DNA metabarcoding reveals modern and past eukaryotic communities in a high-mountain peat bog system

Garcés-Pastor, Sandra [a,b]; Wangensteen, Owen S. [c,d]; Pérez-Haase, Aaron [a,e]; Pèlachs, Albert [f]; Pérez-Obiol, Ramon [g]; Cañellas-Boltà, Núria [h]; Mariani, Stefano [c]; Vegas-Vilarrúbia, Teresa [a].

[a] Department of Evolutionary Biology, Ecology and Environmental Sciences, Universitat de Barcelona, Barcelona, Spain

[b] Current address: Tromsø Museum, UiT The Arctic University of Norway, Tromsø, Norway.

[c] Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, Greater Manchester, UK

[d] Current address: Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway

Center for Advanced Studies of Blanes, Spanish Research Council (CEAB-CSIC), Blanes, Spain

[f] Department of Geography, Universitat Autònoma de Barcelona, Spain

[g] Botany Unit, Department of Animal Biology, Plant Biology and Ecology, Universitat Autònoma de Barcelona, Spain

[h] Institute of Earth Sciences Jaume Almera (ICTJA-CSIC), Barcelona, Spain

Abstract

Peat bogs located in high mountains are suitable places to study local environmental responses to climate variability. These ecosystems host a large number of eukaryotes with diverse taxonomic and functional diversity. We carried out a metabarcoding study using universal 18S and COI markers to explore the composition of past and present eukaryotic communities of a Pyrenean peat bog ecosystem. We assessed the molecular biodiversity of four different moss micro-habitats along a flood gradient in the lentic Bassa Nera system (Central Pyrenees). Five samples collected from different sediment depths at the same study site were also analysed, to test the suitability of these universal markers for studying paleoecological communities recovered from ancient DNA and to compare the detected DNA sequences to those obtained from the modern community. We also compared the information provided by the sedimentary DNA to the reconstruction from environmental proxies such as pollen and macro-remains from the same record. We successfully amplified ancient DNA with both universal markers from all sediment samples, including the deepest one (~10,000 years old). Most of the metabarcoding reads obtained from sediment samples, however, were assigned to living edaphic organisms and only a small fraction of those reads was considered to be derived from paleoecological communities. Inferences from ancient sedimentary DNA were complementary to the reconstruction based on pollen and macro-remains, and the combined records reveal more detailed information. This molecular study yielded promising findings regarding the diversity of modern eukaryotic peat bog communities. Nevertheless, even though information about past communities could be retrieved from sediment samples, preferential amplification of DNA from living communities is a

caveat for the use of universal metabarcoding markers in paleoecology.

keywords: Sedimentary DNA, Community DNA, Peat bog paleoecology, Eukaryotes, Pyrenees

Introduction

Depositional systems located in areas with low anthropogenic impact, such as mountain peat bogs, are invaluable paleoenvironmental archives that enable study of local environmental processes and responses to climate variability (Smol et al. 2001). Communities living in these ecosystems can be considered sentinels of past and current climate shifts. The study of the historical changes in their biodiversity is crucial for understanding the dynamics of ongoing ecological processes driven by climate forcings (Mann 2002). Previous paleoecological studies on peatland communities traditionally used morphological remains of living taxonomic groups and fossil material, such as vascular plants, mosses, microalgae, chironomids and pollen (Charman 2002; Godwin; 1981; Smol et al. 2001). These studies, however, provide a limited sense of the total biodiversity, depending on the availability of taxonomic expertise (Parducci et al. 2015).

Molecular methods that use high-throughput sequencing, such as metabarcoding (Taberlet et al. 2012), are a comprehensive, objective and efficient approach to molecular biodiversity assessment, which can often outperform morphological surveys (Epp et al. 2012). The results of metabarcoding analyses are critically dependent on the choice of metabarcoding marker and, specifically, on the universality or specificity of the primer set (Wangensteen and Turon 2017). Most applications of metabarcoding in paleoecological studies have focused on past vegetation, using chloroplast genetic markers such as trnL or rbcL in lake sediments (Domaizon et al. 2017; Anderson-Carpenter 2011; Jørgensen et al. 2012; Pedersen et al. 2013). Although universal (broad taxonomic range) eukaryotic markers have been

applied successfully to study the community DNA of modern environments such as soils or marine benthos (Young et al. 2014; Guardiola et al. 2016; Wangensteen et al. 2018a,b), they have been applied only rarely to study peat bogs. Recently, Singer et al. (2016) studied the diversity of living Oomycetes in peat bogs using the nuclear 18S rRNA marker. This marker has also been used to study free-living soil Cercozoa (Harder et al. 2016) and microbial eukaryotic communities in lakes (Capo et al. 2015, 2016, 2017). The present study focused on a broad spectrum of eukaryotic diversity, and we used a multi-gene approach that included the 18S marker and the mitochondrial cytochrome c oxidase subunit I (COI), which to our knowledge had never been employed in peatland systems before.

The DNA extracted from sediment samples is a combination of modern DNA from living communities and ancient DNA from the remains of long-dead organisms (Bellemain et al. 2013; Epp et al. 2012; Lejzerowicz et al. 2013b; Pawlowski and Holzmann 2014). In general, sediment DNA studies rely on the assumption that the age of the recovered DNA is the same as the age of the sediments in which it is found, in the absence of contamination. This is generally the case when using specific primers that selectively amplify the DNA from remains of organisms such as vascular plants or photosynthetic microalgae that do not currently live in the deep strata of the sediment. Universal primers, used commonly to amplify the 18S rRNA gene and COI are able to detect a wide range of microbial taxa, many of which may belong to living communities (Guardiola et al. 2015, 2016). Moreover, considering the age of the sediment and DNA degradation rates, the concentration of DNA from living organisms present in sediment samples is expected to overwhelm that of ancient DNA by several

orders of magnitude and should be considered when interpreting results from universal primers (Capo et al. 2015; Torti et al. 2015).

This study is a molecular exploration of the eukaryotic diversity present in moss communities and sediment samples from mountain peat bogs. We selected Bassa Nera, a wetland system in the Central Pyrenees (Pérez-Haase and Ninot 2006; 2017) as a case study. The locality was the subject of several previous taxonomic studies of modern and past plant communities (Pérez-Haase and Ninot 2006; Cañellas-Boltà et al. 2009; Cambra 2015; Garcés-Pastor et al. 2016, 2017), which enabled comparisons between molecular results and those of morphological approaches.

Study site

Bassa Nera is a lentic system located in the peripheral zone of Aigüestortes i Estany de Sant Maurici National Park at an altitude of 1891 m a.s.l (Fig. 1). Previous paleoenvironmental studies reconstructed the development of the modern peat bog from a previous lacustrine environment over the Holocene (Garcés-Pastor et al. 2016, 2017). The vegetation of Bassa Nera catchment forms a complex mosaic that ranges from a moderate-depth water body (z_{max} = 5 m) with flat shores surrounded by *Sphagnum* carpets, to Cyperaceae fens and *Sphagnum* bogs, and subalpine forest of *Pinus uncinata* and *Abies alba* on steeper slopes (Carrillo et al. 2008; Pérez-Haase and Ninot 2017). The main habitats are geogenous fens (*Scheuchzerio palustris-Caricetea fuscae*) and ombrogenous bogs (*Oxycocco palustris-Sphagnetea magellanici*) (Pérez-Haase

et al. 2010). Climate is subalpine with Atlantic influence and mean annual precipitation (1152 mm) is well distributed across the seasons (Ninyerola et al. 2003). Mean annual temperature is 4.25 ºC, January being the coldest month (-3 ºC on average) and July the warmest (14 ºC on average).

Materials and methods

We used an Illumina MiSeq high-throughput sequencer to analyse two metabarcoding markers, 18S and COI, on two sets of samples: (1) modern community samples from several peat bog microenvironments, used to establish occurrence and abundance baselines for a wide array of eukaryotic taxa, needed to characterize the extant diversity of high mountain peat bogs and to monitor future changes in these communities, and (2) sediment samples, used to test the suitability of 18S and COI universal markers to evaluate the past diversity of several eukaryotic groups, accomplished by comparing the results from ancient DNA to paleoenvironmental reconstructions based on morphological remains.

Field sampling and DNA extraction

Four different microhabitats were sampled along a water flooding gradient in August 2016 to characterize the modern communities (Fig. 1). Three replicates of 100 mL of the dominant mosses from each microhabitat were obtained and stored in 96% ethanol. The sampled sites lie next to the mire monitoring plots

used by Pérez-Haase and Ninot (2006) to measure water table depth, so that average moisture conditions, groundwater pH and electrical conductivity are known for these sites (Pérez-Haase and Ninot 2017). The studied microhabitats were: A) Hummock (*Carici fuscae-Sphagnetum magellanici* Bick 1985) B) Carpet (*Sphagno fallacis-Caricetum lasiocarpae* Steffen ex Passarge 1964), C) Fen (*Tofieldio calyculatae-Scirpetum cespitosi* Ballesteros, Baulies, Canalís et Sebastià ex Rivas-Martínez et Costa 1998), and D) Floating mire of *Sphagnum* and *Drosera longifolia* on the *Equisetum fluviatile* lake shore belt (*Equisetetum limosi* Steffen 1931) (Pérez-Haase et al. 2010).

Sediment samples were obtained from core BSN-6 (270 cm long), recovered from a hummock of *Sphagnum magellanicum* and *S. capillifolium* in the littoral of Bassa Nera in 2011 (Pèlachs et al. 2016; Garcés-Pastor et al. 2017). The core was processed and sampled following strict precautions to prevent contamination in the Palaeoecology Laboratory of the Universitat Autònoma de Barcelona. The external surface of the core was discarded following usual paleoecological practices. Then the core was sliced and subsampled with sterilized knives for different variables (Pèlachs et al. 2016). The samples were stored individually in double plastic bags at -20 ºC to prevent external contamination until DNA extraction.

The age-depth model was constructed with seven Accelerator Mass Spectrometry radiocarbon dates, obtained from peat and macro-remains. The 270-cm core spans the last ~10,210 cal years, with an average confidence interval error of ca. 220 yr and a mean sedimentation rate of 0.07 ± 0.21 cm yr^{-1}, ranging from 0.016 to 0.86 cm yr^{-1}. The age-depth model provide a robust chronology for the interpretation of the molecular history (Garcés-Pastor et al.

2017). Five sediment samples were studied from the following depths: 31; 109; 160; 220 and 265 cm, which correspond to 140, 3795, 6165, 8339 and 10,094 cal yr BP, respectively (Electronic Supplementary Material [ESM] Table S1).

DNA extraction was performed at the Department of Evolutionary Biology, Ecology and Environmental Sciences at the University of Barcelona. All extraction procedures were carried out under a laminar-flow cabinet in a dedicated pre-PCR laboratory, following strict precautions. All the equipment was cleaned with 10% sodium hypochlorite solution and rinsed in deionised Milli-Q water between samples. To avoid carryover contamination, the ancient sediment samples were processed before the modern samples. Three different extraction replicates were obtained from each sediment sample. Samples were homogenized using a 600 W hand blender. A fraction of 0.3 g of each homogenized sample was extracted using a Norgen Soil DNA Isolation Plus Kit (www.norgenbiotek.com). An extraction blank consisting in 300 µL of molecular biology-grade water was included in the batch, processed and sequenced along with the rest of the samples. DNA concentrations of the purified DNA extracts were estimated with 1 µL of the final elution, using a high-sensitivity dsDNA assay in a Qubit fluorometer (www.thermofisher.com).

PCR, sequencing and bioinformatics pipelines

Two metabarcoding markers were used to identify a wide taxonomic range of detected eukaryotic taxa. The V7 region of nuclear-encoded ribosomal 18S rRNA gene was amplified using the 18S_allshorts primers (100–110 bp, 5'-

TTTGTCTGSTTAATTSCG-3’ and 5’-TCACAGACCTGTTATTGC-3’) (Guardiola et al. 2015), which are expected to provide information for all eukaryotic groups. The V7 fragment amplified by these primers is about 150 bp shorter than the one targeted by Capo et al. (260 bp, 2016, 2017) and other 18S rRNA regions used in other studies: V9 (180 bp, Singer et al. 2016) and V4 (350 bp, Capo et al. 2015). This is convenient for paleoenvironmental studies in which DNA may be fragmented. This primer set has been used successfully to assess eukaryotic diversity of marine sediments (Guardiola et al. 2015, 2016) and shallow marine hard-bottom communities (Wangensteen et al. 2018a,b). The Leray-XT primer set, a novel degenerated primer set amplifying a 313 bp fragment of the mitochondrial marker COI (miCOIintF-XT 5'-GGWACWRGWTGRACWITITAYCCYCC-3'; Wangensteen et al. 2018b; and jgHCO2198 5'-TAIACYTCIGGRTGICCRAARAAYCA-3'; Geller et al. 2013) was also used. This marker features nearly full amplification coverage for almost all main eukaryotic lineages with the remarkable exceptions of Viridiplantae and Ciliophora (Wangensteen et al. 2018b). The conditions for PCR amplifications, library preparation and sequencing are described in ESM File S1.

The bioinformatic analyses were based on the OBITools software suite (Boyer et al. 2016) and followed similar pipelines used for the same markers in previous works (Guardiola et al. 2016; Wangensteen et al. 2018a, 2018b, Siegenthaler et al. 2019). Results of the Leray-XT primer set applied to unfiltered environmental samples are known to include some bacterial sequences arising from unspecific amplifications. Since our study is specifically focused on eukaryotic diversity, the bioinformatics pipeline for COI included additional steps for removing these bacterial sequences. All bioinformatics steps are described in

detail in ESM File S1.

