Dupont, A.Ö.C.; Griffiths, R.I.; Bell, T.; Bass, D. 2016. **Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs**. *Environmental Microbiology*, 18 (6). 10.1111/1462-2920.13220

Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs.

Dupont AÖC[1], Griffiths RI[2], Bell T[3], Bass D[1,4]

[1] Department of Life Sciences, the Natural History Museum, Cromwell Road, London, UK, SW7 5BD

[2] Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford, UK, OX10 8BB

[3] Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire, UK, SL5 7PY

[4] Cefas, Barrack Road, The Nothe, Weymouth, UK, DT4 8UB

Corresponding author: David BASS; Natural History Museum, department of Life Sciences; Cromwell Road, London SW7 5BD, UK; d.bass@nhm.ac.uk, 0207 942 5387.

Running title: Soil pH and protistan diversity.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1462-2920.13220

INTRODUCTION

Protists are key components of microbial communities in both aquatic and terrestrial environments. They represent 10^4-10^7 individuals per gram of dry soil (Adl and Coleman, 2005; Adl and Gupta, 2006; Bamforth, 2007) and over 50% of total aquatic biomass (Sherr and Sherr, 2002, 2007). The diversity of functional groups (trophic status, free-living vs symbiotic, etc.) makes them major participants of the microbial loop and regulators of biogeochemical flows (Calbet and Landry, 2004). With the continuing development and growing capability of molecular techniques, protist diversity is increasingly revealed as orders of magnitude greater than morphological or even earlier sequence-based assays suggested (Bates *et al.*, 2012). The past ten years have seen extraordinary advances in our knowledge of microbial eukaryotic diversity, primarily through the adoption of molecular tools for phylogenetically-based classification which provides a coherent evolutionary framework to explore diversity. Additionally the routine use of environmental sequencing utilising next generation sequencing technologies has permitted the discovery of many new groups and novel eukaryotic lineages in many different biomes (e.g. Takishita *et al.*, 2007; Jones *et al.*, 2011; de Vargas *et al.*, 2015). However, the challenge of overlaying ecological and biogeographical insight onto this diversity still remains, particularly in the complex and heterogeneous soil environment.

Microscopy-based soil protist diversity studies far outnumber soil eDNA studies, and both lag far behind their marine and freshwater counterparts. Diversity assessments that rely on culturing and/or visual identification have revealed a large diversity of cell forms and taxa dominated by bacterivores, predators, and some autotrophs (inferred from e.g. Stout 1984; Bamforth, 2007; Domonell *et al.*, 2013). These studies often rely on protists capable of growing in culture medium supplemented with bacteria and recognition of visually distinctive (and relatively large) forms, e.g. via liquid

aliquot isolation techniques (Domonell *et al.*, 2013). One consequence of this bias is that naked and testate amoebae, ciliates, some flagellates, diatoms, and green algae dominate the results. In some cases fungi are reported and/or the focus is specifically on heterotrophs/bacterivores. Environmental PCR and sequencing studies (as outlined above) are not taxonomically, ecologically, or visually constrained in the same ways and reveal many novel lineages including parasites (e.g. (Geisen *et al.*, 2015). Non-PCR based metagenome sequencing studies rarely feature protists, mostly focusing on bacteria (Pearce *et al.*, 2012; Fierer *et al.*, 2013) and/or aspects of metabolism. Metatranscriptomic studies (Urich *et al.* 2008; (Lehembre *et al.*, 2013; Turner *et al.*, 2013) interestingly reveal diversity profiles of active soil biota that differ in some important respects to amplicon studies, particularly demonstrating higher diversity and abundance of genetically divergent lineages (including many parasitic lineages) that are underrepresented in amplicon studies at least partly due to negative PCR biases resulting from mismatches between commonly used primer sequences and divergent templates, and amplicon length variation. However, there are so few molecular studies on soil for comparison that it is too early to generalise about soil protistan diversity, particularly because soils are so heterogeneous. Often soil-based molecular studies are primarily concerned with specific ecological situations and focus on broad changes in total eukaryote community structure and rarely look in detail at the validity of protist hits, so usually illuminate protistan diversity and distribution at relatively low taxonomic resolution (Murase and Frenzel, 2008; Turner *et al.*, 2013). Furthermore, most studies use the standard SILVA 18S (Quast *et al.*, 2013) database for eukaryotic taxonomies, while other highly curated ones such as the Protist Ribosomal Reference database (PR2; Guillou *et al.*, 2013) remain fairly unknown or unused.

The Countryside Survey (www.countrysidesurvey.org.uk), a recent multi-sample assessment of bacterial communities across the full spectrum of UK soil types

showed that bacterial community structure was strongly determined by soil variables such as soil pH (Griffiths *et al.*, 2011). Alpha diversity was positively related to pH, with greater diversity in soils of decreasing acidity. However, beta diversity was higher in acidic soils, possibly reflecting greater habitat heterogeneity in those samples. Here we produced and analysed a eukaryotic 454 sequencing SSU rDNA dataset, generated from a subset of the 2007 Countryside Survey samples. We seek to compare community structures across the three soil pH classes (low, medium, high), and contrast with patterns observed in bacterial communities. We also explored which taxonomic groups differ between different soil types, and at what level of taxonomic hierarchy differences are manifest. Finally we examined the reliability of protist taxonomic assignments by comparing the performance of different databases. We used the databases to provide an in depth evaluation of some novel groups, which are highly represented in soil 18S libraries but whose evolutionary affiliations and relationships are yet to be resolved.

RESULTS

Data processing and OTU calling

45,505 quality-filtered sequences were analysed using the QIIME pipeline. After removing singleton and chimeric sequences these were clustered into 2566 OTUs across all 15 samples. Following taxonomic affiliation based on the PR2 database sample CS11 was found to be dominated by fungi (two OTUs accounting for >75% of sequence reads) so this sample was omitted from subsequent analyses. Metazoan and Streptophyta OTUs were also removed, leaving 2284 OTUs representing nine high level protistan taxa (at taxonomic level 2 – see below) (Fig. 2). Amended OTUs and other highly divergent ones are summarised in Table 1.

Taxonomy assignment outputs are presented as an informal taxonomic hierarchy of six or seven levels depending on the reference database used (SILVA119 and PR2

respectively). Level one (L1) specifies the eukaryotic domain and is not discussed further. Subsequent levels range from L2 (approximates to supergroup/phylum) to L6. Our analysis defines OTUs at a higher resolution than this, therefore a single taxonomic profile may apply to more than one OTU. The most highly represented high ranking taxa were opisthokonts (mostly fungi), alveolates (mostly apicomplexans), and rhizarians (subphylum Filosa; (Bass and Cavalier-Smith, 2004) (Fig. 2). The ten most abundant OTUs included five fungi (the coprophile *Lasidiobolus*, *Taphrina*/*Cryptococcus* (possible pathogen), *Bannoa*/*Sporobolomyces* (yeast associated with plant leaf surfaces), *Penicillium* (common soil saprotroph; sometimes plant pathogens), a divergent possibly parasitic apicomplexan (see below), the common soil flagellates *Eocercomonas*, *Sandona*, and *Oikomonas*, and an uncharacterised divergent variosean amoebozoan.

Relationship between community structure and pH

At all taxonomic levels from L2 to OTU, there were significant differences in micro-eukaryote composition between low and high pH soils (P<0.05). This was also the case for low and medium pH soils from levels L3 to OTU (P<0.05; Table 2). There was no significant difference between medium and high pH soils at any taxonomic level. Even at phylum level (L2) the low pH soils have a distinct community structure, being dominated by opisthokonts (with a high representation of fungi), with markedly fewer rhizarian and amoebozoan OTUs than medium and high pH soils (Fig. 2). Lower in the taxonomic hierarchy (L4) differences in other groups in addition to fungi become more apparent. The low pH soils had a significantly lower total OTU count (447; average 146/sample) than medium and high pH (1247 (ave. 478) and 1314 (ave. 398) respectively, although note that high and low pH were represented by five samples and medium by four only). However, beta-diversity of the low pH soils was the highest (3.06) compared to medium (2.61) and high (2.64). Only 11 OTUs (2.5%) were detected in all low pH samples.

Low pH samples correlated positively with the first axis of a principal component analysis (Fig. 3), while medium and high pH ones correlated mostly with the negative first axis, so that samples belonging to low pH cluster together and apart from the rest. The first two axes of the analysis explained over 63% of the variance, although the projection of some samples is rather poor on those axes. Indeed, low pH samples were the strongest contributors for defining the first axis. High and medium pH samples correlate positively to different environmental variables, the strongest being bulk density (BD) and pH (ph_class); low pH samples were positively correlated to moisture and the first axis from a plant detrended correspondence analysis DCA1_2007, see Griffiths *et al.*, (2011). All other variables, although significant, were more weakly correlated ($r^2 < 0.7$).

The SIMPER results in Table 3 show the 30 OTUs contributing most strongly to protistan community differences between the different pH levels. These explained 61% of the differences between medium and high pH and low to medium pH, and 54% of the differences between low and medium pH. Of these 41 OTUs 41% are related to organisms with parasitic lifestyles, 20% related to those with pathogenic/symbiotic lifestyles associated with living plants, 20% to known saprotrophs, 17% bacterivores, and 5% photosynthetic autotrophs. The (putatively) parasitic lineages were dominated by fungi and Apicomplexa (which together accounted for 31 of the 41 OTUs) plus one mesomycetozoean. Other high SIMPER-ranking taxa included Cercozoa (2 OTUs), chlorophytes (2), Amoebozoa (2) and one stramenopile OTU. Other parasites in the taxonomic assignments in addition to those shown in Fig. 3 included other mesomycetozoeans, plasmodiophorids (Neuhauser *et al.* 2014), and kinetoplastids (*Ichthyobodo*-relative).

