Published in Quaternary Science Reviews, Vol. 91, 2014, p. 62-69
which should be used for any reference to this work

Effect of taxonomic resolution on ecological and palaeoecological inference – a test using testate amoeba water table depth transfer functions

Edward A.D. Mitchell [a,*], Mariusz Lamentowicz [b], Richard J. Payne [c], Yuri Mazei [d]

[a] *Laboratory of Soil Biology, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2000 Neuchâtel, Switzerland*
[b] *Department of Biogeography and Palaeoecology, Faculty of Geosciences, Adam Mickiewicz University, Dzięgielowa 27, 61-680 Poznań, Poland*
[c] *Biological and Environmental Science, University of Stirling, Stirling FK94LA, Scotland, UK*
[d] *Department of Zoology and Ecology, Penza State University, Krasnaya str., 40, Penza 440026, Russia*

ABSTRACT

Keywords:
Palaeoecology
Bioindication
Transfer function
Taxonomy
Testate amoebae
Quantitative ecology
Peatland

Sound taxonomy is a major requirement for quantitative environmental reconstruction using biological data. Transfer function performance should theoretically be expected to decrease with reduced taxonomic resolution. However for many groups of organisms taxonomy is imperfect and species level identification not always possible.

We conducted numerical experiments on five testate amoeba water table (DWT) transfer function data sets. We sequentially reduced the number of taxonomic groups by successively merging morphologically similar species and removing inconspicuous species. We then assessed how these changes affected model performance and palaeoenvironmental reconstruction using two fossil data sets.

Model performance decreased with decreasing taxonomic resolution, but this had only limited effects on patterns of inferred DWT, at least to detect major dry/wet shifts. Higher-resolution taxonomy may however still be useful to detect more subtle changes, or for reconstructed shifts to be significant.

1. Introduction

Methodological development, testing and improvement are key to robust palaeoecological reconstruction and are recognised priorities for future research (Seddon et al., 2014). Sound taxonomy is one of the major requirements for quantitative reconstruction of environmental change based on biological data (Birks, 2003), and more generally for bioindication. The performance of transfer functions and bioindicator approaches should thus be expected to decrease if taxonomic resolution is reduced (Nahmani et al., 2006; Heiri and Lotter, 2010). However for many groups of organisms identification to a high taxonomic resolution may not be possible or practical. Some organisms used in palaeoecology show a high degree of intra-specific variability, in others diagnostic features are not preserved, or the criteria for separating species are insufficiently clear to have confidence in species-level identification (Payne et al., 2011). On top of this intrinsic uncertainty considerable palaeoecological data generation is conducted by researchers with limited taxonomic experience (e.g. postgraduate students) for whom high taxonomic resolution identification may be difficult to achieve. A rational response to this uncertainty is therefore for taxa to be identified to a lower taxonomic resolution, either based on formal taxonomic units (e.g. genus or family, as common in pollen analysis) or more informal 'types' based on morphologically-similar taxa (as common in testate amoeba analysis). This situation raises questions about the implications of variability in taxonomic resolution for the quality of data produced, and particularly for the robustness of transfer function results.

Here we use the example of testate amoeba analysis to address the consequences of variability in taxonomic resolution for quantitative inference. Testate amoebae are commonly used in quantitative palaeoecology, most frequently for palaeohydrological reconstruction in peatlands (Charman, 2001; Mitchell et al., 2008a). Several important methodological aspects of testate amoeba analysis in palaeoecology have only recently received attention (e.g. Booth, 2002; Mitchell et al., 2008b; Payne and Mitchell, 2009; Sullivan and Booth, 2011; Payne et al., 2012; Avel and Pensa, 2013). Taxonomic resolution is especially critical in the case of

---

* Corresponding author.
E-mail address: edward.mitchell@unine.ch (E.A.D. Mitchell).

testate amoebae because of the current state of confusion regarding the validity of many taxa. Recent molecular and ultra-structural studies are revealing a wealth of unknown diversity even within apparently well-established taxa (Heger et al., 2011, 2013; Lara et al., 2011; Kosakyan et al., 2013). Confusion exists in all types of testate amoebae including those building agglutinated shells (e.g. *Difflugia* (Mazei and Warren, 2012)), those recycling siliceous plates from their prey (e.g. *Nebela tincta* group (Kosakyan et al., 2013)), and those building their shell from self-secreted idiosomes (e.g. *Euglypha rotunda* group (Wylezich et al., 2002)). This confusion is compounded by the state of testate amoeba taxonomy, which lacks a modern synthesis (excepting Mazei and Tsyganov, 2006, which is only available in Russian) and where for some common genera the most comprehensive papers date from the 1920s (e.g. Deflandre, 1929). Taxonomic uncertainty has widespread implications for the ecology, palaeoecology and biogeography of testate amoebae and other protists (Mitchell and Meisterfeld, 2005; Mitchell et al., 2008a; Heger et al., 2009). As a result of this uncertainty an intermediate taxonomical resolution is usually applied in testate amoeba palaeoecology, with many taxa lumped into “types” (e.g. *Cyclopyxis arcelloides*-type includes *C. arcelloides*, *Cyclopyxis kahli*, *Difflugia globulus*, *Phryganella acropodia*, etc.) (Charman et al., 2000). However the degree to which such types are applied and how they are defined is variable, with unknown consequences for quality and consistency of results.

We conducted a series of numerical experiments on five modern data sets used to build transfer functions for water table depth inference. We gradually reduced the taxonomic resolution by merging morphologically similar species and removing some inconspicuous species that are likely to be overlooked by less-experienced analysts. The performance of transfer functions was then assessed using standard cross-validation metrics. We hypothesised that reducing the taxonomic resolution would cause a decline in model performance and that this decrease would be approximately proportional to the loss of taxonomic richness in the data set.

2. Methods

We used five training sets of peatland testate amoebae and water table depth from the Swiss Alps (Engadine) (Lamentowicz et al., 2010a) Poland (Lamentowicz and Mitchell, 2005; Lamentowicz et al., 2011), Alaska (USA) (Payne et al., 2006), the Jura Mountains of Switzerland and France (Mitchell et al., 1999, 2001) and Greece (Payne and Mitchell, 2007). Together these data sets cover the full fen to bog gradient. Taxonomic resolution in the original studies was generally average to high within the context of such datasets (40–62 taxa overall per data set).

In order to assess the effect of taxonomic resolution on transfer function model performance we developed several taxonomic aggregation scenarios ranging from maximum resolution (the actual resolution used during the microscopic counts) and a series of cases with decreasing taxonomic resolution in which morphologically similar groups of species were gradually merged and some taxa likely to be overlooked by inexperienced analysts were removed. To this aim we divided the taxa into five categories: A) unmistakable (e.g. *Archerella flavum*), B) limited possible confusion (*Assulina* spp.), C) moderately confusing (e.g. *Centropyxis platystoma*,

Table 1
Combined list of testate amoeba taxa from seven modern data sets with indication of identification certainty and pooling in successively broader taxonomic categories. See Supplementary Table 1 for further details.