Removal of edaphic organisms

Sediment DNA (i.e. DNA extracted from sediment samples) is a mixture of DNA from long-dead organisms and from living organisms that are known to dwell in soils/sediments (Fungi, Cercozoa, non-photosynthetic Chrysophyta, Oomycetes, Ciliophora, Nematoda, Annelida, Platyhelminthes and Rotifera; Fierer et al. 2003; Andersen et al. 2013; Asemaninejad et al. 2017). Representatives of these groups may also live in surface bog habitats. Thus, detection of these phyla in sediment samples should be interpreted with caution. High abundances of DNA sequences from these taxa, compared to non-edaphic taxa, in the sediment samples, could be interpreted as the result of amplification of DNA from living organisms in deep soil communities, rather than the prevalence of these taxa in ancient surface communities. To avoid this problem, when comparing ancient and living communities, these groups were removed from our analyses so that only those groups typical of surface peat bog environments, i.e. Bacillariophyta, Arachnida, Insecta, Crustacea, Tracheophyta, Bryophyta, etc., were kept, enabling more reliable reconstructions of past surface communities.

Statistical analyses

To compare modern and past communities, we applied the Jaccard dissimilarity

index of presence/absence. Nonlinear-MDS ordinations were performed with the R package vegan (Oksanen et al. 2018). The significance of dissimilarities between modern and ancient communities was assessed using the function anosim in the same package. The function rarecurve in vegan was used to plot rarefaction curves for every sample to check saturation in MOTU (Molecular Operational Taxonomic Unit) richness.

Given that the decay rate of ancient DNA may differ among different taxonomic groups (Zhu et al. 2005), the relative abundance of reads from ancient taxa shows high levels of uncertainty, so that only presence/absence data were used to compare ancient and living communities. For Viridiplantae, only results from the 18S marker were used to compare modern and ancient plant communities, whereas Arthropoda were compared using the COI marker. Given the low amount of DNA reads from ancient communities, compared to the total number of reads in sediment samples, a threshold of 1 in 10,000 total reads (after the removal of singletons) was used as evidence of presence.

Paleoenvironmental data

To compare the information provided by sedimentary DNA with the palaeoenvironmental reconstruction based on morphological methods, we used the available palynological and macroremain data from the palaeoecological study of Garcés-Pastor et al. (2017). Pollen and macroremain analyses were performed according to standard procedures (Moore et al. 1991; Mauquoy et al. 2010). Details on pollen and macroremain methods can be found in Garcés-

Pastor et al. (2017).

Results

DNA yield and sequencing depth

The DNA concentrations recovered from the sediment samples were in the range from 0.077 to 14.9 ng/µl (ESM Table S1), lower than the DNA extracted from modern samples (4.9 to 31.2 ng/µl). Replicates extracted from the sample at 220 cm depth (8339 cal yr BP) yielded only 0.077 ± 0.008 ng/µl (average ± SD) of DNA. They were, however, included in our analyses because PCR amplifications were successful. Results from this sample should nevertheless be interpreted with caution.

After removal of bacterial sequences and singletons, a total of 3,566,813 DNA sequences (DNA reads) composed the final dataset for the 18S marker. Of those, 2,165,734 reads (60.7%) belonged to modern communities (mean of 180,478 reads per sample) and 1,401,079 reads (39.3%) belonged to the sediment samples (mean of 93,405 reads per sample). For COI, the final dataset included 1,762,447 reads, with 1,140,928 reads from modern communities (mean of 95,077 reads per sample) and 621,519 reads from sediment samples (mean of 41,435 reads per sample). Rarefaction curves per sample (ESM File S2) showed that this sequencing depth approached saturation in the number of MOTUs detected for both markers in all samples, except for sediment samples

from 220 cm depth, probably because of the low number of reads obtained from the low DNA recovered at this depth.

Modern community structure inferred from 18S and COI markers

The relative abundance of DNA reads from 18S and COI analyses showed substantial differences among habitat types and between primers (Fig. 2). The percentage of DNA reads that could not be assigned to a Phylum or lower rank (unassigned Eukarya, Metazoa and Stramenopiles) was higher for COI (20.5%) than for 18S (0.8%). As expected, 18S yielded high abundances of reads from Bryophyta (39.0%), Tracheophyta (15.3%), and Arthropoda (23.0%). Conversely, our COI primer set, which is unable to amplify most Viridiplantae, showed a remarkable dominance of Arthropoda (57.9%). Occasionally, one replicate from a community yielded more DNA reads of a specific Phylum than the other replicates. This was the case for Platyhelminthes in Carpet-1 and Fen-2, and Tracheophyta for Floating-3. These differences were mostly a consequence of high abundance of reads from a particular MOTU in those samples, probably related to the presence of a single large individual in that replicate (see tables in Mendeley Data; DOI: 10.17632/j358x9sjjd for abundance of individual MOTUs).

The relative MOTU richness of the different groups is represented in Fig. 2c and 2d for 18S and COI markers, respectively. A higher dominance of MOTUs with small body size is shown, compared to the barplots of DNA read abundance. A relatively homogenous pattern of relative MOTU richness among the different

modern communities is shown.

Modern versus ancient samples

Removal of edaphic organisms highlights the similarities between modern and sediment samples for both markers (Fig. 3). The 18S marker returned high values for relative MOTU richness of Tracheophyta in the sediment samples. On the other hand, COI detected high numbers of MOTUs from Arthropoda, Rhodophyta and Bacillariophyta.

Non-metric multidimensional scaling ordination for the non-edaphic communities of modern and sediment samples using Jaccard dissimilarities (Fig. 4) highlighted the significant differences between ancient and modern communities for 18S (ANOSIM R=0.98, p-value < 10^{-4}, N=27) and COI markers (ANOSIM R=0.97, p-value <10^{-4}, N=27).

Plant communities

The 18S relative read abundance and relative MOTU richness of plant communities (Viridiplantae) are shown in Fig. 5. The amount of reads assigned to plants was lower in sediment samples (4.5% of total reads, including edaphic taxa) compared to modern samples (69.8%), and it decreased with depth.

Distinct patterns of community structure can be distinguished in modern

communities. Hummock and Carpet are dominated by Sphagnopsida, whereas Fen and Floating have a higher proportion of Bryopsida. ESM Table S2 shows the 20 most abundant MOTUs for each modern community. *Sphagnum* dominated Hummock and Carpet, whereas the most abundant MOTU in Fen and Floating communities was a sequence assigned to Bryopsida (ID = 0.99). The hummock community also contained some Magnoliopsida (*Sanguisorba* and *Parnassia)* and Liliopsida (Poaceae). The carpet community has a higher proportion of Liliopsida (Cyperoideae and Poaceae), whereas Magnoliopsida are represented by Asterales and *Filipendula*. The fen community is mainly composed by Bryopsida, with some Magnoliopsida (*Utricularia*) and Liliopsida (Poaceae). Finally, the floating community has higher amounts of Bryopsida and Droseraceae, followed by *Sphagnum*, Cyperoideae and *Utricularia.*

The sediment samples showed higher abundances of Liliopsida and Magnoliopsida, with some Pinopsida and Zygnematophyceae, whereas Sphagnopsida were surprisingly almost absent (Fig. 5). ESM Table S3 shows the rank of the 20 most abundant MOTUs for sediment samples after removing the edaphic taxa. All samples reflect a relatively high number of Tracheophyta. A shift from Pooideae to Cyperoideae can be observed over time, with Pooideae being more abundant in the oldest samples (265-220 cm) and Cyperoideae dominating samples from 160 to 31 cm. Sample 31 has 14.3% Viridiplantae DNA sequences, dominated by Cyperoideae, some Mesangiospermae (Magnoliopsida) and Bryopsida. Sample 109 (2.6% Viridiplantae) is also dominated by Cyperoideae, Pooideae and Mesangiospermae. Sample 160 (1.8% Viridiplantae) has high prevalence of Cyperoideae and Mesangiospermae, with some Pinidae and Cupressaceae. Sample 220 (11.9% Viridiplantae) is dominated by Pooideae and

Magnoliopsida (rosids and asterids), with some Pinidae and Bryophyta. Sample 265 (0.07% Viridiplantae) is still dominated by Pooideae with some Pinidae and a remarkable abundance of Desmidiales (Zygnematophyceae). ESM Fig. S1 shows the non-metric multidimensional scaling ordination for Viridiplantae (18S marker) in modern and sediment samples using Jaccard dissimilarities. Significant differences were found among modern and sediment communities (ANOSIM $R=0.92$, p-value $< 10^{-4}$, $N=27$).

Arthropod communities

The Arthropoda communities in modern and sediment samples differed appreciably (Fig. 6). The modern samples present high abundance of mites (mainly Oribatida and Trombidiformes), whereas the sediment samples display large inter-sample variability. For instance, sample 265 yielded a larger proportion of Opiliones and Copepoda, whereas sample 220 showed proportions more similar to modern samples. On the other hand, sample 160 featured high abundances of an unassigned arthropod sequence. Finally, samples 31 and 109 have outstanding proportions of aquatic crustaceans (Copepoda and Branchiopoda). Some orders, such as Ostracoda, only appeared in the modern samples. ESM Table S4 shows the rank of the 20 most abundant MOTUs for COI in modern samples. Many of the most abundant MOTUs are Arthropoda, esspecially mites: Oribatida, Sarcoptiformes and Trombidiformes, basal Hexapoda (Collembola), Insecta (Diptera, Coleoptera), Maxillopoda (Cyclopoida, Harpacticoida) and Ostracoda. In some cases, the taxonomy could be assigned

to the species level.

ESM Table S5 shows the rank of the 20 most abundant MOTUs for COI in ancient samples, without the edaphic taxa. Contrary to modern samples, most MOTUs could be identified only to the levels of kingdom to order. From the Arthropoda that could be identified, most were Branchiopoda, Maxillopoda and Arachnida. There is a community shift from Arachnida, Insecta and Collembolla (220, 265) in deeper samples, to Branchiopoda and Maxillopoda in the more recent samples (31, 109, 160). ESM Figure S2 shows the non-metric multidimensional scaling ordination for the Arthropoda (COI) in modern and sediment samples using Jaccard dissimilarities. Significant differences were found among modern and sediment communities for Arthropoda (ANOSIM R=0.89, p-value $<10^{-4}$, N=27).

Comparing sedimentary DNA to pollen and macroremains

The presence/absence patterns of 18S Viridiplantae sequences enabled us to make comparisons with environmental reconstructions from pollen and macroremains (Garcés-Pastor et al. 2017). Conifer and Bryophyta DNA sequences were detected from all sediment samples (Fig. 7). Dicotyledon and monocotyledon richness was higher in sample 220. On the other hand, aquatic green algae were richer in sample 31.

Sample 31 presented high proportions of Cyperoideae and Bryopsida DNA with a MOTU assigned to Mesangiospermae (Magnoliopsida) that also appeared in samples 109 and 160 (BOG2_000000149). It also presented some

Pooideae, Saxifragales, Petrosaviidae and Pinidae and traces of *Vaccinium* sp. (ESM Table S3). The macroremains presented low proportions of *Sphagnum*, with Ericaceae, *Polytrichum* and *Equisetum*. *Pinus,* Ericaceae and Poaceae pollen frequencies were well represented, while Apiaceae had its highest values.

Sample 109 also presented high proportions of Cyperoideae, Pooideae and Mesangiospermae DNA, along with Poaceae and asterids. The presence of Betulaceae, *Betula* and Pinidae is also remarkable, with *Equisetum,* Bryophyta and Desmidiales. *Equisetum* and *Sphagnum* macroremains were also found at this depth. Pollen presented the highest amounts of *Abies*, Poaceae and Cyperaceae. On the other hand, *Pinus* grains were well represented and *Betula* had relatively low frequencies. Some ferns (Monolete-spore and *Selaginella*) were also observed.

Sample 160 had high amounts of Cyperoideae and Mesangiospermae, followed by Pinidae, Cupressaceae and some Betulaceae. It also contained Desmidiales, Rhodophyta and Bryophyta. COI detected a MOTU assigned to Porifera (ESM Table S5). Although freshwater Porifera are rare, their presence was reported in Bassa Nera (Garcés-Pastor et al. 2017). In the morphological paleoenvironmental reconstruction, no *Sphagnum* macroremains were found at this depth, whereas pollen presented high amounts of *Pinus,* with *Betula* and some Poaceae.

Sample 220 had high abundance of Pooideae, Pinidae and Bryopsida DNA, along with some Betulaceae**,** Sapindaceae and *Pinus*. Traces of Cyperoideae and Ericales were also detected. This depth did not present *Sphagnum* or other macroremains. High proportions of *Pinus*, *Betula* and *Corylus* were encountered. Poaceae and Cyperaceae had their lowest values. Some

Botryococcus were observed.

Sample 265 has the lowest DNA abundances, but high proportions of Pooideae and Desmidiales, with some Pinidae and traces of *Betula*, Brassicaceae, *Prunus* and Bryophyta. COI detected a MOTU each of Rhodophyta, Porifera and Bacillariophyceae (ESM Table S5). No macroremains were reported. Pollen of *Betula* and *Artemisia* reached highest values, with some *Pinus* and Poaceae. Some ferns (Monolete-spore) and algae (*Botryococcus* and *Pediastrum*) were also observed.