Some OTUs near the top of the SIMPER table (Table 3) showed striking differences in occurrence between pH levels (i.e. contributing most strongly to community differences). For example, OTU 2542 (most closely matching *Archaeorhizomyces*

finlayi, 98% identity) was strongly present in medium and high pH soils, but absent from all but one low pH sample, in which it was represented by only four sequence reads. Conversely, OTU 2440, also matching *Archaeorhizomyces finlayi* (92% identity) but with a different genotype, was more strongly represented in low pH samples. The sequences from the bacterivores *Sandona*, *Eocercomonas*, and the variosean amoeba lineage Mb5C were markedly more abundant in medium and high than low pH samples. The apicomplexan putative parasite OTUs 2376 and 2342 were also markedly more frequent in medium and high pH soils; 1787 was only found in high pH.

The taxonomic assignments showed a large number of OTUs (311) belonging to Alveolata. 59% of these grouped with parasitic Apicomplexa in a phylogenetic analysis, many of which were phylogenetically divergent (Fig. 4). The majority of the apicomplexan OTUs branched with terrestrial gregarines, but also included deep-branching relatives of lecudinids, *Selenidium*, coccidians, colpodellids, and novel lineages. The rest of the alveolate OTUs grouped with perkinsids and ciliates.

Protist community differences across samples correlated with those of bacteria (Mantel test; r= 0.509, *P*=0.001). To visualise this we plotted the bacterial and prokatyotic ordinations (NMDS; Fig. 3) as well as the pairwise correlations between the prokaryotic and eukaryotic OTUs (Fig. 5). The result showed blocks of positive and negative associations between bacterial and eukaryotic OTUs. Many of these likely reflect the shared constraints of soil pH. The figure also provides candidates for ecological interactions, including potential specialised parasite/host and predator/prey relationships.

Comparison of PR2 and SILVA taxonomy

We compared the taxonomic assignments produced using the same QIIME pipeline

on the whole dataset from two SSU rDNA databases – SILVA 119 (Quast *et al.*, 2013) and PR2 (Guillou *et al.*, 2013). At taxon level 2, which should give the most informative high-level taxonomic overview, the profiles appeared quite different (Fig. 6). This partly resulted from different composition of high-level taxa between databases – for example Stramenopiles (3%), Rhizaria (16%), and Alveolata (24%) were shown separately in the PR2 analysis, but as the supergroup SAR (38%, grouping Stramenopiles, Alveolata and Rhizaria) in Silva. However, the proportions of SAR and Opisthokonta in our results were different, depending on the database used, as some OTUs were accounted for in other groupings. Other differences result from some single lineages being represented at several taxonomic levels in Silva (e.g. BW-dinoclone28, Colponema sp. Peru, LG5-05, RT5iin25) because they are incompletely annotated across levels in the database.

DISCUSSION

We show that soil protist communities differ significantly between soils of different pH classes but to a lesser extent than bacterial communities analysed from the same samples. Low pH soils had markedly different micro-eukaryote assemblages from medium and high pH soils, whereas the latter categories were much more similar to each other. As for bacteria, protistan beta-diversity was also highest at low pH (Griffiths *et al.*, 2011). This might be a trivial expectation if protists were interacting solely with bacteria. However, only a small proportion of the protist taxa most characteristic of protist assemblage differences between the different pH levels were related to bacterivores, like many cercozoan flagellates (Bass *et al.*, 2008; Howe *et al.*, 2009, 2011); the majority were related to parasites (of animals, plants, and other eukaryotic microbes), and protist and fungi otherwise known to interact with plant rhizospheres or phyllospheres (e.g. *Taphrina*, *Polymyxa*, *Archaeorhizomyces*; Table 2). Therefore, the ecological distribution of both above- and below-ground larger organisms appear to play strong roles in the determination of soil protist community

structure, articulated by saprotrophy, coprophily, parasitism, and symbiosis (e.g. ectomycorrhizal fungi and rhizosphere-associated protists). Correlation analyses showed strong variation in co-occurrence between protistan and bacterial OTUs. Negative or positive correlations might simply be explained by shared preference of members of each domain for certain environmental conditions. However other interactions, for example preferential grazing of bacteria by protists (Chrzanowski and Šimek, 1990; Glücksman *et al.*, 2010), antagonistic interactions such as chemical and morphological defence (Jürgens and Matz, 2002), pathogenicity, competition, etc., and synergistic interactions such as trophic cascades (Brussaard, 1997; Corno *et al.*, 2013) offer more biologically complex and powerful explanations for the related responses of both domains to pH level differences in their environment.

Detailed taxonomic interpretation of the OTUs revealed an interesting diversity of novel and recently characterized lineages, many of which appear to be soil specialists, perhaps important in biological processes specific to this habitat. For example, Archaeorhizomycetes, a recently described class of soil fungi (Rosling *et al.*, 2011), was represented by 29 OTUs, some of which contributed relatively strongly to micro-eukaryote assemblage differences between pH classes. At least some Archaeorhizomycetes are associated with plant roots (Rosling *et al.*, 2011). Our data suggests that distribution of members of this group is also influenced by pH, perhaps by being associated with plants characteristic of different soil types. The summary of the most divergent valid OTUs in Table 1 shows that these belong to Cercozoa, many members of which are known to be important in soils (Bass *et al.*, 2008; Howe *et al.*, 2009; A. Howe *et al.*, 2011), Alveolata – most of which are Apicomplexa, shown on Fig. 4 and discussed more below, novel parasitic mycetozoans and putative kinetoplastids, fungi (unsurprisingly; (Richards and Bass, 2005; Bass and Richards, 2011), and amoebozoans, which harbour a large and most

uncharacterized diversity in soils (Berney *et al.*, 2015). One amoebozoan OTU, affiliated to the lineage Mb-5c, is most related to *Arboramoeba*, a very recently described genus of large, network-forming variosean amoebae (Berney *et al.*, 2015), and which was a high-ranking discriminator between low and other pH categories in the SIMPER analysis (Table 3). Thirty other OTUs were also affiliated with *Arboramoeba*. When blastn-searched against the nt database in GenBank, many sequences in Table 1 and other taxonomically uncertain OTUs from this study returned environmental sequences generated by other soil eDNA studies, particularly Lehembre *et al* (2013) and the taxonomically unfortunately mis-annotated study by (Lesaulnier *et al.*, 2008; Bass and Richards, 2011), strongly indicating that many protist lineages found preferentially or exclusively in soils, often phylogenetically distinct from currently characterized lineages, await discovery.

Particularly interesting are five mutually related OTUs which our eukaryote-wide analysis (see Methods) show branch within Labyrinthulea, a class of often fungal-like stramenopiles, many of which are decomposers or parasites. More specifically they are related to two more environmental clades – one from soil, the other soil and freshwater, clustering at the base of the Amphifilidae clade, which apart from the marine *Amphifila marina* comprises all freshwater environmental sequences (Anderson and Cavalier-Smith, 2013; Takahashi *et al.*, 2014). The phylogenetic position of a representative three OTUs from this clade are shown on Fig. 7; although the branch leading to these does not look that long Table 1 shows that these have only 76-78% sequence similarity with the next most closely related sequences in GenBank. This phylogenetic analysis suggests that these organisms may also be filopodial thecate amoebae but their actual phenotype and ecology can only be confirmed when they are directly observed. Other notable highly divergent OTUs in Table 1 include several with no discernable affiliation, some novel putative excavate sequences (OTUs 459, 518, 526?), and endomyxans (OTUs 920 & 1878), which

may be plant or animal parasites or free-living filose/reticulose amoebae (Bass *et al.*, 2009).

Another group of interest that also accounted for many highly divergent OTUs was Apicomplexa (Table 1; Fig. 4), a phylum including a vast diversity of obligate parasites, including the causative agents of malaria, coccidiosis, cryptosporidiosis, and toxoplasmosis. Within Apicomplexa are the Gregarines, unicellular parasites of terrestrial, freshwater, and marine habitats, which form very widely distributed and resistant cysts (Rueckert *et al.*, 2011) and have the largest variation of rDNA evolution rates of any eukaryote group (Cavalier-Smith 2014). Most apicomplexan diversity is thought to be marine (Rueckert *et al.*, 2010), but there is increasing evidence of their extreme (and often separate) diversity in soils (Bates *et al.*, 2012). We detected 147 gregarine OTUs, the majority of which grouped with (but often highly distinctly from) known terrestrial gregarines, which cluster in two clades (Rueckert *et al.*, 2011; Wakeman and Leander, 2013) that in some phylogenetic trees group together (Wakeman and Leander, 2012). Notably, apicomplexan OTUs dominate the diversity detected in sample CS13, including a high representation of OTU 2376, which Fig. 4 shows branches in the Terrestrial Gregarines I clade. Local concentrations of host individuals/material may account for the dominance of gregarines in this sample, which may also be the case to varying extents in other samples.

Apicomplexans provide a good illustration of cases where databases are very incomplete and/or taxonomic marker genes very divergent; for these a taxonomic annotation based on phylogenetic inference is far more informative than sequence affinity measures, and often essential. However, it is important to remember that the resolution of such analyses is limited due to the HTS read lengths. Nonetheless, to

our knowledge Fig. 4 is the first phylogenetic analysis of apicomplexan diversity detected as part of a soil HTS study.

