Published in European Journal of Protistology 58, 175-186, 2017
which should be used for any reference to this work

NAD9/NAD7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene)—A new "Holy Grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)?

Quentin Blandenier[a,b], Enrique Lara[a,*], Edward A.D. Mitchell[a,c], Daniel M.C. Alcantara[b], Ferry J. Siemensma[d], Milcho Todorov[e], Daniel J.G. Lahr[b,*]

[a]*Laboratory of Soil Biodiversity, University of Neuchâtel, Emile-Argand 11, 2000 Neuchâtel, Switzerland*
[b]*Department of Zoology, University of Sao Paulo, 05508-090 Sao Paulo, Brazil*
[c]*Botanical Garden of Neuchâtel, Chemin du Perthuis-du-Sault 58, 2000 Neuchâtel, Switzerland*
[d]*Julianaweg 10, 1241VW Kortenhoef, Netherlands*
[e]*Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, 2 Gagarin St., 1113 Sofia, Bulgaria*

Abstract

Molecular phylogeny is an indispensable tool for assessing evolutionary relationships among protists. The most commonly used marker is the small subunit ribosomal RNA gene, a conserved gene present in many copies in the nuclear genomes. However, this marker is not variable enough at a fine-level taxonomic scale, and intra-genomic polymorphism has already been reported. Finding a marker that could be useful at both deep and fine taxonomic resolution levels seemed like a utopic dream. We designed Amoebozoa-specific primers to amplify a region including partial sequences of two subunits of the mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene (NAD9/NAD7). We applied them to arcellinids belonging to distantly related genera (*Arcella*, *Difflugia*, *Netzelia* and *Hyalosphenia*) and to Arcellinid-rich environmental samples to obtain additional Amoebozoa sequences. Tree topology was congruent with previous phylogenies, all nodes being highly supported, suggesting that this marker is well-suited for deep phylogenies in Arcellinida and perhaps Amoebozoa. Furthermore, it enabled discrimination of close-related taxa. This short genetic marker (ca. 250 bp) can therefore be used at different taxonomic levels, due to a fast-varying intergenic region presenting either a small intergenic sequence or an overlap, depending on the species.

Keywords: *Arcella*; Environmental DNA survey; Intergenic region; Mitochondrion; Molecular barcoding

---

Introduction

Molecular phylogeny is the golden standard for phylogenetic reconstruction as well as species-level taxonomy. However, at both levels, currently available markers have their limits. Universal single-gene markers for deep phylogeny are not readily applicable to several eukaryotic groups such as the entire supergroup Amoebozoa (Lahr et al. 2011b; Tekle et al. 2008, 2016). When available, universal primers generally target the ribosomal small subunit (18S or SSU rDNA gene), but present biases across the large phylogenetic diversity of eukaryotes (Pawlowski et al. 2012). Likewise,

*Corresponding authors.
E-mail addresses: enrique.lara@unine.ch (E. Lara), dlahr@ib.usp.br (D.J.G. Lahr).

species-level markers such as cytochrome oxidase subunit 1 (COI or COX) are generally useful only within small phylogenetic groups and different primer sets need to be designed for each group (Kudryavtsev 2014). Finding a marker that would be useful for both deep phylogenetic and species-level studies is a goal that few taxonomists dare dream about – a true Holy Grail. Here we report the finding of a marker that comes very close to such a dream.

The eukaryotic superclade Amoebozoa comprises mainly free-living, unicellular, heterotrophic protists. Most species are characterized by metabolic cell bodies, often highly variable in their shape, as well as a unique way of moving and feeding by means of cytoplasmic streaming or pseudopodia (Grebecki 1994). One group of Amoebozoa, the Arcellinida, have a test (shell) and belong to a polyphyletic morphological group termed "testate amoebae" (Kosakyan et al. 2016). There are currently about 2,400 described Amoebozoa species, mainly catalogued using morphological approaches (Gray et al. 2004). However, environmental DNA studies have shown that this diversity is greatly underestimated (Berney et al. 2015; Geisen et al. 2014). Underestimates are due to both a lack of universality with general markers as well as low taxonomic resolution at species level. Furthermore, Amoebozoa are usually underrepresented in environmental DNA studies based on 18S rDNA, owing to long 18S rDNA sequence, frequent mismatches in primer regions and presence of introns (Berney et al. 2004; Geisen et al. 2015; Pawlowski et al. 2012). The few existing DNA barcoding studies based on COI have revealed a multitude of genotypes hidden within morpho-species, some of which are differentiable only based on ultrastructural features and are referred to as "pseudo-cryptic species" (Kosakyan et al. 2012; Nassonova et al. 2010; Singer et al. 2015).

The most commonly used eukaryotic barcoding marker is the nuclear gene coding for the 18S rDNA because it is found in the genomes of all eukaryote and possesses a mosaic of highly conserved and variable nucleotide sequences (Pawlowski et al. 2012). Yet, unlike in other groups (e.g. Foraminifera, Morard et al. 2016; Pawlowski and Holzmann 2014), the 18S rDNA gene is not sufficiently variable to resolve interspecies relationships in several taxa including Amoebozoa (Nassonova et al. 2010). Other nuclear markers such as actin have been used to a lesser extent, but this marker is of limited usefulness in the case of Arcellinida due to the presence of paralogs (Lahr et al. 2011b). There are also concerns about the possible existence of paralogs of the 18S rDNA in the nuclear genome itself (Oliverio et al. 2014). On the other hand, Amoebozoa 18S rDNA gene sequences often present high heterogeneity in rates of evolution across sites and between species, and both phenomena impair considerably the establishment of deep phylogenies (Tekle et al. 2008). This problem has also been reported in arcellinids, where fast-evolving taxa can be difficult to place in trees and distort the relationships among internal branches (Kudryavtsev et al. 2009).

Mitochondrial genes are a promising alternative as they normally do not duplicate and evolve generally faster than nuclear genomes (Brown et al. 1979; Fučíková and Lahr 2016; Stoneking 2013); on the other hand, genes coding for some conserved protein genes may be easy to align and conserved enough to infer ancient phylogenetic relationships (Fučíková and Lahr 2016). The commonly used mitochondrial barcoding marker COI, which is now established as the canonical barcoding marker for animals (Hebert et al. 2003) has been used also for various Amoebozoa groups (Kudryavtsev 2014; Nassonova et al. 2010) and in particular for arcellinid testate amoebae (Kosakyan et al. 2012, 2013, 2016; Singer et al. 2015). Although this gene proved to be an excellent marker for fine-level taxonomy, it is too variable to assess ancient dichotomies.

Nicotinamide adenine dinucleotide dehydrogenases genes (NADH) have been hypothesized to be the most conserved group of Amoebozoa mitochondrial genes (Heidel and Glöckner 2008). Therefore, we hypothesized that potential barcode markers could be found in these genes. To date, ten complete amoebozoan species mitochondrial genomes from seven different genera are available in GenBank under "Amoebozoa complete mitochondrion" search (Table 1). These species belong to distantly-related organisms (Lahr et al. 2011a; Nikolaev et al. 2005; Smirnov et al. 2005). As mitochondrial genomes of Amoebozoa show high levels of synteny (Fučíková and Lahr 2016), we designed a set of primers to amplify a presumably highly conserved region in the NADH gene complex, i.e., the end of NAD9 and the beginning of NAD7, in order to obtain a marker that would be conserved enough in evolution to resolve deep phylogenetic relationships but which could be used to discriminate closely related species (due to the redundancy of the genetic code and the putative variability of the intergenic region). We tested amplification of this region with a selection of arcellinid species covering the main phylogenetic lineages, in order to evaluate the efficiency of the proposed marker. In addition, we applied these primers to environmental DNA samples to test whether this marker could be used in environmental barcoding studies.