Discussion

Universal primers are suitable tools to assess modern peat bog communities

Our results suggest that the 18S marker is appropriate to detect and identify a broad range of eukaryotes and assess relative abundances of Viridiplantae DNA in peat bog environments. Because of its low natural variability, however, this marker has lower taxonomic resolution than COI (Anslan and Tedersoo 2015; Wangensteen et al. 2018b). On the other hand, the primer set used to amplify COI proved to be useless to retrieve information about vegetal communities. This primer set is most suitable to assess Metazoa, enabling characterization of many Arthropoda**,** Annelida, and Rotifera to the species level, despite persistent known gaps in reference databases (Murria et al. 2019). COI also yielded a higher proportion of unassigned DNA sequences that could very well correspond to the ones that 18S identified as Cercozoa or Ochrophyta, highlighting some important gaps at the phylum level in current COI reference databases for eukaryotic

groups other than Metazoa (Wangensteen et al. 2018b).

The obtained MOTUs from the modern vegetation communities (ESM Table S2) broadly correspond to the communities observed during the sampling. For the case of the Floating-3 replicate, *Drosera* was recorded during sample processing. *Utricularia* sp. was found in the floating and fen communities and was also observed in the catchment (Pérez-Haase and Ninot 2006, 2017). The MOTUs classified as Petrosaviidae may include DNA sequences of Poaceae, Typhaceae and/or Cyperaceae, which are absent from the 18S reference databases. All modern samples presented Petrosaviidae or Cyperaceae reads, which is coherent with the catchment vegetation. There is a community shift from Sphagnopsida to Bryopsida as samples get closer to the pond. The presence of Tracheophyta such as *Sanguisorba*, *Parnassia* and Violaceae in the Hummock, together with *Filipendula* in the carpet, is typical of less humid microenvironments. On the other hand, the presence of carnivorous *Utricularia* and Droseraceae in the fen and floating communities indicates wetter conditions and probably nutrient deficit (Ellison 2006).

Our COI metabarcoding protocol was able to retrieve a high amount of assigned DNA sequences of Arthropoda from modern communities (ESM Table S4). Oribatida was the most abundant order, and different aquatic mite families, such as Nothridae, Malaconothridae, Camisiidae, and Limnozetidae, dominate each community (Thorp and Covich 2009). Hummock has *Nothrus pratensis*, whereas carpet presents *Tyrphonothrus maior*. In the case of fen and floating communities, there is no lower taxonomic identification for these mites below the order level, probably because of gaps in the reference databases. Diptera and Harpacticoida abundances increase with proximity to the floating area. The only

Diptera with high abundances in the hummock is the chironomid *Limnophyes*. Carpet presents the tabanid *Atylotus fulvus* and the chironomid *Paracricotopus*. The fen community has the ceratopogonids *Stilobezzia ochracea* and *Culicoides kibunensis* and the chironomid *Corynoneura*. The Floating community has a higher abundance of Diptera, the ceratopogonids *Dasyhelea modesta* and *Palpomyia lineata*, and the chironomids *Monopelopia tenuicalcar* and *Polypedilum tritum*. The harpacticoid copepod *Bryocamptus pygmaeus* is found in the carpet, fen and floating communities. This species inhabits freshwater environments in mountain regions and displays wide ecological plasticity (Jersabek et al., 2001). With the 18S marker we also obtained good taxonomic resolution for some Arthropoda (ESM Table S2). As occurs with COI, the order Oribatida showed the highest dominance in all communities, although with lower taxonomic resolution. There are some taxa that could be assigned to genus, such as *Hydrozetes*, an aquatic mite in the fen and floating communities, or the freshwater copepod *Acanthocyclops* in the fen.

The use of both markers enabled us to assess the extant community structure of Bassa Nera. In order to use MOTUs as ecological indicators, high taxonomic resolution, at the genus or species level, is desirable. Therefore, COI would be more suitable than 18S for obtaining detailed ecologically relevant information from arthropod taxa, whereas the better assignment rates of 18S might make this marker more suitable for detecting changes in the relative abundances of higher taxonomic assemblages.

Living edaphic taxa must be removed from metabarcoding results of sediment

samples in paleoecological DNA studies

One major caveat for metabarcoding analysis of sediment samples using universal primers, is to distinguish DNA reads from living soil/sediment communities (edaphic organisms) from those amplified from the remains of long-dead organisms (Bellemain et al. 2013; Coolen and Shtereva 2009; Epp et al. 2012; Lejzerowicz et al. 2013a; Pawlowski et al. 2014). Many groups of organisms are known to dwell in soils and sediments down to several meters depth, such as Fungi, Cercozoa, non-photosynthetic Chrysophytes, Oomycetes, Ciliophora, Nematoda and Annelida (Fierer et al. 2003; Andersen et al. 2013; Asemaninejad et al. 2017) and they were detected in high abundances in this study. After removing the possibly living edaphic taxa, the patterns of relative MOTU richness for 18S and COI markers became more similar between sediment and modern samples (Fig. 3). ANOSIM, however, showed that modern and sediment communities are still significantly different.

Differences in the proportions of detected MOTUs between sediment and modern samples might be a result of differential preservation rates of DNA among different taxa. Our results for read abundance obtained from sediment samples suggest that the 18S rRNA gene fragment is degraded faster for plants than for animals. Moreover, the low detection rate of 18S from Sphagnopsida DNA in sediment samples cannot be a consequence of primer bias or gaps in reference databases, since this taxon was abundantly detected and identified from our modern samples. Epp et al. (2012) also found lower amounts of bryophyte DNA in sediment samples, whereas recent soil samples yielded high abundances. They suggested that bryophytes may contain secondary metabolites that

enhance DNA degradation (Xie and Lou 2009) and this could potentially cause proportionally higher DNA degradation rates compared to other groups. Differential detectability can also depend on the initial DNA abundance. The mitochondrial marker COI has in general high numbers of copies per cell (Pääbo et al. 2004). On the other hand, the copy number of tandem rRNA sequences for 18S present in nuclear genomes may vary considerably between different eukaryotic groups (Zhu et al. 2005). As a result, quantitative comparisons of ancient community structure based on sequence abundances, is generally impossible.

Our results suggest that even presence/absence molecular surveys of sediment communities can differ considerably from their modern counterparts. None of the reconstructed assemblages from ancient communities studied here could be considered to reproduce any modern assemblage (Fig. 4). This suggests that broader spatial and temporal sampling studies should be performed to create modern community-DNA analogues for all Phyla. Moreover, RNA metabarcoding (Guardiola et al. 2016; Lejzerowicz et al. 2013b) would be a suitable technique to be used for assessing only living, or recently dead organisms, whose results could then be compared to the results from total DNA metabarcoding to distinguish living edaphic taxa from ancient DNA remains in sediment samples.

DNA results from universal markers may still be useful for paleoenvironmental

reconstructions

Our results show that the reconstruction obtained from the 18S marker for ancient DNA (Fig. 3) cannot attain fine taxonomic resolution. With some exceptions, most of the recovered sequences could be assigned to the level of family or above. This might constrain the interpretation and comparisons to pollen and macroremain data. We found, however, that the studied paleoenvironmental proxies offer complementary information that could be useful for paleoecological reconstructions, even if our dataset included samples from only five sediment depths.

Alhough the taxonomic resolution of the marker does not enable us to identify to the species level, correspondence between assigned MOTUs and pollen or macroremains were found for many taxa. For example, the *Sphagnum* macroremain proportions could be related to the Bryophyta DNA sequences, and *Polytrichum* macroremains might correspond to the MOTU assigned to Bryopsida. Also, the *Vaccinium* sequence fits within the Ericaceae pollen. The high amounts of *Pinus* and *Abies* pollen match with the DNA sequences of Pinidae, and so on.

Moreover, DNA analyses allow for the detection of many taxa, such as Desmidiales, Streptophytina, Chlorophyceae and Scenedesmaceae, which would be overlooked by pollen analysis. These taxa suggest moist environments, also corroborated by the presence of COI DNA sequences of Bacillariophyceae and Porifera (ESM Table S5).

Despite low DNA concentrations and the low sequencing depth recovered from samples 220 and 265, the detected MOTUs match quite well with the

information recovered from pollen and macroremains for these samples. Our results highlight that the interpretation of ancient sediment DNA does not overlap perfectly with the reconstruction based on pollen and macroremains, but the combination of both reconstructions reveals more detailed information about plant paleocommunities than that achieved by either approach individually (Jørgensen et al. 2012). Pollen analysis may provide information at a more regional level, while macroremains and sediment DNA may provide more local details (Alsos et al. 2018). A higher taxonomic resolution for plant species could probably be obtained from using different metabarcoding markers, such as chloroplast markers (Parducci et al. 2017).

Pros and cons and future improvements in peat bog metabarcoding

Our results suggest that 18S and COI markers are useful to assess the biodiversity of modern peat bog communities, but there is a major caveat in the application of universal eukaryotic metabarcoding markers to sediment samples, related to the high proportion of DNA recovered from living edaphic communities. A multi-marker approach is recommended to cover total community biodiversity (Epp et al. 2012). Although some constraints could be related to limitations of DNA extraction methods or primer specificities, we think that currently, the most significant drawback in the analysis of community and sediment DNA is the lack of complete reference databases. Such collections must contain a broad range of barcode DNA sequences derived from accurately identified species, covering all major lineages of Eukaryota. Nevertheless, DNA identifications can be more

easily standardized and are more traceable and objective in comparison with morphology-based identification approaches (Jørgensen et al. 2012).

In this study, the 18S rRNA gene provided useful information about past plant communities, whereas information from COI was mainly restricted to Metazoa. The use of COI would enable high-resolution taxonomic assignment of animal communities, if a complete reference database were available (Wangensteen and Turon 2017). With the current reference database available for Pyrenean peat bog communities, however, the taxonomic results from COI are just slightly better than those from the 18S marker. This issue will undoubtedly be solved in the future by improving barcoding efforts. To obtain a more detailed description of the vegetation paleoenvironments, it would be desirable to use chloroplast markers, which enable better taxonomic resolution than 18S for higher plants. Markers used in this work provided insights on ancient communities and results that agree broadly with those obtained from morphological analysis of pollen and macroremains. The present work was an exploratory study with 18S and COI markers on a small number of sediment samples. More extensive studies with higher temporal resolution will enable more detailed understanding of the information provided by DNA from past communities.

Current paleoecological studies that rely on morphological remains are based mostly on the identification of vegetal material. Paleoecological studies using other organisms such as Arthropoda or other Metazoa have been limited to the scarce biological traces that remain in the sediment. With the proper analysis of metabarcoding data based on modern analogues, this DNA technique has the potential to offer a new paleoenvironmental multi-proxy approach addressing diverse taxa from the same period. Such an approach would allow for

a better understanding of the relationships between animal and vegetation communities and their response to past climate shifts. The advantage of metabarcoding to study a large number of taxa simultaneously, in the absence of morphological expertise, is obvious in the case of understudied or complex groups.

Although the use of metabarcoding does not depend on taxonomic expertise, it requires bioinformatics skills. The laboratory procedures and data collecting may be considerably shorter than for morphological analyses, but the use of appropriate bioinformatics pipelines and reliable reference databases is crucial for obtaining accurate results. Further investigations are also needed to study how DNA degradation affects the results with respect to markers from different taxa. For example, in this study we found that the DNA from *Sphagnum* and other mosses is probably not well preserved and might be undetectable in ancient samples, with the markers used.

Another limiting factor is the scant knowledge of the autoecology of many small metazoan groups, with some notable exceptions such as chironomids (Tarrats et al. 2017). Once the reference databases are improved and the DNA sequences are assignable to the genus or species level, the ecological interpretation of this data will need current information on species distributions and autoecological preferences. This knowledge would allow the acquisition of reliable ecological information from a number of independent taxonomic sources such as mites, collembola, and many other small arthropods and metazoans, which would provide robust inferences of paleoecological reconstructions from the detection of metazoan species (Pansu et al. 2015).

We have shown that metabarcoding of paleoecological communities using

universal markers is currently limited by the small number of DNA reads obtained from past remains, compared to those derived from living edaphic taxa. This limitation, however, can be easily circumvented by using new ultra-throughput sequencing technologies, such as NovaSeq (Singer et al. 2019), which would increase the sequencing depth per sample by two orders of magnitude, allowing for higher number of reads from ancient remains and thus more robust paleoecological inferences.

This study was the first attempt to sequence DNA in ancient samples from Pyrenean peat bogs. We were able to amplify DNA and get useful sequencing information from samples spanning a period of 10,000 years. Although the number of sediment samples analysed in this preliminary work was small and did not allow to obtain robust inferences, the ancient DNA interpretation was coherent with the pollen and macroremain reconstruction, and the universal markers enabled us to detect organisms that would be difficult to study using conventional paleoecological techniques. These results open the way to more detailed reconstructions of past communities using novel molecular proxies derived from DNA metabarcoding.

Acknowledgements

We thank Professor Xavier Turon for providing us with the 18S primers. We are indebted to Editor Mark Brenner and to three anonymous reviewers for their suggestions, which contributed to improve upon earlier versions of this

manuscript.