Other OTUs putatively from parasites included plant root-infecting plasmodiophorids (27 OTUs), a group that includes the causative agents of clubroot in *Brassica* spp, powdery potato scab, and virus-vectoring parasites (Neuhauser *et al.*, 2014), labyrinthulids other than the divergent group discussed above (87), Mesomycetozoea including 24 ichthyosporean OTUs, many fungi including 105 cryptomycotan and 106 chytrid OTUs, oomycetes and hyphochytrids (17), and single-figure numbers of perkinsid relatives, metamonad gut symbionts, and kinetoplastids. Some further OTUs grouped within or were related to parasitic groups that could not be clearly affiliated, e.g. Holozoa (del Campo *et al.*, 2012), and Endomyxa (including the highly divergent OTUs 920 and 1878; Table 1), which includes predatory and parasitic amoebae (Hess *et al.*, 2012; Berney *et al.*, 2013) and ascetosporean invertebrate parasites (Hartikainen *et al.*, 2014) in addition to plasmodiophorids and their relatives. We also detected and expanded the known diversity of an uncharacterised apicomplexan clade, predatory colpodellids, and novel diversity within perkinsids, which were also earlier thought to be exclusively marine but environmental diversity sequencing studies have also shown to be diverse in freshwater habitats (Bråte *et al.*, 2010). Our evidence suggests that these putative parasites are also frequent in soils, perhaps with small invertebrate or micro-eukaryote hosts. It is clear that parasite/symbiont diversity in soils is highly undersampled and its potential role as a reservoir of pathogens relevant to agriculture, silviculture, aquaculture understudied. The majority of the 'parasitic' OTUs sequenced were clearly distinct from named organisms, and often also from environmental sequences in GenBank (even if they didn't meet the criteria for inclusion in Table 1), and therefore inferring lifestyles of these novel and otherwise unknown organisms should remain tentative until more information is available.

In general, we cannot assume that all members of clades including known parasites are also parasitic, and inferring function based on environmental sequence data/phylogenetic position alone is risky unless the sequence identity to thoroughly characterised lineages is high and appropriately resolving. Groups partly comprising parasites may also include symbionts for which detrimental parasitism (pathology) has not been demonstrated (e.g. some plasmodiophorids), and other trophic strategies – saprotrophism being a frequent example (e.g. oomycetes, fungi, labyrinthulids). Similarly, groups known to be generally bacterivorous based on evidence from culture isolation studies (e.g. cercomonads and glissomonads; Bass *et al* 2009b; Howe *et al* 2009) may also contain lineages with quite different lifestyles (e.g. the algivorous viridiraptorid glissomonads (Hess and Melkonian, 2013).

In terms of general micro-eukaryotic soil diversity our results are in agreement with previous sequencing-based studies, showing a high proportion of fungi, alveolates, and rhizarians. Recent studies (Urich *et al.*, 2008, Geisen *et al.*, 2015) showed a similar diversity profile by sequencing the soil metatranscriptome, (a good indicator of active cells as opposed to dormant or dead forms), and also that parasitic lineages are more abundant than many had assumed. For instance, strongly represented in Urich *et al.* (2008) data were the plasmodiophorid plant parasites, which are not conducive to culturing or cell isolation diversity studies and whose environmental diversity is much greater than host-oriented studies and those of economically important taxa would suggest (Neuhauser *et al.* 2014). Alveolates were also well represented in all sequence based studies; (Bates *et al.*, 2012) noted that a significant proportion of their OTUs affiliated with Apicomplexa. Comparison of DNA- and RNA-derived studies of soil apicomplexans will be important to distinguish between encysted and actively infecting forms (Rueckert *et al.*, 2011).

Even though short HTS-generated sequences have inherently low phylogenetic resolution, a combined approach to their taxonomic affiliation using both sequence similarity matching and phylogenetic analyses can provide more resolution and accuracy than blast-based methods alone. Further biological interpretation is possible via functional inference based on the resulting taxon profiles. We emphasise the need for phylogenetic moderation of raw taxon assignment outputs. It is important to acknowledge the significance of the percentage similarity between query and subject sequences. A SSU rDNA match of 95% or less (which dominate most HTS protistan diversity analyses) to a named database sequence is almost certainly not the species specified in the subject ID (if one is given) and may well not be the same genus. Below 85-90% assignments in the lower half of the taxonomic hierarchy become very doubtful. Here phylogenetic analyses can help, but are limited by both the signal carried by the OTU sequence fragment and database representation of related sequences. Databases themselves also powerfully influence perception of community structures. We directly compared the taxonomic profile outputs of two publically available and commonly used databases, Silva 119 (Quast *et al.*, 2013) and PR2 (Guillou *et al.*, 2013) without any further taxonomic analyses or interpretation, and show that the results differ, at least at some levels of taxonomic resolution. This is due to different taxonomic structures adopted by the two databases, different relative representation of taxonomic groups within them, and incomplete and/or incorrect annotations, e.g. the single lineages BWdinoclone28 and Colponema sp. Peru, appearing as high level lineages because of the absence of a higher level taxonomic structure for them. Their different outputs might misleadingly suggest strong biological differences between communities. The enduring lack of a generally adopted, comprehensive, and uniformly high quality taxonomic database for protists hinders the emergence of a body of data that can be consistently compared across studies.

EXPERIMENTAL PROCEDURES

Sample details; DNA amplification and sequencing

15 soil DNA samples (Fig.1) from the 2007 Countryside Survey (Griffiths *et al.*, 2011) representing 5 replicates each of low (pH 4.23 ±0.23), medium (pH 6.15 ±0.08) and high (pH 8.28 ±0.16) soil pH categories (Fig. 1) Primer sets EukA7F 5'-AACCTGGTTGATCCTGCCAGT-3' (Medlin *et al.*, 1988) and Euk570R 5'-GCTATTGGAGCTGGAATTAC-3' (Weekers *et al.*, 1994) were used to amplify a ~600bp product covering the V1 to V3 region of the 18S rRNA gene. Bacterial 16S rRNA genes were assessed using the primer sets 28F (GAGTTTGATCNTGGCTCAG) and 519R (GTNTTACNGCGGCKGCTG) as described in (Dowd *et al.*, 2008). Amplicons were sequenced in the forward direction by microbial tag-encoded pyrosequencing utilising a Roche 454 FLX instrument (Roche 454 Life Sciences, Branford, CT, USA).

Sequence processing and taxonomic affiliation

The resulting sequences obtained from 454 pyrosequencing were analysed using the QIIME software (Caporaso *et al.*, 2010). Data quality filtering removed sequences with length under 150bp, mean quality score lower than 25, those with no primer or with primer mismatches and with homopolymers over 6 nucleotides. Sample sequences were then de-multiplexed based on their barcode sequences. The subsequent library was assigned into Operational Taxonomic Units (OTU) with Uclust at 97% pairwise sequence similarity and no reverse strand matching. Representative sequences were picked up as the most abundant sequences in each OTU, and an OTU table was generated. Rarefaction of the OTU table was obtained with rarefy() function from the vegan package in R. Samples were rarefied to the level representing the lowest number of sequences across all samples, for both bacterial and prokaryotic OTU tables. Taxonomic assignments were obtained by

BLASTn (Altschul *et al.*, 1997) searches of the representative set against the PR2 database (Protist Ribosomal Reference database, (Guillou *et al.*, 2013) and the SILVA 119 database for 18S data (Quast *et al.*, 2013).

Phylogenetic analyses

Sequence alignments were generated using the e-ins-i algorithm of MAFFT alignment online (Katoh and Standley, 2013). Phylogenetic trees were built using RAxML-BlackBox (Stamatakis *et al.*, 2008) on the Cipres Science Gateway Portal (Miller *et al.*, 2010). The ML analyses used the GTR model with CAT approximation (all parameters estimated from the data); bootstrap values were mapped onto the tree with the highest likelihood value. After taxonomic affiliation OTUs corresponding to metazoans and plants species were removed prior to further analyses. Where Blast matches were below the thresholds specified (e-value <1e-30 and percentage identity 90%) a “No Blast Hit” report was produced. These were blasted separated against the NCBI GenBank nr/nt database and analysed phylogenetically in a RAxML tree of a selection of 500 eukaryotic 18S sequences including representatives of all supergroups as well as phylogenetically poorly resolved lineages, downloaded from GenBank and aligned (results not shown). Where taxonomic affiliation was then possible at some level of the taxonomic hierarchy the taxonomic affiliation results were amended. Highly divergent and/or taxonomically unresolved OTUs are shown in Table 1. In other cases the sequences were clearly not 18S rRNA genes, or were putatively chimeric/artefactual and were therefore removed.

Some OTUs were unassignable using the QIIME pipeline and returned “none” or “no blast hit”. Manual re-blasting showed some of these to be closely related to characterised lineages in well-established groups and the taxon assignments duly amended.

Statistical analyses

Statistical analyses were carried on the R software version 2.15.1 (R Core development Team, 2005), under the Vegan 2.0-8 (Oksanen *et al.*, 2013) and FactoMineR 1.25 (Lê *et al.*, 2008) packages. Similarity percentage (SIMPER) and analysis of similarity (ANOSIM) analyses, using Bray-Curtis dissimilarities, were carried out in the R software, within Vegan.

Acknowledgements

AD holds a PhD studentship supported by The Natural history Museum and University College London. RG acknowledges funding from the UK Natural Environment Research Council (standard grant NE/E006353/1). DB was supported by NERC grants NE/H000887/1 and NE/H009426/. TB is supported by a Royal Society University Research Fellowship.