Material and Methods

Testate amoeba collection and DNA extraction

Testate amoebae were obtained from various environments (i.e., mosses, ponds, bog lakes and sediments) and different geographical origins (Table 2). Both *Arcella* lineages were collected from a pond in Sao Paulo (BR), identified in Utermöhl's plankton chambers (Hydro-Bios, Kiel, Germany) under inverted microscope (Olympus IX81 mounted with Plan Apo objectives (Mitutoyo, Japan) with 40× magnification), picked up with a stirred Pasteur pipette, cleaned several times with sterilized pond water and cultured in 25 ml flasks containing 1 ml of cereal grass media (Lahr and Katz 2009).

Table 1. List of the ten complete mitochondrial genomes.

Species		NCBI reference sequence	References
Acanthamoeba	*castellanii*	NC_001637	Burger et al. 1995
	polyphaga	KP054475	Karlyshev et al. unpublished
Balamuthia	*mandrillaris*	NC_027736	Greninger et al. 2015
Dictyostelium	*citrinum*	DQ336395	Heidel and Glöckner 2008
	discoideum	NC_000895	Ogawa et al. 2000
	fasciculatum	EU275727	Heidel and Glöckner 2008
Phalansterium	*sp.*	KC121006	Pombert et al. 2013
Physarum	*polycephalum*	NC_002508	Takano et al. 2001
Polysphondylium	*pallidum*	NC_006862	Burger et al unpublished
Vermamoeba	*vermiformis*	NC_013986	Bullerwell et al. 2010

Table 2. List of sequenced species, sampling localization and sequences details.

Species	Environments	Localisations and countries	Co-ordinates	Number of cells used per extraction	Sequences length (bp)	AT%	Genbank number
Arcella intermedia	pond	Sao Paulo (BR)	23°57′ S, 46°73′ W	∼40k	253 (+ 5*)	80.7	KY499885
Arcella intermedia laevis	pond	Sao Paulo (BR)	23°57′ S, 46°73′ W	∼40k	253 (+ 5*)	79.5	KY499884
Difflugia nodosa	Bog lake	Ljulin (BG)	42°38′ N, 23°09′ E	5	246 (+1*)	71.3	KY499878
Difflugia nodosa	pond	Groeneveld (NL)	52°21′ N, 5°25’ E	21	245 (+1*)	70.7	KY499879
Difflugia pyriformis	pond	Groeneveld (NL)	52°21′ N, 5°25’ E	28	267 (+1*)	70.9	KY499876
Difflugia pyriformis	ditch	Laegieskamp (NL)	52°28′ N, 5°14’ E	20	267 (+1*)	70.5	KY499877
Env_L_1	*Sphagnum* sp	Le Cachot (CH)	47°00′ N, 6°39′ E	-	251	77.7	KY499887
Env_L_2	*Sphagnum* sp	Le Cachot (CH)	47°00′ N, 6°39′ E	-	260	70.8	KY499880
Env_B_1	*Sphagnum* sp	Le Cachot (CH)	47°00′ N, 6°39′ E	-	320 (+1*)	72.8	KY499888
Hyalosphenia papilio	*Sphagnum* sp	Le Cachot (CH)	47°00′ N, 6°39′ E	3	261 (+1*)	77.9	KY499886
Netzelia oviformis	pond	Neuchâtel (CH)	47°00′ N, 6°95′ E	2	256	77.0	KY499881
Netzelia tuberculata	lake	le Loclat (CH)	47°02′ N, 6°99′ E	4	253 (+1*)	79.8	KY499882
Netzelia wailesi	sediments	Sima-Moor (AT)	47°58′ N, 12°62′ E	3	262 (+1*)	78.6	KY499883

*Number of RNA editing sites.

For the other species, two to 28 cells of similar morphology and originated from the same sample were picked up individually, as explained above, and placed together into Eppendorf tubes containing 100 µl of guanidine buffer (Chomczynski and Sacchi 1987). *Arcella* cultures were spun in microcentrifuge tubes for 10 min at 5000 × *g* and 800 µl of guanidine buffer was added to the resulting pellet. DNA extraction was performed in these tubes as described in Gomaa et al. (2012). DNA was finally re-suspended in sterile demineralized water. Environmental DNA was extracted from *Sphagnum* mosses collected in the peatland Le Cachot in the Swiss Jura Mountains (Table 2). Env_L_1 and Env_L_2 are two environmental clones belonging to a lawn and Env_B_1 to a peatland border as described in Singer et al. (2016). Testate amoeba cells were extracted from 20 g of fresh *Sphagnum* and concentrated by sieving at 150 µm. The resulting fraction was used for global DNA extraction with a MoBio Power Soil® DNA Isolation kit according to the manufacturer's instruction (Kosakyan et al. 2015).

Cells used for DNA extraction were documented using alternatively a Leitz Orthoplan and an Olympus IX81 microscope. Shells were prepared and documented for scanning electron microscopy (SEM) as described previously in Gomaa et al. (2012). The following reference papers were used for species identification: Chardez 1967; Deflandre 1928; Ogden 1979, 1980. Test size, shape and composition, as well as nucleus shape (especially in genus *Difflugia*) were used as criteria for species identification and discrimination.

Primer design, DNA extraction, PCR amplification and sequencing

We designed primers comparing the entire mitochondrial genomes from the following amoebozoans: *Acanthamoeba castellanii* (NC_001637), *Dictyostelium citrinum* (NC_007787), *Dictyostelium discoideum* (NC_000895), *Polyspondylium pallidum* (NC_006862) and *Vermamoeba vermiformis* (NC_013986) (Table 1). *Phalansterium* (KC121006) and *Physarum polycephalum* (NC_002508) mitochondrial genomes were partly unannotated and therefore could not be used for primer design. *Acanthamoeba polyphaga* (KP054475) and *Balamuthia*

mandrillaris (NC_027736) genomes sequences arise after the design of the primer. All the NADH dehydrogenases sequences were obtained by a single polymerase chain reaction (PCR) with two newly design primers NAD9_386F (5′-TGGTTAGAACGAGAAGTTTGGGATATGT-3′) and NAD7_67R (5′-GTGCGCAGCAGGRTGTTGWGGWCC-3′) (list of all primers designed in Supplementary material 1). The DNA was amplified in a total volume of 17 μl with an amplification profile consisting of a 3 min initial denaturation at 94 °C followed by 35 or 40 cycles program of 30 s at 94 °C, 30 s at 61 °C, 1 min at 72 °C and a final elongation of 10 min at 72 °C. Except for *Netzelia tuberculata*, all positive products were cloned using a PCR2.1 Topo TA cloning vector and transformed into *E. coli* TOP10’ One Shots cells (Invitrogen kit) according to the manufacturer’s instructions. *N. tuberculata* was sequencing without any cloning step. The protocol used for amplifying environmental DNA was identical to the one used for isolated cells and *Arcella* cultures. Sequencing of five to ten clones for each amplicon (34 for each environmental clones) was performed with single direction by an initial amplification using a BigDye197 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems), further analysed with an ABI-3130XL DNA sequencer (Applied Biosystems). Sequences were deposited in GenBank with the following accession number: KY499876–KY499888.

