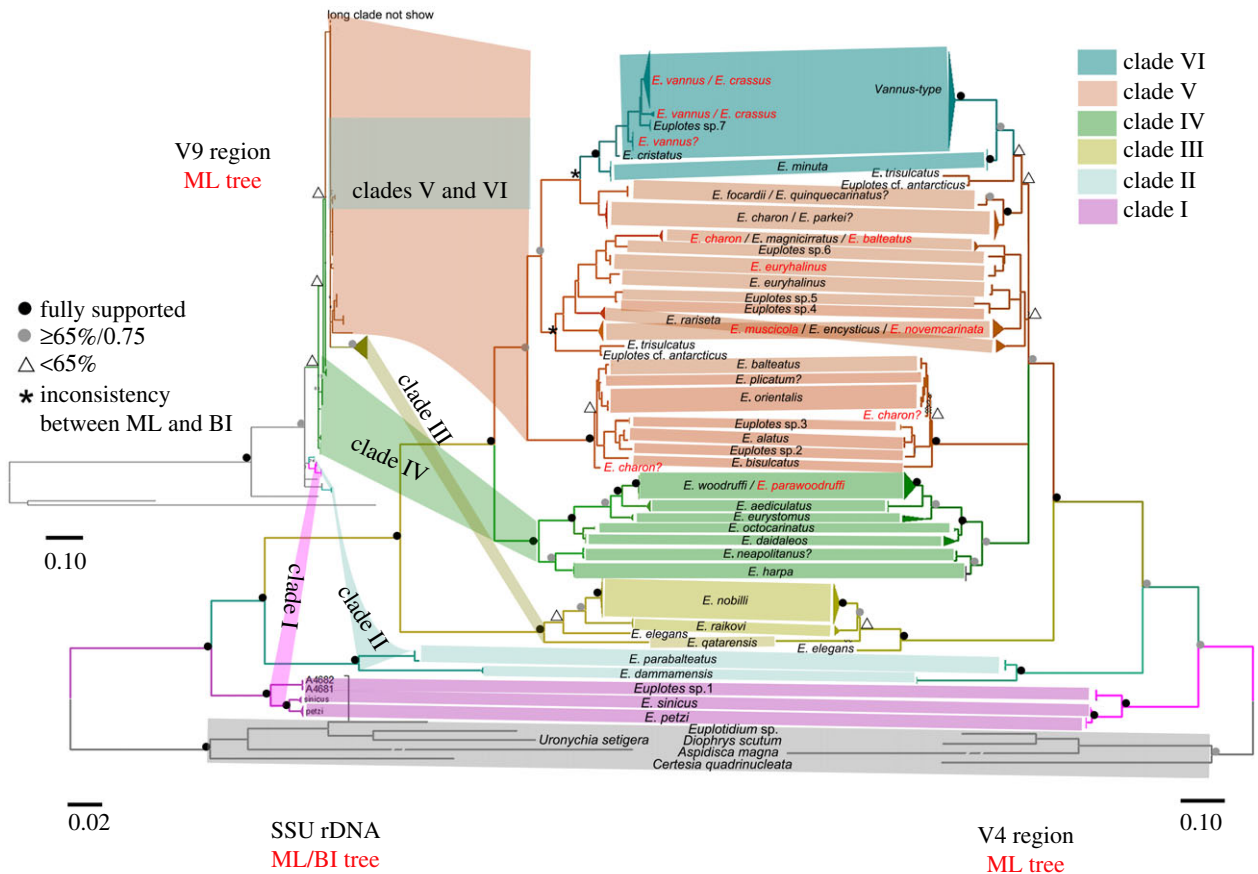
Figure 1. Maximum likelihood (ML) and Bayesian analysis trees of *Euplotes* species based on SSU rDNA and hypervariable region V4 and V9 data. Numbers at nodes represent the bootstrap values. Fully supported (100%/1.00) branches are marked with solid circles. Grey circles indicate bootstraps above 65%/0.75, and bootstraps less than 65% are represented by triangles. Asterisk (*) indicates disagreement between ML and BI analyses. The scale bar corresponds to 10 substitutions per 100 nucleotide positions in V4 and V9 trees, and 2 in SSU trees. (Online version in colour.)

(e) Morphological and ecological characters

A '*Euplotes*' keyword search of PubMed and Google Scholar provided access to all relevant previous publications. The consistency of the resulting publications was then manually checked. Each publication was used to obtain morphological and ecological characters of the species in question, along with the sequence data from the NCBI database. The resulting comprehensive datasets allowed an investigation based on integrative taxonomy to be carried out.

3. Results

The topologies of phylogenetic trees based on SSU rDNA, ITS-5.8S region and LSU rDNA are nearly identical (figures 1–3). All phylogenetic analyses show that *Euplotes* s.l. is monophyletic and comprises two sister groups with full support: a species-poor group represented by the extant species *E. petzi/ E. sinicus* (clade I) and a species-rich group comprising clades II–VI. Within the latter group, clade II (*E. parabalteatus + E. dammaensis*), clade III (*E. raikovi + E. nobilii + E. elegans + E. qatarensis*) and clade IV (*E. harpa + E. daidaleos + E. eurystomus + E. aediculatus + E. woodruffi* (including some wrongly named *parawoodruffi*) + *E. neaplolitanus*?) are all fully supported in the SSU rDNA and LSU rDNA trees. Clades V and VI are assumed to be progenies of a hypothetical common ancestor supported by SSU rDNA and ITS-5.8S region trees. The independence of clade VI (*vannus*-type species) is well supported in the SSU rDNA and LSU rDNA trees (figures 1 and 3). The radiation of the remaining species (clade V) is not well resolved, which is consistent with previous studies [22,35].

