Culturing and environmental DNA sequencing uncover hidden kinetoplastid biodiversity and a major marine clade within ancestrally freshwater *Neobodo designis*

Sophie von der Heyden† and Thomas Cavalier-Smith

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Correspondence
Thomas Cavalier-Smith
tom.cavalier-smith@zoo.ox.ac.uk

Bodonid flagellates (class Kinetoplastea) are abundant, free-living protozoa in freshwater, soil and marine habitats, with undersampled global biodiversity. To investigate overall bodonid diversity, kinetoplastid-specific PCR primers were used to amplify and sequence 18S rRNA genes from DNA extracted from 16 diverse environmental samples; of 39 different kinetoplastid sequences, 35 belong to the subclass Metakinetoplastina, where most group with the genus *Neobodo* or the species *Bodo saltans*, whilst four group with the subclass Prokinetoplastina (*Ichthyobodo*). To study divergence between freshwater and marine members of the genus *Neobodo*, 26 new *Neobodo designis* strains were cultured and their 18S rRNA genes were sequenced. It is shown that the morphospecies *N. designis* is a remarkably ancient species complex with a major marine clade nested among older freshwater clades, suggesting that these lineages were constrained physiologically from moving between these environments for most of their long history. Other major bodonid clades show less-deep separation between marine and freshwater strains, but have extensive genetic diversity within all lineages and an apparently biogeographically distinct distribution of *B. saltans* subclades. Clade-specific 18S rRNA gene primers were used for two *N. designis* subclades to test their global distribution and genetic diversity. The non-overlap between environmental DNA sequences and those from cultures suggests that there are hundreds, possibly thousands, of different rRNA gene sequences of free-living bodonids globally.

INTRODUCTION

Kinetoplastea comprise the mainly free-living bodonid flagellates and parasitic trypanosomatids within the phylum Euglenozoa (Cavalier-Smith, 1981, 2003a, b). Bodonids are common, heterotrophic protozoa that occur in a wide variety of environments, including marine, freshwater and soil, from the tropics to the Arctic (Al-Qassab *et al.*, 2002; Larsen & Patterson, 1990; Patterson & Simpson, 1996; Vørs,

1992, 1993). The bodonids *Bodo saltans*, *Neobodo designis* and *Rhynchomonas nasuta* are among the 20 most commonly seen zooflagellates (Patterson & Lee, 2000); it is often suggested that such protists are cosmopolitan and may have wide ecological tolerances (Atkins *et al.*, 1998, 2000; Finlay, 2002; Patterson & Lee, 2000; Patterson & Simpson, 1996). This frequent assumption of broad physiological tolerance and easy interchange between marine and freshwater environments was recently clearly contradicted for members of the genus *Goniomonas*, heterotrophic cryptomonads with deeply divergent marine and freshwater clades (von der Heyden *et al.*, 2004a). In bodonids, recent evidence for *N. designis* suggested marked divergence between freshwater and marine strains, and all bodonid morphospecies, especially *B. saltans*, exhibited extensive genetic diversity (von der Heyden *et al.*, 2004b). Despite considerable interest in free-living bodonids (Callahan *et al.*, 2002; Doležel *et al.*, 2000; Simpson *et al.*, 2002), their true biodiversity has probably been grossly underestimated by simple light microscopy, as most 'species' are harder to tell apart by using light microscopy than the easily identifiable *B. saltans* (von der Heyden *et al.*, 2004b).

Environmental surveys using general eukaryote 18S rRNA gene primers revealed a larger-than-expected diversity of free-living eukaryotes (Moon-van der Staay et al., 2001; Moreira & López-García, 2002) but, curiously, found very few kinetoplastids – only seven environmental sequences were detected (Atkins et al., 2000; López-García et al., 2003; van Hannen et al., 1999). Rasmussen et al. (2001) used kinetoplastid-specific primers to amplify 24S α-subunit rRNA genes from Danish lake sediments and found 27 of 43 clones sequenced to be different from each other and from known cultures, indicating a much higher diversity of kinetoplastids than realized previously. This suggests, as does the even greater diversity found recently for Cercozoa by using phylum-specific primers (Bass & Cavalier-Smith, 2004), that group-specific primers may be better for elucidating the biodiversity of specific groups than generalized eukaryotic ones. Given the much larger database for 18S than for 24S rRNA genes, culture-independent studies of kinetoplastid diversity using group-specific 18S rRNA gene primers should allow more comprehensive assessment of their global diversity. To achieve this and to study their distribution among different geographical areas and environments, we here use class-specific kinetoplastid PCR primers to amplify 18S rRNA genes from DNA extracted from a wide range of samples (16 environmental libraries yielded 39 new sequences). We also isolated 26 strains corresponding to the accounts of Skuja (1948) and Eyden (1977) of N. designis from both marine and freshwater environments in order to test whether earlier indications of divergence among strains from these contrasting environments (von der Heyden et al., 2004b) are real or an artefact of undersampling. To further test ubiquitous dispersal (Finlay, 2002) and environmental specialization within N. designis, we designed clade-specific PCR primers to amplify a highly variable region of the 18S rRNA gene from two subclades within the N. designis species complex, one putatively exclusively freshwater and one marine.

Our results show that most environmental sequences fall within well-characterized groups within the free-living metakinetoplastids (Bodonida sensu stricto) and that four sequences belong to the early diverging prokinetoplastids [for the recent important distinction between prokinetoplastids and metakinetoplastids and the segregation of prokinetoplastids from 'core bodonids' as the novel order Prokinetoplastida, see Moreira et al. (2004)]. In the present paper, we shall refer collectively to all non-trypanosomatid metakinetoplastids simply as bodonids, rather than the more cumbersome term 'core bodonids' used previously (von der Heyden et al., 2004b), as we prefer to retain the order Bodonida following the removal of prokinetoplastids, with three suborders: Eubodonina (eubodonids: B. saltans), Parabodonina (parabodonids, e.g. the genus Parabodo) and Neobodonina (neobodonids). Within the N. designis morphospecies, we show that strains isolated from different environments fall into exclusively marine and freshwater lineages – they are not distributed randomly between environments (as sometimes assumed; Finlay, 2002;

Patterson & Lee, 2000) – and that N. designis is probably a vast complex of at present morphologically indistinguishable, but genetically diverse, 'species'.

METHODS

Isolation, genomic DNA extraction, PCR amplification and sequencing of kinetoplastid 18S rRNA genes. Details of strains used, their sources and GenBank accession numbers are given in Table 1; Table 2 shows the sources for environmental libraries and the number of different genotypes isolated from each. All bodonid strains were isolated and their DNA was extracted, PCR-amplified and sequenced as described previously (von der Heyden et al., 2004b). Collection and processing of environmental samples were as described by Bass & Cavalier-Smith (2004). Samples were stored at 4 °C immediately after collection and were used within a few hours of collecting, except for samples collected in the South Island of New Zealand, which remained at 4 °C for up to 5 days. The 18S rRNA genes from both cultures and environmental samples were PCR-amplified by using kinetoplastid-specific primers: kineto14F (5′-CTGCCAGTAGTCATATGCTTGTTTCAAGGA-3′ and kineto2026R (5′-GATCCTTCTGCAGGTTCACCTACAGCT-3′). These primers yield a PCR product of approximately 1·9–2·2 kb. PCR amplification of DNA extracted from environmental samples used the same PCR mixture as described previously (von der Heyden et al., 2004b) but, to reduce chimaera formation between different 18S rRNA genes within a sample, a modified cycling program was used: an initial denaturation at 94 °C for 5 min, 35 cycles (94 °C for 30 s, 69 °C for 36 s, 72 °C for 4·5 min) and a final extension at 72 °C for 10 min. Prior to gel purification, four PCR products were pooled for each sample amplified.

For clade-specific amplification of a highly variable region of the 18S rRNA gene for two clades of the N. designis species complex, two primer sets were used: FWb_681F (5′-GGAGTCGGTTACGTCCC-RTCCTCCGRRYCG-3′) and FWb_1244R (5′-TATTCTCTTTTGG-CGGGMTCAGCAAGCGAG-3′) for a freshwater clade (Fig. 7; amplifies 580 bp) and MARa_502F (5′-CAGAGATTTTCAATGGGG-GATATTTAAYG-3′) and MARa_1321R (5′-GGACGTGCTGAGG-ATATTCCCGWTA-3′) for a marine clade (Fig. 8; amplifies 780 bp). The following cycling conditions were used for the PCRs: initial denaturation at 94 °C for 3 min, 35 cycles (94 °C for 30 s, 70 °C for 36 s, 72 °C for 1 min), final extension at 72 °C for 5 min (freshwater clade), or an initial denaturation at 94 °C for 3 min, 35 cycles (94 °C for 30 s, 69 °C for 36 s, 72 °C for 1 min 15 s), final extension at 72 °C for 5 min (marine clade).

All fragments were gel-purified by using a GFX kit (Amersham Biosciences) and cloned into the pCR2.1 vector (Invitrogen) by using TOPO TA cloning kits (Invitrogen). Three to five clones were picked from libraries generated from cultured strains and 12–15 clones were picked for initial screening of environmental libraries; this was later reduced to eight to ten clones. All clones were partially sequenced by using the M13F primer; the sequences were aligned with each other and those of kinetoplastids from our database. Clones were edited by using BioEdit (Hall, 1999) and checked by BLAST (Altschul et al., 1990); all clones from environmental samples proved to be kinetoplastid, showing 100 % amplification specificity. To eliminate sequences represented multiple times, preliminary gamma-corrected BioNJ trees were constructed. To avoid the inclusion of identical or near-identical clones from the same library, those differing by > 10 nt in the first 600–800 nt sequenced were chosen for complete bidirectional sequencing and phylogenetic analysis; by excluding more similar sequences, we probably underestimated the actual diversity of closely

Table 1. New strains of *N. designis* and *Dimastigella* and their place of isolation

All freshwater strains were isolated and grown in soil extract with added salts medium; marine strains were grown in Plymouth Erdschreiber medium (UKNCC, 2001). FW, Freshwater; NI, North Island.