REFERENCES

Adl and Gupta,V. (2006) Protists in soil ecology and forest nutrient cycling. *Canadian Journal of Forest Research* **36**: 18051817.

Adl,S. and Coleman,D. (2005) Dynamics of soil protozoa using a direct count method. *Biology and Fertility of Soils* **42**: 168171.

Anderson, OR and Cavalier-Smith, T (2013) Ultrastructure of Diplophrys parva, a new small freshwater species, and a revised analysis of Labyrinthulea (Heterokonta). *Acta* ….

Bamforth,S. (2007) Protozoa from aboveground and ground soils of a tropical rain forest in Puerto Rico. *Pedobiologia* **50**: 515525.

Bass,D. and Cavalier-Smith,T. (2004) Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY* **54**: 2393–2404.

Bass,D., Chao,E., Nikolaev,S., Yabuki,A., Ishida,K.-I., Berney,C., *et al.* (2008) Phylogeny of novel naked Filose and Reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea revised. *Protist* **160**: 75–109.

Bass,D., Howe,A., Mylnikov,A., Vickerman,K., Chao,E., Smallbone,J., *et al.* (2009) Phylogeny and Classification of Cercomonadida (Protozoa, Cercozoa): Cercomonas, Eocercomonas, Paracercomonas, and Cavernomonas gen. nov. *Protist* **160**: 483521.

Bass,D. and Richards,T. (2011) Three reasons to re-evaluate fungal diversity “on Earth and in the ocean.” *Fungal Biology Reviews* **25**: 159164.

Bates,S., Clemente,J., Flores,G., Walters,W., Parfrey,L., Knight,R., and Fierer,N. (2012) Global biogeography of highly diverse protistan communities in soil. *The ISME Journal* **7**: 652–659.

Berney,C., Geisen,S., Wichelen,J., Nitsche,F., Vanormelingen,P., Bonkowski,M., and Bass,D. (2015) Expansion of the “Reticulosphere”: Diversity of Novel Branching and Network-forming Amoebae Helps to Define Variosea (Amoebozoa). *Protist* **166**: 271–295.

Berney,C., Romac,S., Mahé,F., Santini,S., Siano,R., and Bass,D. (2013) Vampires in the oceans: predatory cercozoan amoebae in marine habitats. *The ISME Journal* **7**: 2387–2399.

Bråte,J., Logares,R., Berney,C., Ree,D., Klaveness,D., Jakobsen,K., and Shalchian-

Tabrizi,K. (2010) Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *The ISME journal* **4**: 1144–53.

Brussaard,L. (1997) Biodiversity and Ecosystem Function in Soil. *Ambio* **26**: 563–570.

Calbet,A. and Landry,M.R. (2004) Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography*.

Del Campo,J., Not,F., Forn,I., Sieracki,M., and Massana,R. (2012) Taming the smallest predators of the oceans. *The ISME Journal* **7**: 351–358.

Caporaso,G., Kuczynski,J., Stombaugh,J., Bittinger,K., Bushman,F., Costello,E., *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**: 335–6.

Chrzanowski,T. and Šimek,K. (1990) Prey-size selection by freshwater flagellated protozoa. *Limnology and Oceanography* **35**:

Corno,G., Villiger,J., and Pernthaler,J. (2013) Coaggregation in a microbial predator–prey system affects competition and trophic transfer efficiency. *Ecology* **94**: 870881.

Domonell,A., Brabender,M., Nitsche,F., Bonkowski,M., and Arndt,H. (2013) Community structure of cultivable protists in different grassland and forest soils of Thuringia. *Pedobiologia* **56**: 17.

Dowd,S., Callaway,T., Wolcott,R., Sun,Y., McKeehan,T., Hagevoort,R., and Edrington,T. (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiology* **8**: 125.

Fierer,N., Ladau,J., Clemente,J., Leff,J., Owens,S., Pollard,K., *et al.* (2013) Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science (New York, N.Y.)* **342**: 621–4.

Geisen, Laros, Vizcaíno, Bonkowski, and Groot (2015) Not all are free-living: high-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Molecular Ecology* **24**: 4556–4569.

Geisen,S., Tveit,A., Clark,I., Richter,A., Svenning,M., Bonkowski,M., and Urich,T. (2015) Metatranscriptomic census of active protists in soils. *The ISME Journal* **9**: 2178–2190.

Glücksman,E., Bell,T., Griffiths,R., and Bass,D. (2010) Closely related protist strains have different grazing impacts on natural bacterial communities. *Environmental Microbiology* **12**: 3105–3113.

Griffiths,R., Thomson,B., James,P., Bell,T., Bailey,M., and Whiteley,A. (2011) The bacterial biogeography of British soils. *Environmental Microbiology* **13**: 1642–1654.

Guillou,L., Bachar,D., Audic,S., Bass,D., Berney,C., Bittner,L., *et al.* (2013) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research* **41**: D597–D604.

Hartikainen,H., Ashford,O., Berney,C., Okamura,B., Feist,S., Baker-Austin,C., *et al.* (2014) Lineage-specific molecular probing reveals novel diversity and ecological partitioning of haplosporidians. *The ISME Journal* **8**: 177–186.

Hess,S. and Melkonian,M. (2013) The Mystery of Clade X: Orciraptor gen. nov. and Viridiraptor gen. nov. are Highly Specialised, Algivorous Amoeboflagellates (Glissomonadida, Cercozoa). *Protist* **164**: 706–747.

Hess,S., Sausen,N., and Melkonian,M. (2012) Shedding Light on Vampires: The Phylogeny of Vampyrellid Amoebae Revisited. *PLoS ONE* **7**:

Howe,A., Bass,D., Chao,E., and Cavalier-Smith,T. (2011) New genera, species, and improved phylogeny of Glissomonadida (Cercozoa). *Protist* **162**: 710–22.

Howe,A., Bass,D., Vickerman,K., Chao,E., and Cavalier-Smith,T. (2009) Phylogeny, taxonomy, and astounding genetic diversity of glissomonadida ord. nov., the dominant gliding zooflagellates in soil (Protozoa: Cercozoa). *Protist* **160**: 159–89.

Howe,A.T., Bass,D., Scoble,J.M., Lewis,R., and Vickerman…,K. (2011) Novel Cultured Protists Identify Deep-branching Environmental DNA Clades of Cercozoa: New Genera< i> Tremula</i>,< i> Micrometopion</i>,< i> Minimassisteria</i …. *Protist*.

Jones,M., Forn,I., Gadelha,C., Egan,M., Bass,D., Massana,R., and Richards,T. (2011) Discovery of novel intermediate forms redefines the fungal tree of life. *Nature* **474**: 200–203.

Jürgens,K. and Matz,C. (2002) Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. *Antonie van Leeuwenhoek* **81**: 413–434.

Katoh,K. and Standley,D. (2013) MAFFT multiple sequence alignment software

version 7: improvements in performance and usability. *Molecular Biology and Evolution*.

Lê,S., Josse,J., and Husson,F. (2008) FactoMineR: an R package for multivariate analysis. *Journal of statistical software*.

Lehembre,F., Doillon,D., David,E., Perrotto,S., Baude,J., Foulon,J., *et al.* (2013) Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes. *Environmental Microbiology* **15**: 2829–2840.

Lesaulnier,C., Papamichail,D., McCorkle,S., Ollivier,B., Skiena,S., Taghavi,S., *et al.* (2008) Elevated atmospheric CO2 affects soil microbial diversity associated with trembling aspen. *Environmental Microbiology* **10**: 926–941.

Medlin,L., Elwood,H.J., Stickel,S., and Sogin,M.L. (1988) The characterization of enzimatically amplified eukaryotic 16S-like rRNA-coding region.

Miller, MA, Pfeiffer, W, and Schwartz, T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *… Workshop (GCE)*.

Murase,J. and Frenzel,P. (2008) Selective grazing of methanotrophs by protozoa in a rice field soil. *FEMS Microbiology Ecology* **65**: 408–414.

Neuhauser,S., Kirchmair,M., Bulman,S., and Bass,D. (2014) Cross-kingdom host shifts of phytomyxid parasites. *BMC Evolutionary Biology* **14**: 33.

Pearce,D., Newsham,K., Thorne,M., Calvo-Bado,L., Krsek,M., Laskaris,P., *et al.* (2012) Metagenomic Analysis of a Southern Maritime Antarctic Soil. *Frontiers in Microbiology* **3**:

Quast,C., Pruesse,E., Yilmaz,P., Gerken,J., Schweer,T., Yarza,P., *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**: D590–D596.

Richards,T. and Bass,D. (2005) Molecular screening of free-living microbial eukaryotes: diversity and distribution using a meta-analysis. *Current opinion in microbiology* **8**: 240–52.

Rosling,A., Cox,F., Cruz-Martinez,K., Ihrmark,K., Grelet,G.-A., Lindahl,B., *et al.* (2011) Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science (New York, N.Y.)* **333**: 876–9.

Rueckert,S., Chantangsi,C., and Leander,B. (2010) Molecular systematics of marine gregarines (Apicomplexa) from North-eastern Pacific polychaetes and nemerteans, with descriptions of three novel species: Lecudina phyllochaetopteri sp. nov.,

Difficilina tubulani sp. nov. and Difficilina paranemertis sp. nov. *International journal of systematic and evolutionary microbiology* **60**: 2681–90.

Rueckert,S., Simdyanov,T., Aleoshin,V., and Leander,B. (2011) Identification of a divergent environmental DNA sequence clade using the phylogeny of gregarine parasites (Apicomplexa) from crustacean hosts. *PloS one* **6**: e18163.