The hypervariable regions reveal more diverse tree topologies (figures 1–3). The V4 region seems to have a similar evolution rate to the SSU rDNA and has a closely matching tree topology for all clades except the *vannus*-type group (clade VI). By contrast, the V9 region fails to resolve the evolutionary relationships of several species with clades V and VI virtually forming a single clade (figure 1). The ITS1 and ITS2 trees each show higher resolution than the ITS-5.8S trees for species of the *vannus*-type group (clade VI). In ITS1 tree, the *E. harpa + E. woodruffi (parawoodruffi) + E. neaplolitanus*? group is placed in clade V, which makes relationships among species of clade V become more tangled (figure 2). The ITS2 tree shows that clade III is sister to clade II, rather than diverging before clades IV, V, VI (figure 2). In the D1−D2 tree, clade IV is divided into two sub-clades, i.e. *E. harpa + E. neaplolitanus*? and *E.woodruffi + E. aediculatus* (figure 3). The positions of *Euplotes* sp. 4 and *Euplotes* sp. 5, which fall into clade V in the SSU rDNA and V4 trees (figure 1), are unstable in the ITS-5.8S rDNA, ITS2, LSU rDNA and D1−D2 trees (figures 2 and 3).

The above phylogenetic relationships of the certain *Euplotes* species were re-examined in a framework of hypothetical evolutionary trajectories of *Euplotes* (figure 4 and the electronic supplementary material, figure S2).

4. Discussion

The goal of species delimitation is determining boundaries between species units. Traditionally, the process of speciation is a stage of evolutionary divergence that accompanies
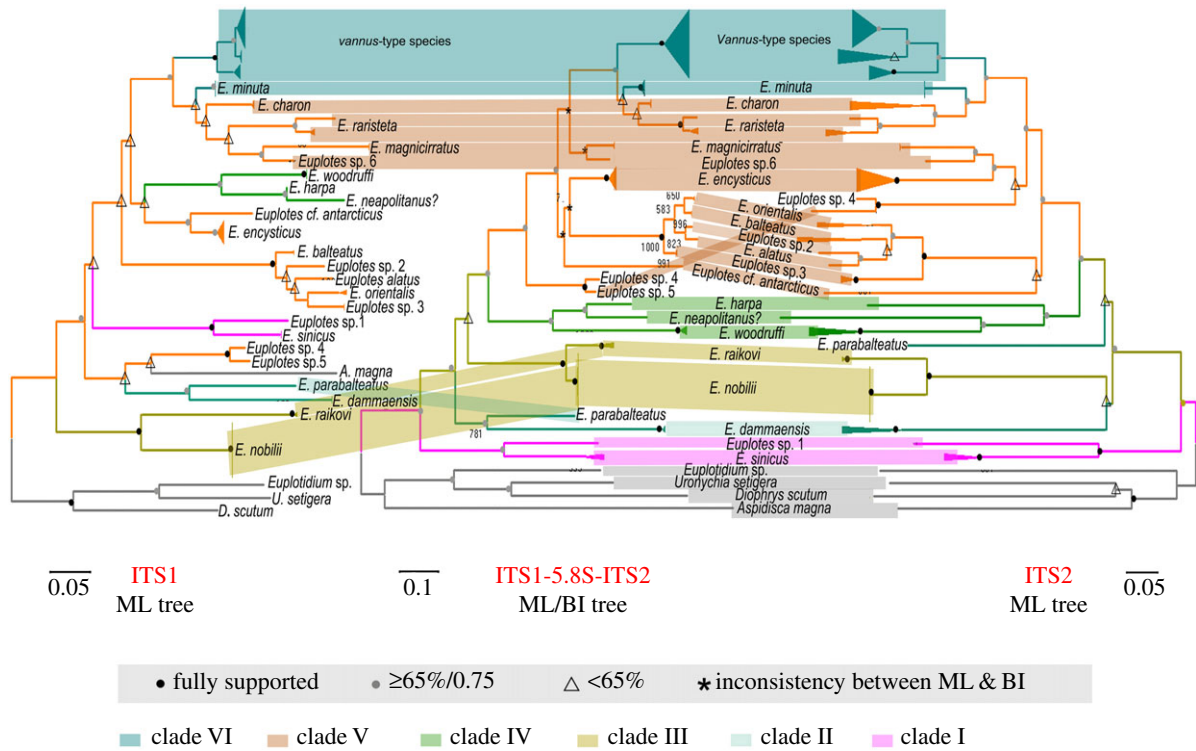
Figure 2. Maximum likelihood (ML) and Bayesian analysis trees of *Euplotes* based on ITS-5.8S rDNA and hypervariable region ITS1 and ITS2 data. Numbers at nodes represent the bootstrap values. Fully supported (100%/1.00) branches are marked with solid circles. Grey circles indicate bootstraps above 65%/0.75, and bootstraps less than 65% are represented by triangles. Asterisk (*) indicates disagreement between ML and BI analyses. The scale bar corresponds to five substitutions per 100 nucleotide positions in ITS1 and ITS2 trees, and 10 in ITS-5.8S trees. (Online version in colour.)
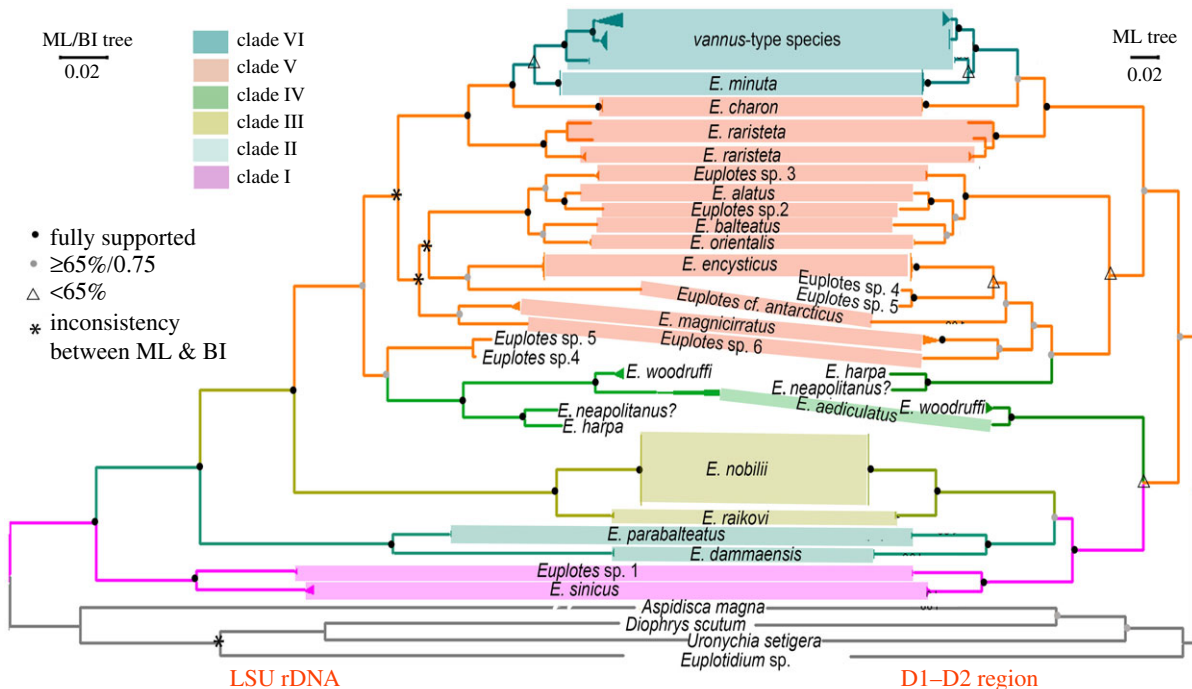
Figure 3. Maximum likelihood (ML) and Bayesian analysis trees of *Euplotes* based on LSU rDNA and hypervariable region D1–D2 data. Numbers at nodes represent the bootstrap values. Fully supported (100%/1.00) branches are marked with solid circles. Grey circles indicate bootstraps above 65%/0.75, and bootstraps less than 65% are represented by triangles. Asterisk (*) indicates disagreement between ML and BI analyses. The scale bar corresponds to two substitutions per 100 nucleotide positions. (Online version in colour.)