Full name	Code	Degree of confusion[a]					Species pooling or deletion in the 6 cases and number of taxa retained					
		A	B	C	D	E	1 = original resolution	2 = pooling of most confusing taxa	3 = Charman & al book	4 = limited resolution	5 = genus level	6 = genus level or less
							100	66	63	34	29	16
Amphitrema stenostoma	Amph sten		x				Amph sten	Amph wrig-type	Amph sten	Amph wrig-type	Amph sp.	Amph sp.
Amphitrema wrightianum	Amph wrig		x				Amph wrig	Amph wrig-type	Amph wrig	Amph wrig-type	Amph sp.	Amph sp.
Arcella artocrea	Arce arto		x				Arce arto	Arce arto	Arce arto	Arce cati-type	Arce sp.	Arce sp.
Arcella catinus	Arce aren			x			Arce aren	Arce cati	Arce cati	Arce cati-type	Arce sp.	Arce sp.
Arcella discoides	Arce disc			x			Arce disc	Arce disc	Arce disc	Arce vulg-type	Arce sp.	Arce sp.
Arcella gibbosa	Arce gibb			x			Arce gibb	Arce gibb	Arce gibb	A. hemi-type	Arce sp.	Arce sp.
Arcella hemisphaerica	Arce hemi			x			Arce hemi	Arce hemi	Arce hemi	A. hemi-type	Arce sp.	Arce sp.
Arcella vulgaris	Arce vulg			x			Arce vulg	Arce vulg	Arce vulg	Arce vulg-type	Arce sp.	Arce sp.
Archerella (Amphitrema) flavum	Amph flav	x					Arch flav	Arch flav	Arch flav	Arch flav	Arch sp.	Arch sp.
Argynnia (Nebela) dentistoma	Nebe dent			x			Argy dent	Argy dent-type	Argy dent-type	Argy dent-type	Argy sp.	Diff sp.
Argynnia (Nebela) vitraea	Nebe vitr			x			Argy vitr	Argy dent-type	Argy dent-type	Argy dent-type	Argy sp.	Diff sp.
Assulina muscorum	Assu musc		x				Assu musc	Assu musc	Assu musc	Assu. sp	Assu. sp	Assu. sp
Assulina scandinavica	Assu scan		x				Assu scan	Assu semi-type	Assu semi-type	Assu. sp	Assu. sp	Assu. sp
Assulina seminulum	Assu semi		x				Assu semi	Assu semi-type	Assu semi-type	Assu. sp	Assu. sp	Assu. sp
Bullinularia indica	Bull indi	x					Bull indi	Bull indi	Bull indi	Bull indi	Bull. sp	Bull. sp
Campascus minuta	Camp minu				x	x	Camp minu	Deleted	Deleted	Deleted	Camp sp.	Deleted
Centropyxis aculeata	Cent acul			x			Cent acul	Cent acul-type	Cent acul-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis aculeata oblonga	Cent ac-o				x		Cent ac-o	Cent acul-type	Cent acul-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis aerophila	Cent a-s			x			Cent aero	Cent aero-type	Cent cass-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis aerophila sphagnicola	Cent aero			x			Cent ae-s	Cent orbi-type	Cent cass-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis cassis	Cent cass			x			Cent cass	Cent aero-type	Cent cass-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis ecornis	Cent ecor			x			Cent ecor	Cent laev-type	Cycl arce-type	Cent laev-type	Cent sp.	Arce sp.
Centropyxis hirsuta	Cent hirs			x			Cent hirs	Cent acul-type	Cent acul-type	Cent aero-type	Cent sp.	Cent sp.
Centropyxis laevigata	Cent laev			x			Cent laev	Cent laev-type	Cent arce-type	Cent laev-type	Cent sp.	Arce sp.
Centropyxis platystoma	Cent plat			x			Cent plat	Cent plat	Cent plat	Cent aero-type	Cent sp.	Cent sp.
Centropyxis spinosa	Cent spin			x			Cent spin	Cent acul-type	Cent acul-type	Cent aero-type	Cent sp.	Cent sp.
Corythion dubium	Cory dubi			x			Cory dubi	Cory dubi-type	Trin-Cory	Cory dubi	Cory sp.	Trin-Cory
Corythion pulchellum	Cory pulc			x		x	Cory pulc	Cory dubi-type	Trin-Cory	Deleted	Cory sp.	Deleted
Cryptodifflugia oviformis	Cryp ovif				x	x	Cryp ovif	Cryp sp.	Cryp ovif	Deleted	Cryp sp.	Deleted
Cyclopyxis arcelloides	Cycl arce				x		Cycl arce	Cycl arce-type	Cycl arce-type	Cycl sp.	Cycl sp.	Cycl sp.
Cyclopyxis eurystoma	Cycl eury			x			Cycl eury	Cycl eury-type	Cycl arce-type	Cycl sp.	Cycl sp.	Cycl sp.

(*continued on next page*)

Table 1 (*continued*)