Sherr,E.B. and Sherr,B.F. (2002) Significance of predation by protists in aquatic microbial food webs. *Antonie van Leeuwenhoek* **81**: 293–308.

Sherr and Sherr (2007) Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Marine Ecology Progress Series* **352**: 187–197.

Stamatakis,A., Hoover,P., and Rougemont,J. (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Systematic biology* **57**: 758–71.

Takahashi,Y., Yoshida,M., Inouye,I., and Watanabe,M. (2014) Diplophrys mutabilis sp. nov., a New Member of Labyrinthulomycetes from Freshwater Habitats. *Protist* **165**: 50–65.

Takishita,K., Tsuchiya,M., Kawato,M., Oguri,K., Kitazato,H., and Maruyama,T. (2007) Genetic Diversity of Microbial Eukaryotes in Anoxic Sediment of the Saline Meromictic Lake Namako-ike (Japan): On the Detection of Anaerobic or Anoxic-tolerant Lineages of Eukaryotes. *Protist* **158**: 51–64.

Turner,T., Ramakrishnan,K., Walshaw,J., Heavens,D., Alston,M., Swarbreck,D., *et al.* (2013) Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *The ISME Journal* **7**: 2248–2258.

De Vargas,C., Audic,S., Henry,N., Decelle,J., Mahé,F., Logares,R., *et al.* (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605.

Wakeman,K. and Leander,B. (2013) Molecular Phylogeny of Marine Gregarine Parasites (Apicomplexa) from Tube-forming Polychaetes (Sabellariidae, Cirratulidae, and Serpulidae), Including Descriptions of Two New Species of Selenidium. *Journal of Eukaryotic Microbiology* **60**: 514–525.

Wakeman,K. and Leander,B. (2012) Molecular Phylogeny of Pacific Archigregarines (Apicomplexa), Including Descriptions of Veloxidium leptosynaptae n. gen., n. sp., from the Sea Cucumber Leptosynapta clarki (Echinodermata), and Two New Species of Selenidium. *Journal of Eukaryotic Microbiology* **59**: 232–245.

Weekers, Gast, Fuerst, and Byers (1994) Sequence variations in small-subunit

ribosomal RNAs of Hartmannella vermiformis and their phylogenetic implications.

FIGURE LEGENDS

Figure 1. The countryside Survey 2007 sampling strategy across the UK. Yellow markers represent low pH samples, red ones indicate medium pH soil samples and green ones high pH samples.

Figure 2. Soil microbial diversity comparisons according to pH, per sample (bars) and pH category (pie-charts), for both supergroup/phylum (L2) and class/order (L4) levels.

Figure 3. Relationships between soil variables and microbial communities.
3a. Individuals' factor map of a principal component analysis (PCA) groups samples belonging to high pH (red) and medium (green) pH soils together, but apart from low pH (yellow) ones.
3b. The variables' factor map of the PCA correlates low pH samples positively to moisture (first axis), while medium and high pH ones correlate mostly with bulk density (BD).
3c,d. Bacterial and protistan OTUs ordination (respectively) according to pH groups. Although protestant OTUs cluster together according to the group they belong – high, medium or low – this is much clearer for the bacterial ones. Indeed, the latter separate clearly according to pH groups, while medium and high pH protist OTUs do not separate as clearly from each other.

Figure 4. Maximum Likelihood SSU rDNA phylogeny showing phylogenetic position of non-ciliate alveolates detected in this study. The parasitic apicomplexans occupy all branches above the dinoflagellates, syndinians, and ellobiopsids clade. Maximum Likelihood bootstrap values given where >60%. OTUs produced by this study shown

in bold. Numbers associated with vertical lines marking groups to the right of the tree indicate the total number of OTUs called by the taxonomic annotation pipeline (see Methods); those with < 2% sequence from another OTU were omitted from the tree.

Figure 5. Bacterial-eukaryote correlation matrix. Shades of blue squares indicate positive correlation between bacterial (columns) and eukaryote (rows) OTUs, while red ones indicate negative correlations.

Figure 6. Taxonomic assignment comparisons between PR2 and Silva119 SSU rDNA databases for supergroup/phylum levels.

Figure 7. Maximum Likelihood SSU rDNA phylogeny of Amphifilidae, Thraustochytriidae, and Amphitremida (Labyrinthulea, Stramenopiles), showing novel divergent soil clade detected in this study (shown in bold). This clade contains two more sequences that were omitted from the analyses as they were significantly shorter than the others. Maximum Likelihood bootstrap values given where >75% or useful for interpretation.

TABLE LEGENDS

Table 1. The most divergent 18S rDNA sequences detected in this study. Most of these were unassigned to any taxon by the QIIME procedure. The sequences are too short to be robustly resolved phylogenetically, however assignations in the Group column were estimated by their branching positions in a pan-eukaryote tree (see Methods). OTU 526 is probably chimeric. Most sequences in this table had 85% or less similarity to taxonomically characterised sequences in GenBank. In cases where

this value is >85% the corresponding match to the most probable hit (in most cases an environmental sequence) was 90% or less. In one case (OTU 947) the best match was to a named specimen in GenBank.

Table 2. Analysis of Similarity (ANOSIM) between pH levels at different taxonomic levels (based on the PR2 database). R-statistic (R) and p-values (p) for each pH level comparison are given (L: low pH; M: medium pH; H: high pH); micro-eukaryotic community composition between pH levels is significantly different when $p \leq 0.05$ (L-H all levels, L-M from taxonomic level 3).

Table 3. Similarity percentages analyses **(**SIMPER) of micro-eukaryote community differences between soil pH levels (Low-High (LH), Low-Medium (LM), Medium-High (MH)) and ranking of most influential species in the difference of compositions between pH levels.
The number following the pH level comparison code is the ranking of that OTU relevant to that comparison, e.g. LH1 is the OTU contributing most strongly to the community difference between low and high pH soils.

Supplementary OTU table. Representative set of sequences (as described in Methods) with the respective OTU number, the code of the sequence representative of the OTU and taxonomic affiliation obtained with the Protist Ribosomal Reference (PR2) database.

Table 1

OTU	Group	Closest named match on Genbank	Greatest similarity % match to GenBank sequences Env.	Accession	Charact.	Accession
1528	Cercozoa; Filosa	Placocista	96	FO181529	85	GQ144680
2308	Cercozoa; Filosa	Paulinella	93	JX456225	82	X81811
945	Cercozoa; Filosa	Gynmophrys (= Limnofila)	89	EU567223	88	FJ973365
920	Cercozoa; Endomyxa	Clathrina (env = Opisthokonta)	82	GQ844577	83	AM180960
1878	Cercozoa; Endomyxa	Metabolomonas	86	AB526173	85	HM536167
1190	Alveolata (see Fig. 4)	Gregarina	87	JN846840	84	JQ970325
334	Alveolata (see Fig. 4)	Gregarina	87	JN846840	86	JQ970325
1002	Alveolata (see Fig. 4)	Gregarina	78	JN846839	88	JQ970325
2298	Alveolata (see Fig. 4)	Gregarina	76	JN846839	75	JQ970325
529	Alveolata (see Fig. 4)	Apicomplex sp. 1	88	JN846840	87	KC890798
2360	Alveolata (see Fig. 4)	Apicomplex sp. 1	88	JN846840	88	KC890798
1689	Alveolata (see Fig. 4)	Diophrys	83	EF024740	82	EU267930
947	Alveolata (see Fig. 4)	Eimeria			89	GU479633
2554	Alveolata (see Fig. 4)	Colpodella	89	AB970393	88	AY234843
1031	Diplophrys/stramenopile	Amphifilidae sp.	78	EF023442	72	AB856528
1297	Diplophrys/stramenopile	Amphifilidae sp.	78	EF023442	72	AB856528
2291	Diplophrys/stramenopile	Amphifilidae sp.	78	EF023658	72	AB856528

328	Diplophrys/stramenopile	Amphifilidae sp.	76	KC454889	73 AB856528
1179	Diplophrys/stramenopile	Amphifilidae sp.	76	KC454889	73 AB856528
829	?	Pilobolus	91	AB970383	72 DQ211050
526	Excavata?	(Petalomonas)	77	JX069065	78 AF386635
459	Excavata?	Ichthyobodo	86	EU860484	79 KC208028
518	Excavata	Notosolenus	81	FO181403	81 KC990930
1021	?	Halichondria	87	HQ910364	81 KC899029
450	?	Halichondria	91	HQ910364	84 KC899029
630	?	Halichondria	90	HQ910364	84 KC899029
1510	Fungi	Alternaria	88	EF023366	87 KJ489375
1645	Fungi	Schizangiella	88	JX003447	88 AF368523
2122	Mesomycetozoea	Fabomonas	94	AB510393	82 JQ340335
505	Amoebozoa/Fungi	Monoblepharis	87	EF023424	88 KJ668082
51	Amoebozoa	Ceratiomyxella	88	AM409569	87 FJ544419
1824	Amoebozoa?	Glaucocystis	89	AM409569	87 X70803