References

Alsos IG, Lammers Y, Yoccoz NG, Jørgensen T, Sjögren P, Gielly L, Edwards ME (2018). Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. PloS one, 13(4), e0195403

Andersen R, Chapman S, Artz R (2013) Microbial communities in natural and disturbed peatlands: A review. Soil Biol Biochem 57: 979–994. doi:10.1016/j.soilbio.2012.10.003

Anderson-Carpenter L (2011) Ancient DNA from lake sediments: bridging the gap between paleoecology and genetics. BMC Evol Biol 11: 30

Anslan S, Tedersoo L (2015) Performance of cytochrome c oxidase subunit I (COI), ribosomal DNA Large Subunit (LSU) and Internal Transcribed Spacer 2 (ITS2) in DNA barcoding of Collembola. Eur J Soil Biol 69: 1-7

Asemaninejad A, Thorn R, Lindo Z (2017) Vertical distribution of fungi in hollows and hummocks of boreal peatlands. Fungal Ecol 27:59-68

Bellemain E, Davey ML, Kauserud H, Epp LS, Boessenkool S, Coissac E, Gemi J, Edwards M, Willersley E, Gussarova G, Taberlet P, Haile J, Brochmann C (2013) Fungal palaeodiversity revealed using high- throughput metabarcoding of ancient DNA from arctic permafrost. Environ Microbiol 15: 1176-1189

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) obitools: a unix-inspired software package for DNA metabarcoding. Mol Ecol 16(1): 176-182

Cambra J (2015) Micro-scale distribution of algae in a Pyrenean peat-bog, Spain. Hidrobiológica 25: 213-222

Cañellas-Boltà N, Rull V, Vigo J, Mercadé A (2009) Modern pollen-vegetation relationships along an altitudinal transect in the central Pyrenees (southwestern

Europe). Holocene 19: 1185-1200

Capo E, Debroas D, Arnaud F, Perga ME, Chardon C, Domaizon I (2017) Tracking a century of changes in microbial eukaryotic diversity in lakes driven by nutrient enrichment and climate warming. Environ microbiol 19: 2873-2892

Capo E, Debroas D, Arnaud F, Guillemot T, Bichet V, Millet L, Lejzerowicz F (2016) Long- term dynamics in microbial eukaryotes communities: a palaeolimnological view based on sedimentary DNA. Mol Ecol 25: 5925-5943

Capo E, Debroas D, Arnaud F, Domaizon I (2015) Is planktonic diversity well recorded in sedimentary DNA? Toward the reconstruction of past protistan diversity. Microb Ecol 70: 865-875

Carrillo E, Brugués M, Carreras J, Cros RM, Ferré A, Ninot JM, Pérez-Haase A, Ruiz E (2008) Singularitat de la vegetació de les reserves integrals de Trescuro i d'Aiguamòg. In: Jornades sobre recerca al Parc Nacional d'Aigüestortes i Estany de Sant Maurici. 25–27 October. Vall de Boí, Barruera, pp 177–192

Charman D (2002) Peatlands and environmental change. John Wiley & Sons Ltd., Chichester, UK

Coolen MJ, Shtereva G (2009) Vertical distribution of metabolically active eukaryotes in the water column and sediments of the Black Sea. FEMS Microbiol Ecol 70: 525-539

Domaizon I, Winegardner A, Capo E, Gauthier J and Gregory-Eaves I (2017) DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. J Paleolimnol 58: 1-21

Ellison AM (2006) Nutrient limitation and stoichiometry of carnivorous plants. Plant Biol 8: 740-747

Epp LS, Boessenkool S, Bellemain EP, Haile J, Esposito A, Riaz T, Erseus C, Gusarov

VI, Edwards ME, Johnsen A, Stenøien HK (2012) New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. Mol Ecol 21: 1821-1833

Fierer N, Schimel JP, Holden PA (2003) Variations in microbial community composition through two soil depth profiles. Soil Biol Biochem 35: 167-176

Garcés-Pastor S, Cañellas-Boltà N, Clavaguera A, Calero MA, Vegas-Vilarrúbia T (2016) Vegetation shifts, human impact and peat bog development in Bassa Nera pond (Central Pyrenees) during the past millennium. Holocene 27: 553-565

Garcés-Pastor S, Cañellas-Boltà N, Pèlachs A, Soriano JM, Pérez-Obiol R, Pérez-Haase A, Calero MA, Andreu O, Escolà N, and Vegas-Vilarrúbia T (2017). Environmental history and vegetation dynamics in response to climate variations and human pressure during the Holocene in Bassa Nera, Central Pyrenees. Palaeogeogr Palaeocl 479: 48-60

Geller J, Meyer C, Parker M, Hawk H (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol Ecol 13: 851–61

Godwin H (1981) The archives of the peat bogs. Cambridge University Press, Cambridge, UK

Guardiola M, Uriz M J, Taberlet P, Coissac E, Wangensteen O S, and Turon X (2015). Deep-sea, deep-sequencing: metabarcoding extracellular DNA from sediments of marine canyons. PloS one 10: e0139633

Guardiola M, Wangensteen O, Taberlet P, Coissac E (2016) Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. PeerJ 4 e2807

Harder CB, Rønn R, Brejnrod A, Bass D, Al-Soud WA, Ekelund F (2016) Local diversity

of heathland Cercozoa explored by in-depth sequencing. ISME J 10: 2488

Jersabek C, Brancelj A, Stoch F, Schabetsberger R (2001) Distribution and ecology of copepods in mountainous regions of the Eastern Alps. Hydrobiologia 453: 309-324

Jørgensen T, Haile J, Möller P, Andreev A (2012) A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. Mol Ecol 21: 1989-2003

Lejzerowicz F, Esling P, Majewski W, Szczucinski W, Decelle J, Obadia C, Martines Arbizu P, Pawlowski J (2013a) Ancient DNA complements microfossil record in deepsea subsurface sediments. Biol Letters 9:20130283 DOI 10.1098/rsbl.2013.0283.

Lejzerowicz F, Voltsky I, and Pawlowski J (2013b) Identifying active foraminifera in the Sea of Japan using metatranscriptomic approach. *Deep Sea Research Part II:* Topical Studies in Oceanography 86: 214-220

Mann M (2002) The value of multiple proxies. Science 297: 1481-1482

Mauquoy D, Hughes P, van Geel B (2010) A protocol for plant macrofossil analysis of peat deposits. Mires Peat 7

Moore PD, Webb JA and Collinson ME (1991) Pollen Analysis. Blackwell, Oxford

Múrria C, Väisänen LOS, Somma S, Wangensteen OS, Arnedo MA, Prat N (in press) Towards an Iberian DNA barcode reference library of freshwater macroinvertebrates and fishes. Limnetica

Ninyerola M, Pons X, Roure JM (2003) Atles Climàtic Digital de Catalunya. Barcelona: Universitat Autònoma de Barcelona

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018)

vegan: community ecology package. R package version 2.5-3

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler, J, Rohland N, Hofreiter M (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38: 645-679

Pansu J, Giguet-Covex C, Ficetola G, Gielly L, Boyer F, Zinger L, Choler P (2015) Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA. Mol Ecol 24: 1485-1498

Parducci L, Bennett K, Ficetola G, Alsos I, Suyama Y, Wood JR, Pedersen MW (2017) Ancient plant DNA in lake sediments. New Phytol 214: 924-942

Parducci L, Väliranta M, Salonen JS, Ronkainen T, Matetovici I, Fontana SL, Eskola T, Sarala P, Suyama Y (2015) Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. Phil Trans R Soc B 370: 20130382

Pawlowski J, Esling P, Lejzerowicz F, Cedhagen T, Wilding TA (2014) Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Mol Ecol 14:1129-1140

Pawlowski J, Holzmann M (2014). A plea for DNA barcoding of foraminifera. J Foramin Res 44: 62-67

Pedersen M, Ginolhac A, Orlando L, Olsen J, Andersen K, Holm J, Kjær KH (2013) A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. Quaternary Sci Rev *75*: 161-168

Pèlachs A, Pérez-Obiol R, Soriano JM, Pérez-Haase A (2016) Dinàmica de la vegetació, contaminació ambiental i incendis durant els últims 10.000 anys a la Bassa Nera (Val d'Aran). X Jornades sobre Recerca al Parc Nacional d'Aigüestortes i Estany

de Sant Maurici. Vall de Boí, Barruera, pp 75–87

Pérez-Haase A, Ninot JM (2006) Caracterització florística i ecològica de les molleres de la Nassa Nera (Aiguamòg). VII Jornades sobre Recerca al Parc Nacional d'Aigüestortes i Estany de Sant Maurici, Generalitat de Catalunya, Barcelona, pp 193-213

Pérez-Haase A, Ortuño E, JM (2010) Diversitat de comunitats vegetals a les molleres de la Vall d'Aran (Pirineus centrals). Acta Bot Barcinonensis 53: 61-112

Pérez-Haase A, Ninot JM (2017) Hydrological heterogeneity rather than water chemistry explain high plant diversity and uniqueness of a Pyrenean mixed mire. Folia Geobot 1:18

Siegenthaler A, Wangensteen OS, Benvenuto C, Campos J, Mariani S (2919) DNA metabarcoding unveils multiscale trophic variation in a widespread coastal opportunist. Mol Ecol 28: 232-249

Singer D, Lara E, Steciow MM, Seppey CV, Paredes N, Pillonel A, Oskazo T, Belbahri L (2016) High-throughput sequencing reveals diverse oomycete communities in oligotrophic peat bog micro-habitat. Fungal Ecol 23: 42-47

Singer G, Fahner NA, Barnes J, McCarthy A, Hajibabaei M (2019) Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. Sci Rep 9: 5991

Smol JP, Birks HJB, Last WM (2001) Tracking Environmental Change Using Lake Sediments. Volume 3: Terrestrial, Algal and Siliceous Indicators, Developments in Paleoenvironmental Research. Kluwer Academic Publishers, Dordrecht, Netherlands

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 21: 2045-

2050

Tarrats P, Cañedo-Argüelles M, Rieradevall M, Prat N. (2017) Chironomid communities as indicators of local and global changes in an oligotrophic high mountain lake (Enol Lake, Northwestern Spain). J Limnol 76: 355-365

Thorp J, Covich A (eds) (2009) Ecology and classification of North American freshwater invertebrates. Academic Press, New York

Torti A, Lever MA, Jørgensen BB (2015) Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Mar Genom 24: 185-196

Wangensteen OS, Turon X (2017) Metabarcoding techniques for assessing biodiversity of marine animal forests. In: Rossi S, Bramanti L, Gori A, Orejas C (eds), Marine Animal Forests, The Ecology of Benthic Biodiversity Hotspots. Springer International Publishing, Switzerland, pp 445-473

Wangensteen OS, Cebrian E, Palacín C, Turon X (2018a). Under the canopy: community-wide effects of invasive algae in marine protected areas revealed by metabarcoding. Mar Pollut Bull 127: 54-66

Wangensteen OS, Palacín C, Guardiola M, Turon X (2018b) DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ 6: e4705

Xie C, Lou H (2009) Secondary metabolites in bryophytes: an ecological aspect. Chem Biodivers 6: 303-312

Young JM, Weyrich LS, Cooper A (2014) Forensic soil DNA analysis using high-throughput sequencing: a comparison of four molecular markers. Forensic Sci Int Genet 13: 176-184

Zhu F, Massana R, Not F (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol

850 52:79-92

Figure Legends

Fig. 1 Sampling points of the modern and sediment samples of Bassa Nera. **A**) Location of the study area. **B**) Topographic map of the region surrounding Bassa Nera **C**) Location of the sampling points (black dots) and core extraction (star)

Fig. 2 Patterns of relative abundance of DNA reads (**a, b**) and relative MOTU richness (**c, d**) per sample using 18S (**a, c**) and COI (**b, d**) markers in the four modern communities

Fig. 3 Relative richness of non-edaphic taxa from sediment and modern samples, according to the detected presence of MOTUs of 18S (a) and COI (b) markers, after removal of edaphic taxa

Fig. 4 Non-metric multidimensional scaling ordination using Jaccard index analysis to presence/absence dataset dissimilarities with non-edaphic MOTUs of samples for 18S (**a**) and COI **(b**) markers

Fig. 5 Relative abundance or DNA reads (**a**) and relative richness of MOTUs (**b**) for the divisions of Viridiplantae detected using 18S marker

Fig. 6 Relative abundance or DNA reads (**a**) and relative richness of MOTUs (**b**) for

orders of the phylum Arthropoda detected by COI marker

Fig. 7 Diagram with the presence/absence data of detected DNA sequences of Viridiplantae and the abundances of pollen and macroremains from the morphological study by Garcés-Pastor et al. (2017)

Fig. 1. Sampling points of the modern and sedimentary samples of Bassa Nera. A) Location of the study area. B) Topographic map of