Full name	Code	Degree of confusion[a]					Species pooling or deletion in the 6 cases and number of taxa retained					
		A	B	C	D	E	1 = original resolution	2 = pooling of most confusing taxa	3 = Charman & al book	4 = limited resolution	5 = genus level	6 = genus level or less
							100	66	63	34	29	16
Cyclopyxis kahli	Cycl kahl			x			Cycl kahl	Cycl eury-type	Cycl arce-type	Cycl sp.	Cycl sp.	Cycl sp.
Cyphoderia ampulla	Cyph ampu		x				Cyph ampu	Cyph ampu-type	Cyph ampu-type	Cyph ampu-type	Cyph sp.	Cyph sp.
Difflugia ampullula	Diff ampu				x		Diff ampu	Diff ampu-type	Diff ampu-type	Diff ampu-type	Diff. sp	Diff sp.
Difflugia avellana	Diff avel				x		Diff avel	Diff luci-type	Diff luci-type	Diff luci-type	Diff. sp	Diff sp.
Difflugia baccillariarum	Diff bacc			x			Diff bacc	Diff bacc-type	Diff bacc-type	Diff baci-type	Diff. sp	Diff sp.
Difflugia bacilifera	Diff baci			x			Diff baci	Diff baci-type	Diff baci	Diff baci-type	Diff. sp	Diff sp.
Difflugia elegans	Diff eleg			x			Diff eleg	Diff bacc-type	Diff bacc-type	Diff baci-type	Diff. sp	Diff sp.
Difflugia gassowski	Diff gass				x		Diff gass	Diff baci-type	Diff oblo-type	Diff baci-type	Diff. sp	Diff sp.
Difflugia globulosa	Diff glob			x			Diff glob	Diff glob	Cycl arce-type	Phry acro-type	Diff. sp	Diff sp.
Difflugia leidyi	Diff leid	x					Diff leid	Diff luci-type	Diff leid	Diff luci-type	Diff. sp	Hele sp.
Difflugia lucida	Diff luci				x		Diff luci	Diff luci-type	Diff luci-type	Diff luci-type	Diff. sp	Hele sp.
Difflugia oblonga	Diff oblo			x			Diff oblo	Diff oblo	Diff oblo-type	Diff baci-type	Diff. sp	Diff sp.
Difflugia penardi	Diff pena				x		Diff pena	Diff pris-type	Diff luci-type	Diff luci-type	Diff. sp	Diff sp.
Difflugia pristis	Diff pris				x	x	Diff pris	Diff pris-type	Diff pris-type	Diff luci-type	Diff. sp	Deleted
Difflugia pulex	Diff pule				x	x	Diff pule	Diff pule-type	Diff pule-type	Diff luci-type	Diff. sp	Deleted
Difflugia rubescens	Diff rube			x			Diff rube	Diff rube	Diff rube	Diff baci-type	Diff. sp	Diff sp.
Euglypha ciliata	Eugl cili			x			Eugl cili	Eugl cili-type	Eugl cili-type	Eugl cili-type	Eugl sp.	Eugl sp.
Euglypha ciliata glabra	Eugl ci-g				x		Eugl ci-g	Eugl cili-type	Eugl cili-type	Eugl cili-type	Eugl sp.	Eugl sp.
Euglypha compressa	Eugl comp			x			Eugl comp	Eugl comp	Eugl comp	Eugl cili-type	Eugl sp.	Eugl sp.
Euglypha cristata	Eugl cris		x			x	Eugl cris	Eugl cris	Eugl cris	Eugl tube-type	Eugl sp.	Eugl sp.
Euglypha filifera	Eugl fili			x			Eugl fili	Eugl rotu-type	Eugl cili-type	Eugl tube-type	Eugl sp.	Eugl sp.
Euglypha laevis	Eugl laev				x	x	Eugl laev	Eugl rotu-type	Eugl rotu-type	Eugl tube-type	Eugl sp.	Eugl sp.
Euglypha rotunda	Eugl rotu				x	x	Eugl rotu	Eugl rotu-type	Eugl rotu-type	Eugl tube-type	Eugl sp.	Eugl sp.
Euglypha strigosa	Eugl stri			x			Eugl stri	Eugl stri	Eugl stri	Eugl cili-type	Eugl sp.	Eugl sp.
Euglypha tuberculata	Eugl tube				x		Eugl tube	Eugl tube	Eugl tube	Eugl tube-type	Eugl sp.	Eugl sp.
Heleopera petricola	Hele petr			x			Hele petr	Hele petr-type	Hele petr	Hele petr-type	Hele sp.	Hele sp.
Heleopera rosea	Hele rose		x				Hele rose	Hele rose	Hele rose	Hele rose	Hele sp.	Hele sp.
Heleopera sphagni	Hele spha		x				Hele spha	Hele spha	Hele spha	Hele petr-type	Hele sp.	Hele sp.
Heleopera sylvatica	Hele sylv			x			Hele sylv	Hele petr-type	Hele sylv	Hele petr-type	Hele sp.	Hele sp.
Hyalosphenia elegans	Hyal eleg	x					Hyal eleg	Hyal eleg	Hyal eleg	Hyal eleg	Hyal sp.	Nebe sp.
Hyalosphenia minuta	Hyal minu			x		x	Hyal minu	Deleted	Deleted	Deleted	Hyal sp.	Nebe sp.
Hyalosphenia papilio	Hyal papi	x					Hyal papi	Hyal papi	Hyal papi	Hyal papi	Hyal sp.	Nebe sp.
Hyalosphenia subflava	Hyal subf		x				Hyal subf	Hyal subf	Hyal subf	Hyal subf	Hyal sp.	Nebe sp.
Lagenodifflugia vas	Lage vas		x				Lage vas	Lage sp.	Lage sp.	Diff baci-type	Lege sp.	Diff sp.
Lesquereusia epistomium	Lesq epis			x			Lesq epis	Lesq epis	Lesq epis	Lesq sp.	Lesq sp.	Lesq sp.
Lesquereusia modesta	Lesq mode			x			Lesq mode	Lesq spir-type	Lesq mode	Lesq sp.	Lesq sp.	Lesq sp.
Lesquereusia spiralis	Lesq spir			x			Lesq spir	Lesq spir-type	Lesq spir	Lesq sp.	Lesq sp.	Lesq sp.
Nebela bohemica	Nebe bohe			x			Nebe bohe	Nebe tinc-type	Nebe coll-type	Nebe tinc-type	Nebe sp.	Nebe sp.
Nebela carinata	Nebe cari		x				Nebe cari	Nebe cari-type	Nebe cari	Nebe cari-type	Nebe sp.	Nebe sp.
Nebela collaris	Nebe coll			x			Nebe coll	Nebe coll	Nebe coll-type	Nebe mili-type	Nebe sp.	Nebe sp.
Nebela galeata	Nebe gale		x				Nebe gale	Nebe gale	Nebe tubu-type	Nebe cari-type	Nebe sp.	Nebe sp.
Nebela marginata	Nebe marg		x				Nebe marg	Nebe cari-type	Nebe marg	Nebe cari-type	Nebe sp.	Nebe sp.
Nebela militaris	Nebe mili	x					Nebe mili	Nebe mili-type	Nebe mili	Nebe mili-type	Nebe sp.	Nebe sp.
Nebela minor	Nebe mino			x			Nebe mino	Nebe tinc-type	Nebe mino	Nebe tinc-type	Nebe sp.	Nebe sp.
Nebela parvula	Nebe parv			x			Nebe parv	Nebe tinc-type	Nebe parv	Nebe tinc-type	Nebe sp.	Nebe sp.
Nebela penardiana	Nebe pena			x			Nebe pena	Nebe pena-type	Nebe tubu-type	Nebe cari-type	Nebe sp.	Nebe sp.
Nebela tincta	Nebe tinc			x			Nebe tinc	Nebe tinc-type	Nebe tinc-type	Nebe tinc-type	Nebe sp.	Nebe sp.
Nebela tincta major	Nebe t-ma			x			Nebe t-ma	Nebe tinc-type	Nebe tinc-type	Nebe tinc-type	Nebe sp.	Nebe sp.
Nebela tubulosa	Nebe tubulo			x			Nebe tubulo	Nebe pena-type	Nebe tubu-type	Nebe cari-type	Nebe sp.	Nebe sp.
Padaungiella (Nebela) lageniformis	Nebe lage		x				Nebe lage	Nebe lage	Nebe lage-type	Nebe mili-type	Nebe sp.	Nebe sp.
Padaungiella (Nebela) tubulata	Nebe tubula			x			Nebe tubula	Nebe wail-type	Nebe lage-type	Nebe mili-type	Nebe sp.	Nebe sp.
Padaungiella (Nebela) wailesii	Nebe wail			x			Nebe wail	Nebe wail-type	Nebe lage-type	Nebe mili-type	Nebe sp.	Nebe sp.
Paraquadrula irregularis	Para irre	x					Para irre	Para irre	Para irre	Deleted	Para sp.	Deleted
Phryganella acropodia	Phry acro				x		Phry acro	Phry acro-type	Phry acro-type	Phry acro-type	Phry sp.	Diff sp.
Phryganella paradoxa	Phry para				x	x	Phry para	Diff pule-type	Diff pule-type	Diff luci-type	Phry sp.	Deleted
Physochila (Nebela) griseola	Nebe gris			x			Phys gris	Phys gris	Phys gris	Phys gris	Phys sp.	Diff sp.
Placocista spinosa	Plac spin		x				Plac spin	Plac spin-type	Plac spin-type	Plac spin-type	Plac sp.	Eugl sp.
Placocista spinosa hyalina	Plac sp-h			x			Plac sp-h	Plac spin-type	Plac spin-type	Plac spin-type	Plac sp.	Eugl sp.
Pseudodifflugia fulva	Pseu fulv				x	x	Pseu fulv	Diff pule-type	Pseu fulv	Diff luci-type	Pseu sp.	Deleted
Quadrulella symmetrica	Quad syme	x					Quad syme	Quad syme	Quad syme	Quad syme	Quad sp.	Quad sp.
Sphenoderia fissirostris	Sphe fiss			x		x	Sphe fiss	Sphe lent-type	Sphe lent-type	Sphe lent-type	Sphe sp.	Deleted
Sphenoderia lenta	Sphe lent		x				Sphe lent	Sphe lent-type	Sphe lent-type	Eugl tube-type	Sphe sp.	Eugl sp.
Tracheleuglypha dentata	Trac dent			x		x	Trac dent	Trac dent-type	Sphe lent-type	Eugl tube-type	Trac sp.	Eugl sp.
Trigonopyxis arcula	Trig arcu	x					Trig arcu	Trig arcu	Trig arcu	Trig arcu	Trig sp.	Trig sp.
Trinema complanatum	Trin comp			x			Trin comp	Trin comp-type	Trin-Cory	Trin sp.	Trin sp.	Trin-Cory
Trinema enchelys	Trin ench			x			Trin ench	Trin line-type	Trin-Cory	Trin sp.	Trin sp.	Trin-Cory
Trinema lineare	Trin line			x		x	Trin line	Trin line-type	Trin-Cory	Trin sp.	Trin sp.	Trin-Cory
Trinema penardi	Trin pena			x			Trin pena	Trin comp-type	Trin-Cory	Trin sp.	Trin sp.	Trin-Cory
Trinema sp.	Trin sp		x			x	Trin sp.	Trin sp.	Trin-Cory	Trin sp.	Trin sp.	Trin-Cory