Name	Place	FW/marine	GenBank no.
N. designis			
TETOTO 1	Hair weed from stream, TeToto Gorge, New Zealand (NI)	FW	AY753603
TETOTO 2	Weed from stream, TeToto Gorge, New Zealand (NI)	FW	AY753606
TETOTO 3	Soil, TeToto George, New Zealand (NI)	FW	AY753620
TETOTO 4	Sediment from stream, TeToto Gorge, New Zealand (NI)	FW	AY753619
ROTA 1	Sediment, Lake Rotawiti, New Zealand (NI)	FW	AY753605
ROTA 2	Sediment, Lake Rotawiti, New Zealand (NI)	FW	AY753607
DUNE 1	Dry dune sand, Raglan, New Zealand (NI)	FW	AY753622
DUNE 2	Dry dune sand, Raglan, New Zealand (NI)	FW	AY753623
AB	Stream sediment, Amoundi Beach, Crete	FW	AY753604
UW	Soil, University of Waikato, New Zealand (NI)	FW	AY753621
RAGLAN 1	*Zostera* patch, Raglan Harbour, Raglan, New Zealand (NI)	Marine	AY753626
RAGLAN 2	Algae, Raglan Harbour, New Zealand (NI)	Marine	AY753627
RAGLAN 3	*Zostera* and sea star patch, Raglan, New Zealand (NI)	Marine	AY753629
RAGLAN 4	Sediment, Raglan Harbour, New Zealand (NI)	Marine	AY753630
AUCKLAND	Organic sediment, Auckland, New Zealand (NI)	Marine	AY753628
THAILAND (CCAP 1951/1)	Brown algae, Sri Racha, Thailand	Marine	AY753632
PANAMA	Sediment, Bejuco river, Panama	FW	AY753617
DEVON 1	Sediment, River Exe, Devon	FW	AY753612
DEVON 2	Sediment and algae, River Exe, Devon	FW	AY753611
PP	Sediment, Priest Pot, UK	FW	AY753614
JC11	Top layer of sediment, Priest Pot, Cumbria, UK	FW	AY753608
JC12	Top layer of sediment, Priest Pot, Cumbria, UK	FW	AY753613
JC16	Top layer of sediment, Priest Pot, Cumbria, UK	FW	AY753615
JC17	Top layer of sediment, Priest Pot, Cumbria, UK	FW	AY753609
OXFORD	Sediment, River Thames, Oxford, UK	FW	AY753616
NIVA	Sediment, Nivå Bay, Denmark	Marine	AY753602
Dimastigella trypaniformis			
SHI	Soil, Sourhope, UK	FW	AY753598
SHII	Soil, Sourhope, UK	FW	AY753997
KV1	Soil, Sourhope, UK	FW	AY753999
Neobodo sp.			
Pan 1	Bejuco river, Panama	FW	AY753624
Pan 2	Lake, Camino de Cruces, Panama	FW	AY753625
OX	Sediment, River Thames, Oxford, UK	FW	AY753631
SF	Sediment, Pond, Golden Gate Park, San Francisco, CA, USA	FW	AY753610
DEV1	Algae from rock pool, Lynmouth, UK	Marine	AY753601
DEV 2	Algal scrape, Lynmouth, UK	Marine	AY753600
RAG1	*Zostera* patch and sediment, Raglan, New Zealand	Marine	AY753618

related kinetoplastid sequences. Bidirectional sequencing was carried out by using BigDye terminators with separation on an automated ABI 3730 sequencer (both from Applied Biosystems).

Sequence analysis and phylogenetic methods. Before carrying out phylogenetic analyses, all sequences were checked for chimaera formation. Firstly, the whole dataset was submitted for analysis by BELLEROPHON (Huber *et al.*, 2004), which compares sequences within an alignment with each other to find chimaeras from separate parent sequences. Secondly, each environmental 18S rRNA gene sequence was analysed separately in three parts, each of about 710 nt, by constructing gamma-corrected neighbour-joining trees. The position of each taxon was then checked on the separate trees;

additionally, each of the three sections of the 18S rRNA gene was analysed by using BLAST (Altschul *et al.*, 1990). Chimaeras were detected easily by using these combined methods; three definite chimaeras (7 %) were found and excluded from subsequent analyses. This compares favourably with figures given by Berney *et al.* (2004) and Robison-Cox *et al.* (1995), who found 14 and 4·1–20 % of sequences to be chimaeric in analyses of eukaryote and prokaryote environmental surveys, respectively.

Analyses used PAUP* v. 4.0b10 (Swofford, 2001) with the general time-reversible (GTR) substitution model. To allow for rate variation across sites, all methods assumed a gamma distribution (Γ) and proportion of invariable sites (I), estimated by MODELTEST v. 3.06 (Posada &

Table 2. Environmental library description, including source and the number of different genotypes isolated from each

Note that only the most abundant genotypes will have been detected from each library; also, only clones differing in single sequence-reaction tests by >10 nt were sequenced fully – this probably underestimated the actual diversity of closely related sequences in each sample. FW, Freshwater; NI, North Island; NP, National Park; SI, South Island.

Primers/library name	Source	FW/Marine	No. genotypes
Kineto14F + Kineto2026R			
L1	Muddy sand and diatom film, Vancouver, BC, Canada	Marine	1
L2	Sediment, Harrison Hot Springs, BC, Canada	FW	2
L3	Muddy ditch by road, Capira, Panama	FW	1
L4	Submerged stones, Bejuco River, Panama	FW	1
L5	Sediment/algae from lake, Camino de Cruces, NP Panama	FW	4
L6	Leaf litter in rainforest, Cerro Campana NP, Panama	FW	3
L7	Sediment/algae from very eutrophic lake, Gamboa, Panama	FW	1
L8	Scrape from rocks, edge of sea, nr Portobelo, Panama	Marine	5
L9	Algae from river, Wilderness, Western Cape, South Africa	FW	1
L10	Sediment and water, Sakhalin island, Russia	FW	3
L11	Algae and mud, Okavango Delta, Botswana	FW	5
L12	Moss on tree, Rotorua, New Zealand (NI)	FW	3
L13	Algae and water, Harris Tarn, New Zealand (SI)	FW	3
L14	Algae and water, Raglan, New Zealand (NI)	Marine	1
L15	Stream, Amoundi Beach, Crete	FW	2
L16	Sediment, Priest Pot, Cumbria, UK	FW	3
FWb_681F + FWb_1244R			
FW1	Loamy soil, Anaura Bay, New Zealand (NI)	FW	1
FW2	Moss, Anaura Bay, New Zealand (NI)	FW	3
FW3	Sediment, small tarn on Harris Saddle, New Zealand (SI)	FW	3
FW4	Sediment, small tarn on Harris Saddle, New Zealand (SI)	FW	2
FW5	Stream, nr TeToto Gorge, New Zealand (NI)	FW	1
FW6	Stones and sediment, Hamilton Lake, New Zealand (NI)	FW	3
FW7	Water column, Hamilton Lake, New Zealand (NI)	FW	2
FW8	Soil, Singapore Botanic Gardens, Singapore	FW	1
FW9	Sediment and algae, Priest Pot, Cumbria, UK	FW	3
FW10	Sediment, Priest Pot, Cumbria, UK	FW	2
FW11	1·5 m depth water column, Priest Pot, Cumbria, UK	FW	2
FW12	2·0 m depth water column, Priest Pot, Cumbria, UK	FW	3
FW13	Sediment, Grasmere Lake, Cumbria, UK	FW	1
FW14	Sediment and algae, Grasmere Lake, Cumbria, UK	FW	1
FW15	Moss, Buttermere Lake, Cumbria, UK	FW	1
FW16	Eutrophic pond next to Panama canal, Gamboa, Panama	FW	1
FW17	Sediment and rock scrapings, Bejuco river, Panama	FW	1
FW18	Sediment and rock scrapings, Bejuco river, Panama	FW	5
FW19	Sediment, Harrison Hot Springs, BC, Canada	FW	2
FW20	Freshwater stream, Amoundi Beach, Crete	FW	4
MARa_502F + MARa_1321R			
M1.1	Scrapings from coastal rocks, nr Portobelo, Panama	Marine	1
M2.3	Scrapings from coastal rocks, nr Portobelo, Panama	Marine	1
M3.4	Sediment and algae, Raglan, New Zealand	Marine	1

Crandall, 1998). All heuristic analyses used tree bisection–reconnection for branch rearrangement. Dataset 1 [50 taxa, 1326 positions; prokinetoplastids and metakinetoplastids, using two euglenozoan environmental sequences as outgroup; the closest outgroup found in our previous analyses (von der Heyden *et al.*, 2004b)] was analysed by using maximum likelihood (ML) (GTR + Γ + I; four discrete rate categories) with a heuristic search with 10 random-addition sequence replicates; heuristic bootstrap analyses used 50 resamplings with one addition-sequence replicate. Datasets 2 (149 taxa, 1598 positions; all available core bodonids and trypanosomatids; prokinetoplastids were excluded to allow inclusion of more nucleotide positions) and 3 (76 taxa, 1720 positions; all available *N. designis* sequences and other clade 1 taxa as outgroups) were analysed by using weighted least squares and a heuristic search of 100 random stepwise-addition sequence replicates; bootstrap analyses used 500 resamplings with one addition-sequence replicate, but each replicate was restricted to

10 min. Dataset 4 (37 taxa, 1745 positions; all available *B. saltans* sequences and a selection of other bodonids as outgroup) was analysed by using ML (GTR+Γ+I; four discrete rate categories) with a heuristic search with 10 random-addition sequence replicates; heuristic bootstrap analyses used 75 resamplings with one addition-sequence replicate. Dataset 5 (100 taxa, 546 positions; freshwater subclade of *N. designis*) was analysed by using distance (BioNJ) methods; bootstrap analyses used 1000 BioNJ replicates. Dataset 6 (50 taxa, 728 positions; the marine subclade of *N. designis*) was analysed by using the minimum-evolution method, with a heuristic search of 50 random stepwise-addition sequence replicates; bootstrap analyses used 250 resamplings with one addition-sequence replicate. Alignments are available on request.

Light and electron microscopy of neobodonid strains. Live cells of three genetically divergent *N. designis* strains (freshwater, JC11 and Panama; marine, Thai) were examined by phase-contrast light microscopy with a ×100 objective on a Zeiss Axiovert 200 microscope and photographed with a Roper Scientific Coolsnap HQ camera. For electron microscopy (carried out by Mr Ken Clarke, CEH Windermere, UK), live cells were fixed in 2 % glutaraldehyde in 0·1 M sodium cacodylate buffer for 60 min at 15 °C. Samples were centrifuged, washed in the same buffer and fixed for a further 30 min in 2 % osmium tetroxide in 0·1 M cacodylate buffer; samples were washed three times in 0·1 M cacodylate buffer and resuspended. Suspended cells (5 μl) were placed onto Formvar-coated grids, dried at room temperature and negatively stained with 1 % uranyl acetate in aqueous solution containing 0·001 % bacitracin.