Click here to download Figure Fig 1 - study site.tif

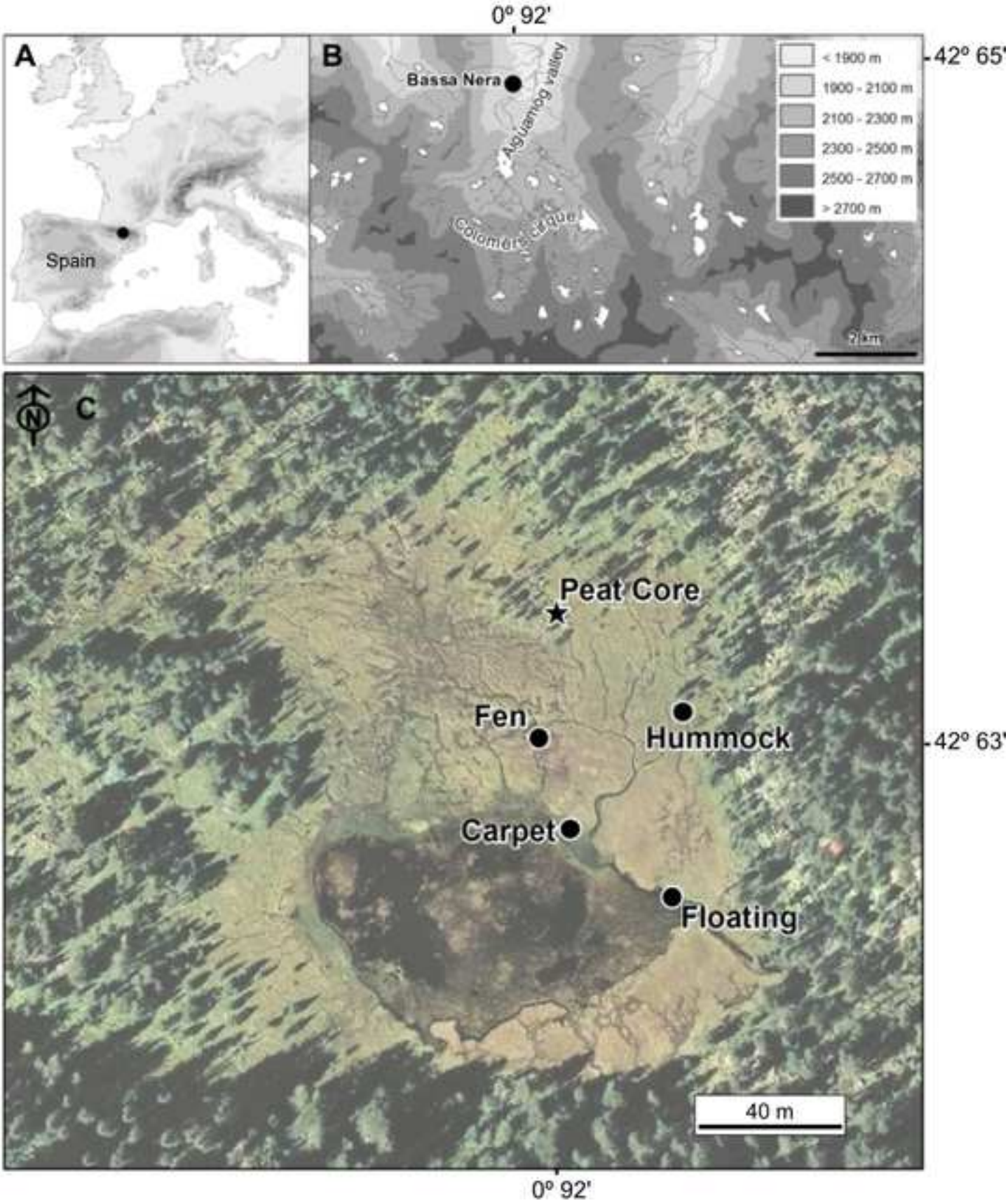


Fig. 2. Patterns of relative abundance of DNA reads (a, b) and relative MOTU richness (c, d) per sample using 18S (a, c) and COI (b, d) markers in the four modern

Click here to download Figure Figure 2 - modern communities.tif

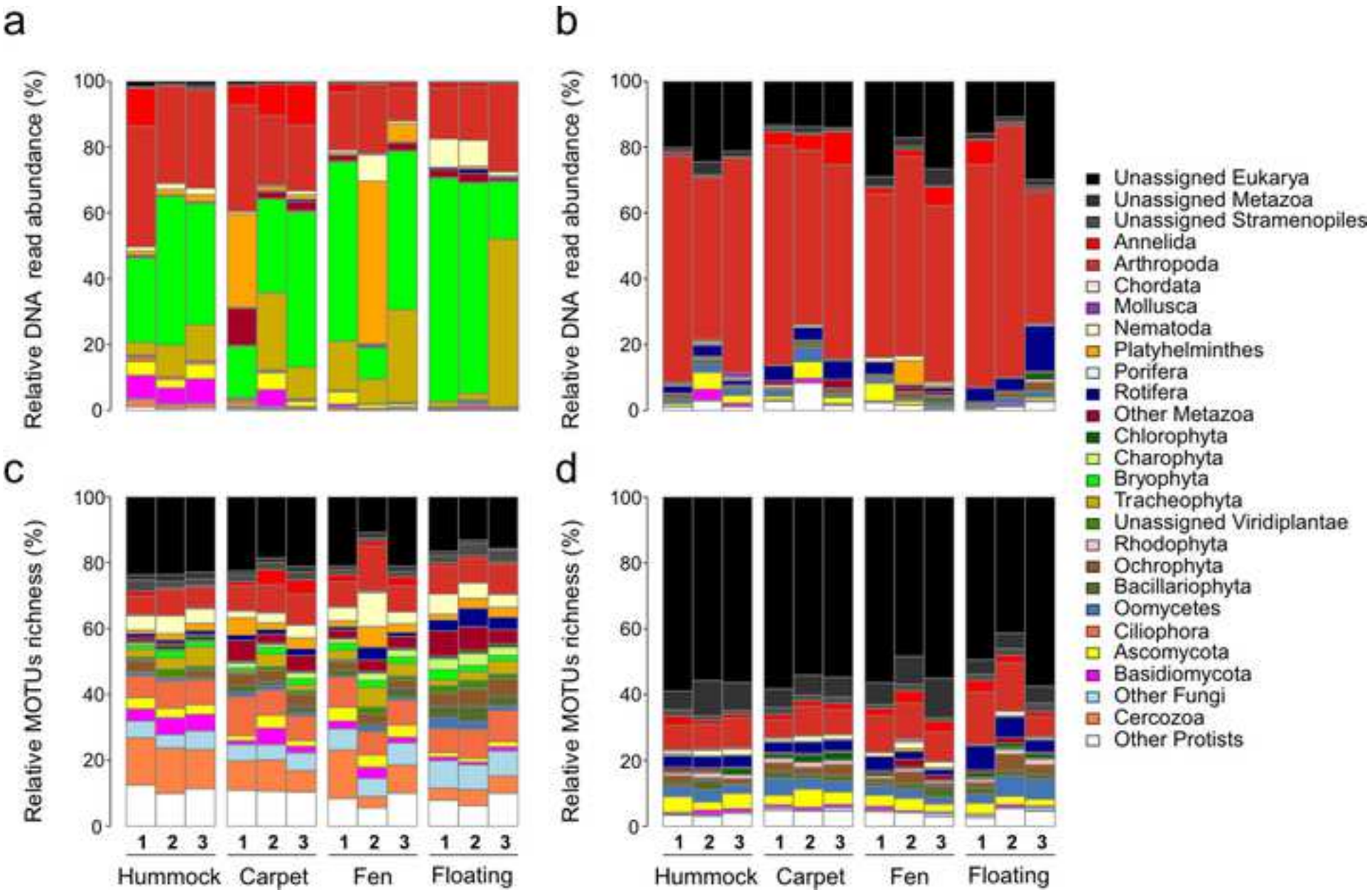


Fig. 3. Relative richness of non-edaphic taxa from sedimentary and modern samples, according to the detected presence of MOTUs of 18S (a) and COI (b) markers, after

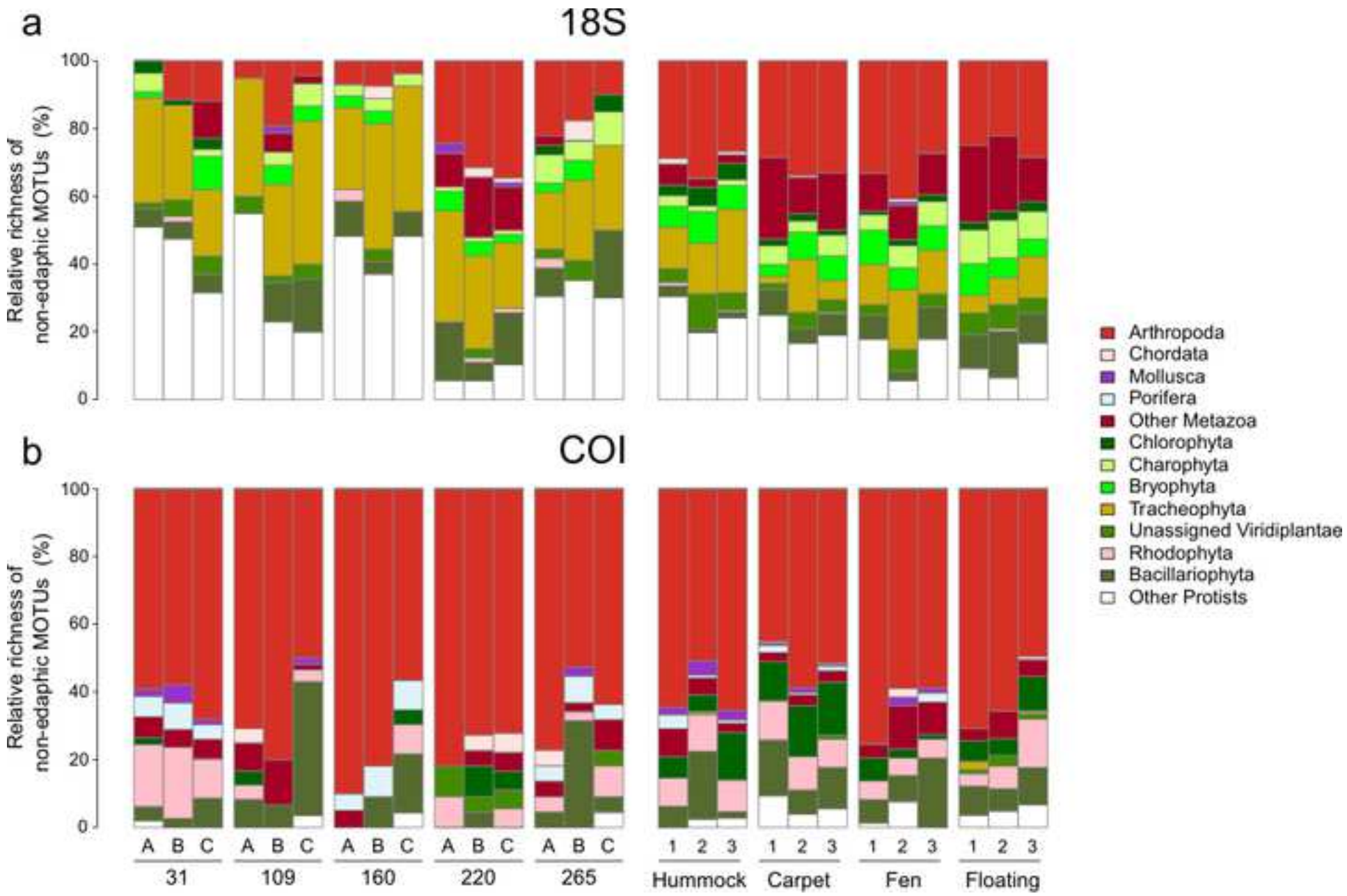


Fig. 4. Non-metric multidimensional scaling ordination using Jaccard index analysis to presence/absence dataset dissimilarities with non-edaphic MOTUs of samples for 18S

Click here to download Figure Figure 4 - barplot non-edaphic diversity of motus MetaBog.tif