ecological, physiological and/or genetic variations [36]. Thus, integrated information must be taken into account in order to facilitate accurate species identification and for the determination of species diversity and ecological function. For unicellular eukaryotes such as ciliates, new species are established by alpha taxonomic research that invariably includes descriptions of their morphology [1]. Morphological characters, however, cannot be the sole basis for defining species boundaries of unicellular eukaryotes owing to their adaptive phenotypic plasticity in response to different environmental fluctuations. In addition, genetic divergence alone cannot provide a comprehensive species definition without phenotypic
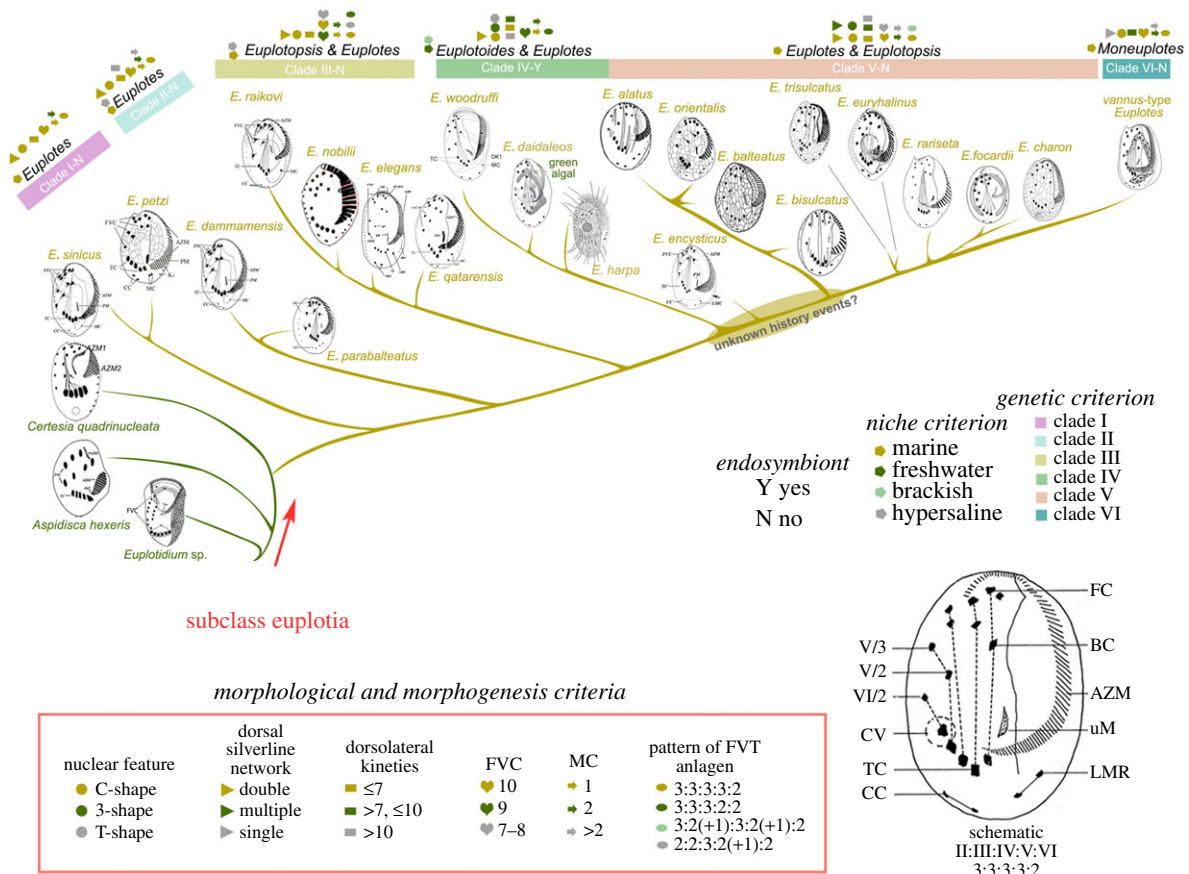
Figure 4. Possible evolutionary trajectories of different *Euplotes* lineages. Morphological and ecological characteristics are indicated by different colours and shapes respectively. The schematic drawing of *Euplotes* is given to demonstrate classical diagnostic characters. Abbreviations: BC, buccal cirrus; CC, caudal cirri; CV, contractile vacuole; FC, frontal cirri; FVT, frontoventral transverse cirri; LMR, left marginal cirral row; TC, transverse cirri; UM, undulating membrane; V/2, V/3 and VI/2, frontoventral cirri V/2, V/3 and VI/2. Dotted lines in the schematic drawing connect cirri originating from the same anlage. (Online version in colour.)

and evolutionary perspectives [37]. We therefore conducted a case study of species delimitation focused on *Euplotes* species using an integrative approach that combined phylogenetic, morphological and ecological data.