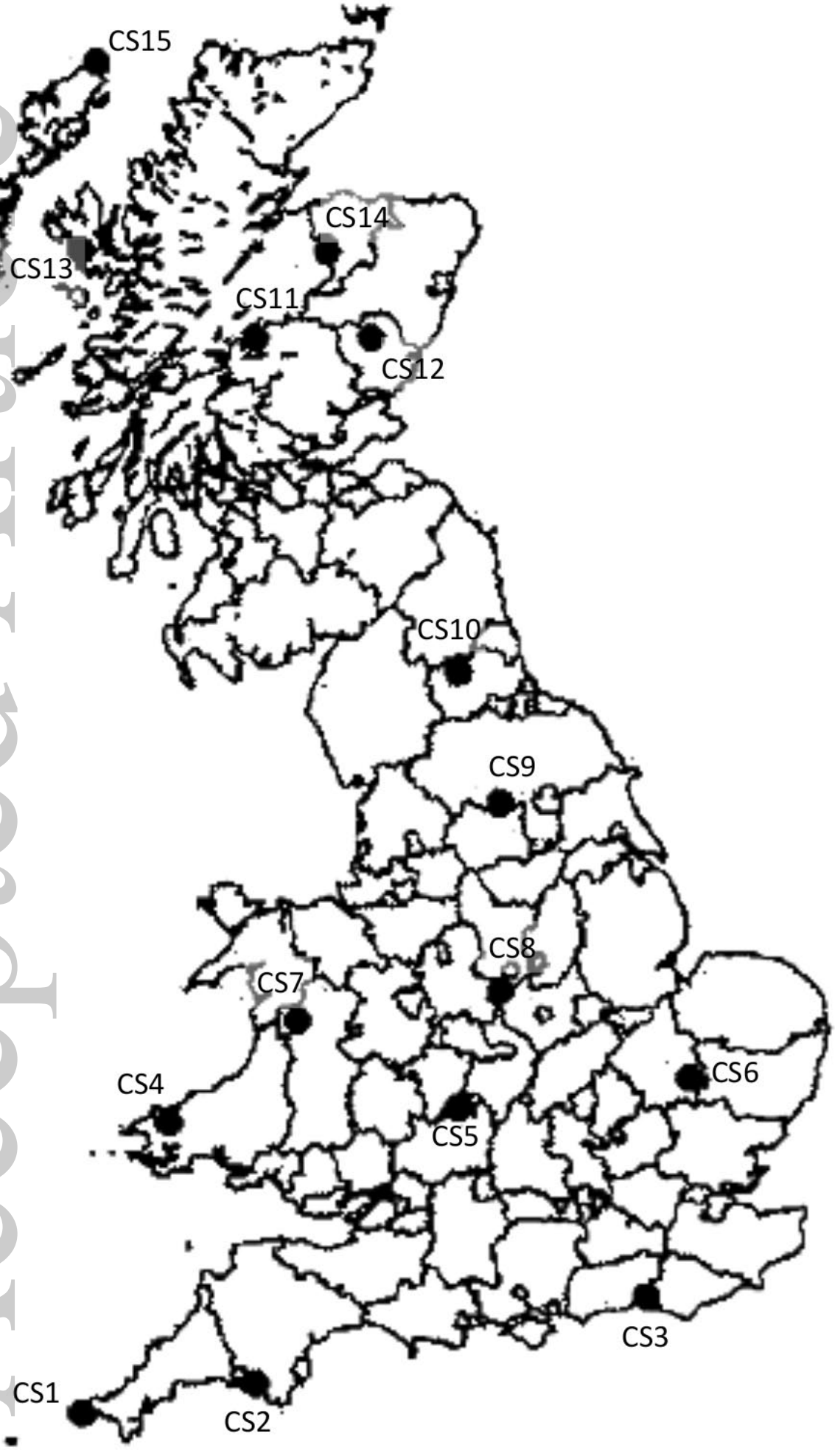
Table 2. ANOSIM comparisons between pH levels at different taxonomic levels (according to PR2 database). R-statistic (R) and p-values (p) for each pH level comparison (L: low pH; M: medium pH and H:high pH).

	pH comparison					
	L-M		M-H		L-H	
Level	R	p	R	p	R	p
L1						
L2	0.4375	0.053	0.1313	0.195	0.444	0.019
L3	0.45	0.044	0.2313	0.148	0.452	0.035
L4	0.45	0.0288	0.2313	0.114	0.452	0.024
L5	0.5438	0.021	0.3438	0.052	0.504	0.024
L6	0.5625	0.032	0.275	0.065	0.62	0.01
OTU	0.7188	0.021	0.05625	0.719	0.57	0.022

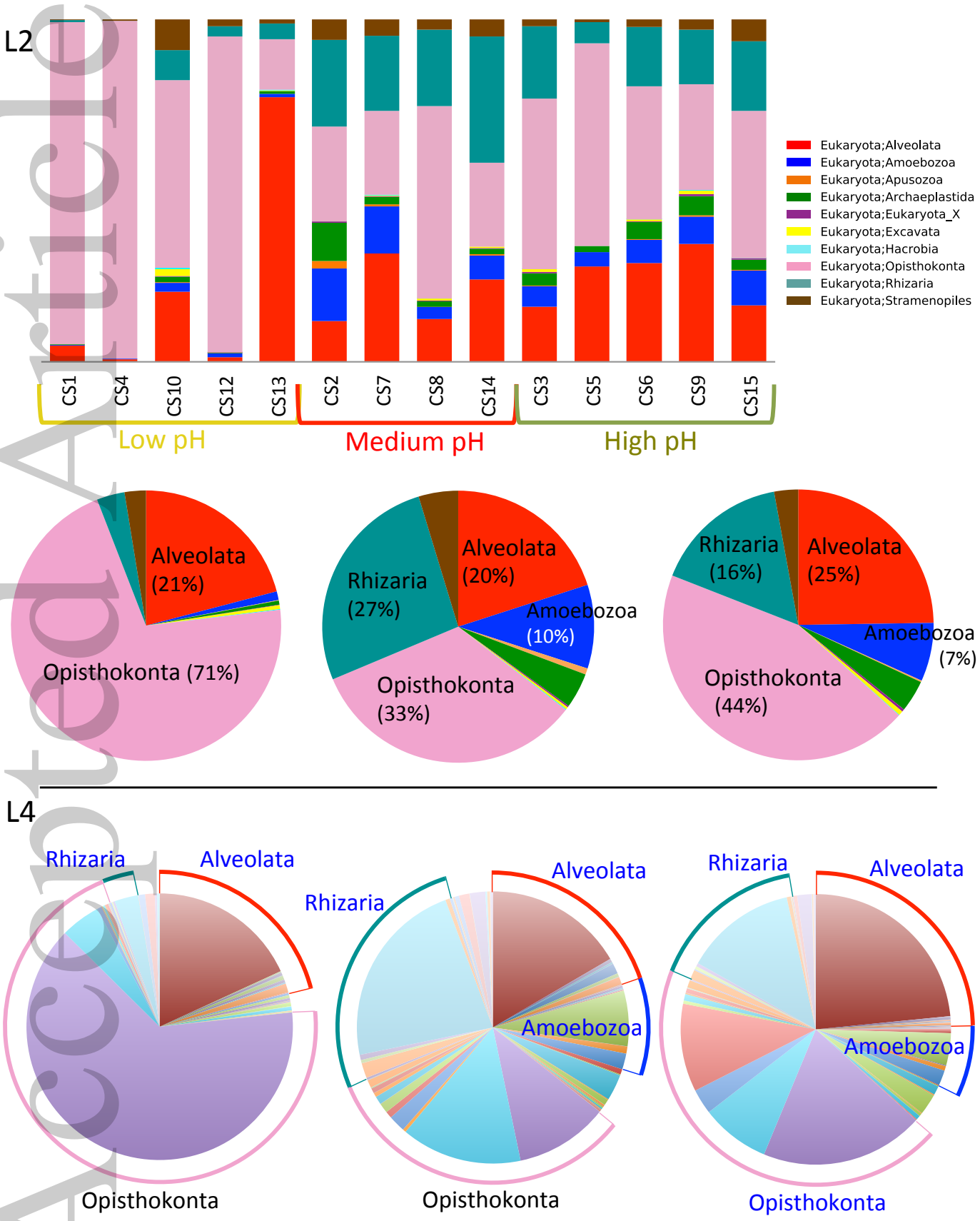
Table 3

soil type comparisons			OTU No.	Taxonomic affiliation	identity % to sequence database	Accession No.
LH1	LM4	MH2	2376	Alveolata, Apicomplexa, Gregarines_XX	98	EF024723
	LM2	MH3	2542	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	98	JF836020
LH8	LM1	MH1	280	Opisthokonta, Fungi, Ascomycota, Pezizomycetes, Lasiobolus ciliatus	100	DQ646532
LH2	LM5	MH4	962	Rhizaria, Cercozoa, Glissomonadida, Sandonidae_X	100	EU646934
LH3	LM3	MH6	1801	Opisthokonta, Fungi, uncharacterised	100	EF023474
LH6	LM7		1787	Alveolata, Apicomplexa, Gregarines_XX	90	EF024723
LH4	LM6	MH11	147	Opisthokonta, Fungi, Ascomycota, Pezizomycotina, Penicillium sp.	100	GU190185
LH7	LM8	MH7	2342	Alveolata, Apicomplexa, Gregarines_XX	95	GQ462637
LH5	LM13	MH5	1052	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Mrakia frigida	100	AB032665
LH9	LM9	MH9	38	Rhizaria, Cercozoa, Cercomonadida, Eocercomonas sp.	100	EF023536
LH10	LM11	MH8	612	Amoebozoa, Variosea, Mb5C-lineage	100	AB425950
	LM10	MH10	2197	Opisthokonta, Fungi, Ascomycota, Taphrinomycotina, Taphrina johansonii	92	AJ495835
LH11	LM12		163	Opisthokonta, Fungi, Chytridiomycota, Rhyzophidiales_X	99	GQ995433
LH13		MH13	1691	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Cryptococcus dimennae	100	AB032627
LH14	LM15		2135	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Catathelasma ventricosum	98	DQ435811
	LM16	MH14	2440	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	95	GQ404765
LH12	LM22	MH12	342	Alveolata, Apicomplexa, Gregarines_XX	98	EF024723
LH18		MH15	809	Rhizaria, Cercozoa, Plasmodiophorida, Polymyxa graminis	100	AF310898
LH15	LM14	MH22	2539	Alveolata, Apicomplexa, Gregarines_XX	100	EF024926