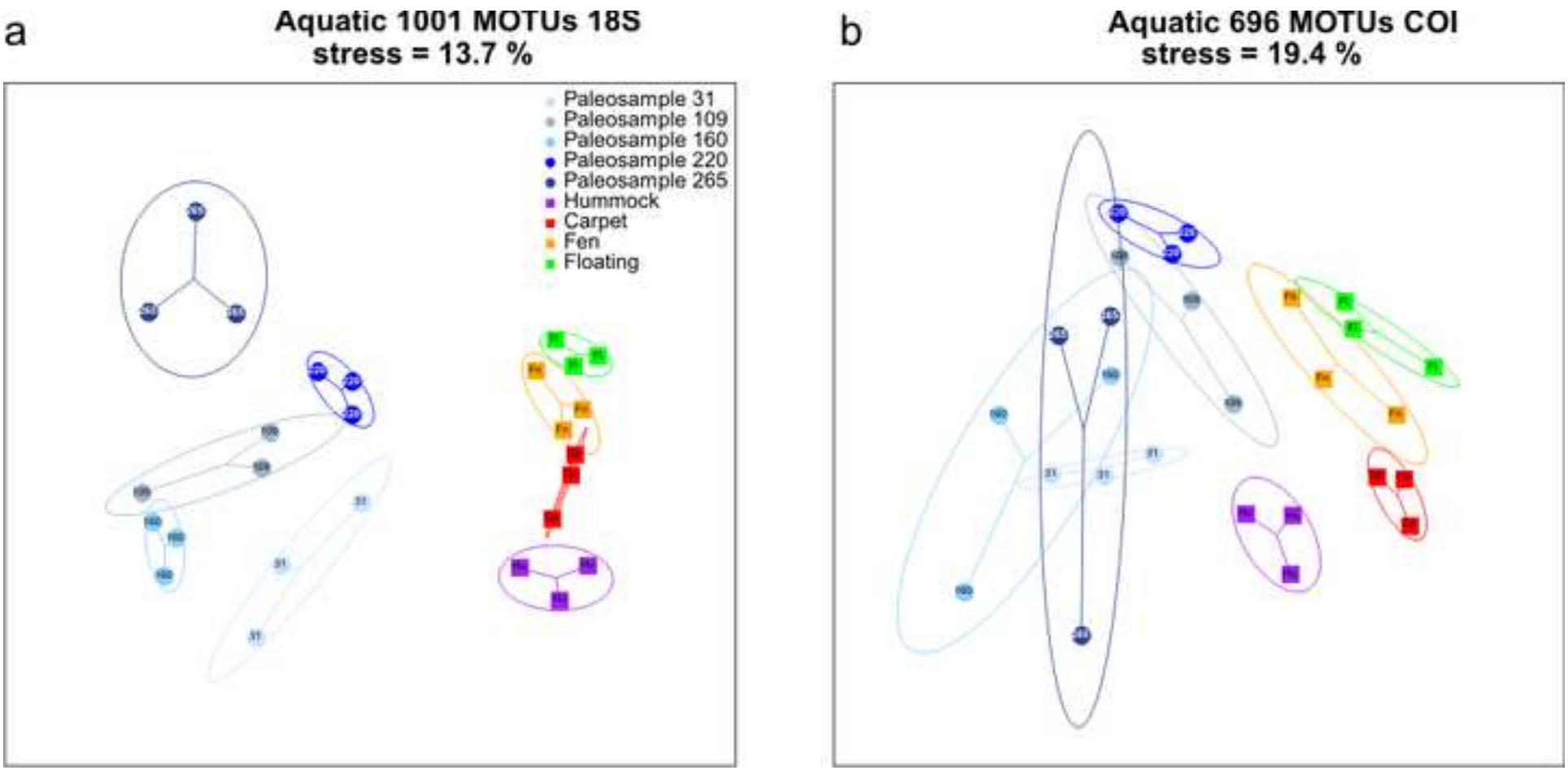


Fig. 5. Relative abundance or DNA reads (a) and relative richness of MOTUs (b) for the divisions of Viridiplantae detected using 18S marker

Click here to download Figure Figure 5 - barplot Viridiplantae biomass and motus MetaBog.tif

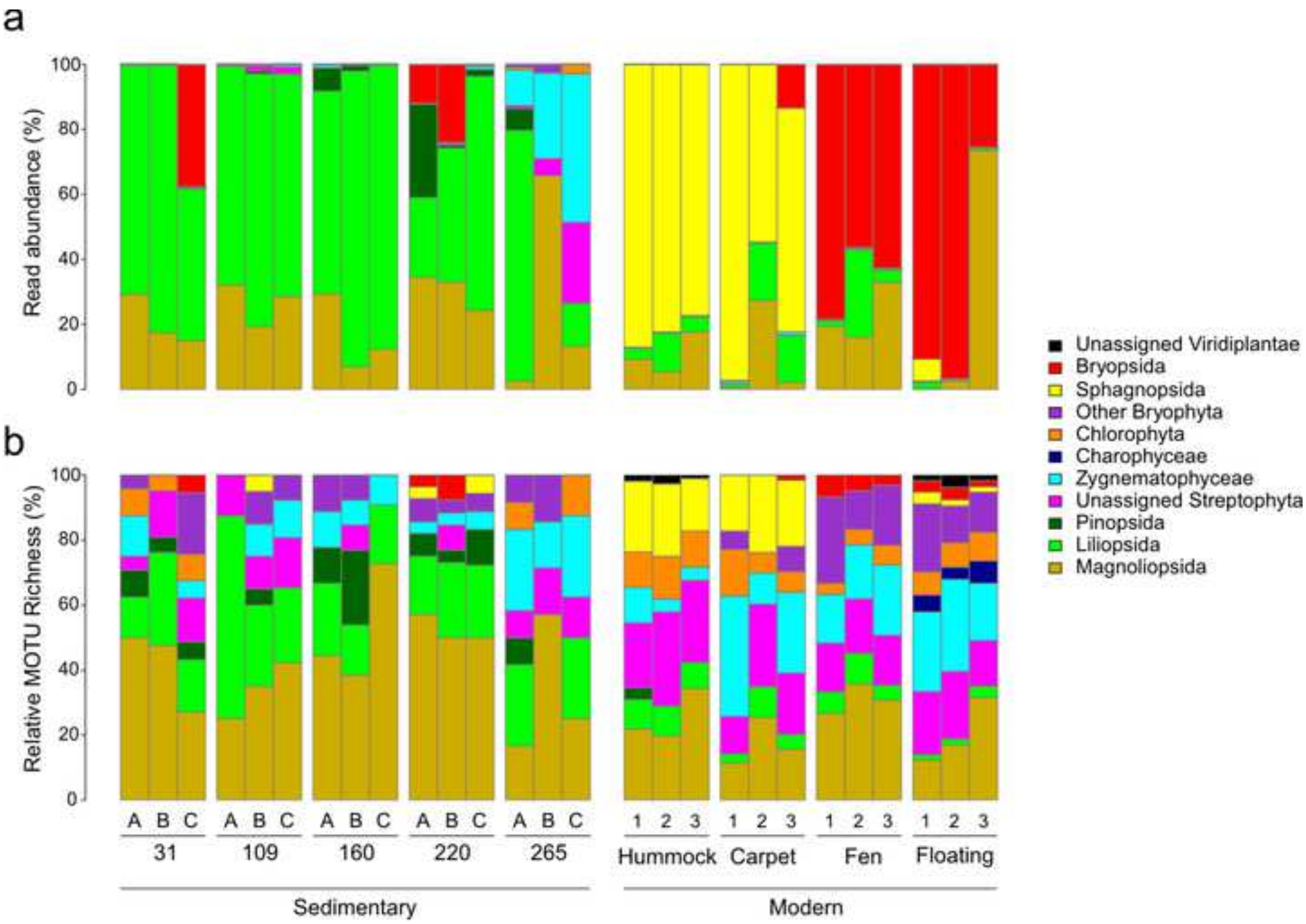


Fig. 6. Relative abundance or DNA reads (a) and relative richness of MOTUs (b) for orders of the phylum Arthropoda detected by COI marker

Click here to download Figure Figure 6 - barplot Arthropoda biomass and motus MetaBog.tif

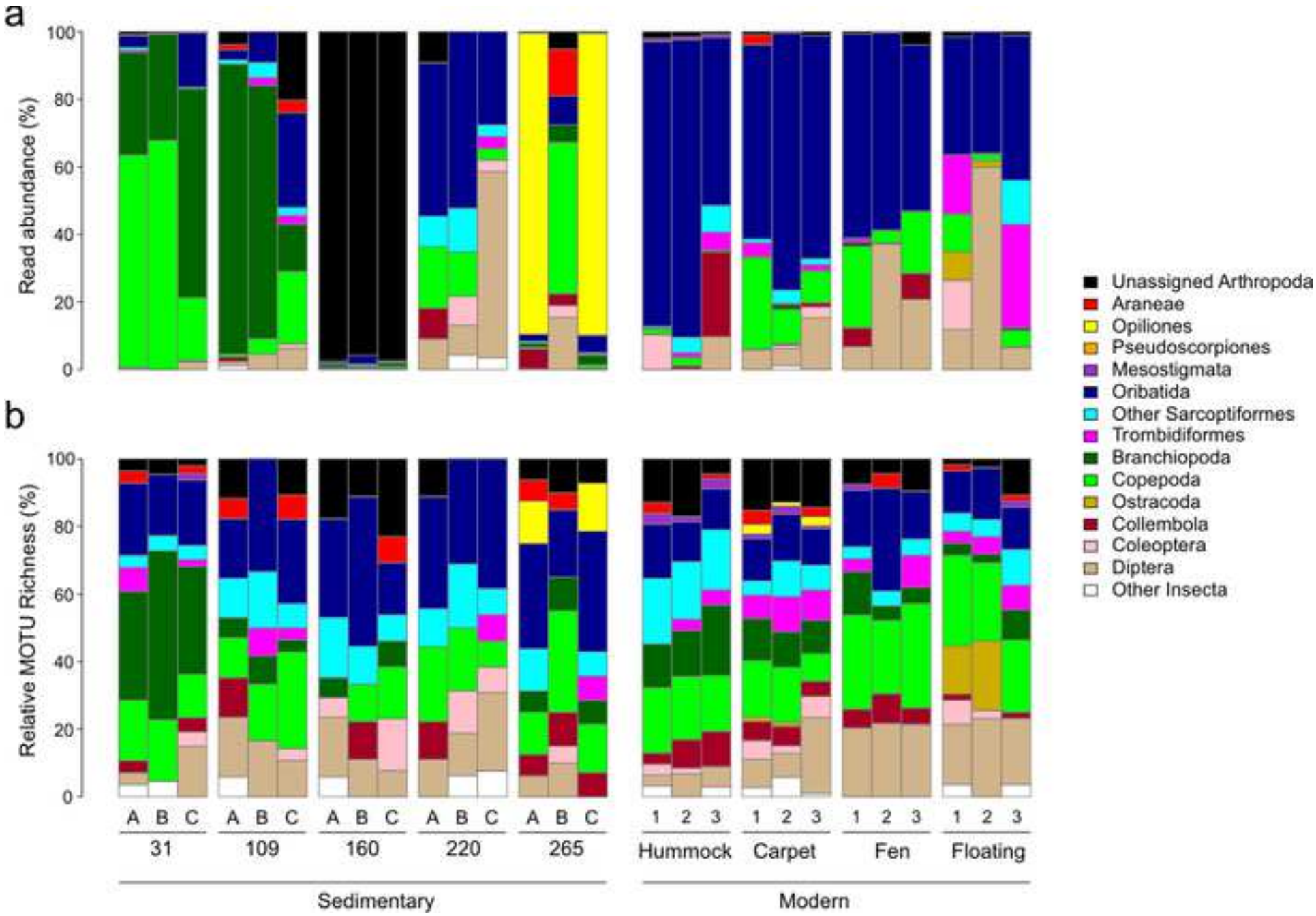


Fig. 7. Diagram with the presence/absence data of detected DNA sequences of Viridiplantae and the abundances of pollen and macroremains from the morphological

Click here to download Figure Fig 7 - pollen, macroremains and DNA.tif

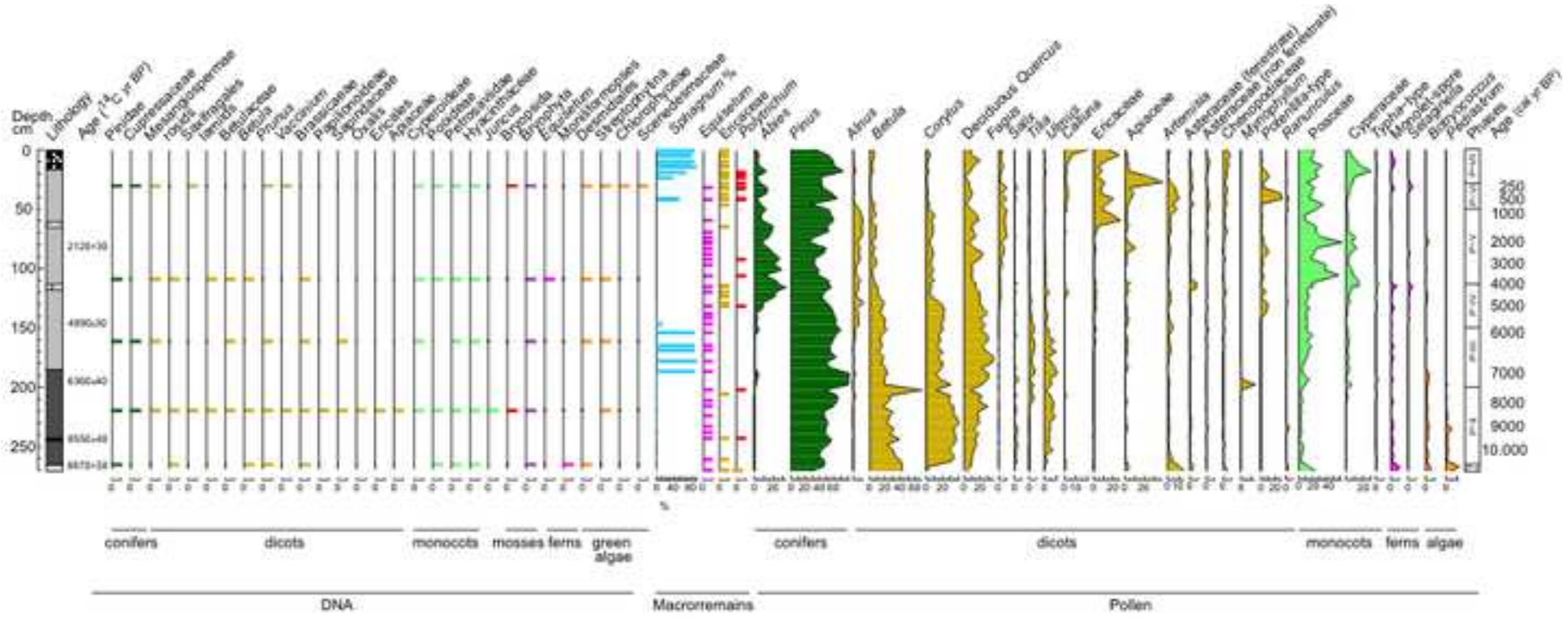


Figure S1 (supplementary material)