(a) Species delimitation could be confounded by phenotypic plasticity and indistinguishable morphology

Morphological characters, including those used for species or even genus separation, are often variable within *Euplotes*. Among these are the marginal cirri (MC) and FVC, the dargyrome and the macronucleus. Different combinations of numbers and arrangement of frontoventral, ventral, caudal (FVC pattern) and marginal cirri on the ventral surface were considered meaningful for the division of *Euplotes* into several genera [21]. The dargyrome pattern is routinely used for species circumscription and has proved especially useful for separation of taxa with similar FVC patterns. The shape of the macronucleus is taxonomically informative in the genus *Euplotes*, and sometimes unique to a single species (e.g. *E. woodruffi*). These features, along with the body shape, size and colour, have traditionally formed the basis for the alpha-taxonomy of *Euplotes* and its relatives [13,21]. The above-mentioned characters, however, are not always stable. Dargyrome patterns, for example, which have been regarded as the gold standard for species separation, do not reflect genuine variations in the structure and assemblage of the cortical alveoli [14,21,38]. Gates & Curds [19] carried out a quantitative

analysis of the dargyrome in clonal cultures of *Euplotes* and found that the three subcategories of the 'double' dargyrome pattern, i.e. 'double-patella', 'double-eurystomus' and 'complex', are variable within a clone and therefore should not be used in assessing taxonomic affinities. Furthermore, Tuffrau *et al.* [20] found that a marine tropical *Euplotes* exhibits an extensive range of variability of the dargyrome, from single to complex type. Body shape and size of *Euplotes* species vary considerably under different nutritional conditions, even within clonal cultures, and therefore are of limited taxonomic value [13,14]. Predator-induced morphological defence responses have been documented in several species, e.g. *E. octocarinatus*, *E. patella* and *E. aediculatus*, resulting in dramatic changes in size and/or body shape in the presence of certain predators [15,16]. Finally, the stability of cirral patterns has also been brought into question. Gates [17] found that certain FVC patterns may be variable within both *E. harpa* and *E. minuta*, after being cultured in different conditions of temperature and salinity, and Hufnagel & Torch [18] observed intra-clonal dimorphism of caudal cirri in *E. vannus*.

It is reported that some morphological appearance could be changed by habitat filtering as a response to environmental variations, but it is difficult to determine at which point this variation in the phenotype resulted in the establishment of a new species [39]. For instance, on the one hand we observed morphological similarity between *E. crassus* and *E. vannus* even though their genotypes differ significantly, but on the other hand phenotypic plasticity in the form of different sizes (large versus small) of the adoral zone of membranelles (AZM) in cells of *E. charon* that have identical SSU rDNA

sequences (data not shown). Conversely, distinct SSU rDNA genotypes were also detected among different individuals in clonal cultures of *E. charon* (electronic supplementary material, figure S1) owing to intra-individual SSU rDNA polymorphism [40]. That is, some morphological characters (e.g. size of AZM) should not be used for species delimitation as they may vary as a result of environmental or other factors rather than being genetically determined. By contrast, genetic variation might not result in morphological differentiation, although it could reveal signatures of adaptation dynamics during speciation [36].

From these data, we were unable to identify causal relationships between morphological differentiation and genetic variation that accompanies speciation. There should be a database that includes two or more criteria such as morphological characters, genetic markers and ecological features to deepen our understanding of which species has different phenotypic properties, and which genotype could vary within populations, individuals and clones under different environmental conditions. An integrative approach such as this could be used to minimize subjective judgements of species delimitation.

(b) Adaptive strategies and their influence on speciation interpretation

Ciliates are known to have been in existence since before the Precambrian and have therefore had many years to evolve adaptive strategies for survival in a wide range of different ecological niches [41]. Both genetic and ecological processes are involved in speciation but their effects are different [42,43]. It has been assumed that genetics provide the framework for the evolutionary trajectory of ciliates, while environmental factors influence genetic divergence among related species and of populations of the same species [37].

Euplotes sinicus and *E. petzi* (clade I) completed the speciation process following adaptation to their environmental niches in the polar regions. The psychrophile *E. petzi* comprises at least two distinct genotypes, one each in the Arctic and Antarctic, because gene flow between the two is restricted by geographical barriers [44]. *Euplotes sinicus* and *E. petzi* maintained slow rates of evolution, their short branches on phylogenetic trees suggesting they are genetically similar to their common ancestor (figures 1–3). The ancestors of clades II, III and IV may have undergone ecological speciation and evolved independently. The hypersaline environment probably resulted in the allopatric speciation of *E. dammanmensis* and caused the divergence of the common ancestor of clade II [43]. The common ancestor of clade III might well have existed before the great transition of marine life to land [45]. Possibly as a result of environmental changes during this period, this species evolved different adaptive strategies resulting in different patterns of cortical development and adaptation to new ecological niches (figure 4). Other adaptations included the ability to form cysts and the harbouring of endosymbionts [46,47]. Within this scenario, one member of clade III, harbouring symbionts of different kinds compared to those of its close relatives, may have adapted to brackish water and become the ancestor of clade IV. Previous studies have found that *Euplotes* species harbouring *Polynucleobacter*-like bacteria represent a distinct monophyletic group, suggesting that they arose from a common ancestor [4]. This may then have

been followed by multiple speciation events in different freshwater niches because it has been reported that all freshwater species of *Euplotes* harbour *Polynucleobacter*-like bacteria [4].