[a] A: Unmistakable, B: Limited possible confusion, C: Moderately confusing, D: Very confusing, E: Easily overlooked.

Centropyxis aerophila, *C. aerophila* var. *sphagnicola*, *Centropyxis ecornis*), D) very confusing (e.g. *Euglypha tuberculata*, *E. rotunda*), E) easily overlooked (e.g. *Corythion pulchellum*, *Trinema lineare*, *Cryptodifflugia oviformis*, *Sphenoderia fissirostris*). Based on this, we gradually reduced the number of species in 6 steps: 1: Original taxonomic resolution (100 taxa in total, excluding eight taxa identified only to genus level – with the exception of *Trinema* sp. in the Jura data set which were mostly *T. lineare* – and 30 taxa recorded in fewer than three samples), 2: Pooling most confusing taxa only (66 taxa), 3: Pooling into types specified by the widely-used identification guide of Charman et al. (2000) (63 taxa), 4: Limited taxonomic resolution (34 taxa), 5: Genus level (29 taxa) and 6: Genus level or lower (16 taxa). The full list of taxa observed in all studies, together with the way each taxon was treated (i.e. with which species each taxon was grouped and which ones may be overlooked) are listed in Table 1 with full details of the rationale given in Supplementary Table 1. Representative taxa are illustrated in Fig. 1. This approach resulted from a consensus among the authors who have both their own experience of learning the taxonomy as well as teaching it to undergraduate and graduate students.

We then used these different data sets and assessed the performance of transfer function models (in all cases weighted