Alignment and phylogenetic analysis

Obtained sequences (environmental, cultures and isolates-derived) were aligned against all Amoebozoan sequences found in GenBank that included the fragment of interest, plus the oomycete *Phytophthora ramorum* NC_009384 as outgroup. The sequences were aligned manually using BioEdit software (Hall 1999). We built a Maximum Likelihood phylogenetic tree with the software MEGA6 (Tamura et al. 2013). T92 + G was selected as the best model using the Akaike criteria as implemented in the MEGA package. Robustness of the nodes was calculated by performing 500 bootstraps replicates.

Results

Structure of the sequences

We used the primers combination which worked the best (i.e., NAD9_386F–NAD7_67R) among all possible (Supplementary material 1) on the collected species (Fig. 1). We obtained 13 sequences from at least ten different arcellinid testate amoeba taxa (Table 2) from four distantly related families. Four sequences belonged to genus *Difflugia* (two isolates of *D. nodosa* and two of *D. pyriformis*), two from clonal cultures of genus *Arcella* (*A. intermedia* and *A. intermedia laevis*), three to genus *Netzelia* (*N. oviformis*, *N tuberculata* and *N. wailesi*), one to *Hyalosphenia papilio* and three environmental clones sequence type (Env_L_1, Env_L_2 and Env_B_1) from *Sphagnum* samples.

The sequences ranged from 245 to 320 bp in length. NAD9 formed the main part of the sequences with a length between 210 bp (both *Arcella* lineages) and 226 bp (both *D. nodosa*) whereas NAD7 was much shorter, between 42 bp (both *Difflugia* and *Netzelia* species) and 61 bp (*Hyalosphenia papilio*). The AT ratio was high, ranging from 70.5% to 80.7% (Table 2) which is similar to *Acanthamoeba castellanii* (70.6%), *Dictyostelium citrinum* (73%) and *Polysphondylium pallidum* (75.8%). However, the AT:GC ratio on the priming region was not too unbalanced; furthermore, the high melting temperature of the primers allowed high annealing temperatures in the PCR cycling program, which reduced self-annealing problems.

The relative configurations of the NAD9 and NAD7 genes fell into three different categories (Fig. 2). In the first case (both *D. pyriformis* isolates, *N. oviformis*, *N. wailesi* and Env_B_1) there was an intergenic region with specific sequences. This nucleotide indel is very short (i.e., 2–6 bp) and is made solely of A and T except for the Env_B_1 with a particularly long intergenic region (i.e., 46 bp) composed of all four nucleotides. In the second case (*N. tuberculata* and Env_L_2) the two genes are perfectly contiguous without any indels or overlaps. In the last case, the two genes overlapped, either only by 1 bp for *Arcella* (the **A** nucleotide at the end of TA**A** corresponded to the beginning of **A**TG), 7 bp for Env_L_1 as well as *Hyalosphenia papilio*, or 16 bp for both *D. nodosa*. All NAD9 sequences possessed a TGA codon (canonically a stop codon, Table 3). This codon was located in nucleotide position 195 for all species, except for the sequence of Env_L_2 where it was present in position 204 and for both *Arcella* and *Netzelia* which had two TGA codons in position 195 and 204. The insertion of single nucleotides in the sequences, corresponding to several species or groups of species, suggests frequent mitochondrial RNA editing, as demonstrated in other Amoebozoa (Traphagen et al. 2010).

Phylogenetic relationships among species

The phylogenetic tree (Fig. 3) showed that all our sequences branched together, forming a robust arcellinid clade with 85% bootstrap support (B). Another robust clade (94% B) included all pyriform *Difflugia*. Likewise, both *Arcella* branch together robustly as well as both *Netzelia* (Table 5), all five sequences forming another robust clade (88% B). *Hyalosphenia papilio* and Env_L_1 were very similar (differing by 14 bp), and branch together along with Env_B_1 which is more basal (Fig. 3). Finally, Env_L_2 position is weakly supported (<50% B) and is sometimes basal to the whole arcellinid tree (data not show).

The sequences we obtained using the NADH dehydrogenases primers allowed us to discriminate accurately all species, even in the case of closely-related species. The two

Table 3. Amino acid sequence of NAD9 revealing the TGA codon (*) and the RNA editing sites (bold letter). The Ω character represent the TAA stop codon.

Nucleotide position	10 20 30 40 50 60 70
Difflugia pyriformis Groeneveld (NL)	GLKFI LHLDL RRTLTDYGFK GHPLCKDYPL CGYVELRYDD I HHLI CS**A**PV EVS QGFRCFK FSNP* KTL CI F S **Ω**
Difflugia pyriformis Laegieskamp (NL)	GLKFI LHLDL RRI LTDYGFK GHPLCKDYPL CGYI ELRYDD I HHLI CL**A**PV EVS QGFRCFK FGNP* KTL CI F S **Ω**
Difflugia nodosa Ljulin (BG)	GLKFI LHTDL RRI LTDYGFK GFPLRKEFPL S**G**YVELHYDD I YQSI SVAPL EVS QGFRNHK YKNP* RNL CI MNKI**Ω**
Difflugia nodosa Groeneveld (NL)	GLNFI LHTDL RRI LTDYGFK GHPLRKESPL NGYVELRYDD I YQSI NVKPI EVS QVFRNY**K** YVNP* KNS CT T S KI**Ω**
Netzeliaoviformis Neuchâtel (CH)	CI **K**FI LHNDL RRI LTDYGFH GHPLRKDYPL I GYVELYYND LT QSI VFTSV EQMQGFRFFK FDNP* FN* K**Ω**
Netzelia tuberculataLe Loclat (CH)	CLKFI LHNDL RRI LTDYGFH GHPLRKDYPL I GYVELFYND LI QSI VFTNV EQMQGFRFFK FDNP* FN* K**Ω**
Netzelia wailesi Sima-Moor (AT)	CI KFLLNNDL RRI LTDYGFH GYPLRKDYPL I GYVELHYND LMQS VI F**Y**CV EQMQGFRFFK FENP* FN* KL **Ω**
Arcella intermedia laevis Sao Paulo (BR)	GLKF**Y**LHSDL RRI LNDY**G**VK GFPLRKDFP**L** I GFI ELFYDD VNQSI I I **E**CV EQT QSFRFFK FENP* **Q**N* KV **Ω**
Arcella intermedia Sao Paulo (BR)	GLKF**Y**LHNDL RRI LNDY**G**FK GFPLRKDFP**L** I GFI ELFYDD VNQSI I I **E**CV EQMQS YRFFK FENP* **Q**N* KV **Ω**
Hyalosphenia papilio Le Cachot (CH)	GVKFLLHLDL RRLLTDYGFK GHPLRKDFPL I GYLEI RYDD I L KI I I SE**S**V ETS QSYRVYR FI NP* FKWSF **Ω**
Env_L_1 Le Cachot (CH)	GVKFLLHSDL RRLLTDYGFK GHPLRKDFPL I GYLEI RYDD I L KI I I SESV ETS QSYRVYR FI NP* FKWSF **Ω**
Env_L_2 Le Cachot (CH)	GCRFLLHEDL RRI LTDYGFE GFPLRKDFPL VGYFEI RYEN TT SFI SRDPL EAS QALRI FF FDNPWLL* KY **Ω**
Env_B_1 Le Cachot (CH)	GI KFLLHGDL RRLLTDYGFR GHPLRKDFPL I GYFEARYDD I L KGI FTE**S**V ETAQVYRRYK FI NP* AKL NL F **Ω**

Table 4. List of the RNA editing sites.