The relationships between adaptive strategies and speciation became complex in the species-rich clades V and VI. Their marine ancestors might have undergone different degrees of ecological speciation and gene flow [43] as a result of which their descendants became diverse and their phylogenetic relationships became tangled [35] (figures 1–3). Over time, the members of this clade gradually occupied a broad range of habitats and niches (for details, see the electronic supplementary material, table S1), probably resulting in geographical isolation and increased genetic divergence [48]. Support for this hypothesis is seen in *E. charon* which has relatively high genetic differentiation (electronic supplementary material, figure S1) and worldwide distribution [1]. Furthermore, sympatry between recently split species may also have arisen through range changes occurring after speciation [49]. This has resulted in difficulties in species delimitation owing to similarities in their morphology and adaptive strategies. Consequently, errors in the identification of such species seem to be frequent, e.g. both *E. novemcarinatus* (HM140402) and *E. muscorum* (HM140407) are populations of *E. encysticus* and *E. encysticus*, respectively, and *E. charon* (AJ305249) might be closely related to species in clade V. The identity of these species has been previously questioned [24,35]. Another example includes the delimitation of *E. encysticus* and its separation from *E. novemcarinatus* (=*E. moscicola*?) and *E. muscorum*, because these three species have similar morphologies and low genetic differentiation [26,50]. Members of clade VI, the *vannus*-type *Euplotes*, may be undergoing incipient speciation because this group includes numerous phylotypes with little morphological diversification [36]. Delimitation of these species is therefore challenging and is discussed in detail in the following section.

(c) Taxonomic controversy and cryptic species in *vannus*-type *Euplotes* species

Taxonomic controversy focused on *vannus*-type *Euplotes* exemplifies a more general difficulty of delimiting species boundaries [12]. In this group, *E. vannus*, *E. crassus* and *E. minuta* possess a single-type dargyrome and a cirrotype-10 pattern [12,21,51]. *Euplotes minuta* can be easily separated from the other two species by its small body size [13,51,52]. Our phylogenetic trees show that *E. minuta* diverged early within clade VI (figures 1–3). Based on their original diagnoses, *E. vannus* and *E. crasssus* can be easily separated by their body size (large versus intermediate, respectively). Tuffrau, however, reported *E. vannus* (not *E. crassus*) to be the species with intermediate dimensions [53]. The distinction between these two species was subsequently thrown further into doubt [51,53]. In addition, there is evidence that *E. crassus* and *E. vannus* can interbreed under laboratory conditions [51]. The boundary between *E. crassus* and *E. vannus* might also be blurred by potential hybridization and introgression even though interbreeding between them may be restricted in nature.

In this investigation, we sampled 12 populations of '*Euplotes vannus*', 11 from China and one from Saudi Arabia, (electronic supplementary material, table S1). Their rDNA sequences and all available sequences of *E. vannus* and *E. crassus* in GenBank,

formed three distinctly tangled clades in the ITS1/ITS2 phylogenetic trees, and four in the SSU/LSU/D1−D2 trees (figures 1−3). It is therefore likely that *E. vannus* and *E. crassus* are undergoing incipient sympatric speciation [36], and cryptic species exist in *vannus*-type group, which are difficult to recognize as they exhibit only minor morphological and/or physiological differences.

(d) Phylogenetic networks reveal the evolutionary trajectory of *Euplotes* species

The phylogenetic networks reveal that two main sister groups (clade I, clades II−VI) of *Euplotes* share a common ancestor (electronic supplementary material, figure S2). Therefore, the most shared character states between these two sister groups are presumed to be plesiomorphies of the common ancestor. These include the presence of 10 FVC, 1−2 marginal cirri, double dargyrome, C-shaped macronucleus, 3 : 3 : 3 : 3 : 2 frontoventral transverse anlagen pattern, an absence of endosymbionts and a marine habitat (figure 4).

Euplotes sinicus and *E. petzi* are thought to be the closest relatives of the common ancestor of *Euplotes* [24,44] (figure 4 and electronic supplementary material, figure S2). The psychrophilic *E. petzi* was first recorded as an Antarctic species [54], and later found in Arctic areas [44]. By contrast, *E. sinicus* is eurythermic and has only been recorded in China. Considering the distributions and adaptations to different temperatures of these two species, we speculate that they may have evolved from their common ancestor as a result of allopatric speciation [43].

Within clades II−VI, the ancestor of clade II seems to be clearly separated from its sister group (clades III−VI), and then gave rise to *E. dammanmensis* and *E. parabalteatus*, probably as a result of habitat specialization (hypersaline versus marine, respectively). Up to this point, the plesiomorphies of the common ancestor of all *Euplotes* species may have evolved a slight variation of the double dargyrome and a higher number of marginal cirri (clade II, figure 4).

The ancestor of clade III might have been a cosmopolitan species, its dispersal across different environments resulting in adaptive radiation among its progeny (figure 4). Differentiation among *E. qatarensis*, *E. elegans*, *E. raikovi* and *E. nobilii* became pronounced, probably owing to their environmental heterogeneity (hypersaline lagoon, anoxic water, sediment and polar habitats). Furthermore, they also evolved different adaptive strategies such as cyst formation or the establishment of stable symbiotic relationships with bacteria [46,47,55]. The double dargyrome and C-shaped macronucleus were maintained and new morphological characteristics developed in *E. elegans* and *E. raikovi* such as reduced numbers of FVC and the absence of cirrus VI/2 [25,55].

The ancestor of clade IV might have been cosmopolitan and its adaptation to different environments (freshwater, marine) might thus have enabled it to survive significant evolutionary events such as the great transition. Some of its progeny probably evolved adaptive strategies such as the formation of symbiotic relationships with bacteria or algae, and reduced numbers of FVC. Species in clade IV (including *E. encysticus*) are characterized by the possession of nine FVC and of endosymbiotic bacteria or algae in all species except *E. harpa* [56]. Additionally, most species are found in freshwater habitats, although *E. woodruffi* and *E. harpa* live in brackish waters and *E. encysticus* inhabits fresh and brackish

waters [26]. It has been demonstrated, however, that *E. woodruffi* is capable of salinity adaptation [57], which may have resulted in its separation from its ancestral lineage. The distinctly characteristic T-shaped macronucleus also supports the validity of *E. woodruffi*.