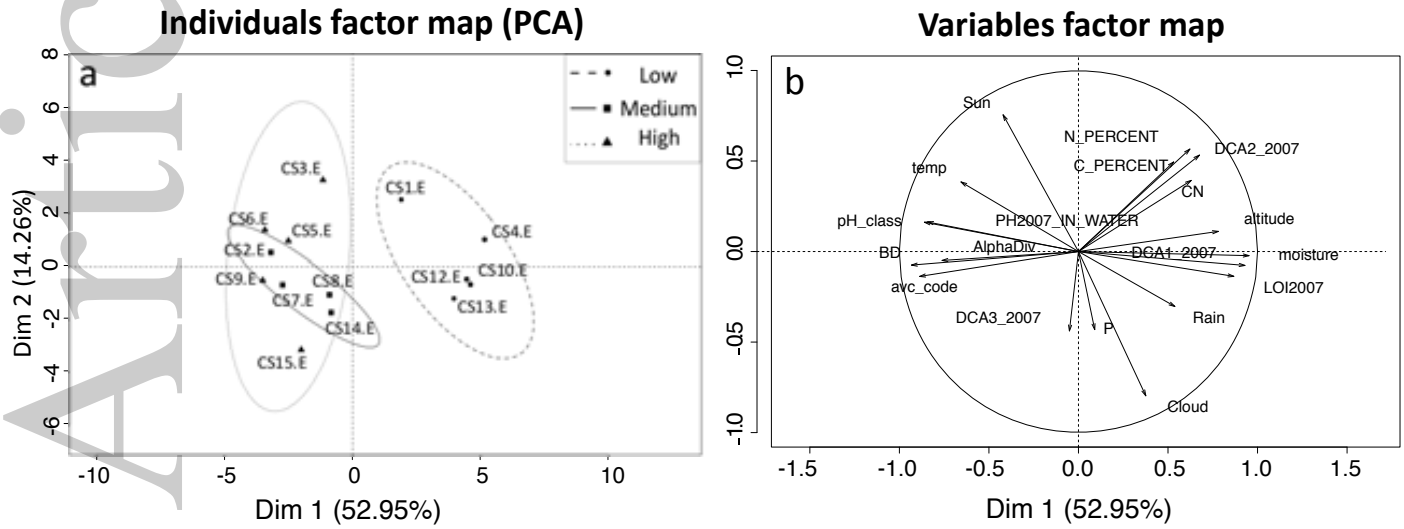
	LM20	MH17	216	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	98	JF836020
LH16	LM23		1353	Alveolata, Apicomplexa, Gregarines_XX	93	EF024926
	LM19	MH20	2157	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Camarophyllopsis hymenocephala	99	DQ444862
	LM21	MH19	554	Opisthokonta, Fungi, Ascomycota, Archaeorhizomyces finlayi	95	JF836020
LH23	LM18	MH21	2501	Rhizaria, Cercozoa, Cercomonadida, Paracercomonas sp.	100	AM114800
LH29	LM17	MH16	738	Amoebozoa,Tubulinea, Nolandellidae_X	99	EF023499
LH17	LM29	MH18	2412	Archaeplastida, Chlorophyceae, Oedocladium prescottii	100	DQ078298
LH19	LM24		1850	Alveolata, Apicomplexa, Gregarines_XX	97	EF024723
LH20		MH23	777	Opisthokonta, Fungi, Mortierellales, Mortierella sp.	100	EF023700
LH21		MH25	2187	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Asterotremella longa	97	AB035586
LH22		MH24	2565	Alveolata, Ciliophora, Litostomatea, Enchelys polynucleata	99	DQ411861
LH24			2024	Alveolata, Apicomplexa, Gregarines, Ascogregarina taiwanensis	90	DQ462455
LH25			2194	Alveolata, Apicomplexa, Coccidia, Cryptosporidium serpentis	94	AF093500
LH26	LM25		1039	Opisthokonta, Fungi, Ascomycota, Pezizomycotina, Verticillium albo-atrum	100	ABPE01001453
	LM26		2069	Opisthokonta, Fungi, Cryptomycota_X	100	AB695466
	LM28	MH26	2321	Opisthokonta, Mesomycetozoa, Ichthyosporea, Ichthyophonida sp.	100	AJ130859
LH28	LM27		283	Opisthokonta, Fungi, Chytridiomycota, Chytridiomycotina, Rhyzophidiales_X	98	DQ244005
		MH28	2276	Stramenopiles, Chrysophyceae-Synurophyceae, Clade-C_X	100	EF023425
LH27	LM30		2360	Alveolata, Apicomplexa, Gregarines_XX	88	KC890798
LH30		MH27	970	Archaeplastida, Chlorophyceae, Sphaeropleales_X	100	EF023843
		MH29	448	Opisthokonta, Fungi, Basidiomycota, Pucciniomycotina, Bannoa sp.	98	DQ631899
		MH30	422	Opisthokonta, Fungi, Basidiomycota, Agaricomycotina, Austropaxillus sp.	99	DQ534673