Species	Nucleotide editing	Positions	Resulting codon
Difflugia pyriformis Groeneveld (NL)	G	144	A (**G**CT)
Difflugia pyriformis Laegieskamp (NL)	G	144	A (**G**CT)
Difflugia nodosa Ljulin (BG)	A	182	K (AAA)
Difflugia nodosa Groeneveld (NL)	G	97	G (GGT)
Netzelia tuberculata Le Loclat (CH)	A	11	A (AAA)
Netzelia wailesi Sima-Moor (AT)	C	146	Y (TA**C**)
Arcella intermedia laevis Sao Paulo (BR)	T;T;A;A;A;T	15;55;89;143;196;246	Y (**T**AT); G (GG**T**); L (TT**A**); E (GAA); Q (CAA); L (**T**AA)
Arcella intermedia Sao Paulo (BR)	T;T;A;A;A;T	15;55;89;143;196;246	Y (**T**AT); G (GG**T**); L (TT**A**); E (GAA); Q (CAA); L (**T**AA)
Hyalosphenia papilio Le Cachot (CH)	T	147	S (**T**CA)
Env_B_1 Le Cachot (CH)	T	147	S (**T**CA)

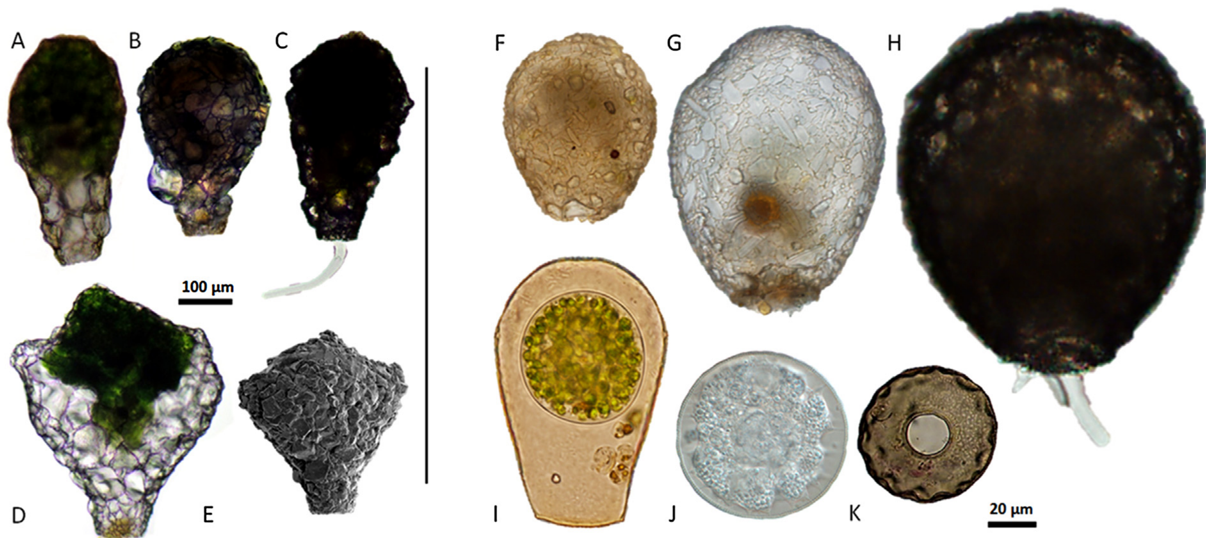

Fig. 1. Light (**AK**) and scanning electron (**E**) micrographs of the ten analysed species: **A.** *Difflugia pyriformis* from Groeneveld (NL); **B–C.** *D. pyriformis* from Laegieskamp (NL), in **C** an active cell showing a pseudopod; **D.** *D. nodosa* from Groeneveld (NL); **E.** *D. nodosa* from Ljulin (BG); **F.** *Netzelia wailesi* from Sima-Moor (AT); **G.** *N. oviformis* from Neuchâtel (CH); **H.** *N. tuberculata* from le Loclat (CH) showing pseudopods; **I.** *Hyalosphenia papilio* from Le Cachot (CH); **J.** *Arcella intermedia laevis* from Sao Paulo (BR); **K.** *A. intermedia* from Sao Paulo (BR) with characteristic depressions on the test. **J** and **K** *Arcella* pictures were taken by A. Porfirio at the University of Sao Paulo. Scale bar: **A–E** = 100 µm, **F–K** = 20 µm.

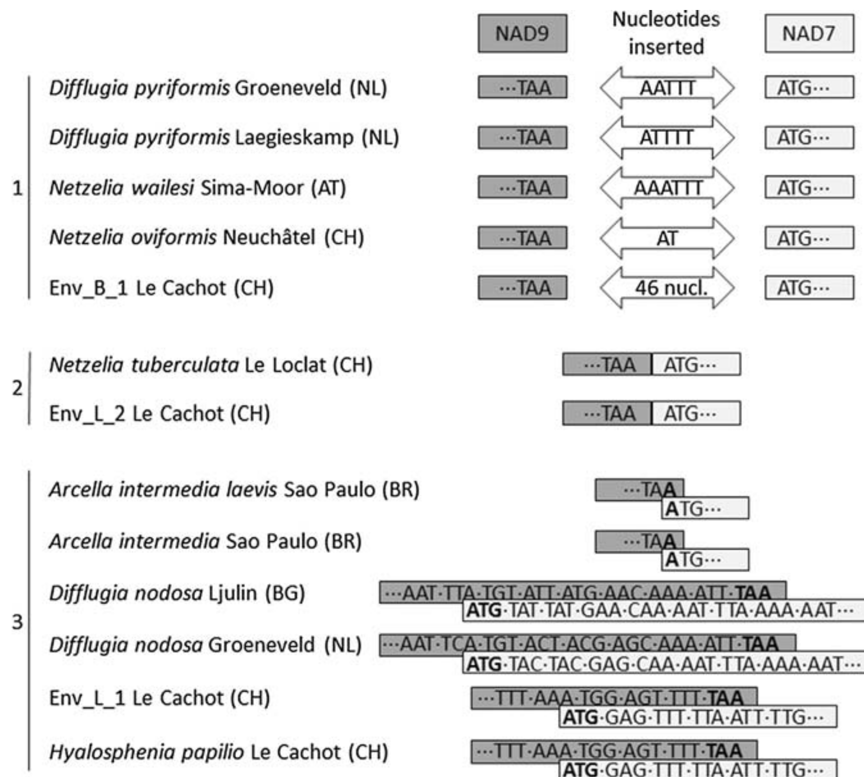

Fig. 2. Configuration of the three types of intergenic region. The dark grey box represents the NAD9 gene and the light grey one the NAD7. The three dots (···) express a continuation of the gene, whereas the single dot (·) expresses a fictive space between two codons. The first case represents the nucleotides insertion model, the second the contiguously genes alignment and finally the overlapping one.

isolates identified as *D. nodosa* differed by 14% in their NAD9/NAD7 sequences and the two *D. pyriformis* isolates by 6%. Similarly, the three *Netzelia* species diverged by, respectively, 9, 16 and 17% (Table 5).