Click here to download attachment to manuscript Fig S1 - Viridiplantae.tif

Click here to view linked References

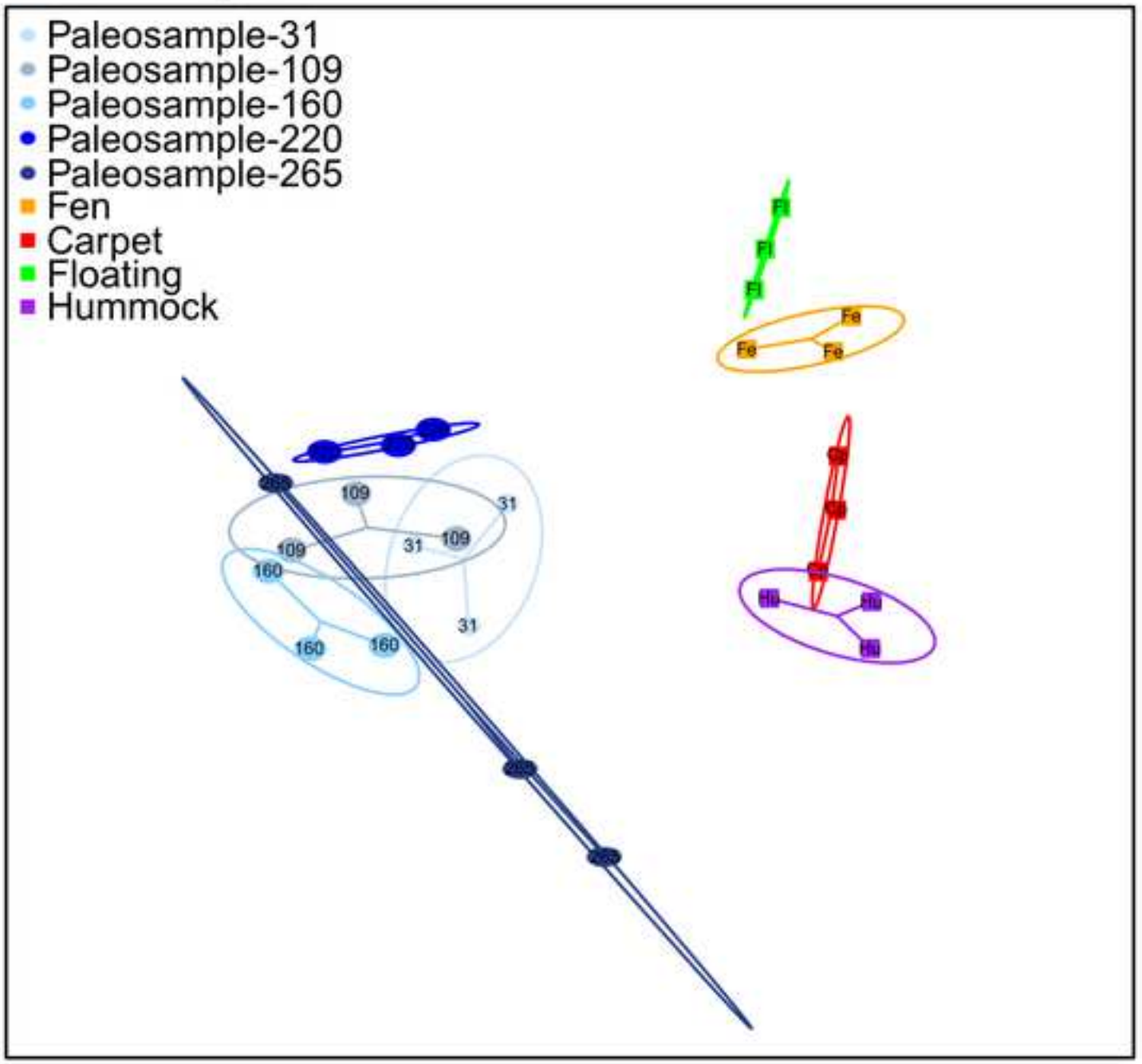
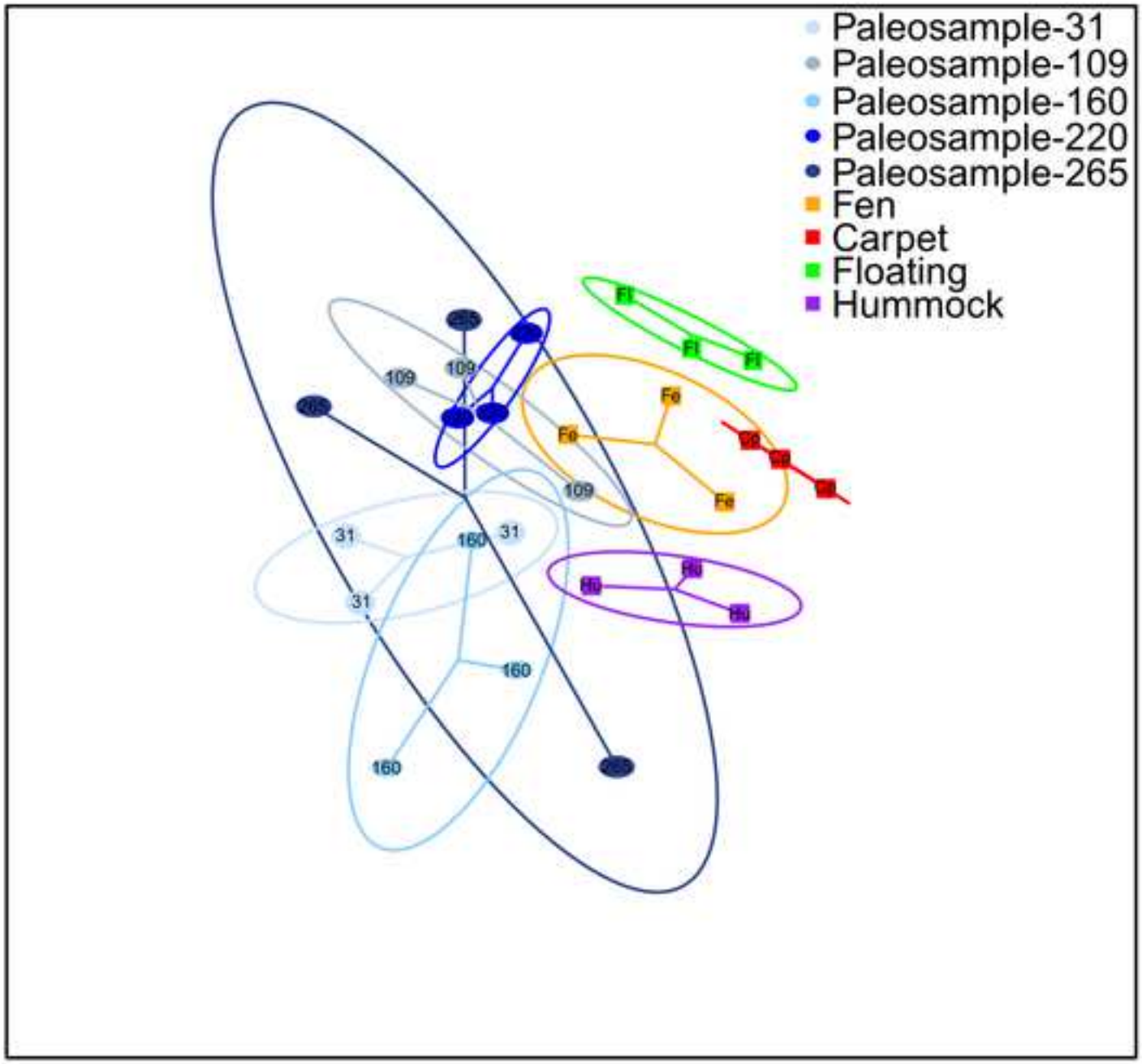


Figure S2 (supplementary material)

Click here to download attachment to manuscript Fig S2 - Arthropoda.tif

Click here to view linked References



Supplementary table S1. DNA concentration of sedimentary and modern samples from Bassa Nera peat bog extracted with Norgen Soil DNA Isolation Plus Kit.

Community	Depth (cm)	Inferred age (cal yr BP)	DNA concentration (ng/µl, mean ± SD)
Sedimentary	31	140	14.9 ± 1.8
Sedimentary	109	3795	2.83 ± 0.44
Sedimentary	160	6165	1.16 ± 0.29
Sedimentary	220	8339	0.077 ± 0.008
Sedimentary	265	10094	0.98 ± 0.14
Hummock	0 - 2	modern	31.2 ± 9.0
Carpet	0 - 2	modern	5.3 ± 4.7
Fen	0 - 2	modern	4.9 ± 0.7
Floating	0 - 2	modern	5.6 ± 2.8

Table S2. 20 most abundant 18S MOTUs for modern samples. Best id = Best identity

Hummock			Carpet			Fen			Floating		
Scientific name	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**
Sphagnum	1	211999	*Sphagnum*	1	144277	Bryopsida	0.99	212126	Bryopsida	0.99	249266
Desmonomata	1	82196	*Bothrioplana semperi*	1	62581	*Bothrioplana semperi*	0.99	77189	Droseraceae	1	126639
Tectocepheus sarekensis	1	40149	*Rhynchoscolex simplex*	1	25935	*Utricularia*	1	76531	Anystina	0.93	37122
Hygrocybe	0.99	20472	Cyperoideae	1	22051	Brachypylina	1	24896	Desmonomata	1	28406
Acrogalumna longipluma	1	20360	asterids	1	18330	Desmonomata	1	16016	*Rhabdolaimus aquaticus*	1	27565
Cernosvitoviella atrata	1	19699	Bryopsida	0.99	16462	Poaceae	0.99	14507	*Hydrozetes lacustris*	1	17888
Sanguisorba	1	17635	*Aeolosoma sp. GG-2011*	0.98	11302	*Rhabdolaimus aquaticus*	1	11974	*Enochrus quadripunctatus*	1	9841
Poaceae	0.99	17069	Poaceae	0.99	10116	*Hydrozetes lacustris*	1	11882	Podocopida	1	8806
Hydrophilinae	1	12190	Harpacticoida	1	10086	*Geocentrophora sphyrocephala*	1	9342	*Sphagnum*	1	8590
Brachypylina	1	11678	*Cernosvitoviella atrata*	1	8355	*Calyptostoma velutinus*	0.94	8145	*Calyptostoma velutinus*	0.94	4164
Fungi	1	11436	Fungi	1	7457	Harpacticoida	0.96	7488	Tubificina	1	3032
Agaricomycetes	1	7709	Tubificina	1	7060	Podoplea	0,88	5362	Cyperoideae	1	2639
Violaceae	1	4612	*Filipendula vulgaris*	0.99	6371	Fungi	1	4934	Parasitengona	0,94	2503
Helicoon fuscosporum	1	4525	Brachypylina	1	6083	Harpacticoida	1	4382	Macrobiotidae	0.99	2456
Catenulida	0.95	4026	*Chamaedrilus cognettii*	1	5502	Lumbriculidae	1	4115	Zygoptera	1	2238
Steganacaridae	1	3886	Chaetonotidae	1	5376	*Limnognathia maerski*	1	2516	*Utricularia*	1	2126
Harpacticoida	0.98	3423	Entelegynae	0.98	4964	Acanthocyclops	1	1862	Harpacticoida	0.98	1949
Parnassia	1	2957	Tubificina	0.99	4753	*Aeolosoma sp. GG-2011*	0.98	1804	Chaetonotidae	1	1931
Prismatolaimus	1	2944	*Lepidochaetus zelinkai*	1	4494	Leotiomycetes	0.98	1737	*Lumbriculus*	1	1886
Fungi	1	2765	Naididae	0.97	4472	Peniophorella	0.99	1724	Brachypylina	1	1584