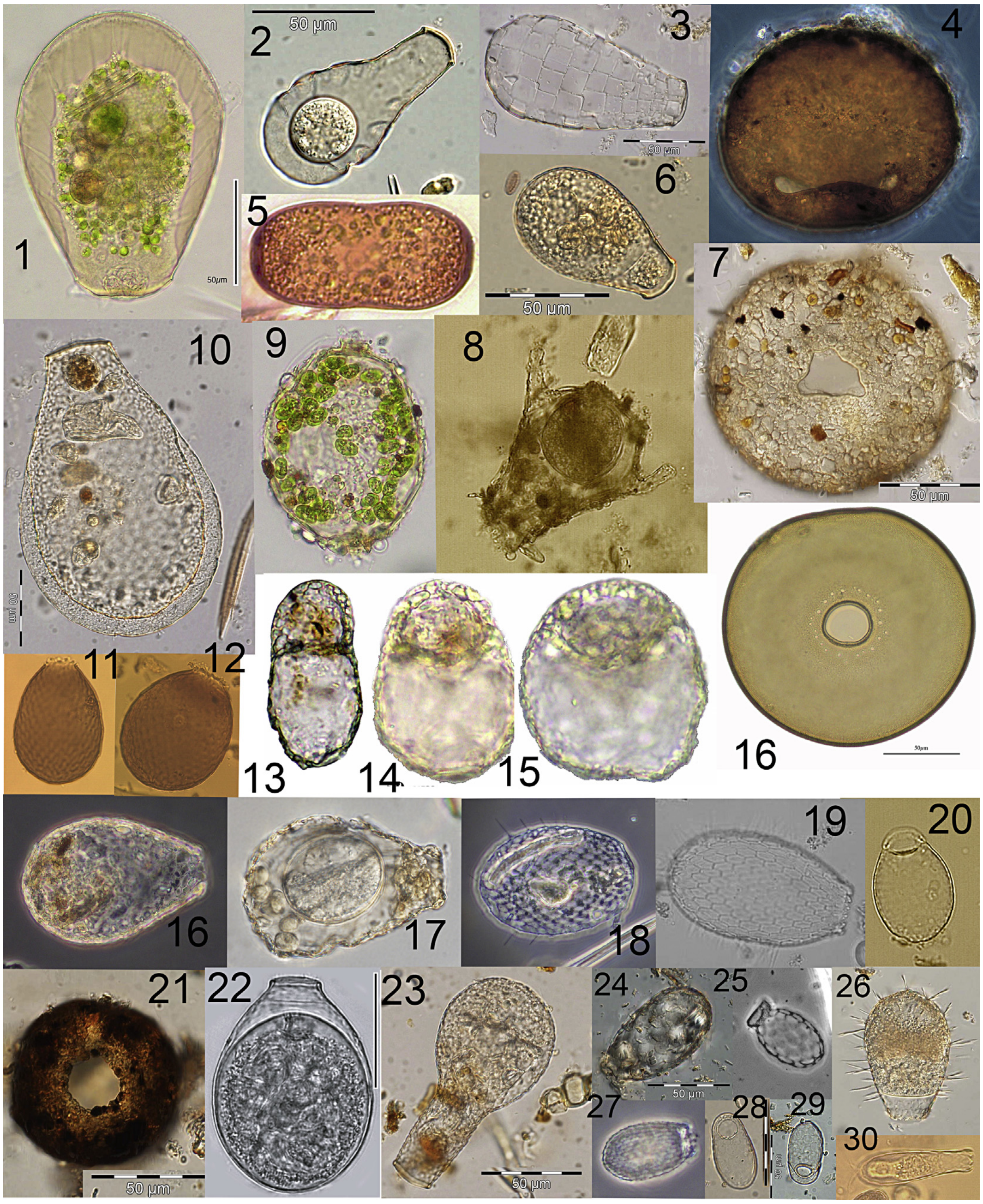

Fig. 1. Illustrations of testate amoeba taxa corresponding to the five categories of identification confidence. 1: *Hyalosphenia papilio*, 2: *Hyalosphenia elegans*, 3: *Quadrulella symmetrica*, 4: *Bullinularia indica* 5: *Archerella flavum*, 6: *Nebela militaris*, 7: *Trigonopyxis arcula*, 8: *Difflugia leidyi*, 9: *Amphitrema wrightianum*, 10: *Nebela carinata*, 11: *Assulina muscorum*, 12: *A. seminulum*; 13: *Centropyxis platystoma*, 14: *Centropyxis aerophila*, 15: *Centropyxis aerophila var. sphagnicola*, 16: *Arcella artocrea*; 16: *Argynnia (Nebela) dentistoma*, 17: *Physochila (Nebela) griseola*; 18: *Plycocista spinosa*, 19: *Euglypha ciliata*, 20: *Corythion dubium*, 21: *Cyclopyxis kahli*, 22: *Nebela tincta*, 23: *Padaungialla (Nebela) wailesii*, 24: *Tracheleuglypha dentata*, 25: *Sphenoderia fissirostris*, 26: *Euglypha ciliata*, 27: *Euglypha laevis*, 20: *Trinema enchelys*, 29: *Trinema complanatum*, 30: *Euglypha cristata*.

averaging with classical deshrinking) using standard cross-validation metrics: R^2 between observed and predicted values (R^2) and root mean squared average error of prediction (RMSEP). We hypothesised that reducing the taxonomic resolution would cause an approximately proportional decrease in transfer model performance (i.e. decreases in R^2 and increase in RMSEP).

We then tested the effect of taxonomic aggregation on palaeoecological inference, using the same pooling approach. Taxonomic resolution as well as the length of the water table depth gradient varied among the seven data sets. Both were highest for the Engadine and Poland training sets and these were therefore chosen to assess the effect of taxonomic resolution on palaeoecological interpretation using a high-resolution palaeoecological record from Mauntschas, Engadine (Lamentowicz et al., 2010b; van der Knaap et al., 2011) and a second data set from Tuchola mire, northern Poland (Lamentowicz et al., 2008).

3. Results

Pooling and removal of species resulted in a gradual decrease in overall taxonomic richness (down to 84%, 80%, 52%, 42%, and 27% in cases 2–6) and average richness per sample in the seven data sets (down to 94%, 92%, 76%, 67%, and 53% in cases 2–6, Fig. 2, Supplementary Table 2). Effects on the two palaeoecological data sets were similar (Supplementary Table 2).

The effect of taxonomic pooling and removal of species on model performance are given in Fig. 3 with further details given in Supplementary Table 3. The performance of all five transfer functions declined with decreasing taxonomic richness: Boot R^2 decreased on average by 16% while RMESP increased on average by 21% from maximum resolution to case 6. Best performing models were in cases 1–3 and worst performing in cases 4 and 6 (Supplementary Table 3).

Overall with increasing taxonomic pooling the values for inferred DWT gradually diverge from those predicted using the full taxonomic resolution (Supplementary Fig. 1). However, taxonomic pooling and removal of species did not dramatically modify the patterns of reconstructed DWT in the two studied cores (Fig. 4a and b). However the peaks were more marked with case 6 than with case 1 in both data sets. Furthermore a ca 20 cm shift in reconstructed DWT appeared in the upper part of the Mauntschas core with the lower resolution model indicating wetter conditions. The prediction error remained relatively stable from case 1 to case 4 but then increased substantially in cases 5 and 6. It should be noted that in both the Engadine and Tuchola records the prediction errors were larger than the differences in inferred DWT among the cases.

4. Discussion

Decreasing taxonomic resolution had a relatively limited effect on ecological inference. This result agrees with a previous report of strong correlation between taxonomic richness assessed at species and genus level (Wilkinson and Davis, 2000). This could suggest that there is a high degree of ecological redundancy in the response of species to water table depth among peatland testate amoeba taxa. In a totally different ecological context a similar observation was made between taxonomic resolution of tropical trees and the correlation with environmental variables (concentration of Ca, K, Mg, Na) assessed by Mantel test and redundancy analyses (Cayuela et al., 2011). Nevertheless, the lack of clear reduction in transfer-function performance we observed with reduced taxonomic resolution is counter-intuitive and contradicts observations available for other groups such as macro-invertebrates where the proportion of specialists (i.e. most valuable bioindicators) dropped from 84% at species level resolution to 43% at family level resolution (Nahmani et al., 2006).

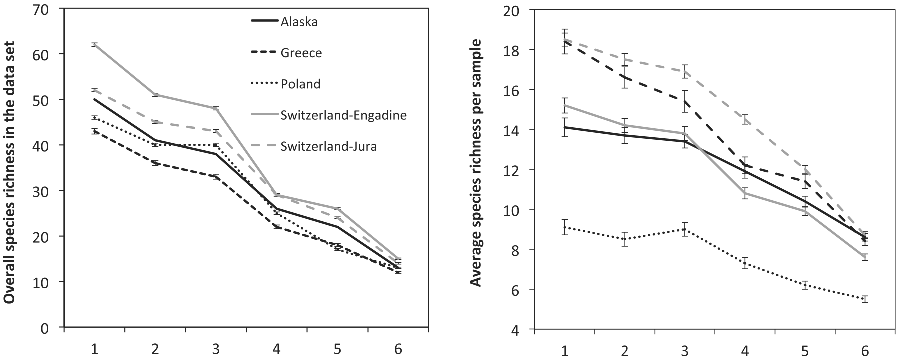

Fig. 2. Effect of taxonomic filtering on overall species richness in each data set (top left), average species richness per sample (top right), species richness in each data set as % of overall richness (bottom left) and average species richness per sample as % of potential maximum for the data set (bottom right).