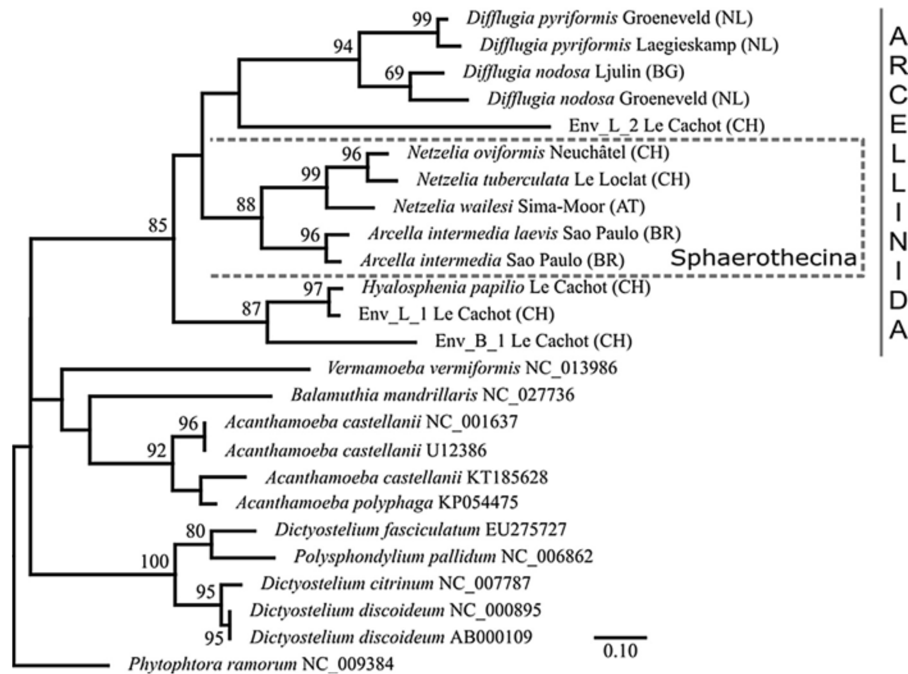

Fig. 3. Phylogenetic tree of Arcellinida based on NAD9-NAD7 genes sequences. The tree was rooted with the Oomycete *Phytophthora ramorum* and based on a maximum likelihood analysis. Numbers at the nodes indicate the resulting bootstraps (only values above 60% were kept). The scale bar indicates 0.1 sequence divergence.

Discussion

NAD9/NAD7 is an accurate fragment for species identification and phylogenetic reconstructions

We describe here a pair of primers that work reliably across all Arcellinida. The designed primers are, at first sight, incongruent with "canonical" practices, particularly GC content and length (Dieffenbach et al. 1993). This is because the amoebozoan mitochondrial genome is GC-poor (see Table 2), stable in terms of coding capacity and gene order, but present highly divergent nucleotide sequences, impossible to align outside coding regions (Glöckner 2013), resulting in lesser potential primer hybridization regions. Our two primers NAD9_386F and NAD7_67R successfully amplified the target genes for different taxa spread across the arcellinid tree.

The resulting fragment presents the advantage of being small (i.e., less than 320 bp) and is therefore easy to amplify. This short size may likewise allow their retrieval using high throughput sequencing (HTS) approaches, opening the way for the construction of a large Arcellinida sequence dataset, which has been a challenge so far owing to the difficulty to amplify the commonly used markers such as 18S rDNA. Indeed, environmental sequencing is very useful to strengthen the topology of phylogenetic trees and to discover unknown clades or new taxa within known clades. Here, we obtained an environmental sequence (Env_L_1) that was closely related to *Hyalosphenia papilio*. The *Sphagnum* sample from which it derived contained many cells of *Nebela guttata*, also belonging to Hyalospheniidae (Kosakyan et al. 2013). It is therefore most likely that the original organism from which the sequence derived belonged to this family. The other environmental clones encountered could not be related to any known sequence. Env_B_1, a clone from a peatland border, robustly branched with Hyalospheniidae. It belonged, likely, to the sister group of the clade; a good candidate could be the group of *Bullinularia*, which has been previously shown to robustly branch with that family (Gomaa et al. 2012). The third environmental clone Env_L_2 belongs also to Arcellinida but represents most likely a group unrelated to those represented in our study. After improving the method and develop a confident database, this marker might be used in environmental diversity and ecological studies in which Amoebozoa are typically poorly covered (Geisen et al. 2015; Pawlowski et al. 2012).

On the other hand, NAD9/NAD7 may be used also for discriminating closely-related species at fine taxonomic levels. Notably, the two *Difflugia nodosa* isolates could be separated on the basis of their sequences (Table 5). These organisms were morphologically similar, with a large and robust pyriform test compressed in cross-section and vacuoles that were filled with numerous zoochlorellae in agreement with the description of Penard (1890). According to Penard (1890), the size ranges from 300 to 380 μm, yet the specimens from Groeneveld were considerably larger (483 μm in average), while the ones from Ljulin fitted the original description (353 μm). This taxon represents a case where molecular methods reveal hitherto unknown diversity. Similar instances are frequent in Arcellinida, as within the *Nebela collaris* complex (Kosakyan et al. 2012, 2013) or *Difflugia tuber-*

spinifera with or without spines (Gomaa et al. 2015). Thus, our study brings further support to the idea that the diversity of arcellinids has been underestimated by morphology-based taxonomy (Kosakyan et al. 2016; Porfirio et al. 2016, this issue).

Detailed study will most likely show that *D. nodosa* represents a complex of different molecular (and possibly morphological) species. This may be also valid for *D. pyriformis*, which belongs to a controversial group, the *Difflugia pyriformis* complex (Perty 1849). Indeed, the morphological variation among this taxon is high and the descriptions are unclear (Mazei and Warren 2014). A more detailed morphological and molecular study of this species complex would be needed to proceed to any taxonomic action. Our identifications of the analysed material were tentative as neither of our isolates corresponded in their dimensions to the original description of *D. pyriformis*. Indeed, cells from the Groeneveld isolate measured 302–410 μm and those from Laegieskamp 423 μm as compared to 140–200 μm for the original description (Perty 1849).

The discrimination power of NAD9/NAD7 is due partly to the fast evolution of the third position of codons, but also to the varying size and sequences in the intergenic space. In summary, our new genetic marker contains both slow- and fast-evolving regions that are useful for phylogeny.

Three possibilities exist in the genetic frontier: insertion of an intergenic region, contiguously or overlap

All mitochondrial genes of Amoebozoa are closely packed and the overlap we observed is widespread in Amoebozoa (Burger et al. 1995; Ogawa et al. 2000). The intergenic space region can also possess a short sequence inserted (Fig. 2) as shown in the two isolates of *Difflugia pyriformis*, *Netzelia wailesi*, *N. oviformis* and the Env_B_1 sequence as well as for *Dictyostelium discoideum*, *Polysphondylium pallidum* and *Vermamoeba* vermiformis. It is mostly solely composed of A and T repeats reaching up to six bp in *Netzelia wailesi*. Env_B_1 is the only current exception with a large nucleotides insertion (i.e., 46 bp) including two C and seven G base pairs. The region varies systematically even within one genus and each of the species surveyed in this study presented a different sequence. The case of *Netzelia tuberculata* shows a different configuration than the two others *Netzelia* as the two genes are perfectly contiguous. One could notice that the number of nucleotides inserted among this genus is diminishing from six in the case of *N. wailesi*, to two for *N. oviformis* and nothing for *N. tuberculata* which is more closely related to the previous species (Fig. 3). However, in cases with overlap, the length of the "common" region is stable within genus; this is observed in both *D. nodosa* isolates, both *Arcella*, as well as Env_L_1 and *Hyalosphenia papilio*. It is noteworthy to mention that *Acanthamoeba castellanii* (NC_001637) and *A. polyphaga* (KP054475) possess also together an identical overlapping of 13 nucleotides (data not shown). This length can be proposed, therefore, as a deeper phylogenetic marker to be used for differentiating isolates at the genus level. This proximity suggests that these two genes need to be nearby to be efficient; this could be due to the existence of a single promoter for the regulation of the mitochondrial DNA genes transcription (Ogawa et al. 2000).