Species of clades V and VI retain some characteristics of their ancestors, such as 10 FVC and a marine habitat (figure 4 and electronic supplementary material, figure S2). Among the extant species within these two clades, multiple- and single-type dargyrome patterns evolved independently within both the *E. encysticus* + *E. muscicola* + *E. novemcarinata* group and the *vannus*-type group, although most have a double dargyrome pattern (figure 4). Clade VI is often considered as a conserved group with single-type dargyrome pattern and containing morphologically indistinguishable species with considerable genetic diversification [51]. This might be explained by the fact that members of clade VI are usually globally distributed and many selective forces might contribute to their genetic diversification. Cryptic or incipient species might exist in this group and species delimitation should be carefully considered owing to uncompleted speciation.

(e) Potential of hypervariable regions for studies of molecular taxonomy and ecology in *Euplotes*

Species delimitation is fundamental to understanding biodiversity and ecology. Using hypervariable regions of the SSU rDNA as a DNA barcode for discriminating species, combined with HTS, has accelerated studies of bacterial communities [58]. By contrast, the application of this methodology to protists is still in its infancy [59]. Previous work has suggested that nuclear rDNA, including SSU rDNA, LSU rDNA and ITS-5.8S rDNA, and mitochondrial cytochrome *c* oxidase subunit I (COI) sequences are potential candidate barcoding gene markers for the molecular taxonomy of ciliates [29]. Although relatively few ciliate genera have been investigated, findings to date suggest that suitable gene markers for DNA barcoding and species delimitation might differ among ciliate groups [29,60].

For molecular ecological studies, the precondition is that there are universal primers to amplify the candidate marker that represents the species diversity. COI is the universal DNA barcoding gene for Metazoa. However, it has proved difficult to design universal COI primers for all groups of ciliates so this gene is not regarded as suitable for metabarcoding, even though it can be used for species identification in particular genera [29]. Suitable alternative regions of the remaining nuclear rDNA should be determined. Modern HTS platforms (e.g. Illumina Miseq PE300) can only sequence gene markers with appropriate length (less than or equal to 550 bp). Therefore, the initial task in molecular microbial ecology is to seek suitable proxies for nuclear DNA from a broad range of taxa.

The present investigation showed that the hypervariable V4 region of SSU rDNA and D1−D2 region of LSU rDNA could resolve evolutionary relationships among *Euplotes* species better than other regions of rDNA (figures 1 and 3) and the D1−D2 region has the potential to differentiate morphospecies and cryptic species among the *vannus*-type group (figure 3). Additionally, their lengths are suitable for HTS sequencing. This is consistent with previous investigations focusing on other ciliate groups, i.e. *Frontonia* [29] and Tintinnida [61], which concluded that D1−D2 performs best for

8

rspb.royalsocietypublishing.org Proc. R. Soc. B 285: 20172159

differentiating between closely related species. Hitherto, theD1–D2 region of the LSU rDNA seemed to be most promising gene marker for species delimitation within ciliates ([29,60], present investigation). However, the D1–D2 region has yet to be assessed for metabarcoding analyses with HTS technology owing to the lack of sufficient reference sequences in public databases. Obviously, LSU rDNA reference databases within an integrative taxonomic framework are urgently needed for improving our understanding of ciliate diversity and ecology.

The V4 region of SSU rDNA, for which public databases include sequences of far more species, is the most widely used in current metabarcoding analyses of ciliates [59,60]. Furthermore, previous studies revealed that the V9 region could provide similar taxon discrimination as the V4 region [59,60]. The present study, however, shows that V9 is not effective in resolving evolutionary relationships among *Euplotes* species because this region contains insufficient phylogenetic signal, resulting in tangled relationships between related species in V9 region trees. By contrast, the V4 region was phylogenetically informative and the derived evolutionary relationships were largely congruent with that using the full-length SSU rDNA. Currently, the V4 region is the most suitable alternative for studies of DNA taxonomy in high-throughput DNA sequencing techniques even though it cannot provide high resolution of relationships among cryptic and incipient species.

References

1. Lynn DH. 2008 *The ciliated protozoa: characterization, classification, and guide to the literature*. Berlin, Germany: Springer.

2. Song WB, Warren A, Hu X. 2009 *Free-living ciliates in the Bohai and Yellow seas*. Beijing, People's Republic of China: Science Press.

3. Vaiopoulou E, Gikas P. 2012 Effects of chromium on activated sludge and on the performance of wastewater treatment plants: a review. *Water Res.* **46**, 549–570. (doi:10.1016/j.watres.2011.11.024)

4. Vannini C, Ferrantini F, Ristori A, Verni F, Petroni G. 2012 Betaproteobacterial symbionts of the ciliate *Euplotes*: origin and tangled evolutionary path of an obligate microbial association. *Environ. Microbiol.* **14**, 2553–2563. (doi:10.1111/j.1462-2920.2012.02760.x)

5. Pucciarelli S, Devaraj RR, Mancini A, Ballarini P, Castelli M, Schrallhammer M, Petroni G, Miceli C. 2015 Microbial consortium associated with the Antarctic marine ciliate *Euplotes focardii*: an investigation from genomic sequences. *Microb. Ecol.* **70**, 484–497. (doi:10.1007/s00248-015-0568-9)

6. Boenigk J, Ereshefsky M, Hoef-Emden K, Mallet J, Bass D. 2012 Concepts in protistology: species definitions and boundaries. *Eur. J. Protistol.* **48**, 96–102. (doi:10.1016/j.ejop.2011.11.004)

7. De Queiroz K. 2007 Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886. (doi:10.1080/10635150701701083)

8. De Vargas C et al. 2015 Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**, 1261605. (doi:10.1126/science.1261605)

9. Caron DA. 2013 Towards a molecular taxonomy for protists: benefits, risks, and applications in plankton ecology. *J. Eukaryot. Microbiol.* **60**, 407–413. (doi:10.1111/jeu.12044)

10. Luo X, Gao F, Yi Z, Pan Y, Al-Farraj SA, 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.

11. Liu W et al. 2017 Diversity of free-living marine ciliates (Alveolata, Ciliophora): faunal studies in coastal waters of China during the years 2011–2016. *Eur. J. Protistol.* **61**, 424–438. (doi:10.1016/j.ejop.2017.04.007)

12. Caprette CL, Gates MA. 1994 Quantitative analyses of interbreeding in populations of vannus-morphotype *Euplotes*, with special attention to the nominal species *E. vannus* and *E. crassus*. *J. Eukaryot. Microbiol.* **41**, 316–324. (doi:10.1111/j.1550-7408.1994.tb06084.x)

13. Curds CR. 1975 A guide to the species of the genus *Euplotes* (Hypotrichida, Ciliatea). *Bull. Br. Mus. Nat. Hist. (Zool.)* **28**, 1–61.