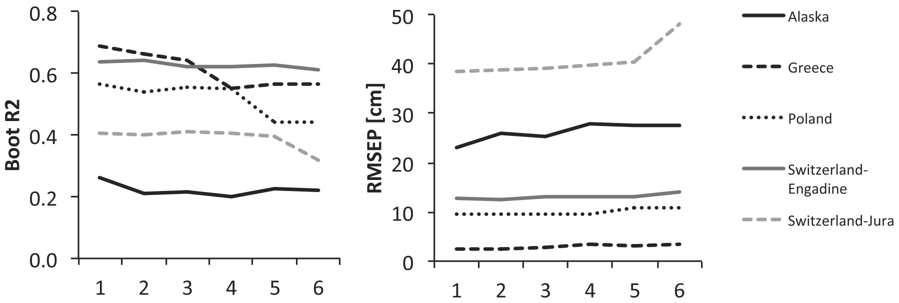

Fig. 3. Effect of taxonomic filtering on the performance metrics of seven testate amoeba depth to water table transfer function models.

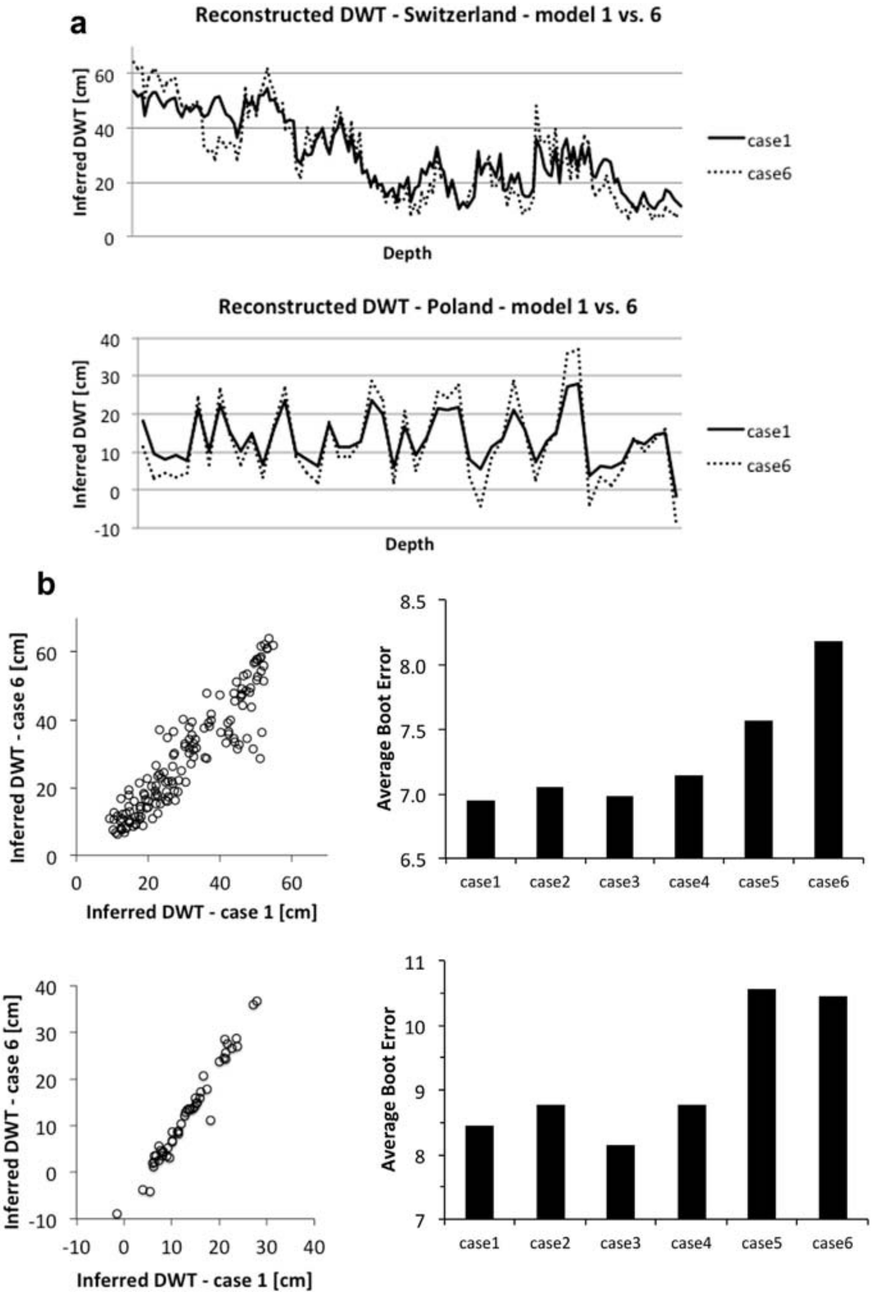

Fig. 4. a. Effect of taxonomic filtering on reconstructed depth to water table from the Mauntschass core using the Engadine transfer function (Engadine, (Lamentowicz et al., 2010b; van der Knaap et al., 2011)) (top) and the Tuchola mire core from Poland (Lamentowicz et al., 2008) using the Polish transfer function (bottom), showing the comparison of DWT pattern for case 1 vs. case 6. The taxonomic filtering was the same for the modern and paleo data sets. b. Effect of taxonomic filtering on reconstructed depth to water table from the Mauntschass core using the Engadine transfer function (Engadine, (Lamentowicz et al., 2010b; van der Knaap et al., 2011)) (top) and the Tuchola mire core from Poland (Lamentowicz et al., 2008) using the Polish transfer function (bottom), showing the comparison of case 1 vs. case 6 and change in average Bootstrapped error of prediction from case 1 to case 6.

However although the rather drastic reduction in taxonomic resolution did not cause the models to lose all predictive power, the quality of inference nevertheless decreased. Inferred DWT values from the two studied paleoenvironmental records showed shifts towards more extreme values and in one case a ca 20 cm discrepancy between two models. Such a "shift" would clearly have been interpreted and considered as a substantial "event" in the palaeoecological history of the site. This observation calls for caution and suggests that species pooling at least beyond a certain level will affect the quality of (palaeo)ecological inference from testate amoebae.