Arcellinids are using the same genetic code as *Acanthamoeba castellanii* and present RNA editing evidences

Both universal and alternative genetic codes have been identified in mitochondria (Heidel and Glöckner 2008). In Amoebozoa mitochondrial genomes, examples of both genetic codes have also been mentioned. Indeed, the social amoeba *Dictyostelium* (Cole and Williams 1994), *Polysphondylium* and *Vermamoeba* use the universal code (i.e., TGA as a stop codon) whereas, *Acanthamoeba* uses the alternative option (i.e., TGA coding for tryptophan). All our Arcellinida mitochondrial sequences seem to employ an alternative genetic code. Indeed, the NAD9 genes are not supposed to end with a TGA codon (at position 195) given the presence of a tryptophan (W) located in the same position for Env_L_2 (Table 3), *Dictyostelium*, *Polysphondylium* and *Balamuthia*. The same case happens in position 204 for Env_L_2, both *Arcella* and *Netzelia* which possess also a TGA codon whereas a tryptophan is present at the homologous position in *Hyalosphenia* and Env_L_1, suggesting the use of a non-universal code in their mitochondria. As *Vermamoeba* and Arcellinids are tubulinids and use different genetic codes, it seems that a transition took place within this clade. It is not yet clear where it occurred in the phylogeny of the group, as we do not yet have enough data to place this event; further data from more taxa will be needed to show if this can be considered as characteristic of Arcellinida. The same codon usage can be found in *Acanthamoeba,* likely resulting from convergence. As we have never found a TAG codon and all TGAs are coding for tryptophan, it is possible that arcellinids have only maintained the TAA stop codon in the mitochondria. This is not true for the nucleus, demonstrated by other arcellinid genes that apparently use TGA as stop (Lahr et al. 2011a,b, 2013). However, it is still necessary to verify if RNA editing is not being used in these non-canonical TGA positions.

RNA editing seems to occur in arcellinids as well as for other Amoebozoa (Table 4), as already inferred for COI (Oliverio et al. 2015). NAD RNA editing is thus consistent with previous knowledge. We propose a nucleotide addition hypothesis (Table 4) based on sequence similarities. However, the exact location of editing sites remains difficult to infer in Env_L_1 and *Hyalosphenia papilio*. We observe also a narrow phylogenetic distribution of this process, especially for the *Arcella* lineages which shared the six same nucleotides added to the same loci, but also for both *Difflugia pyriformis* isolates which possessed in common a G addition in position 145 (Table 4). We did not observe nucleotide indels or substitutions (Gray 2012).

NAD9/NAD7 is a reliable phylogenetic marker for the Arcellinida

Our phylogenetic analysis with marker NAD9/NAD7 supports the monophyly of Arcellinida robustly (85% B) and separates genera congruently to the current literature (Gomaa et al. 2015; Lara et al. 2008). Such a result suggests that the

Table 5. Nucleotide similarity for NAD9-NAD7 genes [%].

	D. pyriformis G.	*D. pyriformis L.*	*D. nodosa* (BG)	*D. nodosa* (NL)	*N. oviformis*	*N. tuberculata*	*N. wailesi*	*A. intermedia l.*	*A. intermedia*	*H. papilio*	Env_L_1	Env_L_2	Env_B_1
D. pyriformis Groeneveld (NL)	–												
D. pyriformis Laegieskamp (NL)	**94**	–											
D. nodosa Ljulin (BG)	78	**77**	–										
D. nodosa Groeneveld (NL)	75	75	**86**	–									
N. oviformis Neuchâtel (CH)	67	65	64	64	–								
N. tuberculata Le Loclat (CH)	70	67	67	67	**91**	–							
N. wailesi Sima-Moor (AT)	66	64	66	67	**84**	**83**	–						
*A. intermedia laevis*Sao Paulo (BR)	68	68	70	68	76	75	75	–					
A. intermedia Sao Paulo (BR)	67	67	70	67	76	77	76	**94**	–				
H. papilio Le Cachot (CH)	67	67	67	67	69	69	67	70	72	–			
Env_L_1 Le Cachot (CH)	67	67	66	67	67	69	67	70	73	**96**	–		
Env_L_2 Le Cachot (CH)	64	64	65	61	60	61	59	66	66	64	64	–	
Env_B_1 Le Cachot (CH)	64	63	64	66	67	67	67	69	69	76	76	61	–

new marker may be useful also for deep phylogeny. Our survey on different genera is encouraging and indicates that the addition of further sequences from an even broader phylogenetic sampling (e.g., *Nebela*, *Centropyxis*, *Heleopera,* etc.) will be useful to reconstruct the Arcellinida tree.

Our new primers were efficient to distinguish closely-related species as expected with two well-conserved mitochondrial genes. It is generally believed that mitochondrial genomes evolve faster than nuclear genomes (Heidel and Glöckner 2008). COI has indeed proven useful to discriminate closely-related species in Hyalospheniids, but failed to resolve more basal nodes (Kosakyan et al. 2013). The nuclear ribosomal internal transcribed spacer (ITS1/ITS2), possibly suffers from the same disadvantages, and may be present in multiple copies in the genome as suspected for other amoebozoans (Nassonova et al. 2010). Actin, a commonly used protein-coding nuclear marker (e.g., Kudryavtsev et al. 2011; Kudryavtsev and Pawlowski 2015; Nikolaev et al. 2005) has been shown to present high paralogy in Arcellinida genomes (Lahr et al. 2011b). Finally, Arcellinids possess highly variable 18S rDNA genes generating often long branches in the trees (Kudryavtsev et al. 2009, 2011) and frequent mutations in the priming sites, precluding the amplification of full length sequences. Furthermore, the frequent presence of introns and insertions in many taxa (Gomaa et al. 2012, 2015; Lara et al. 2008) renders alignments even more difficult. For instance, the information provided for a comprehensive phylogeny of the group including sequences of more than 2 kb in length had to be reduced to 700 bp alignable molecular characters. In comparison, NAD9/NAD7, despite its short length, contains enough phylogenetic information to retrieve the relationships between groups that can be observed with 18S rDNA. Indeed, results recently obtained by other means such as the monophyly of the Sphaerothecina (Kosakyan et al. 2016), as well as monophyly of the vase-shaped members in genus *Difflugia* are robustly supported. The node supporting the monophyly of Arcellinida is also unambiguous. In addition, branch lengths in the NAD9/NAD7 tree seem more even than in 18S rDNA phylogeny (compare to Gomaa et al. 2012; Kudryavtsev et al. 2011), suggesting that reconstructions inferred from this fragment should be less prone to long-branch attraction. From this viewpoint, NAD9/NAD7 appears as an attractive example of a gene that can be used both for barcoding and relatively deep phylogeny. However, as for any mitochondrial marker, the use of NAD9/NAD7 is limited to species living under aerobic conditions; parasitic taxa such as the amitochondrial amoebozoan *Entamoeba histolytica* lack a mitochondrial genome (Loftus et al. 2005) and will therefore not be covered by our proposed mitochondrial marker, and this will be a limitation for environmental DNA surveys.