RESULTS

Our kinetoplastid-specific primers amplified the 18S rRNA gene from 16 diverse environmental samples, yielding sequences from all major free-living clades within the subclass Metakinetoplastina except for parabodonids (unsurprisingly, not from any trypanosomatids or other parasites) and also from Prokinetoplastida, which suggests that there were no gross amplification artefacts amplifying any one clade preferentially (Fig. 1). Nine environmental libraries from temperate environments and seven from tropical environments were constructed (Table 2). Most

new sequences belong to neobodonids (clade 1: *Neobodo, Rhynchobodo, Rhynchomonas, Dimastigella, Cruzella, Cryptaulaxoides*-like), but eubodonids (clade 3: *B. saltans*) are also well-represented. In total, we aligned 75 new bodonid sequences (including three *Dimastigella* and seven sequences from strains that group closely as sister to *Cryptaulaxoides*-like) to all previously published ones. There is extensive genetic diversity in a number of lineages, notably *N. designis, B. saltans* and the genus *Rhynchobodo*. Of the 39 different environmental sequences found, four group with the early-diverging prokinetoplastids with good support (77 %) (Fig. 2). Two of these (L4.9 and L12.4) group as sister to the *Ichthyobodo/Perkinsiella* symbiont clade with moderate (66 %) bootstrap support. Two other sequences (L7.7 and L8.10) group with the unknown deep-sea hydrothermal-vent environmental sequence AT4-56. All other environmental sequences generated in this study group within core bodonids: all group within clades that contain isolates of known morphology (Fig. 3). Clade 1 (neobodonids) contains the majority of new environmental sequences (64 %). Clade 3 (eubodonids) contains 26 %; no environmental sequences were found in clade 2 (parabodonids), which mainly consists of parasitic species. Fig. 1 shows the percentages of environmental sequences within the major clades of kinetoplastids and neobodonids. All three major clades have strong bootstrap support (Fig. 3); clade 1 (Neobodonina; 86 %), clade 2 (Parabodonina; 91 %), clade 3 (Eubodonina; 100 %). Clade 1 is further subdivided into five clades, labelled 1A–1E (von der Heyden *et al.*, 2004b). Fig. 3 also includes sequences for three additional *Dimastigella trypaniformis* strains (kindly donated by Professor K. Vickerman, Glasgow, UK) and seven newly isolated strains that group as sister to *Cryptaulaxoides*-like in clade 1D.

We isolated 26 strains of bodonids that were indistinguishable in morphology from *N. designis* (Eyden, 1977; Skuja, 1948) and found that they fall into clade 1C and can be subdivided into six moderately to robustly supported major lineages (designated Nd1–Nd6; Fig. 4) that consist almost exclusively of marine or freshwater isolates and/or environmental sequences, and that, within each clade, there is remarkable genetic diversity among even closely related strains. Fig. 5 shows light micrographs of three genetically divergent strains of *N. designis* to illustrate their similarity. Fig. 6 shows electron micrographs of one marine and one freshwater *N. designis* strain, whose cilia were examined for the presence of mastigonemes.

By using still more-specific primers designed to amplify only a narrow array of genotypes within one marine and one freshwater clade (indicated by two boxes in Fig. 4), we further explored the genetic diversity of two subclades of *N. designis*. We obtained 42 new sequences from 20 environmental libraries within the freshwater clade (Table 1, Fig. 7); the freshwater-specific primers amplified products only from freshwater samples. Bootstrap support for lineages within this subclade is generally low; support for the clade itself is high (Fig. 7; 90 %). Although we screened
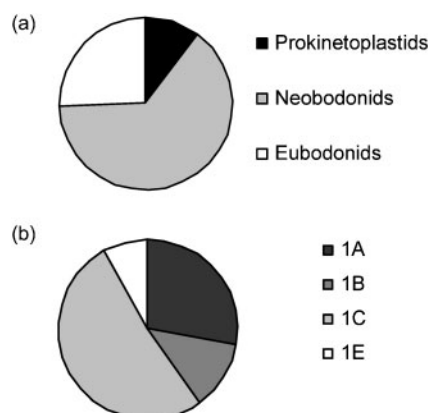
Fig. 1. Pie charts showing the percentages of environmental sequences recovered for (a) each clade of prokinetoplastids and metakinetoplastids and (b) subclades of clade 1 for which the majority of environmental sequences were found.
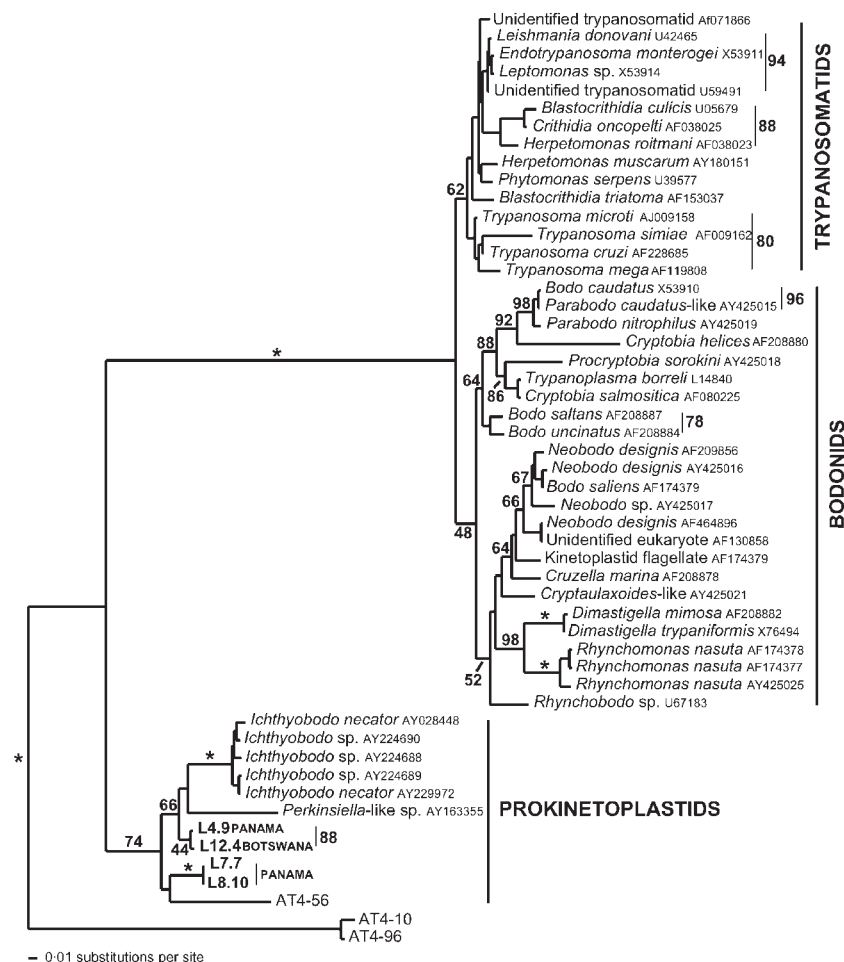
Fig. 2. ML tree of 50 18S rRNA gene sequences of kinetoplastids and two outgroup lineages using 1326 positions ($a = 0.522712$, $i = 0.305469$, $-lnL = 10526.49$). This tree included only the new prokinetoplastid environmental sequences; the new bodonid environmental sequences are shown in Fig. 3. The tree is rooted as described previously (López-García et al., 2003; Moreira et al., 2004; von der Heyden et al., 2004b). New sequences are shown in bold. Bootstrap percentages from 50 replicates are given at major nodes; percentages $> 43\%$ are shown. Asterisks denote 100% bootstrap support.

150 DNA samples, most did not show amplified bands on agarose gels and we could construct only three libraries for the marine subclade, suggesting either that there are fewer neobodonid species in marine environments or that they are less abundant, and therefore were not recovered by using our methodology. At least four lineages can be distinguished within the marine clade, each with robust bootstrap support (Fig. 8; 95–100%), but overall bootstrap support for this major marine clade is low (52% in Fig. 4; 33% in Fig. 8). Two lineages contain strains exclusively from either the northern or southern hemispheres; one contains strains from both hemispheres and one is represented by only one environmental sequence.

In a separate ML analysis of *B. saltans* isolates and environmental sequences (Fig. 9), the clade of sequences from Priest Pot, Cumbria, UK (L16.16 and L16.17) did not branch as deeply as in the distance tree in Fig. 3, but within a bigger clade of northern-hemisphere sequences. An environmental sequence from Panama (L5.5) is the most-divergent eubodonid sequence.

DISCUSSION

Our results suggest that the major clades that constitute overall kinetoplastid diversity have probably now been revealed, as all new environmental sequences fall within groups for which the morphology has been at least partially characterized; there is, however, great genetic diversity within individual clades and between morphologically identical strains, suggesting that there are large numbers