The importance of taxonomic resolution for palaeoecological inference and bioindication may differ among taxonomic groups and even within groups depending on the question of interest or characteristics of the environment. For example aggregation of marine benthic nematode community data to genus level did not affect ecological interpretation but further aggregation did (Somerfield and Clarke, 1995). Correlation between marine macrozoobenthos faunal patterns assessed using highest vs. lowest taxonomic resolution was shown to be higher in polluted habitats than in less polluted areas (Olsgard et al., 1998). This result may simply be due to higher species-richness in pristine habitats and hence the fact that more information is lost with lumping as compared to the species-poor more polluted areas. In line with this idea, aggregation of marine macro-zoobenthos data to family level had a higher impact on correlation between similarity matrices of communities from the more diverse sub-littoral habitat than on the less diverse intertidal habitat (Somerfield and Clarke, 1995).

Reduced taxonomic resolution in palaeoecological analysis of testate amoebae is potentially attractive as this could allow quicker counting, more consistent results between analysts and quicker learning for inexperienced analysts. Payne et al. (2011) have highlighted the possibility for modest confusion of taxa between training sets and palaeoecological data to produce entirely different reconstructions. Our results show much more modest difference and suggest that it is likely to be much safer to group potentially-confused data than to risk confusing them. Our results also show that the groupings suggested by Charman et al. (2000), widely used by peatland palaeoecologists, although rather drastic in some respects (e.g. the very broad *C. arcelloides* type) lead to little or no deterioration in transfer function performance. Nevertheless our results do show loss of performance at the most extreme level of grouping (to genus or below). Such grouping is not routinely conducted and our results provide sufficient cause for concern to suggest that this should not be carried out. Our results support current practise by suggesting that grouping of easily confused taxa is unlikely to significantly compromise palaeoecological reconstruction. Our results further suggest that even a further stage of grouping (taxa which might be confused) is unlikely to have major negative effects suggesting new possibilities for the combination of existing datasets and construction of supra-regional transfer functions.

In the light of current taxonomic studies revealing a wealth of unsuspected diversity in testate amoebae (Heger et al., 2010; Heger et al., 2011; Kosakyan et al., 2012, 2013; Heger et al., 2013), it would be interesting to perform the opposite experiment, aiming at higher taxonomic resolution than has been generally been applied. If such “cryptic” (morphologically undistinguishable) or “pseudo-cryptic” (with very subtle morphological differences) species also differed in their ecological optima the performance of transfer function might be substantially improved. An associated cost, however, would be that analysts would truly need to be expert taxonomists and analyses would take more time, but depending on the research question this could be justified.

Acknowledgements

This work was funded by Swiss NSF projects no. 205321–109709/1 and 205321–109709/2, the Scientific & Technological Cooperation Programme Switzerland–Russia faculty exchange project, and the Science and Technology Cooperation Program Switzerland–Russia grant IZLR Z3_128338 to E. Mitchell. R. Payne is supported by an Impact Fellowship from the University of Stirling. Yu. Mazei is supported by the grant of the President of Russian Federation (MD-4435.2014.4) and the grant of the Russian Foundation for Basic Research (13-04-00542). Project supported by a grant PSPB-013/2010 from Switzerland through the Swiss Contribution to the enlarged European Union.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.quascirev.2014.03.006.

References

Avel, E., Pensa, M., 2013. Preparation of testate amoebae samples affects water table depth reconstructions in peatland palaeoecological studies. Est. J. Earth Sci. 62, 113–119.

Birks, H.J.B., 2003. Quantitative palaeoenvironmental reconstructions from Holocene biological data. In: Mackay, A., Battarbee, R.W., Birks, H.J.B., Oldfield, F. (Eds.), Global Change in the Holocene. Arnold, pp. 107–123.

Booth, R.K., 2002. Testate amoebae as paleoindicators of surface-moisture changes on Michigan peatlands: modern ecology and hydrological calibration. J. Paleolimnol. 28, 329–348.

Cayuela, L., de la Cruz, M., Ruokolainen, K., 2011. A method to incorporate the effect of taxonomic uncertainty on multivariate analyses of ecological data. Ecography 34, 94–102.

Charman, D.J., 2001. Biostratigraphic and palaeoenvironmental applications of testate amoebae. Quat. Sci. Rev. 20, 1753–1764.

Charman, D.J., Hendon, D., Woodland, W.A., 2000. The Identification of Testate Amoebae (Protozoa: Rhizopoda) in Peats. Quaternary Research Association, London.

Deflandre, G., 1929. Le genre *Centropyxis* Stein. Arch. Protistenkd. 67, 322–375.

Heger, T.J., Mitchell, E.A.D., Leander, B., 2013. Holarctic phylogeography of the testate amoeba *Hyalosphenia papilio* (Amoebozoa: Arcellinida) reveals extensive genetic diversity explained more by environment than dispersal limitation. Mol. Ecol. 22, 5172–5184.

Heger, T.J., Mitchell, E.A.D., Ledeganck, P., Vincke, S., Van De Vijver, B., Beyens, L., 2009. The curse of taxonomic uncertainty in biogeographical studies of free-living terrestrial protists: a case study of testate amoebae from Amsterdam Island. J. Biogeogr. 36, 1551–1560.

Heger, T.J., Mitchell, E.A.D., Todorov, M., Golemansky, V., Lara, E., Leander, B.S., Pawlowski, J., 2010. Molecular phylogeny of euglyphid testate amoebae (Cercozoa: Euglyphida) suggests transitions between marine supralittoral and freshwater/terrestrial environments are infrequent. Mol. Phylogenet. Evol. 55, 113–122.

Heger, T.J., Pawlowski, J., Lara, E., Leander, B.S., Todorov, M., Golemansky, V., Mitchell, E.A.D., 2011. Comparing potential COI and SSU rDNA barcodes for assessing the diversity and phylogenetic relationships of cyphoderiid testate amoebae (Rhizaria: Euglyphida). Protist 162, 131–141.

Heiri, O., Lotter, A.F., 2010. How does taxonomic resolution affect chironomid-based temperature reconstruction? J. Paleolimnol. 44, 589–601.

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.

Lamentowicz, Ł., Gąbka, M., Rusinska, A., Sobczynski, T., Owsianny, P.M., Lamentowicz, M., 2011. Testate amoeba (Arcellinida, Euglyphida) ecology along a poor-rich gradient in fens of western Poland. Int. Rev. Hydrobiol. 96, 356–380.

Lamentowicz, M., Lamentowicz, Ł., van der Knaap, W.O., Gąłka, M., Mitchell, E.A.D., 2010a. Contrasting species-environment relationships in communities of testate amoebae, bryophytes and vascular plants along the fen-bog gradient. Microb. Ecol. 59, 499–510.

Lamentowicz, M., Mitchell, E.A.D., 2005. The ecology of testate amoebae (Protists) in Sphagnum in north-western Poland in relation to peatland ecology. Microb. Ecol. 50, 48–63.