Finding a marker that is useful for both deep phylogenetic and species-level studies has long been all but an unrealistic dream – indeed a potential barcoding “Holy Grail”. NAD9/NAD7 is as close to such a dream as can be!

Acknowledgements

This study was supported by Funds Wüthrich et Mathey-Dupraz; EL was supported by a grant from the Swiss SNF (310003A_143960), and DJGL by a FAPESP Young Investigator Award (#2013/04585-3). We thank A. Pillonel for technical help, D. Singer for support and fruitful discussions, and Angie Opitz for the samples from Sima-Moor which contained *Netzelia wailesi*. We are finally grateful to A. Porfirio and C. Duckert for their help concerning cultures, isolation, identification and micrographs.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejop.2016.12.002.

References

Berney, C., Fahrni, J., Pawlowski, J., 2004. How many novel eukaryotic ‘kingdoms’? Pitfalls and limitations of environmental DNA surveys. BMC Biol. 2, 1.

Berney, C., Geisen, S., Van Wichelen, J., Nitsche, F., Vanormelingen, P., Bonkowski, M., Bass, D., 2015. Expansion of the ‘reticulosphere’: diversity of novel branching and network-forming amoebae helps to define Variosea (Amoebozoa). Protist 166, 271–295.

Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial-DNA. Proc. Natl. Acad. Sci. U. S. A. 76, 1967–1971.

Bullerwell, C.E., Burger, G., Gott, J.M., Kourennaia, O., Schnare, M.N., Gray, M.W., 2010. Abundant 5S rRNA-like transcripts encoded by the mitochondrial genome in amoebozoa. Eukaryotic Cell 9, 762–773.

Burger, G., Plante, I., Lonergan, K.M., Gray, M.W., 1995. The mitochondrial-DNA of the ameboid protozoan, *Acanthamoeba-castellanii* – complete sequence, gene content and genome organization. J. Mol. Biol. 245, 522–537.

Chardez, D., 1967. *Difflugia oblonga* Ehrenberg et ses variétés. Bull. Inst. Agron. Stat. Rech. Gembloux, new series 2, 589–595.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate phenol chloroform extraction. Anal. Biochem. 162, 156–159.

Cole, R.A., Williams, K.L., 1994. The dictyostelium-discoideum mitochondrial genome—a primordial system using the universal code and encoding hydrophilic proteins atypical of metazoan mitochondrial-DNA. J. Mol. Evol. 39, 579–588.

Deflandre, G., 1928. Le genre Arcella Ehrenberg: Morphologie-Biologie. Essai phylogénétique et systématiqe. Arch. Protistenkd. 64, 152–287.

Dieffenbach, C.W., Lowe, T.M.J., Dveksler, G.S., 1993. General concepts for PCR primer design. PCR Methods Appl. 3, S30–S37.

Fučíková, K., Lahr, D.J.G., 2016. Uncovering cryptic diversity in two amoebozoan species using complete mitochondrial genome sequences. J. Eukaryotic Microbiol. 63, 112–122.

Geisen, S., Kudryavtsev, A., Bonkowski, M., Smirnov, A., 2014. Discrepancy between species borders at morphological and molecular levels in the genus *Cochliopodium* (Amoebozoa, Himatismenida), with the description of *Cochliopodium plurinucleolum* n. sp. Protist 165, 364–383.

Geisen, S., Tveit, A.T., Clark, I.M., Richter, A., Svenning, M.M., Bonkowski, M., Urich, T., 2015. Metatranscriptomic census of active protists in soils. ISME J. 9, 2178–2190.

Glöckner, G., 2013. Genome analysis of social amoebae. In: Romeralo, M., Baldauf, S., Escalante, R. (Eds.), Dictyostelids: evolution, genomics and cell biology. Springer, Berlin, Heidelberg, pp. 35–47.

Gomaa, F., Todorov, M., Heger, T.J., Mitchell, E.A.D., Lara, E., 2012. SSU rRNA phylogeny of Arcellinida (Amoebozoa) reveals that the largest Arcellinid genus, *Difflugia* leclerc 1815, is not monophyletic. Protist 163, 389–399.

Gomaa, F., Yang, J., Mitchell, E.A.D., Zhang, W.J., Yu, Z., Todorov, M., Lara, E., 2015. Morphological and molecular diversification of asian endemic *Difflugia tuberspinifera* (Amoebozoa, Arcellinida): a case of fast morphological evolution in protists? Protist 166, 122–130.

Gray, M.W., 2012. Evolutionary origin of RNA editing. Biochemistry 51, 5235–5242.

Gray, M.W., Lang, B.F., Burger, G., 2004. Mitochondria of protists. Annu. Rev. Genet. 38, 477–524.

Grebecki, A., 1994. Membrane and cytoskeleton flow in motile cells with emphasis on the contribution of free-living amoebae. Int. Rev. Cytol. 148, 37–80.

Greninger, A.L., Messacar, K., Dunnebacke, T., Naccache, S.N., Federman, S., Bouquet, J., Mirsky, D., Nomura, Y., Yagi, S., Glaser, C., 2015. Clinical metagenomic identification of *Balamuthia mandrillaris* encephalitis and assembly of the draft genome: the continuing case for reference genome sequencing. Genome Med. 7, 113, http://dx.doi.org/10.1186/s13073-015-0235-2.

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

Hebert, P.D.N., Ratnasingham, S., deWaard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B 270, S96–S99.

Heidel, A.J., Glöckner, G., 2008. Mitochondrial genome evolution in the social amoebae. Mol. Biol. Evol. 25, 1440–1450.

Kosakyan, A., Gomaa, F., Lara, E., Lahr, D.J.G., 2016. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. Eur. J. Protistol. 55, 105–117, http://dx.doi.org/10.1016/j.ejop.2016.02.001.

Kosakyan, A., Gomaa, F., Mitchell, E.A.D., Heger, T.J., Lara, E., 2013. Using DNA-barcoding for sorting out protist species complexes: a case study of the *Nebela tincta-collaris-bohemica* group (Amoebozoa; Arcellinida, Hyalospheniidae). Eur. J. Protistol. 49, 222–237.

Kosakyan, A., Heger, T.J., Leander, B.S., Todorov, M., Mitchell, E.A.D., Lara, E., 2012. COI barcoding of nebelid testate amoebae (Amoebozoa: Arcellinida): extensive cryptic diversity and redefinition of the hyalospheniidae schultze. Protist 163, 415–434.

Kosakyan, A., Mulot, M., Mitchell, E.A.D., Lara, E., 2015. Environmental DNA COI barcoding for quantitative analysis of protists communities: a test using the *Nebela collaris* complex (Amoebozoa; Arcellinida; Hyalospheniidae). Eur. J. Protistol. 51, 311–320.

Kudryavtsev, A., 2014. *Paravannella minima* n. g. n. sp. (Discosea, Vannellidae) and distinction of the genera in the vannellid amoebae. Eur. J. Protistol. 50, 258–269.

Kudryavtsev, A., Pawlowski, J., 2015. *Cunea* ng (Amoebozoa, Dactylopodida) with two cryptic species isolated from different areas of the ocean. Eur. J. Protistol. 51, 197–209.

Kudryavtsev, A., Pawlowski, J., Hausmann, K., 2009. Description and phylogenetic relationships of *Spumochlamys perforata* n. sp. and *Spumochlamys bryora* n. sp. (Amoebozoa, Arcellinida). J. Eukaryot. Microbiol. 56, 495–503.

Kudryavtsev, A., Wylezich, C., Pawlowski, J., 2011. *Ovalopodium desertum* n. sp. and the phylogenetic relationships of Cochliopodiidae (Amoebozoa). Protist 162, 571–589.