Fig. 3. Distance tree of 149 18S rRNA gene sequences of all available bodonids and a selection of 13 trypanosomatids used as outgroup, based on 1598 positions (weighted least squares power 2; GTR+Γ+I model; $a = 0.492434$, $i = 0.360997$). Newly sequenced taxa are shown in bold. Sequences from environmental libraries are given by the letter L followed by their library designation (refer to Table 2 for library source). Bootstrap support values are from 500 replicates (restricted to 10 min per replicate); asterisks denote 100% support; where no values are given, support is $< 50\%$. Refer to von der Heyden et al. (2004b) for GenBank accession numbers for additional taxa used in these analyses. The branch length of *Cryptobia helicis* has been shortened by 50% in order to fit.
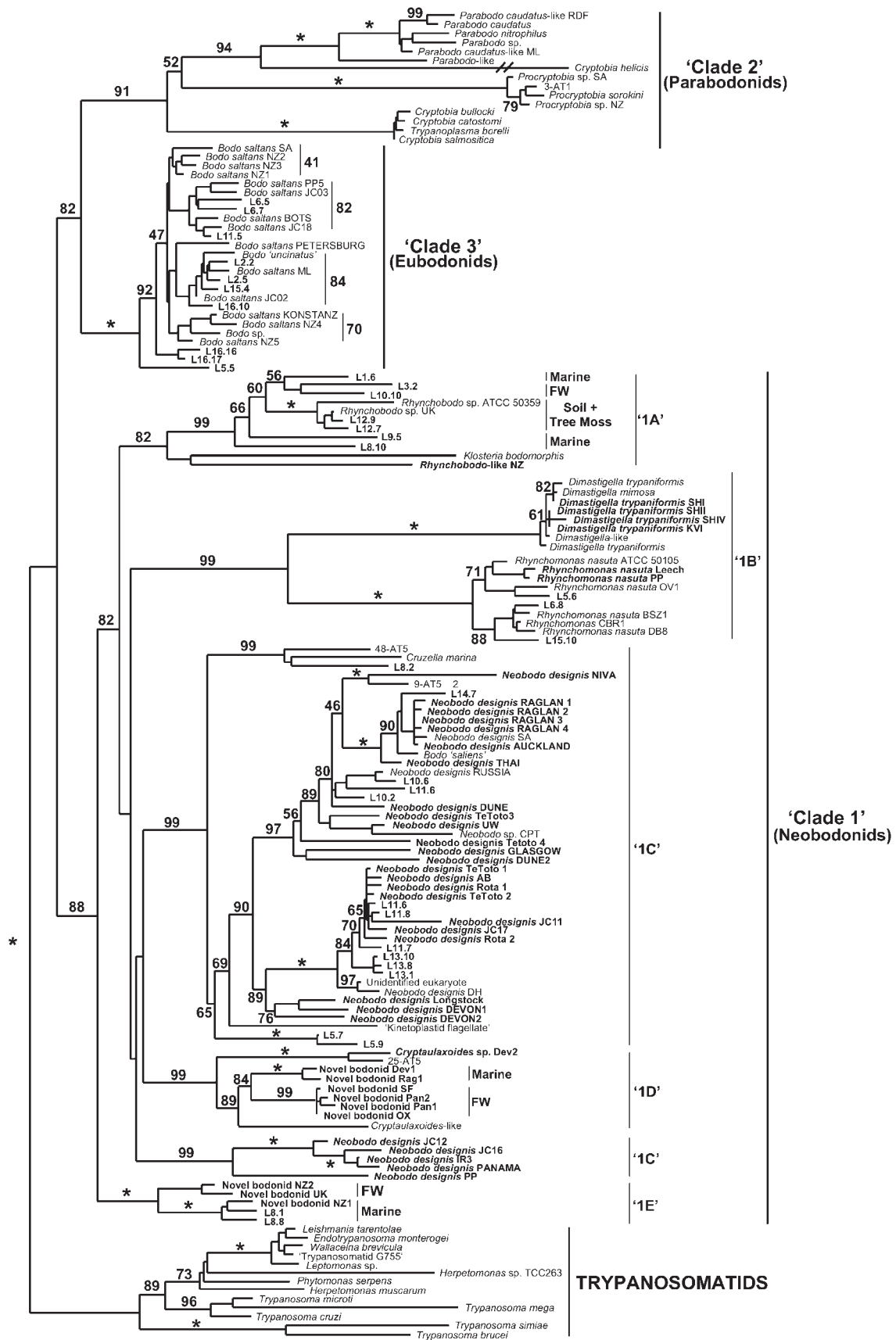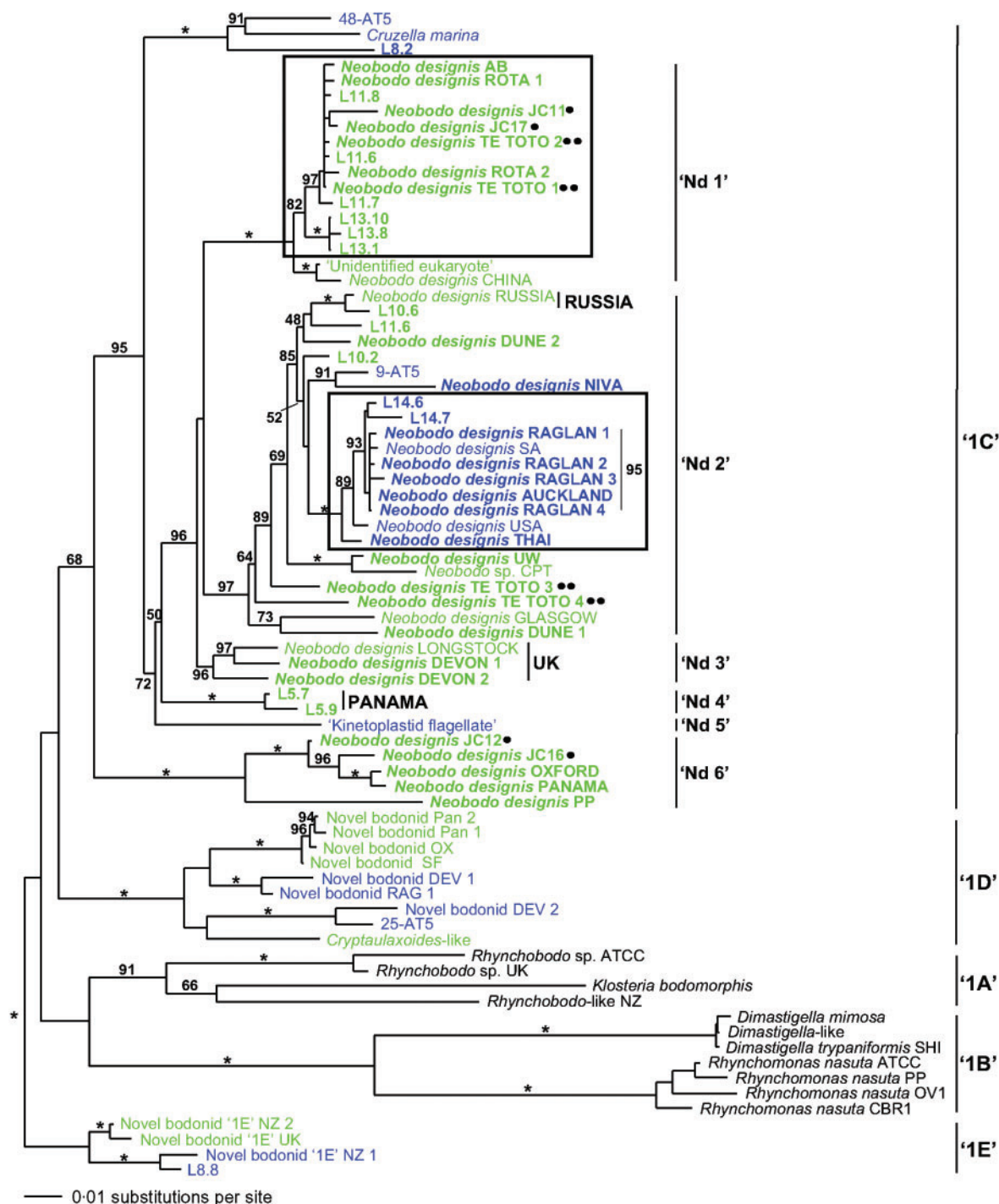
Fig. 4. Distance tree of 76 18S rRNA gene sequences of *N. designis* strains and other clade 1 outgroups using 1720 positions (weighted least squares power 2; GTR+Γ+I model; $a=0.554782$, $i=0.444142$). Lineages marked in green denote sequences from freshwater; those in blue are marine. New *N. designis* sequences only are shown in bold. Bootstrap percentages from 500 replicates (restricted to 10 min per replicate) using the WLS2 method are shown; asterisks denote 100% support; where no values are given, support is <48%. Boxes mark the clades for which clade-specific primers were designed. Dots after lineage names refer to those strains that were collected from the same geographical place at the same time; one dot represents strains from Priest Pot, Cumbria, UK; two dots mark strains from TeToto Gorge, New Zealand.
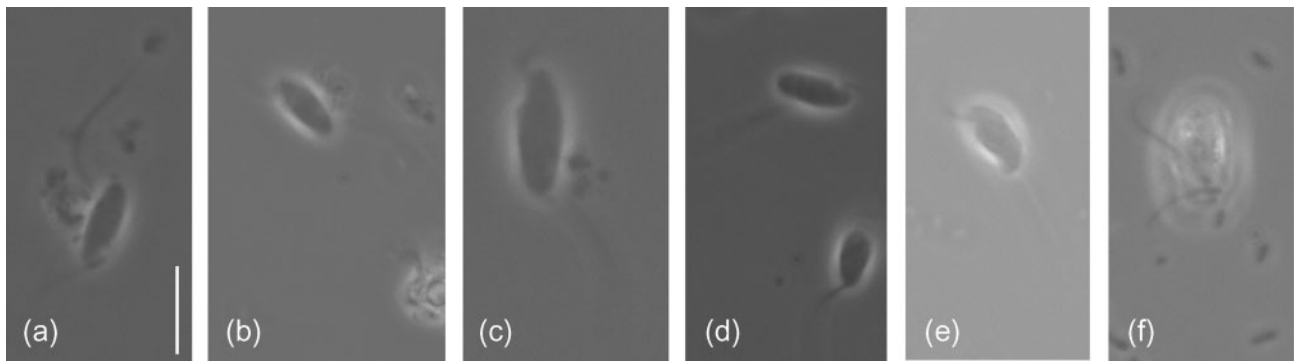
Fig. 5. Phase-contrast light micrographs of living cells of three genetically divergent strains of *N. designis* to illustrate their practical inseparability by light microscopy. (a, b) *N. designis* JC11 (freshwater); (c, d) *N. designis* Panama (freshwater); (e, f) *N. designis* Thai (marine). Bar, 8 μm.

of cryptic individual species within bodonids and prokinetoplastids and that the majority of lineages are species complexes that need a great deal more detailed study.
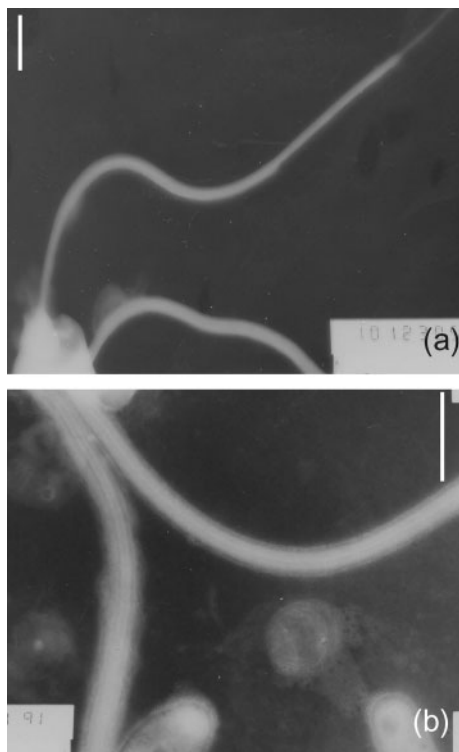
Diversity of prokinetoplastids

We obtained four new sequences that group within the early-diverging prokinetoplastids from samples from

Fig. 6. Electron micrographs of two strains of *N. designis* (one marine and one freshwater) showing both the anterior and posterior cilia; no mastigonemes were found on either strain. (a) *N. designis* JC11 (freshwater); the acronematic tip of the posterior cilium can be seen clearly. (b) *N. designis* Thai (marine). Bars, 1 μm.