Lamentowicz, M., Obremska, M., Mitchell, E.A.D., 2008. Autogenic succession, land-use change, and climatic influences on the Holocene development of a kettle-hole mire in Northern Poland. Rev. Palaeobot. Palynol. 151, 21–40.

Lamentowicz, M., van der Knaap, W.O., Lamentowicz, Ł., van Leeuwen, J.F.N., Mitchell, E.A.D., Goslar, T., Kamenik, C., 2010b. A near-annual palaeohydrological study based on testate amoebae from a sub-alpine mire: surface wetness and the role of climate during the instrumental period. J. Quat. Sci. 25, 190–202.

Lara, E., Heger, T.J., Scheihing, R., Mitchell, E.A.D., 2011. COI gene and ecological data suggest size-dependent high dispersal and low intra-specific diversity in free-living terrestrial protists (Euglyphida; *Assulina*). J. Biogeogr. 38, 640–650.

Mazei, Y., Tsyganov, A., 2006. Freshwater Testate Amoebae. KMK Sci Press, Moscow.

Mazei, Y., Warren, A., 2012. 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 1: species with shells that are pointed aborally and/or have aboral protuberances. Protistology 7, 121–171.

Mitchell, E.A.D., Buttler, A.J., Warner, B.G., Gobat, J.M., 1999. Ecology of testate amoebae (Protozoa : Rhizopoda) in *Sphagnum* peatlands in the Jura mountains, Switzerland and France. Ecoscience 6, 565–576.

Mitchell, E.A.D., Charman, D.J., Warner, B.G., 2008a. Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. Biodivers. Conserv. 17, 2115–2137.

Mitchell, E.A.D., Meisterfeld, R., 2005. Taxonomic confusion blurs the debate on cosmopolitanism versus local endemism of free-living protists. Protist 156, 263–267.

Mitchell, E.A.D., Payne, R.J., Lamentowicz, M., 2008b. Potential implications of differential preservation of testate amoeba shells for paleoenvironmental reconstruction in peatlands. J. Paleolimnol. 40, 603–618.

Mitchell, E.A.D., van der Knaap, W.O., van Leeuwen, J.F.N., Buttler, A., Warner, B.G., Gobat, J.M., 2001. The palaeoecological history of the Praz-Rodet bog (Swiss Jura) based on pollen, plant macrofossils and testate amoebae (Protozoa). Holocene 11, 65–80.

Nahmani, J., Lavelle, P., Rossi, J.P., 2006. Does changing the taxonomical resolution alter the value of soil macroinvertebrates as bioindicators of metal pollution? Soil Biol. Biochem. 38, 385–396.

Olsgard, F., Somerfield, P.J., Carr, M.R., 1998. Relationships between taxonomic resolution, macrobenthic community patterns and disturbance. Mar. Ecol. Prog. Ser. 172, 25–36.

Payne, R., Mitchell, E.A.D., 2007. Testate amoebae-environment relationships and a hydrological transfer function from mires in the Central Rhodope Mountains, Greece. Protist 158, 159–171.

Payne, R.J., Kishaba, K., Blackford, J.J., Mitchell, E.A.D., 2006. Ecology of testate amoebae (Protista) in south-central Alaska peatlands: building transfer-function models for palaeoenvironmental studies. Holocene 16, 403–414.

Payne, R.J., Lamentowicz, M., Mitchell, E.A.D., 2011. The perils of taxonomic inconsistency in quantitative palaeoecology: experiments with testate amoeba data. Boreas 40, 15–27.

Payne, R.J., Mitchell, E.A.D., 2009. How many is enough? Determining optimal count totals for ecological and palaeoecological studies of testate amoebae. J. Paleolimnol. 42, 483–495.

Payne, R.J., Mitchell, E.A.D., Hung, N.-V., Gilbert, D., 2012. Can pollution bias peatland paleoclimate reconstruction? Quat. Res. 78, 170–173.

Seddon, A.W.R., Mackay, A.W., Baker, A.G., Birks, H.J.B., Breman, E., Buck, C.E., Ellis, E.C., Froyd, C.A., Gill, J.L., Gillson, L., Johnson, E.A., Jones, V.J., Juggins, S., Macias-Fauria, M., Mills, K., Morris, J.L., Nogués-Bravo, D., Punyasena, S.W., Roland, T.P., Tanentzap, A.J., Willis, K.J., Aberhan, M., van Asperen, E.N., Austin, W.E.N., Battarbee, R.W., Bhagwat, S., Belanger, C.L., Bennett, K.D., Birks, H.H., Bronk Ramsey, C., Brooks, S.J., de Bruyn, M., Butler, P.G., Chambers, F.M., Clarke, S.J., Davies, A.L., Dearing, J.A., Ezard, T.H.G., Feurdean, A., Flower, R.J., Gell, P., Hausmann, S., Hogan, E.J., Hopkins, M.J., Jeffers, E.S., Korhola, A.A., Marchant, R., Kiefer, T., Lamentowicz, M., Larocque-Tobler, I., López-Merino, L., Liow, L.H., McGowan, S., Miller, J.H., Montoya, E., Morton, O., Nogué, S., Onoufriou, C., Boush, L.P., Rodriguez-Sanchez, F., Rose, N.L., Sayer, C.D., Shaw, H.E., Payne, R., Simpson, G., Sohar, K., Whitehouse, N.J., Williams, J.W., Witkowski, A., 2014. Looking forward through the past: identification of 50 priority research questions in palaeoecology. J. Ecol. 102, 256–267.

Somerfield, P.J., Clarke, K.R., 1995. Taxonomic levels, in marine community studies, revisited. Mar. Ecol. Prog. Ser. 27, 113–119.

Sullivan, M.E., Booth, R.K., 2011. The potential influence of short-term environmental variability on the composition of testate amoeba communities in sphagnum peatlands. Microb. Ecol. 62, 80–93.

van der Knaap, W.O., Lamentowicz, M., van Leeuwen, J.F.N., Hangartner, S., Leuenberger, M., Mauquoy, D., Goslar, T., Mitchell, E.A.D., Lamentowicz, L., Kamenik, C., 2011. A multi-proxy, high-resolution record of peatland development and its drivers during the last millennium from the subalpine Swiss Alps. Quat. Sci. Rev. 30, 3467–3480.

Wilkinson, D.M., Davis, S.R., 2000. Rapid assessments of microbial biodiversity using relationships between genus and species richness. Studies on testate amoebae. Acta Protozool. 39, 23–26.

Wylezich, C., Meisterfeld, R., Meisterfeld, S., Schlegel, M., 2002. Phylogenetic analyses of small subunit ribosomal RNA coding regions reveal a monophyletic lineage of euglyphid testate amoebae (order Euglyphida). J. Eukaryot. Microbiol. 49, 108–118.