Lahr, D.J.G., Grant, J.R., Katz, L.A., 2013. Multigene phylogenetic reconstruction of the Tubulinea (Amoebozoa) corroborates four of the six major lineages, while additionally revealing that shell composition does not predict phylogeny in the Arcellinida. Protist 164, 323–339.

Lahr, D.J.G., Grant, J., Nguyen, T., Lin, J.H., Katz, L.A., 2011a. Comprehensive phylogenetic reconstruction of amoebozoa based on concatenated analyses of SSU-rDNA and actin genes. PLoS One 6, e22780.

Lahr, D.J.G., Katz, L.A., 2009. Reducing the impact of PCR-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity DNA polymerase. Biotechniques 47, 857–863.

Lahr, D.J.G., Nguyen, T.B., Barbero, E., Katz, L.A., 2011b. Evolution of the actin gene family in testate lobose amoebae (Arcellinida) is characterized by two distinct clades of paralogs and recent independent expansions. Mol. Biol. Evol. 28, 223–236.

Lara, E., Heger, T.J., Ekelund, F., Lamentowicz, M., Mitchell, E.A.D., 2008. Ribosomal RNA genes challenge the monophyly of the hyalospheniidae (Amoebozoa: Arcellinida). Protist 159, 165–176.

Loftus, B., Anderson, I., Davies, R., Alsmark, U.C.M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillén, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri Jr, W.A., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., Hall, N., 2005. The genome of the protist parasite *Entamoeba histolytica*. Nature 433, 865–868.

Mazei, Y., Warren, A., 2014. A survey of the testate amoeba genus *Difflugia* leclerc, 1815 based on specimens in the E. Penard and C. G. Ogden collections of the natural history museum, London. Part 2: species with shells that are pyriform or elongate. Protistology 8, 39.

Morard, R., Escarguel, G., Weiner, A.K., André, A., Douady, C.J., Wade, C.M., Darling, K.F., Ujiié, Y., Seears, H.A., Quillévéré, F., De Garidel-Thoron, T., De Vargas, C., Kucera, M., 2016.

Nomenclature for the nameless: a proposal for an integrative molecular taxonomy of cryptic diversity exemplified by planktonic foraminifera. Syst. Biol., syw031.

Nassonova, E., Smirnov, A., Fahrni, J., Pawlowski, J., 2010. Barcoding amoebae: comparison of SSU, ITS and COI Genes as tools for molecular identification of naked lobose amoebae. Protist 161, 102–115.

Nikolaev, S.I., Mitchell, E.A.D., Petrov, N.B., Berney, C., Fahrni, J., Pawlowski, J., 2005. The testate lobose amoebae (order Arcellinida Kent, 1880) finally find their home within Amoebozoa. Protist 156, 191–202.

Ogawa, S., Yoshino, R., Angata, K., Iwamoto, M., Pi, M., Kuroe, K., Matsuo, K., Morio, T., Urushihara, H., Yanagisawa, K., Tanaka, Y., 2000. The mitochondrial DNA of *Dictyostelium discoideum*: complete sequence, gene content and genome organization. Mol. Gen. Genet. 263, 514–519.

Ogden, C.G., 1979. Comparative morphology of some pyriform species of Difflugia (Rhizopoda). Arch. Protistenkd. 122, 143–153.

Ogden, C.G., 1980. Notes on some Difflugiidae from Norfolk (Rhizopodea, Protozoa). Bull. Brit. Mus. (Nat. Hist.) Zool. 39, 125–138.

Oliverio, A.M., Lahr, D.J.G., Grant, J., Katz, L.A., 2015. Are microbes fundamentally different than macroorganisms? Convergence and a possible case for neutral phenotypic evolution in testate amoeba (Amoebozoa: Arcellinida). R. Soc. Open Sci. 2 (12), 150414.

Oliverio, A.M., Lahr, D.J.G., Nguyen, T., Katz, L.A., 2014. Cryptic diversity within morphospecies of testate amoebae (Amoebozoa: Arcellinida) in New England bogs and fens. Protist 165, 196–207.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A.M., Gile, G.H., Holzmann, M., Jahn, R., Jirku, M., Keeling, P.J., Kostka, M., Kudryavtsev, A., Lara, E., Lukes, J., Mann, D.G., Mitchell, E.A.D., Nitsche, F., Romeralo, M., Saunders, G.W., Simpson, A.G.B., Smirnov, A.V., Spouge, J.L., Stern, R.F., Stoeck, T., Zimmermann, J., Schindel, D., de Vargas, C., 2012. CBOL Protist Working Group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol. 10, e1001419.

Pawlowski, J., Holzmann, M., 2014. A plea for DNA barcoding of foraminifera. J. Foram. Res. 44, 62–67.

Penard, E., 1890. Études sur les Rhizopodes d'eau douce. Imprimerie Aubert-Schuchardt.

Perty, M., 1849. Mikroskopische Organismen der Alpen und der italienischen Schweiz. Mittheilungen der Naturforschenden Gesellschaft in Bern 164, 153–176.

Pombert, J.-F., Janouškovec, J., Smirnov, A., Gray, M., James, E., Keeling, P.J., 2013. The complete mitochondrial genome from an unidentified Phalansterium species. Protist Genom. 1, 25–32.

Porfirio et al., 2016, this issue.

Singer, D., Kosakyan, A., Pillonel, A., Mitchell, E.A.D., Lara, E., 2015. Eight species in the *Nebela collaris* complex: *Nebela gimlii* (Arcellinida, Hyalospheniidae), a new species described from a Swiss raised bog. Eur. J. Protistol. 51, 79–85.

Singer, D., Lara, E., Steciow, M.M., Seppey, C.V., Paredes, N., Pillonel, A., Oszako, T., Belbahri, L., 2016. High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat. Fungal Ecol. 23, 42–47.

Smirnov, A., Nassonova, E., Berney, C., Fahrni, J., Bolivar, I., Pawlowski, J., 2005. Molecular phylogeny and classification of the lobose amoebae. Protist 156, 129–142.

Stoneking, M., 2013. Mitochondrial Genome: Evolution, eLS. John Wiley & Sons, Ltd.

Takano, H., Abe, T., Sakurai, R., Moriyama, Y., Miyazawa, Y., Nozaki, H., Kawano, S., Sasaki, N., Kuroiwa, T., 2001. The complete DNA sequence of the mitochondrial genome of *Physarum polycephalum*. Mol. Gen. Genet. 264, 539–545.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tekle, Y.I., Anderson, O.R., Katz, L.A., Maurer-Alcala, X.X., Romero, M.A.C., Molestina, R., 2016. Phylogenomics of 'Discosea': a new molecular phylogenetic perspective on Amoebozoa with flat body forms. Mol. Phylogenet. Evol. 99, 144–154.

Tekle, Y.I., Grant, J., Anderson, O.R., Nerad, T.A., Cole, J.C., Patterson, D.J., Katz, L.A., 2008. Phylogenetic placement of diverse amoebae inferred from multigene analyses and assessment of clade stability within 'Amoebozoa' upon removal of varying rate classes of SSU-rDNA. Mol. Phylogenet. Evol. 47, 339–352.

Traphagen, S.J., Dimarco, M.J., Silliker, M.E., 2010. RNA editing of 10 *Didymium iridis* mitochondrial genes and comparison with the homologous genes in *Physarum polycephalum*. RNA 16, 828–838.