Panama and Botswana (Fig. 2). Two (L7.7 and L8.10) form a substantially divergent sister group to an environmental sequence from deep-sea hydrothermal-vent samples (AT4-56), which is a lineage not yet characterized by morphology. Two others (L4.9 and L12.4) from Panama and Botswana group as sister to *Ichthyobodo*/*Perkinsiella*. It is not known whether these environmental lineages represent true free-living prokinetoplastids or are dispersal phases for parasitic species such as *Ichthyobodo*. It is interesting to note, however, that all environmental sequences related to the parasitic prokinetoplastids have been found from warm tropical environments or deep-sea hydrothermal vents, suggesting that they may be more prolific in these conditions, perhaps because of a larger pool and broader spectrum of hosts to parasitize (Moreira & López-García, 2003).

Diversity of free-living metakinetoplastids (Bodonida *sensu stricto*)

Clade 1 (neobodonids) contains the majority of sequences obtained directly from environmental samples (64 %) and the majority of these are found in clades containing isolates of *Neobodo* (Figs 1 and 3). Clade 3 (eubodonids) contains 26 % of all recovered environmental sequences. As we obtained sequences from the two clades that contain the majority of free-living kinetoplastids, it is unlikely that this is caused by a PCR amplification bias towards them, but may indicate a genuinely greater abundance of these clades over others in environmental samples. This does not mean that other kinetoplastids (e.g. *Parabodo*) may not have been present in the samples, but that they are much scarcer and our methods may not have been sensitive enough to reveal them in all samples. Moreover, *N. designis* and *B. saltans*, for which we also found a high number of environmental sequences, are two of the 20 most commonly observed zooflagellates (Patterson & Lee, 2000), so it is no surprise that they occurred most often. However, the ease of detecting these morphospecies microscopically and by culturing and environmental DNA sequencing could largely be because each comprises such a large number of different strains, rather than because any one strain is abundant or
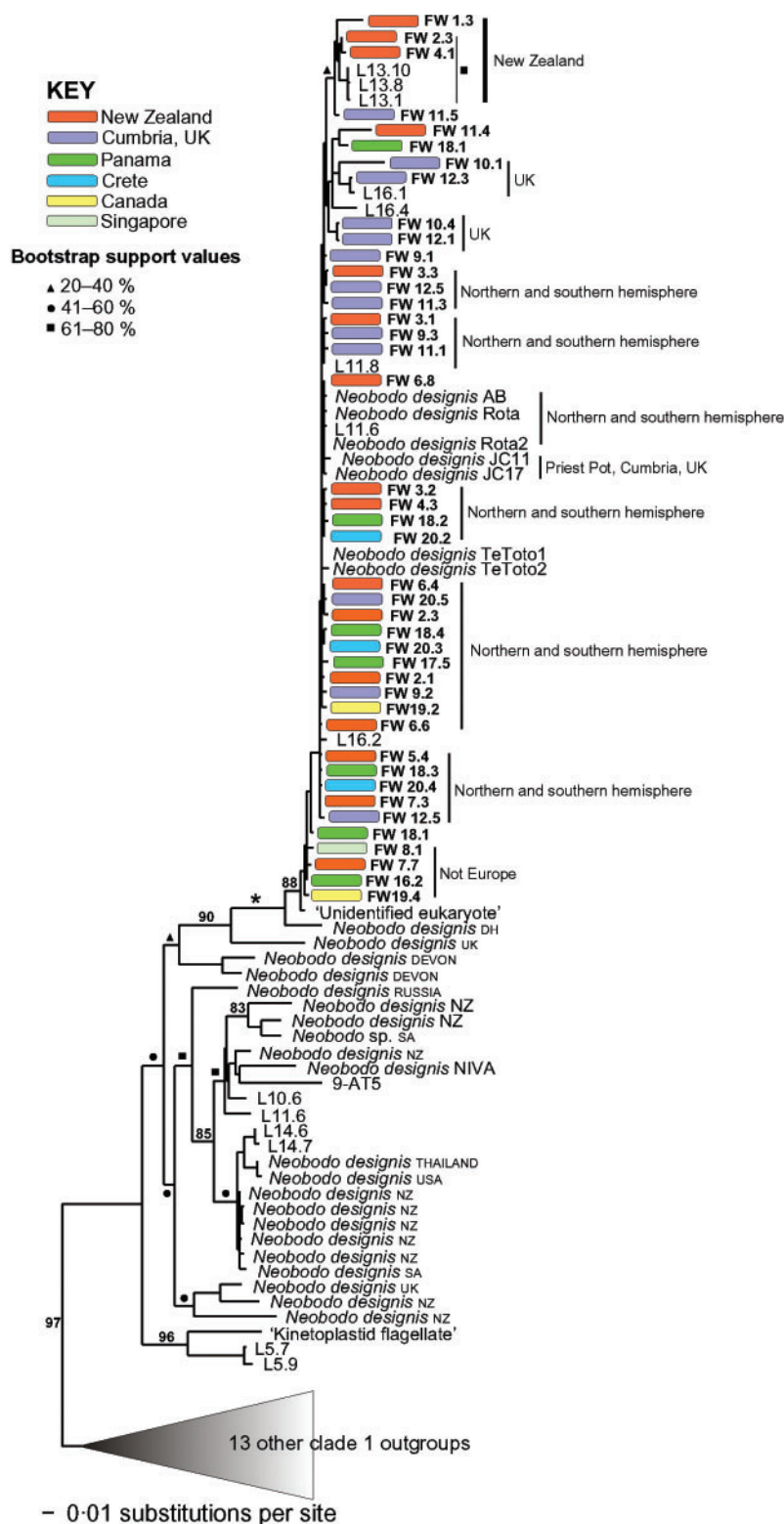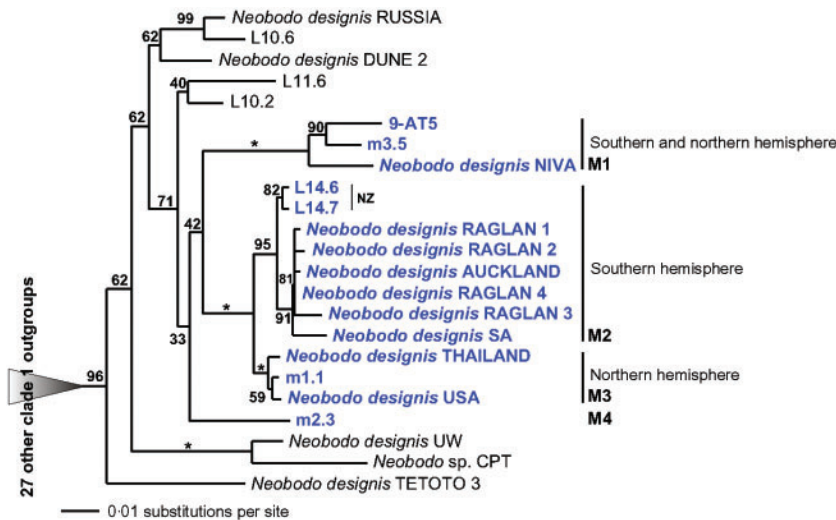
Fig. 7. Distance tree of 100 18S rRNA gene sequences of clade 1 neobodonids based on 546 positions (BioNJ; GTR+Γ+I model; a=0·743192, i=0·302549). The individual environmental sequences are colour-coded to describe their geographical origin, as shown in the key. Bootstrap support values from 1000 BioNJ replicates >20% are shown; symbols denoting bootstrap support values are explained in the key.

widespread. We never identified the same sequence twice from different libraries, suggesting that the total number of strains is very large.

Clade 1A (*Rhynchobodo*, *Klosteria*) contains 28% of environmental sequences, all of which group within the

Rhynchobodo lineage, indicating that members of this genus are also fairly common in the environment. Clade 1B (*Dimastigella/Rhynchomonas*) contains 12% of sequences found from environmental samples, all of which fall in the *R. nasuta* clade. Clade 1C (*Neobodo/Cruzella*) contains 52% of all environmental sequences, indicating

Fig. 8. Distance tree of 50 18S rRNA gene sequences of marine neobodonids and outgroups based on 728 positions (minimum evolution; GTR+Γ+I model; $a=0.591899$, $i=0.406543$). Bootstrap replicates are from 250 heuristic resamplings; supports >40% are shown. The marine clade is coloured blue; clades M1–M4 and their geographical origins are shown.

that representatives of this clade are the most commonly occurring of all kinetoplastids and that they are found in a
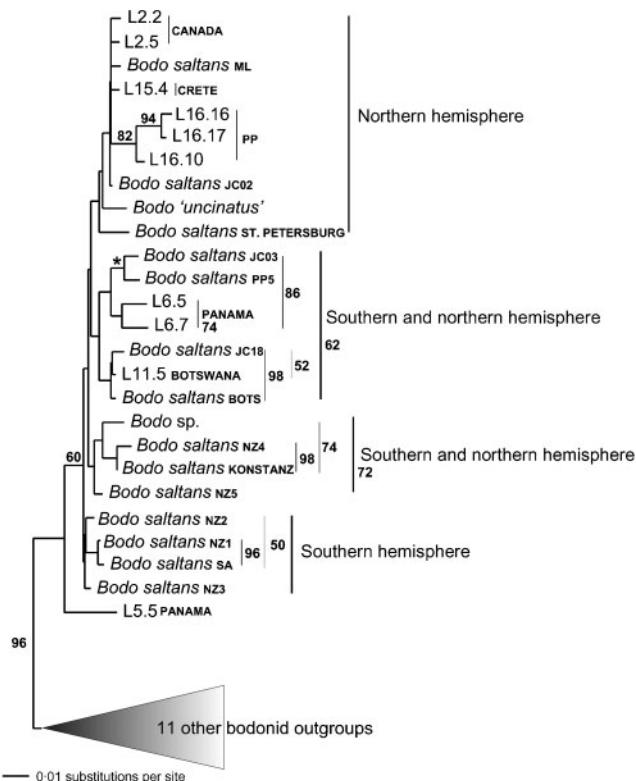
Fig. 9. ML tree of 37 18S rRNA gene sequences comprising clade 3 (*B. saltans* species complex) using 1745 positions (GTR+Γ+I model; $a=0.533397$, $i=0.374946$, $-lnL=9907.62873$). The tree is rooted on 11 other bodonid outgroups. Bootstrap support percentages from 75 replicates >46% are given; an asterisk denotes 100% support. '*Bodo* sp.' and '*B. uncinatus*' are probably *B. saltans* (Moreira *et al.*, 2004; von der Heyden *et al.*, 2004b). Strain JC02 is now CCAP 1907/6.

wide variety of environmental conditions. Clades 1D (novel bodonid '1D'; including *Cryptaulaxoides*-like) and 1E (novel bodonid lineage) contain no and two (8%) environmental sequences, respectively, suggesting that they are far less common in number and/or abundance than other clades. However, our sequences of seven new cultures (six of which are sister to a clade containing *Cryptaulaxoides*-like) of clade 1D indicate for the first time that this clade is comparable in phylogenetic depth to the major *Neobodo* clade and, like it, has marine and freshwater subclades (Fig. 4).