14. Gates MA. 1978 Morphometric variation in the hypotrich ciliate genus *Euplotes*. *J. Eukaryot. Microbiol.* **25**, 338–350.

15. Kuhlmann H-W, Heckmann K. 1985 Interspecific morphogens regulating prey-predator relationships in protozoa. *Science* **227**, 1347–1349. (doi:10.1126/science.227.4692.1347)

16. Kuhlmann H-W, Heckmann K. 1994 Predation risk of typical ovoid and 'winged' morphs of *Euplotes* (Protozoa, Ciliophora). *Hydrobiologia* **284**, 219–227. (doi:10.1007/BF00006691)

17. Gates MA. 1977 Analysis of positional information applied to cirral patterns of the ciliate *Euplotes*. *Nature* **268**, 362–364. (doi:10.1038/268362a0)

18. Hufnagel LA, Torch R. 1967 Intraclonal dimorphism of caudal cirri in *Euplotes vannus*: cortical determination. *J. Eukaryot. Microbiol.* **14**, 429–439.

19. Gates MA, Curds CR. 1979 The dargyrome of the genus *Euplotes*. *Bull. Br. Mus. Nat. Hist. (Zool.)* **35**, 127–200.

20. Tuffrau M, Fryd-Versavel G, Tuffrau H, Génermont J. 2000 Description of *Euplotes versatilis* n. sp., a marine tropical ciliate exhibiting an unusually extensive phenotypic plasticity. *Eur. J. Protistol.* **36**, 355–366. (doi:10.1016/S0932-4739(00)80041-3)

21. Borror AC, Hill BF. 1995 The order Euplotida (Ciliophora): taxonomy, with division of *Euplotes* into several genera. *J. Eukaryot. Microbiol.* **42**, 457–466. (doi:10.1111/j.1550-7408.1995.tb05891.x)

22. Yi Z, Song WB, Clamp JC, Chen Z, Gao S, Zhang Q. 2009 Reconsideration of systematic relationships within the order Euplotida (Protista, Ciliophora) using new sequences of the gene coding for small-subunit rRNA and testing the use of combined data sets to construct phylogenies of the *Diophrys*-complex. *Mol. Phylogenet. Evol.* **50**, 599–607. (doi:10.1016/j.ympev.2008.12.006)

23. Chen X, Zhao Y, Al-Farraj SA, Saleh A-Q, El-Serehy HA, Shao C, Al-Rasheid KA. 2015 Taxonomic descriptions of two marine ciliates, *Euplotes dammamensis* n. sp. and *Euplotes balteatus* (Dujardin, 1841) Kahl, 1932 (Ciliophora, Spirotrichea, Euplotida), collected from the Arabian Gulf, Saudi Arabia. *Acta Protozool.* **52**, 73–87.

24. Jiang J, Zhang Q, Warren A, Al-Rasheid KA, Song W. 2010 Morphology and SSU rRNA gene-based phylogeny of two marine *Euplotes* species, *E. orientalis* spec. nov. and *E. raikovi* (Ciliophora,

Euplotida). *Eur. J. Protistol.* **46**, 121–132. (doi:10.1016/j.ejop.2009.11.003)

25. Jiang J, Zhang Q, Hu X, Shao C, Al-Rasheid KA, Song WB. 2010 Two new marine ciliates, *Euplotes sinicus* sp. nov. and *Euplotes parabalteatus* sp. nov., and a new small subunit rRNA gene sequence of *Euplotes rariseta* (Ciliophora, Spirotrichea, Euplotida). *Int. J. Syst. Evol. Microbiol.* **60**, 1241–1251. (doi:10.1099/ijs.0.012120-0)

26. Fan X, Huang J, Lin X, Li J, Al-Rasheid KA, Hu X. 2010 Morphological and molecular characterization of *Euplotes encysticus* (Protozoa: Ciliophora: Euplotida). *J. Mar. Biol. Assoc. UK* **90**, 1411–1416. (doi:10.1017/S002531541000038X)

27. Gao F *et al.* 2016 The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). *Sci. Rep.* **6**, 24874. (doi:10.1038/srep24874)

28. Gao F *et al.* 2017 Systematic studies on ciliates (Alveolata, Ciliophora) in China: progress and achievements based on molecular information. *Eur. J. Protistol.* **61**, 409–423. (doi:10.1016/j.ejop.2017.04.009)

29. Zhao Y, Yi Z, Gentekaki E, Zhan A, Al-Farraj SA, Song WB. 2016 Utility of combining morphological characters, nuclear and mitochondrial genes: an attempt to resolve the conflicts of species identification for ciliated protists. *Mol. Phylogenet. Evol.* **94**, 718–729. (doi:10.1016/j.ympev.2015.10.017)

30. Stamatakis A. 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)

31. Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772. (doi:10.1038/nmeth.2109)

32. Ronquist F *et al.* 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542. (doi:10.1093/sysbio/sys029)

33. Ronquist F, Huelsenbeck JP. 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. (doi:10.1093/bioinformatics/btg180)

34. Huson DH, Bryant D. 2006 Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267. (doi:10.1093/molbev/msj030)

35. Petroni G, Dini F, Verni F, Rosati G. 2002 A molecular approach to the tangled intrageneric relationships underlying phylogeny in *Euplotes* (Ciliophora, Spirotrichea). *Mol. Phylogenet. Evol.* **22**, 118–130. (doi:10.1006/mpev.2001.1030)

36. Dobzhansky T. 1940 Speciation as a stage in evolutionary divergence. *Am. Nat.* **74**, 312–321. (doi:10.1086/280899)

37. Schlegel M, Meisterfeld R. 2003 The species problem in protozoa revisited. *Eur. J. Protistol.* **39**, 349–355. (doi:10.1078/S0932-4739(04)70109-1)

38. Gates MA. 1988 Dorsal kinetosomal distributions in the ciliate genus *Euplotes*. *J. Eukaryot. Microbiol.* **35**, 506–510.