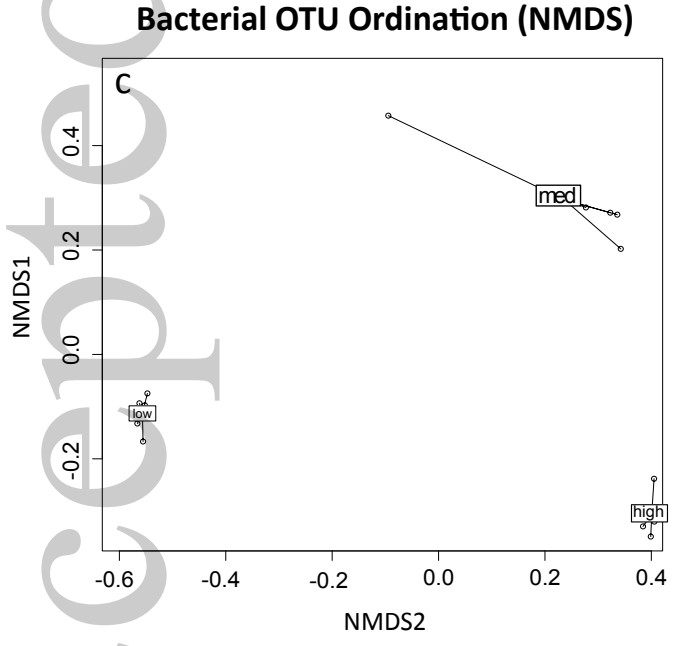


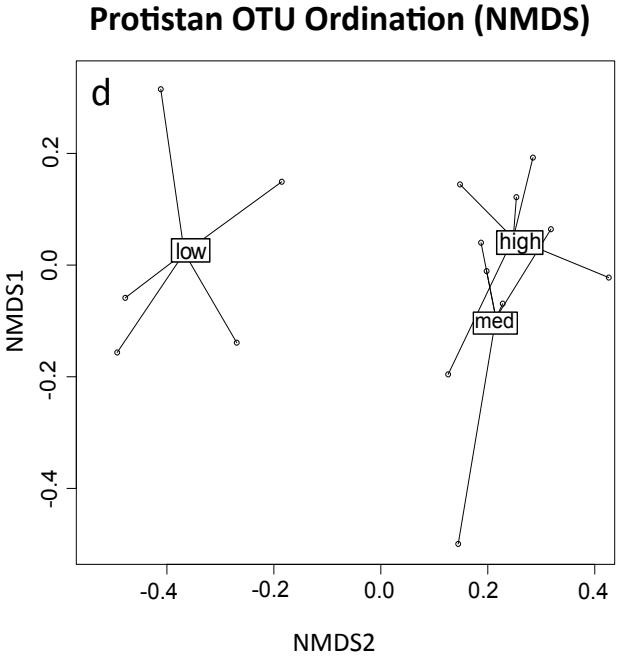
EMI_13220_F1.eps



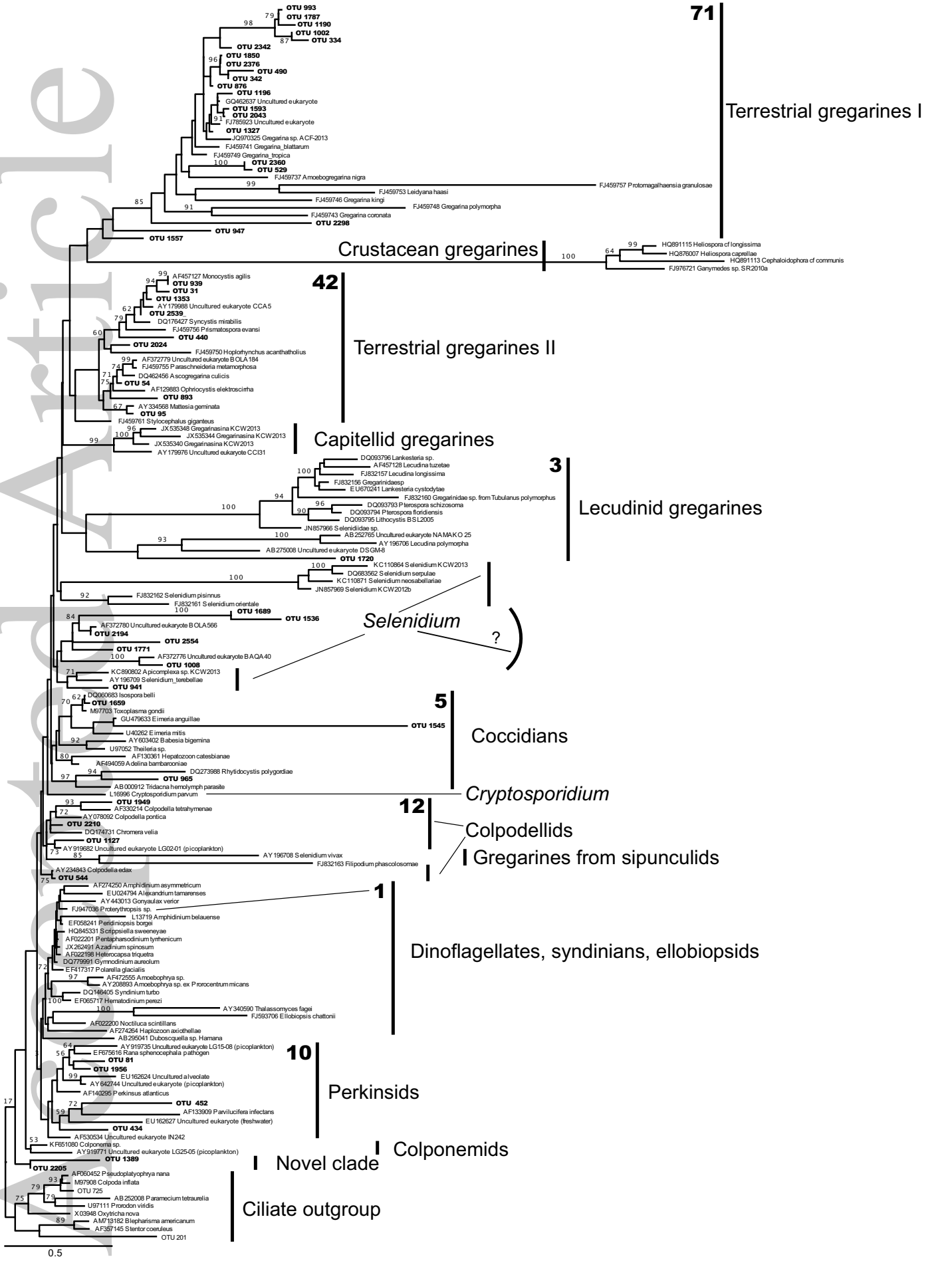
EMI_13220_F2.pdf



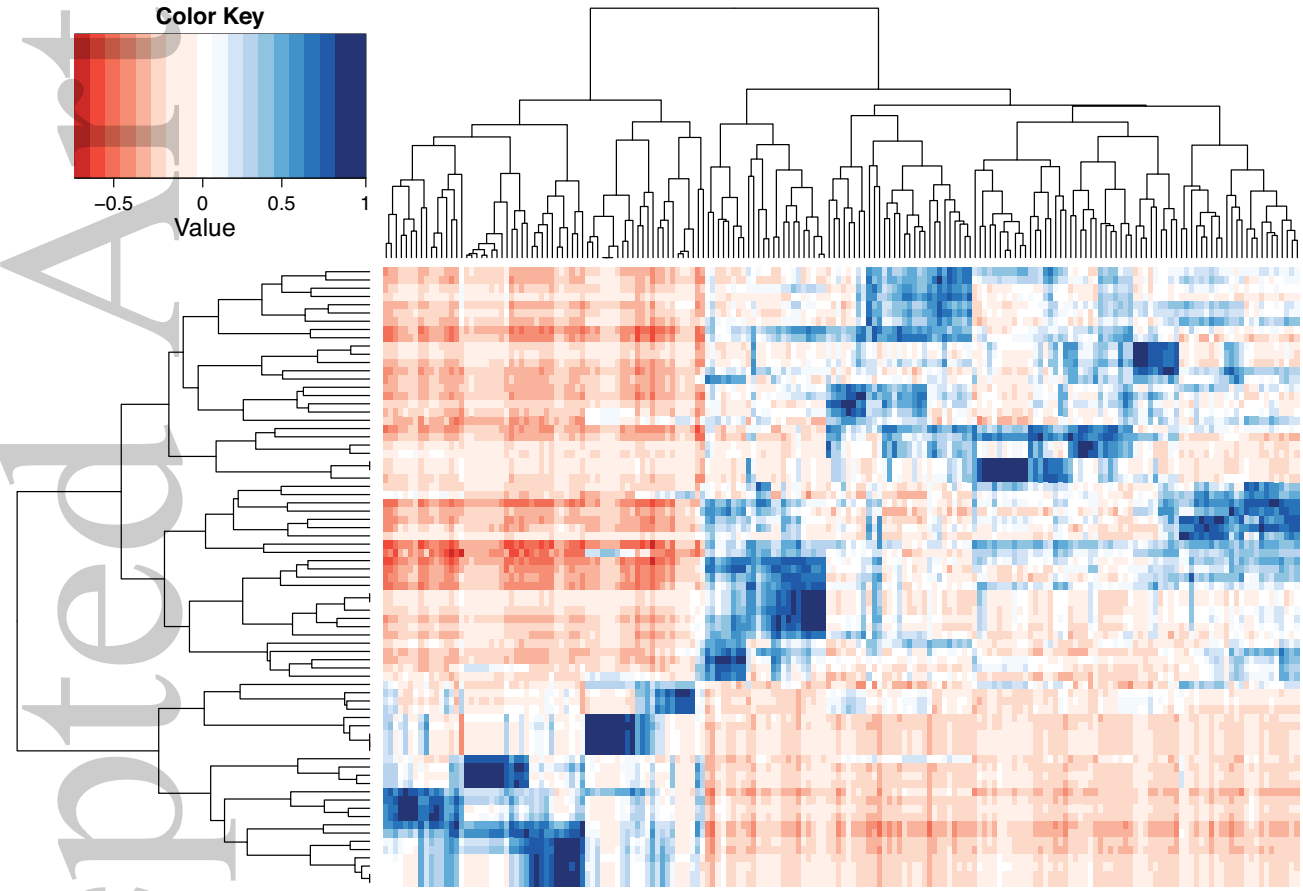




EMI_13220_F3.pdf

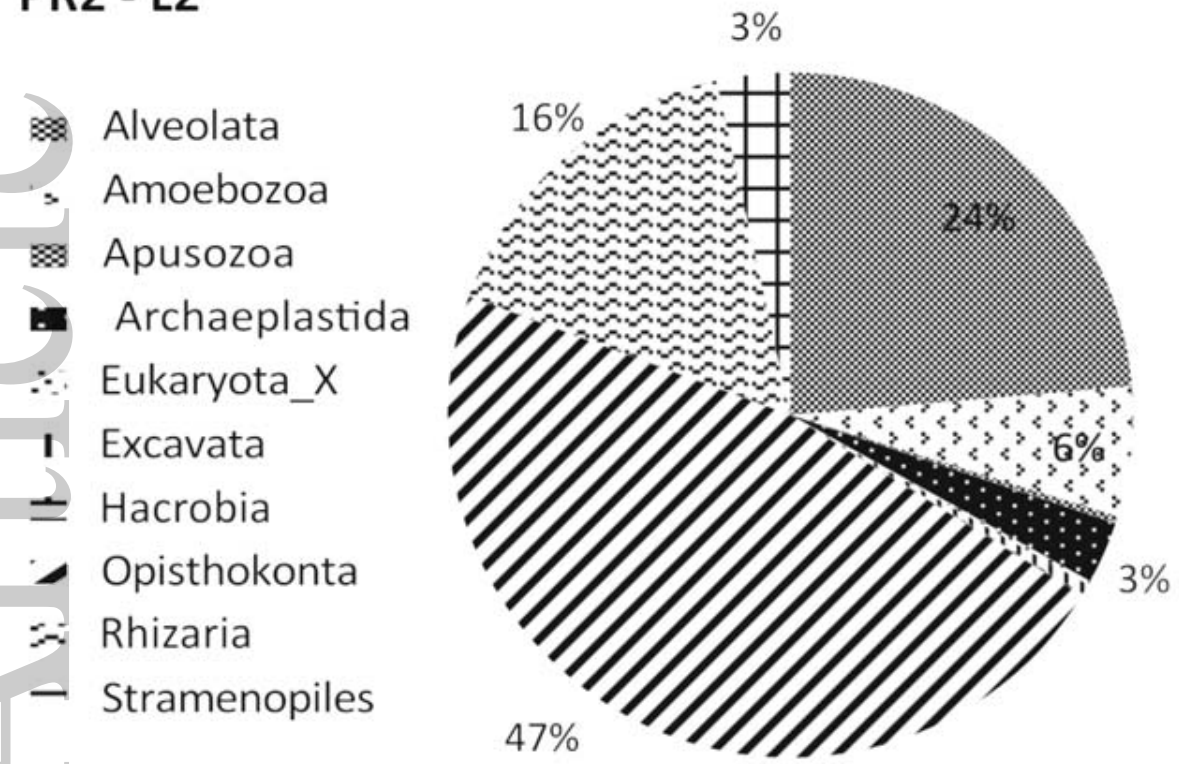


EMI_13220_F4.pdf



EMI_13220_F5.pdf

PR2 - L2



SILVA119 - L2