We found no environmental sequences for clade 2 (parabodonids), which suggests that free-living members are quite rare, at least in our samples. Clade 3 (eubodonids) contains 26% of all environmental sequences; most of these fall into well-defined subclades within the *B. saltans* species complex. Sequence L5.5 from lake sediment in Panama (Table 2) is sister to all others (Figs 3 and 7) and need not necessarily have the same morphology. As the *B. saltans* species complex is studied in more detail, it will probably need to be divided into several species (Callahan *et al.*, 2002; von der Heyden *et al.*, 2004b) and identifying the Panama lineage might provide valuable clues as to the ancestry of the *B. saltans* species complex.

A previous study investigated kinetoplastid diversity in a Danish lake sediment by using PCR primers specific for the 24S α-subunit of rRNA (Rasmussen *et al.*, 2001) and found a high diversity (27 different clones of a total of 43) of kinetoplastid sequences. They found that the majority of their sequences grouped within the free-living bodonids, but also that some grouped as sister to the parasitic trypanosomatids, leading to the conclusion that there might be free-living, parasitic kinetoplastids in Danish lake sediments. However, their phylogenetic trees had no representative sequence of *B. saltans*, used one very long-branch outgroup (*Euglena gracilis*) and no gamma correction. It is therefore very likely that their 'free-living parasitic Kinetoplastida' may simply be *B. saltans*, which sometimes

groups as sister to the trypanosomatids [see also Simpson *et al.* (2002)]. It was not made clear whether their 27 different clones were all from one library or from several and, therefore, whether this number of clones per library is significantly higher than we found in our 16 libraries (only one to six different genotypes were recovered from each library; see Table 2). It is also possible that their standard PCR methods may have increased chimaera formation and therefore presented a deceptively high diversity, a problem that has been underestimated in studies of eukaryotic diversity to date (Berney *et al.*, 2004; Cavalier-Smith, 2004).

Freshwater and marine diversification in *N. designis* and other bodonids

Fig. 4. shows that morphotypes of *N. designis* fall into exclusively marine or freshwater clades; there are seven to nine freshwater clades (two being from environmental DNA only, so *N. designis* morphology was not confirmed) and one major marine clade, plus two further small, but clearly distinct, marine clades (one containing *Cruzella*) that mainly include environmental sequences of unknown morphology (but might, when characterized, prove to be like *N. designis*). The tree topology of clade 1C and the distribution of marine and freshwater strains across it suggest strongly that the ancestral *N. designis* morphotype was originally a freshwater organism and that there were only three effective colonizations of the sea: one by the major marine *N. designis* clade, one by the clade containing *Cruzella* and one by the unique lineage simply labelled 'kinetoplastid flagellate'. This suggests that the vast majority of strains of *N. designis* at least are not tolerant to a wide variety of environmental conditions, contrary to previous suggestions (Arndt *et al.*, 2000; Atkins *et al.*, 2000), but that each strain is restricted to either freshwater or marine environments. As clade C largely comprises *N. designis* morphotypes, it is clear that *Cruzella* evolved from such an ancestor. A separate Bayesian analysis of the Fig. 4 dataset showed exactly the same overall topology, except for clades Nd3 and Nd4 being weak sisters [posterior probability (PP)=0·72] and clades 1D and 1F being sisters, even more weakly (PP=0·5). On that Bayesian tree (not shown), holophyly of the major marine clade was supported moderately well (PP=0·86). Holophyly of clades Nd1, Nd2 and Nd3, and the sister relationship of Nd1 and Nd2, were supported strongly (PP=1·0). In addition, two nodes within the basal freshwater part of Nd2 received PP support of 1·0 and another received support of 0·99. Thus, the Bayesian analysis also fully supported the derived nature of all marine *Neobodo* clades.

It is important to make a crude estimate of the age of clade C but, unfortunately, there is almost no fossil record for kinetoplastids or other excavate protozoa. However, from molecular tree topology and proportions, it would appear that excavates in general and Euglenozoa in particular arose early in the evolution of bikont eukaryotes. Assuming that the primary eukaryotic divergence was indeed between unikonts and bikonts (Stechmann & Cavalier-Smith, 2003),

it is probable that bikonts are at least 800 million years old, as eukaryotes are at least that old (Cavalier-Smith, 2002; Porter & Knoll, 2000). Given the difficulty of resolving the branching order of the three main euglenozoan groups, it is likely that kinetoplastids diverged from euglenoids and diplonemids at fairly similar times early in the diversification of the Euglenozoa. It would therefore be surprising if the separation of Kinetoplastea was much younger than about 600 million years. As the basal kinetoplastid is significantly later than the basal node for Euglenozoa, applying a crude molecular clock to Fig. 1 in the study by von der Heyden *et al.* (2004b) yields a rough estimate of 360 million years for the age of the kinetoplastid cenancestor. As the depth of clade C of Fig. 3 and in Fig. 4 of von der Heyden *et al.* (2004b) is roughly one-tenth that of the whole kinetoplastid clade, applying a crude clock gives a rough estimate of 36 million years for the *Neobodo* morphotype. Another way of making an estimate would be to assume that the early Cretaceous *Paleoleishmania* in amber (Poinar & Poinar, 2004) really refers to a single species and is also identified correctly as a trypanosomatid, which it may be, but need not be: how can one really rule out a mixture of other symbionts? As the *Neobodo* clade is internally markedly deeper on all our rRNA trees than the trypanosomatid clade, this would imply that *Neobodo* is older than 100 million years (the minimum age of the amber fossils) if this molecule were clock-like. However, if trypanosomatids are sisters of core bodonids, as in Fig. 4 of von der Heyden *et al.* (2004b), it would appear from that tree that 18S rRNA has been evolving about twice as fast in core bodonids as in trypanosomatids since they diverged from their common ancestor, so a minimum figure of approximately 50 million years would be a more reasonable estimate for the age of *Neobodo*. Although it is hazardous to estimate ages from molecular trees, as the rRNA gene is an appallingly unreliable 'clock' (Cavalier-Smith, 2002, 2004), it is unlikely that an estimate of 35–50 million years old would be in error by more than two- or threefold. Thus, for tens of millions of years, the ancestrally freshwater *N. designis* complex probably only colonized the sea successfully about three times, although there may be other, as yet undiscovered marine *Neobodo* lineages. From Fig. 4, the depth of the major marine clade appears to be at least one-quarter that of the *Neobodo* clade, suggesting that this major invasion of the marine habitat took place over 9 million years ago and no others have occurred since. The depth of the *B. saltans* clade is only about one-third that of *Neobodo* but, as the overall branch lengths of the corresponding part of the tree are similarly different, this probably reflects differences in the rate of rRNA evolution, rather than a significantly different age. Yet *B. saltans* is essentially always freshwater [although we isolated one strain from decaying seaweed on a beach (von der Heyden *et al.*, 2004b) and *B. saltans* was reported by Vørs (1992) in the Gulf of Finland].

There is other evidence supporting our conclusion that effective interchange between freshwater and marine

habitats is very rare in bodonids. Consider the culture experiments of Mylnikov & Gräfenhan (reviewed by Arndt *et al.*, 2000) and Ekelund (2002). The former transferred marine and freshwater *B. saltans* and *N. designis* and other protists gradually into less/more-saline conditions and found that they adapted to and grew in these. Both *B. saltans* and *N. designis* were transferred successfully into the other types of medium. Other bodonids (*R. nasuta*, *Parabodo nitrophilus*, *Parabodo caudatus*) were able to withstand some, but not a total, change in salinity. Ekelund (2002) carried out similar experiments using 11 soil flagellates (including *R. nasuta*, *P. caudatus*, *B. saltans* and *N. designis*) and showed that none of the flagellates thrived in 15 parts per thousand (p.p.t.) NaCl. In contrast to Mylnikov & Gräfenhan (reviewed by Arndt *et al.*, 2000), whose cultures may have been genetically mixed, Ekelund (2002) used clonal cultures and found that freshwater *N. designis* only tolerated an increased salinity up to 3 p.p.t., even when transferred gradually into higher-salinity media. Our demonstration of the great genetic diversity of morphologically indistinguishable *Neobodo* and *B. saltans* shows that physiological experiments on non-clonal cultures of bodonids (or any other protist) are ecologically meaningless. Experiments where flagellates are transferred gradually into higher/lower-salinity media also may not realistically represent many environmental situations where microbes are naturally dispersed among habitats. There are very few cases in which a small pool of sea water reduces salinity gradually every few days, thereby allowing its inhabitants to adapt slowly to the changing conditions, and then becomes freshwater. Although rivers perpetually discharging into the sea will often put freshwater microbes into gradually increasing salinity, microbes dispersed by wind, precipitation and/or animal vectors will often land in either an environment that they are capable of living in or one in which they will probably perish and not be able to compete with strains that have adapted genetically over millions of years to that habitat.

The freshwater/marine divergence is particularly prominent among strains of *N. designis*, probably because we have sufficient samples to demonstrate this. Although other bodonids do not show such clear-cut divisions between freshwater and marine lineages, the deep clade 1D that is sister to *Cryptaulaxoides*-like, as well as the novel bodonid lineage 1E, each have distinct marine and freshwater subclades (Fig. 4). Within *Rhynchobodo/Klosteria*, marine and freshwater/soil strains group separately and there is a well-supported clade comprising strains from soil and tree moss (Fig. 3). However, within the genera *Rhynchomonas*, *Parabodo* and *Procryptobia*, marine and freshwater samples intermingle freely with each other; within *Rhynchomonas*, a strain (BSZ1) from an Atlantic deep-sea hydrothermal vent even groups with an environmental sequence (L6.8) from the Panamanian rainforest, albeit rather weakly! These apparently mixed groupings may, however, be artefacts of undersampling (similar to what appeared to be true for *Neobodo* when we only had a few strains), hiding ecological

specialization; much more work is needed to explore the diversity and distribution of all these clades.

We showed previously that, in the cryptomonad genus *Goniomonas*, there is an even more fundamental divergence between marine and freshwater strains (von der Heyden *et al.*, 2004a) and that there are also distinct marine and freshwater clades within centrohelid heliozoa (von der Heyden, 2004). Lee & Patterson (1998), using cluster analysis on diverse protistan community assemblages, showed that communities that are geographically close together do not cluster together, but that that those from similar habitats do, and that there is little overlap in community assemblage between marine and freshwater sites. This further strengthens the hypothesis that not all protists are able to move freely between these types of habitats, but are instead restricted to a physiological/ecological range of variables.