39. Weisse T. 2008 Distribution and diversity of aquatic protists: an evolutionary and ecological perspective. *Biodivers. Conserv.* **17**, 243–259. (doi:10.1007/s10531-007-9249-4)

40. Wang C, Zhang T, Wang Y, Katz LA, Gao F, Song WB. 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. (doi:10.1098/rspb.2017.0425)

41. Bosak T, Macdonald F, Lahr D, Matys E. 2011 Putative Cryogenian ciliates from Mongolia. *Geology* **39**, 1123–1126. (doi:10.1130/G32384.1)

42. Lewontin RC. 1974 *The genetic basis of evolutionary change*. New York, NY: Columbia University Press.

43. Rundle HD, Nosil P. 2005 Ecological speciation. *Ecol. Lett.* **8**, 336–352. (doi:10.1111/j.1461-0248.2004.00715.x)

44. Di Giuseppe G, Erra F, Frontini FP, Dini F, Vallesi A, Luporini P. 2014 Improved description of the bipolar ciliate, *Euplotes petzi*, and definition of its basal position in the *Euplotes* phylogenetic tree. *Eur. J. Protistol.* **50**, 402–411. (doi:10.1016/j.ejop.2014.05.001)

45. Sues H-D, Fraser NC. 2013 *Triassic life on land: the great transition*. New York, NY: Columbia University Press.

46. Fotedar R, Stoeck T, Filker S, Fell JW, Agatha S, Al Marri M, Jiang J. 2016 Description of the halophile *Euplotes qatarensis* nov. spec. (Ciliophora, Spirotrichea, Euplotida) isolated from the hypersaline Khor Al-Adaid Lagoon in Qatar. *J. Eukaryot. Microbiol.* **63**, 578–590. (doi:10.1111/jeu.12305)

47. Schrallhammer M, Schweikert M, Vallesi A, Verni F, Petroni G. 2011 Detection of a novel subspecies of *Francisella noatunensis* as endosymbiont of the ciliate *Euplotes raikovi*. *Microb. Ecol.* **61**, 455–464. (doi:10.1007/s00248-010-9772-9)

48. Foissner W, Chao A, Katz LA. 2007 Diversity and geographic distribution of ciliates (Protista: Ciliophora). *Biodivers. Conserv.* **17**, 345–363. (doi:10.1007/s10531-007-9254-7)

49. Barraclough TG, Vogler AP. 2000 Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* **155**, 419–434.

50. Song WB. 1991 The infraciliature and silverline system of *Euplotes novemcarinata* Wang, 1930 (Ciliophora, Hypotrichida). *J. Ocean U. China* **21**, 76–84.

51. Valbonesi A, Ortenzi C, Luporini P. 1988 An integrated study of the species problem in the *Euplotes crassus-minuta-vannus* group. *J. Eukaryot. Microbiol.* **35**, 38–45.

52. Dini F. 1984 On the evolutionary significance of autogamy in the marine *Euplotes* (Ciliophora: Hypotrichida). *Am. Nat.* **123**, 151–162. (doi:10.1086/284194)

53. Tuffrau M. 1960 Révision du genre *Euplotes*, fondée sur la comparaison des structures superficielles. *Hydrobiologia* **15**, 1–77. (doi:10.1007/BF00048080)

54. Wilbert N, Song WB. 2008 A further study on littoral ciliates (Protozoa, Ciliophora) near King George Island, Antarctica, with description of a new genus and seven new species. *J. Nat. Hist.* **42**, 979–1012. (doi:10.1080/00222930701877540)

55. Julian Schwarz M, Zuendorf A, Stoeck T. 2007 Morphology, ultrastructure, molecular phylogeny, and autecology of *Euplotes elegans* Kahl, 1932 (Hypotrichida; Euplotidae) isolated from the anoxic Mariager Fjord, Denmark. *J. Eukaryot. Microbiol.* **54**, 125–136. (doi:10.1111/j.1550-7408.2007.00243.x)

56. Achilles-Day UE, Pröschold T, Day J. 2008 Phylogenetic position of the freshwater ciliate *Euplotes daidaleos* within the family of Euplotidae, obtained from small subunit rDNA gene sequence. *Denisa* **23**, 411–416.

57. Dai R, Xu K, He Y. 2013 Morphological, physiological, and molecular evidences suggest that *Euplotes parawoodruffi* is a junior synonym of *Euplotes woodruffi* (Ciliophora, Euplotida). *J. Eukaryot. Microbiol.* **60**, 70–78. (doi:10.1111/jeu.12007)

58. Caron DA, Worden AZ, Countway PD, Demir E, Heidelberg KB. 2008 Protists are microbes too: a perspective. *ISME J.* **3**, 4–12. (doi:10.1038/ismej.2008.101)

59. Stoeck T, Bass D, Nebel M, Christen R, Jones MD, Breiner HW, Richards TA. 2010 Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**, 21–31. (doi:10.1111/j.1365-294X.2009.04480.x)

60. Santoferrara LF, McManus GB, Alder VA. 2013 Utility of genetic markers and morphology for species discrimination within the order Tintinnida (Ciliophora, Spirotrichea). *Protist* **164**, 24–36. (doi:10.1016/j.protis.2011.12.002)

61. Stoeck T, Przybos E, Dunthorn M. 2014 The D1-D2 region of the large subunit ribosomal DNA as barcode for ciliates. *Mol. Ecol. Resour.* **4**, 458–468. (doi:10.1111/1755-0998.12195)