Our results also show that strains of *N. designis* that diverged tens of millions of years ago can coexist at the same site: we isolated several strains of *N. designis* from the same geographical area at the same time (four strains from sediment in Priest Pot, Cumbria, UK, and TeToto Gorge, New Zealand; marked on Fig. 4 with one and two dots, respectively) that actually group in very different positions on the tree. It is an open question whether this coexistence between very distinct strains reflects differences in ecological adaptations among the strains; however, given their very deep divergence on the rRNA tree, it would be surprising if they do not also differ in hundreds of protein-coding genes and we suspect that at least some such differences are likely to be adaptive. We also suspect that thin-section ultrastructural studies might reveal morphological differences among the strains. Even more detailed light microscopy would probably reveal at least some quantitative differences; previously, we noted that one freshwater strain, *Neobodo* sp. Cape Town (CPT; von der Heyden *et al.*, 2004b), appeared to be morphologically distinct from typical *N. designis* in having a longer anterior cilium and an anteriorly more pointed, squatter and smaller body. However, it died before we were able to take measurements or record these differences photographically. Although some bodonids [*B. saltans* (Brooker, 1971), *Cephalothamnium* (Vickerman, 1990) and *Rhynchomonas* (Swale, 1973)] have ciliary hairs (mastigonemes), we found none in either *Neobodo* strain, confirming the observations of Eyden (1977). If the classification of *Cephalothamnium* is correct (Vickerman, 1990), at least one member of each of the three bodonid suborders would have ciliary hairs; possibly, therefore, their absence in *N. designis* is a secondary loss.

Huge genetic variation and 'species' number of bodonids

The degree of genetic variation within many bodonid clades is comparatively greater than that of whole orders of higher eukaryotes. Within *N. designis*, no 18S rRNA gene sequences yet found are identical; this degree of variation in the 18S

rRNA gene of higher eukaryotes would correspond to separate species. Throughout the protist world, morphospecies are now being found to comprise many cryptic species complexes. Cryptic species, such as the majority of bodonids, have also been reported within Cercozoa (Kühn *et al.*, 2000), Foraminifera (de Vargas *et al.*, 1999; Kucera & Darling, 2002; Pawlowski & Holzmann, 2002), the spumellarid radiozoan *Thalassicolla nucleata* (Zettler *et al.*, 1998), ciliates (Nanney *et al.*, 1989, 1998; Fokin *et al.*, 2004), dinoflagellates (John *et al.*, 2003) and certain types of picoplankton (Potter *et al.*, 1997). The almost-universal presence of such cryptic variation makes the use of the morphospecies in discussions of protist biodiversity and dispersal (Finlay, 2002; Finlay *et al.*, 1996) of limited value as, in bodonids and many other protists, there is no reason to think that named morphospecies have much real biological significance. The conventional 'species' boundaries may largely reflect our ignorance and current inability to subdivide them into more biologically meaningful units.

Cryptic speciation in all lineages and the relatively long evolutionary history of free-living kinetoplastids have produced a much larger number of species than realized previously. Although there are descriptions of over 200 nominal species, many have not been based on pure cultures and are practically unrecognizable; there may be different stages of the life cycle of single species and, hence, the actual number of morphotypes may have been overestimated (Zhukov, 1991). Recent protozoologists can confidently recognize only about 15 'species' within 11 genera. For example, the three 'species' of *Rhynchomonas* [*R. nasuta*, *Rhynchomonas mutabilis* and *Rhynchomonas metabolita*, reviewed by Zhukov (1971)] are currently all attributed to only one species, *R. nasuta*, as the only real differences between them were in cell size and 'it is possible that there are physiological or ecological species within the morphological species' (Zhukov, 1991). The molecular evidence presented here indicates that many of the currently recognizable morphotypes may contain hundreds of distinct lineages; much further work is needed to determine whether each of them represents a different evolutionary peak adapted to particular habitat features, be these the type of prey utilized, the amount of moisture in the surroundings, sediment/soil-grain size or coarseness or other environmental variables. Recent experiments by Smirnov & Thar (2003) studied the microspatial distribution of amoebae and confirmed the presence of microhabitats, which were occupied selectively by amoebic species; similar ecological work is needed for the bodonid cryptic 'species'. In contrast to ciliates, where breeding experiments have shown that there are numerous cryptic biological species within a common morphotype (Fokin *et al.*, 2004; Nanney *et al.*, 1989, 1998), it is not known whether most bodonids are sexual or asexual, so we have no idea whether the biological species concept applies to them. Nanney *et al.* (1989, 1998) suggested that 'genetically divergent tetrahymenids are morphologically identical because of stabilizing selection that preserves a morphological phenotype which probably represents some adaptive peak'. We predict that future studies will reveal important ecological and physiological differences among the lineages comprising rather ancient morphospecies, such as *B. saltans*, *N. designis* and *R. nasuta*, and that, one day, they will be divided into numerous 'species'. The fact that they are currently morphologically indistinguishable does not mean that they are really morphologically identical. Detailed electron microscopy or sophisticated light microscopy with special staining methods might possibly reveal significant morphological differences.

Are bodonids really all cosmopolitan?

von der Heyden *et al.* (2004b) showed some evidence for a restricted biogeographical distribution of strains from the northern and southern hemispheres in clade 3 (the *B. saltans* species complex). Addition of 10 environmental sequences confirms that the four clades recovered previously are still present and fall into northern-hemisphere, southern-hemisphere and two clades of a mixture of hemispheres (Fig. 9), suggesting that some strains of *B. saltans* may have a restricted biogeographical distribution and are not ubiquitously distributed globally. This conclusion must still be tentative because much more sampling is needed for a decisive test. However, in certain exceptionally well-sampled morphospecies of pelagic Foraminifera, the study of the geographical distribution of genotypes with distinctive rRNA genes shows that some are not cosmopolitan, as was assumed incorrectly when only morphospecies were studied (Darling *et al.*, 2004). In the very well-sampled *Paramecium aurelia* species complex, some biological species are cosmopolitan, but others clearly are not (Fokin *et al.*, 2004). In the *Alexandrium tamarense* complex, there also appear to be geographically restricted clades (John *et al.*, 2003). Thus, allopatric speciation in protists may not be as rare as sometimes claimed (Finlay & Fenchel, 2004). As Patterson & Lee (2000) and Coleman (2002) pointed out and Finlay & Fenchel (1999, 2004) also recognized, if morphospecies are seriously undersplit taxonomically, this can inflate the appearance of cosmopolitanism substantially. Our genetic studies of bodonids indicate strongly that many morphospecies have been grossly undersplit taxonomically and show that fine genetic discrimination, combined with even more extensive sampling, must be a central feature of future biogeographical studies of zooflagellates. We suspect that such studies will show that at least some bodonid genotypes are indeed cosmopolitan, as predicted by Finlay & Fenchel (1999, 2004) and Finlay *et al.* (2004), but will reveal others with lower dispersal/colonization potential that are not.

Clade-specific analyses of freshwater and marine neobodonids

The genetic diversity found for a part of freshwater *N. designis* subclade 1 by using specific PCR primers (marked by a box on Fig. 4) was somewhat greater than was retrieved for primers that sampled across a whole group [see Bass &

Cavalier-Smith (2004) for discussion of this]. We found that, even for very closely related sequences (i.e. FW3.3/FW12.5 or FW18.4/FW20.3), there is at least a difference of 2 nt in this highly variable region of the 18S rRNA gene sequence and that, over the whole length of the gene, this difference is probably even greater. This suggests that there are probably hundreds of individual lineages, even within this very narrow subclade of *N. designis*. Clearly, much more sampling of all bodonid lineages (and probably all protozoan lineages) is required in order to ascertain their true global diversity, as we did not retrieve the same sequence from different parts of the globe. Although we did not find any confirming instance of cosmopolitanism for even one genotype, there is little evidence of endemicity or restricted biogeography within individual clades. We found one clade containing sequences exclusively from New Zealand, two exclusively from the UK and one exclusively from Priest Pot (Cumbria, UK); all other clusters of sequences are from very different geographical areas (Fig. 7).

Within the marine clade are four subclades (Fig. 8; labelled M1–M4); two of these (M2 and M3) are composed entirely of sequences from either the northern or southern hemisphere; M4 contains only one environmental sequence from Panama and M1 contains a sequence from a deep-sea hydrothermal vent (9-AT5), an environmental sequence from New Zealand and an *N. designis* isolate from Denmark (Fig. 8). Thus, *N. designis*, like *B. saltans*, shows several clades that are suggestive of restricted biogeographical distribution. These would merit more intensive sampling to determine whether this *prima facie* evidence of endemicity is genuine, as similar evidence appears to be for some foraminifera, or simply a consequence of the presently unavoidable undersampling of their immense genotypic diversity.

ACKNOWLEDGEMENTS

REFERENCES

Al-Qassab, S., Lee, W. J., Murray, S., Simpson, A. G. B. & Patterson, D. J. (2002). Flagellates from stromatolites and surrounding sediments in Shark Bay, Western Australia. *Acta Protozool* 41, 91–144.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403–410.

Arndt, H., Dietrich, D., Auer, B., Cleven, E.-J., Gräfenhan, T., Weitere, M. & Mylnikov, A. P. (2000). Functional diversity of heterotrophic flagellates in aquatic ecosystems. In *The Flagellates: Unity, Diversity and Evolution*, pp. 240–268. Edited by B. S. C. Leadbeater & J. C. Green. London: Taylor & Francis.

Atkins, M. S., Anderson, O. R. & Wirsen, C. O. (1998). Effect of hydrostatic pressure on the growth rates and encystment of flagellated protozoa isolated from a deep-sea hydrothermal vent and a deep shelf region. *Mar Ecol Prog Ser* 171, 85–95.

Atkins, M. S., Teske, A. P. & Anderson, O. R. (2000). A survey of flagellate diversity at four deep-sea hydrothermal vents in the Eastern Pacific Ocean using structural and molecular approaches. *J Eukaryot Microbiol* 47, 400–411.

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

Berney, C., Fahrni, J. & Pawlowski, J. (2004). How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC Biol* 2, 13. http://www.biomedcentral.com/1741-7007/2/13

Brooker, B. E. (1971). Fine structure of *Bodo saltans* and *Bodo caudatus* (Zoomastigophora, Protozoa) and their affinities with the Trypanosomatidae. *Bull Brit Mus Nat Hist* 22, 89–102.

Callahan, H. A., Litaker, R. W. & Noga, E. J. (2002). Molecular taxonomy of the suborder Bodonina (order Kinetoplastida), including the important fish parasite, *Ichthyobodo necator*. *J Eukaryot Microbiol* 49, 119–128.

Cavalier-Smith, T. (1981). Eukaryote kingdoms: seven or nine? *Biosystems* 14, 461–481.

Cavalier-Smith, T. (2002). The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* 52, 7–76.

Cavalier-Smith, T. (2003a). Protist phylogeny and the high-level classification of Protozoa. *Eur J Protistol* 39, 338–348.

Cavalier-Smith, T. (2003b). The excavate protozoan phyla Metamonada Grassé emend. (Anaeromonadea, Parabasalia, Carpediemonas, Eopharyngia) and Loukozoa emend. (Jakobea, Malawimonas): their evolutionary affinities and new higher taxa. *Int J Syst Evol Microbiol* 53, 1741–1758.

Cavalier-Smith, T. (2004). Only six kingdoms of life. *Proc R Soc Lond B Biol Sci* 271, 1251–1262.

Coleman, A. W. (2002). Microbial eukaryote species. *Science* 297, 337.

Darling, K. F., Kucera, M., Pudsey, C. J. & Wade, C. M. (2004). Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics. *Proc Natl Acad Sci U S A* 101, 7657–7662.

de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W. & Pawlowski, J. (1999). Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. *Proc Natl Acad Sci U S A* 96, 2864–2868.

Doležel, D., Jirků, M., Maslov, D. A. & Lukeš, J. (2000). Phylogeny of the bodonid flagellates (Kinetoplastida) based on small-subunit rRNA gene sequences. *Int J Syst Evol Microbiol* 50, 1943–1951.

Ekelund, F. (2002). Tolerance of soil flagellates to increased NaCl levels. *J Eukaryot Microbiol* 49, 324–328.

Eyden, B. P. (1977). Morphology and ultrastructure of *Bodo designis* Skuja 1948. *Protistologica* 13, 169–179.

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science* 296, 1061–1063.

Finlay, B. J. & Fenchel, T. (1999). Divergent perspectives on protist species richness. *Protist* **150**, 229–233.

Finlay, B. J. & Fenchel, T. (2004). Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist* **155**, 237–244.

Finlay, B. J., Corliss, J. O., Esteban, G. & Fenchel, T. (1996). Biodiversity at the microbial level: the number of free-living ciliates in the biosphere. *Q Rev Biol* **71**, 221–237.

Finlay, B. J., Esteban, G. F. & Fenchel, T. (2004). Protist diversity is different? *Protist* **155**, 15–22.

Fokin, S. I., Przyboś, E., Chivilev, S. M., Beier, C. L., Horn, M., Skotarczak, B., Wodecka, B. & Fujishima, M. (2004). Morphological and molecular investigations of *Paramecium schewiakoffi* sp. nov. (Ciliophora, Oligohymenophorea) and current status of distribution and taxonomy of *Paramecium* spp. *Eur J Protistol* **40**, 225–243.

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.

Huber, T., Faulkner, G. & Hugenholtz, P. (2004). BELLEROPHON; a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**, 2317–2319.

John, U., Fensome, R. A. & Medlin, L. K. (2003). The application of a molecular clock based on molecular sequences and the fossil record to explain biogeographic distributions within the *Alexandrium tamarense* "species complex" (Dinophyceae). *Mol Biol Evol* **20**, 1015–1027.

Kucera, M. & Darling, K. F. (2002). Cryptic species of planktonic foraminifera: their effect on palaeoceanographic reconstructions. *Philos Trans R Soc Lond A* **360**, 695–718.

Kühn, S., Lange, M. & Medlin, L. K. (2000). Phylogenetic position of *Cryothecomonas* inferred from nuclear-encoded small subunit ribosomal RNA. *Protist* **151**, 337–345.

Larsen, J. & Patterson, D. J. (1990). Some flagellates (Protista) from tropical marine sediments. *J Nat Hist* **24**, 801–937.

Lee, W. J. & Patterson, D. J. (1998). Diversity and geographic distribution of free-living heterotrophic flagellates – analysis by PRIMER. *Protist* **149**, 229–244.

López-García, P., Philippe, H., Gail, F. & Moreira, D. (2003). Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Natl Acad Sci U S A* **100**, 697–702.

Moon-van der Staay, S. Y., De Wachter, R. & Vaulot, D. (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**, 607–610.

Moreira, D. & López-García, P. (2002). The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol* **10**, 31–38.

Moreira, D. & López-García, P. (2003). Are hydrothermal vents oases for parasitic protists? *Trends Parasitol* **19**, 556–558.

Moreira, D., López-García, P. & Vickerman, K. (2004). An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. *Int J Syst Evol Microbiol* **54**, 1861–1875.

Nanney, D. L., Meyer, E. B., Simon, E. M. & Preparata, R. M. (1989). Comparison of ribosomal and isozymic phylogenies of tetrahymenine ciliates. *J Protozool* **36**, 1–8.

Nanney, D. L., Park, C., Preparata, R. & Simon, E. M. (1998). Comparison of sequence differences in a variable 23S rRNA domain among sets of cryptic species of ciliated protozoa. *J Eukaryot Microbiol* **45**, 91–100.

Patterson, D. J. & Lee, W. J. (2000). Geographic distribution and diversity of free-living heterotrophic flagellates. In *The Flagellates: Unity, Diversity and Evolution*, pp. 269–287. Edited by B. S. C. Leadbeater & J. C. Green. London: Taylor & Francis.

Patterson, D. J. & Simpson, A. G. B. (1996). Heterotrophic flagellates from coastal marine and hypersaline sediments in Western Australia. *Eur J Protistol* **32**, 423–448.

Pawlowski, J. & Holzmann, M. (2002). Molecular phylogeny of foraminifera: a review. *Eur J Protistol* **38**, 1–10.

Poinar, G., Jr & Poinar, R. (2004). *Paleoleishmania proterus* n. gen., n. sp., (Trypanosomatidae: Kinetoplastida) from Cretaceous Burmese amber. *Protist* **155**, 305–310.

Porter, S. M. & Knoll, A. H. (2000). Testate amoebae in the Neoproterozoic Era: evidence from vase-shaped microfossils in the Chuar Group, Grand Canyon. *Paleobiology* **26**, 360–385.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **9**, 817–818.

Potter, D., Lajeunesse, T. C., Saunders, G. W. & Anderson, R. A. (1997). Convergent evolution masks extensive biodiversity among marine coccoid picoplankton. *Biodivers Conserv* **6**, 99–107.

Rasmussen, L. D., Ekelund, F., Hansen, L. H., Sørensen, S. J. & Johnsen, K. (2001). Group-specific PCR primers to amplify 24S α-subunit rRNA genes from Kinetoplastida (Protozoa) used in denaturing gradient gel electrophoresis. *Microb Ecol* **42**, 109–115.

Robison-Cox, J. F., Bateson, M. M. & Ward, D. M. (1995). Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. *Appl Environ Microbiol* **61**, 1240–1245.

Simpson, A. G. B., Lukeš, J. & Roger, A. J. (2002). The evolutionary history of kinetoplastids and their kinetoplasts. *Mol Biol Evol* **19**, 2071–2083.

Skuja, H. (1948). Taxonomie des Phytoplanktons einiger Seen in Uppland, Schweden. *Symb Bot Upsal* **9**, 1–399 (in German).

Smirnov, A. & Thar, R. (2003). Spatial distribution of Gymnamoebae (Rhizopoda, Lobosea) in brackish-water sediments at the scale of centimeters and millimeters. *Protist* **154**, 359–369.

Stechmann, A. & Cavalier-Smith, T. (2003). The root of the eukaryote tree pinpointed. *Curr Biol* **13**, R665–R666.

Swale, E. M. F. (1973). A study of the colourless flagellate *Rhynchomonas nasuta* (Stokes) Klebs. *Biol J Linn Soc* **5**, 255–264.

Swofford, D. L. (2001). PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods), v. 4.0b6. Sunderland, MA: Sinauer Associates.

UKNCC (2001). *Catalogue of the UK National Culture Collection – List of Algae and Protozoa*. Oban, UK: United Kingdom National Culture Collection.

van Hannen, E. J., Mooij, W., van Agterveld, M. P., Gons, H. J. & Laanbroek, H. J. (1999). Detritus-dependent development of the microbial community in an experimental system: qualitative analysis by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **65**, 2478–2484.

Vickerman, K. (1990). Phylum Zoomastigina, class Kinetoplastida. In *Handbook of Protoctista*, pp. 215–238. Edited by L. Margulis, J. O. Corliss, M. Melkonian & D. J. Chapman. Boston: Jones & Bartlett.

von der Heyden, S. (2004). *Testing ubiquitous dispersal and freshwater/marine divergence within free-living protist groups*. DPhil thesis, University of Oxford.

von der Heyden, S., Chao, E. E. & Cavalier-Smith, T. (2004a). Genetic diversity of goniomonads: an ancient divergence between marine and freshwater species. *Eur J Phycol* **39**, 343–350.

von der Heyden, S., Chao, E. E., Vickerman, K. & Cavalier-Smith, T. (2004b). Ribosomal RNA phylogeny of bodonid and diplonemid flagellates and the evolution of Euglenozoa. *J Eukaryot Microbiol* **51**, 402–416.

Vørs, N. (1992). Heterotrophic amoebae, flagellates and heliozoa from the Tvärminne area, Gulf of Finland, in 1988–1990. *Ophelia* **36**, 1–109.

Vørs, N. (1993). Marine heterotrophic amoebae, flagellates and heliozoa from Belize (Central America) and Tenerife (Canary Islands), with descriptions of new species, *Luffisphaera bulbochaete* N. Sp., *L. longihastis* N. Sp., *L. turriformis* N. Sp. and *Paulinella intermedia* N. Sp. *J Eukaryot Microbiol* **40**, 272–287.

Zettler, L. A. A., Anderson, O. R. & Caron, D. A. (1998). Insights on the diversity within a ''species'' of *Thalassicolla* (Spumellarida) based on 16S-like ribosomal RNA gene sequencing. *J Eukaryot Microbiol* **45**, 488–496.

Zhukov, B. F. (1971). A key to colourless free-living flagellates of the suborder Bodonina Hollande. In *Biology and Productivity of Freshwater Organisms*, pp. 241–283. Edited by M. M. Kamshilov. Leningrad: Science Press (in Russian).

Zhukov, B. F. (1991). The diversity of bodonids. In *The Biology of Free-living Heterotrophic Flagellates*, pp. 177–184. Edited by D. J. Patterson & J. Larsen. Oxford: Clarendon Press.