Table S3. 20 most abundant 18S MOTUs for sedimentary samples

31 cm			109 cm			160 cm			220 cm			265 cm		
Scientific name	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**
Cyperoideae	1	28063	Cyperoideae	1	2955	Cyperoideae	1	6802	Bicosoecida gen. 1	0.9	2335	Dinophyceae	0.87	3999
Mesangiospermae BOG2_000000149	0.97	9089	Pooideae	1	1795	Bicosoecida gen. 1 sp. EK-2010a	0.92	5505	Pooideae	1	619	Heterophryidae	0.89	3965
Bryopsida	0.99	7058	Mesangiospermae BOG2_000000149	0.97	1564	Bicosoecida	0.82	1645	Pinidae	1	377	Pooideae	1	80
Pooideae	1	776	Poaceae	0.99	178	Mesangiospermae BOG2_000000149	0.97	748	Bryopsida	0.99	277	Desmidiales	0.94	50
Paramonas globosa	0.87	585	Bicosoecida gen. 1 sp. EK-2010a	0.92	151	*Paramonas globosa*	0.87	554	*Navicula*	1	216	Bicosoecida gen. 2 sp. EK-2010a	0.85	46
Paramonas globosa	0.87	547	asterids	1	143	Bicosoecida gen. 1 sp. EK-2010a	0.9	412	Salamandroidea	1	127	Adeleidae	0.91	19
Chaetonotus acanthodes	1	281	Bicosoecida gen. 1 sp. EK-2010a	0.88	97	Bicosoecida	0.82	369	rosids	0.99	114	Polypodiidae	1	17
Saxifragales	0.98	267	*Equisetum arvense*	1	96	Salamandroidea	1	132	rosids	1	110	rosids	1	16
Bicosoecida gen. 1 sp. EK-2010a	0.92	218	*Paramonas globosa*	0.87	84	Pinidae	1	79	asterids	1	109	Petrosaviidae	1	13
Poaceae	0.99	202	rosids	1	80	Cupressaceae	0.99	39	Desmonomata	0.99	101	Dysderidae	1	13
Gregarinasina	0.89	171	Petrosaviidae	1	76	*Paramonas globosa*	0.91	27	Fragilariaceae	0	91	Salamandroidea	1	12
Pyrenomonadales	0.77	140	Sapindales	1	43	Bicosoecida gen. 1 sp. EK-2010a	0.88	22	Betulaceae	96	76	*Paramonas globosa*	0.89	9
Pinidae	1	121	Streptophytina	0.89	41	Desmidiales	0.91	21	Petrosaviidae	1	74	rosids	0.99	8
Nuclearia	0.86	88	Desmidiales	0.94	29	Mesangiospermae BOG2_000077141	0.97	19	Mesangiospermae	1	61	Pinidae	1	8
Nuclearia	0.76	82	*Navicula*	1	27	Dysteridae	1	13	Sapindales	0.99	47	Bacillariophyta	0.88	7
Chaetonotidae	1	79	Betulaceae	1	24	Mesangiospermae BOG2_000010361	0.96	12	*Rhizidiomyces apophysatus*	1	40	*Prunus*	1	6
Prunus	1	63	Brassicaceae	1	20	Navicula	1	9	Cymbellales	0.9	37	Petrosaviidae	0.98	5
Microdalyellia	1	62	rosids	0.99	18	*Paramonas globosa*	0.92	9	Chaetonotidae	0.93	37	*Navicula*	1	5
Chaetonotidae	1	59	Petrosaviidae	0.98	17	Lauraceae	1	7	Papilionoideae	1	34	*Sinella curviseta*	1	5
Paramonas	0.88	45	Pinidae	1	15	Sapinadaceae	1	7	Pinus	0.99	33	Atripliceae	1	4

Table S4. 20 most abundant COI MOTUs for modern samples

Hummock			Carpet			Fen			Floating		
Scientific name	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**
Nothrus pratensis	0.99	47933	Malaconothridae	0.84	108107	Malaconothridae	0.84	10258	Malaconothridae	0.84	64168
Malaconothridae	0.84	32706	*Tyrphonothrus maior*	1	75834	Sarcoptiformes	0.85	4409	Trombidiformes	0.81	30644
Poduroidea	0.8	16875	Maxillopoda	0.82	38042	*Stilobezzia ochracea*	1	2286	*Dasyhelea modesta*	0.99	17267
Limnophyes sp.7SW	1	7517	*Atylotus fulvus*	0.98	15599	Maxillopoda	0.82	2210	Rotifera	0.78	10853
Microtrombidiidae	0.87	3937	Sarcoptiformes	0.85	14431	*Corynoneura*	0.99	1432	*Enochrus ochropterus*	0.99	10006
Sarcoptiformes	0.83	2588	Scheloribatidae	0.88	13342	Mycobatidae	0.91	1332	Sarcoptiformes	0.89	8325
Sordariomycetes	0.86	2489	*Paracricotopus*	1	10232	Eukaryota	0.79	1293	Eukaryota	0.76	5803
Tectocepheus	0.88	2324	Pristina	0.85	9395	Isotomidae	0.98	1220	Podocopida	0.88	5649
Neelipleona	0.89	2227	*Didymium*	0.85	7599	Platyhelminthes	0.84	968	*Lumbriculus variegatus*	0.99	5612
Anacaena lutescens	1	1973	Malaconothridae	0.99	7191	*Bryocamptus pygmaeus*	0.98	896	Eukaryota	0.73	5250
Eukaryota	0.76	1779	Sarcoptiformes	0.82	6714	Neocopepoda	0.82	725	*Palpomyia lineata*	1	5134
Planorbidae	0.8	1574	Eukaryota	0.79	6290	*Culicoides kibunensis*	0.97	636	*Monopelopia tenuicalcar*	1	4652
Eukaryota	0.78	1568	*Bryocamptus pygmaeus*	0.98	6275	*Malaconothrus*	0.84	565	Eukaryota	0.76	3972
Adineta	0.9	1437	*Murrayon pullari*	0.99	6041	*Leohumicola*	0.9	546	Eukaryota	0.78	3687
Eukaryota	0.71	1390	*Cognettia glandulosa B SM2014*	0.98	5978	Harpacticoida	0.84	518	*Lecane cornuta*	0.87	3630
Eukaryota	0.8	1342	Eukaryota	0.78	5767	Ploima	0.83	442	*Polypedilum tritum*	0.98	2703
Eukaryota	0.76	1311	Trombidiformes	0.81	5726	Sordariomycetes	0.88	440	Trebouxiophyceae	0.79	2605
Eukaryota	0.75	1148	Philodinidae	0.9	4961	Pristina	0.85	417	Ochrophyta	0.75	2287
Eukaryota	0.79	1114	Bdelloidea	0.9	4741	Cyclopoida	0.86	316	*Bryocamptus pygmaeus*	0.98	2225
Leohumicola	0.9	1085	Leotiomycetes	0.88	4325	Eukaryota	0.75	288	Maxillopoda	0.84	2216

Table S5. 20 most abundant COI MOTUs for sedimentary samples

31 cm			109 cm			160 cm			220 cm			265 cm		
Scientific name	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**	**Scientific name**	**Best id**	**Total reads**
Bacillariophyceae	0.81	4774	Branchiopoda	0.73	251	Arthropoda	0.82	2867	*Psychoda alternata*	0.93	15	Rhodophyta	0.79	9248
Porifera	0.82	2032	Bacillariophyta	0.81	51	Porifera	0.75	333	*Tyrphonothrus maior*	1	9	Porifera	0.75	1841
Planorbidae	0.8	1897	*Navicula*	0.84	24	Arthropoda	0.76	41	Malaconothridae	0.84	6	Opiliones	0.77	540
Maxillopoda	0.81	1028	*Thalassionema*	0.86	14	Branchiopoda	0.73	31	Rhodophyta	0.81	6	Bacillariophyceae	0.81	93
Branchiopoda	0.73	396	Bacillariophyta	0.83	14	Arthropoda	0.8	15	*Nothrus pratensis*	0.99	5	*Ceratophysella denticulata*	0.83	25
Florideophyceae	0.84	362	*Sellaphora*	0.85	13	Bacillariophyta	0.84	13	Sarcoptiformes	0.85	3	Opiliones	0.77	24
Branchiopoda	0.73	166	Bacillariophyceae	0.82	13	*Eimeria*	0.72	12	*Murrayon pullari*	0.99	3	Mollusca	0.72	15
Branchiopoda	0.72	127	Bacillariophyceae	0.82	12	Naviculaceae	0.83	9	Rotifera	0.82	3	Branchiopoda	0.75	13
Porifera	0.82	110	*Haslea*	0.86	11	Bacillariophyceae	0.82	7	*Coccomyxa*	0.76	2	*Tyrphonothrus maior*	1	11
Rhodophyta	0.81	80	Bacillariophyceae	0.83	11	*Tyrphonothrus maior*	1	6	Bacillariophyceae	0.85	2	Araneae	0.92	8
Pyropia	0.82	78	*Nitzschia*	0.83	10	Harpacticoida	0.83	6	Crotoniidae	0.9	2	Diptera	0.9	8
Banchiopoda	0.75	74	*Sellaphora*	0.84	10	*Murrayon pullari*	0.99	6	Sarcoptiformes	0.85	2	Harpacticoida	0.79	8
Tyrphonothrus maior	1	73	Bacillariophyceae	0.83	10	*Malaconothrus*	0.87	4	Limoniidae	0.94	2	*Haslea*	0.84	7
Pyropia	0.83	58	Bacillariophyta	0.83	10	Malaconothridae	0.84	4	Harpacticoida	0.84	2	Naviculaceae	0.83	7
Maxillopoda	0.82	57	Arthropoda	0.76	10	*Othius angustus*	0.99	4	Rhodophyta	0.8	1	*Sellaphora*	0.84	7
Branchiopoda	0.73	54	*Sellaphora*	0.87	9	*Nothrus pratensis*	0.99	3	*Acutodesmus*	0.75	1	Harpacticoida	0.82	7
Rhodophyta	0.83	44	Bacillariophyta	0.85	9	Asplanchna	0.81	3	*Sellaphora*	0.86	1	*Bos*	0.99	7
Rhodophyta	0.82	44	*Tyrphonothrus maior*	1	9	Rhodophyta	0.99	2	*Banksinoma*	0.95	1	Platyhelminthes	0.75	7
Branchiopoda	0.74	41	*Ovatella vulcani*	0.92	9	*Dysdera*	0.86	2	Malaconothridae	0.99	1	Porifera	0.82	7
Maxillopoda	0.81	35	Stylochoidea	0.78	9	*Malaconothrus mollisetosus*		2	Sarcoptiformes	0.82	1	Porifera	0.82	7

Supplementary file S1. Methods used for PCR amplification, library preparation, sequencing and bioinformatics pipelines

PCR amplifications

The V7 region of nuclear-encoded ribosomal 18S rRNA gene was amplified using the 18S_allshorts primers (5'-TTTGTCTGSTTAATTSCG-3' and 5'-TCACAGACCTGTTATTGC-3') (Guardiola et al. 2015), which provide information for all eukaryotic groups. The Leray-XT primer set, a novel degenerated set amplifying a 313 bp fragment of the mitochondrial marker COI (miCOIintF-XT 5'-GGWACWRGWTGRACWITITAYCCYCC-3'; Wangensteen et al. 2018b; and jgHCO2198 5'-TAIACYTCIGGRTGICCRAARAAYCA-3'; Geller et al. 2013) was also used. This marker features nearly full amplification coverage for almost all main eukaryotic lineages with the exception of Viridiplantae and Ciliophora (Wangensteen et al. 2018b).

The PCR amplifications were performed at the dedicated environmental DNA laboratory at the University of Salford. 8-base sample-specific tags for identifying the multiplexed samples and a variable number (2-4) of leading random bases, for increasing DNA sequence diversity, were attached to the metabarcoding primers. The amplification mix for the 18S_allshorts primers included 10 µl of AmpliTaq Gold DNA polymerase master mix (Applied Biosystems), 1 µl of each 5 µM forward and reverse 8-base tagged primers, 3 µg of bovine serum albumin and 5 ng of extracted DNA in a total volume of 20 µl per sample. The PCR conditions consisted in a first denaturation step of 10 min at 95 °C and then 45 cycles of denaturation at 95 °C for 30 s, annealing at 45 °C for 30 s and elongation at 72 °C for 30 s. For the amplification of COI using the Leray-XT primers, the mix included 10 µl of AmpliTaq Gold DNA polymerase master mix (Applied Biosystems), 1 µl of each 5 µM forward and reverse 8-base tagged primers, 3 µg of bovine serum albumin and 5 ng of extracted DNA in a total volume of 20 µl per sample. The PCR profile for COI included 10 min at 95 ºC, 35 cycles of 94 ºC 1 min, 45 ºC 1 min and 72 ºC 1 min, and 5 min at 72 ºC. The concentration of the DNA recovered from one of the sedimentary depths was too low, thus 0.5 ng of template DNA was used instead of 5 ng for the PCR of this sample replicates.

PCR products pooling and library preparation

After PCR, the PCR products were multiplexed into two libraries (one per marker) along

with other samples from an unrelated project and these pools were purified and concentrated using Minelute PCR purification columns (QIAGEN). The number of samples sequenced for this study were 28 per marker: 12 amplifications from modern communities (four communities with three ecological replicates per community), 15 amplifications from sedimentary samples (five depths with three extraction replicates per depth) and the extraction blank. A PCR blank using molecular biology grade water as template was run along with the samples, but was not include in neither of the pools, since no amplification bands were observed by electrophoresis in agarose gel. The total number of samples per multiplexed library (including samples for this study and the unrelated project) was 83. Two Illumina libraries were built from the amplicon pools using the NEXTflex PCR-free DNA library prep kit (www.biooscientific.com). Both libraries were sequenced together in a single run of Illumina MiSeq using v3 chemistry 2x250 bp paired-end.

Bioinformatics pipelines and statistical analyses

The bioinformatics analyses were based on the OBITools software suite (Boyer et al., 2016) and followed the same pipelines used for these markers in previous works (Guardiola et al. 2016; Wangensteen et al. 2018a, 2018b), with the exception of the MOTU clustering step. Briefly, the paired-reads were assembled using illuminapairedend. The resulting aligned datasets with alignment quality score > 40 were demultiplexed using ngsfilter, and the 28 samples belonging to this study were selected for further processing. A length filter (obigrep) was applied to the assigned reads (75-180 bp for 18S and 300-320 bp for COI). The obtained reads were dereplicated using obiuniq and chimeric DNA sequences were removed with vsearch (Rognes et al. 2016) using the uchime_denovo algorithm. Individual sequences were clustered into molecular operational taxonomic units (MOTUs) using the step-by-step aggregation clustering algorithm implemented in SWARM v2 (Mahé et al. 2015) with a resolution of d=1 for 18S and d=13 for COI. These values for d have been previously used for similar metabarcoding datasets (Wangensteen and Turon 2017; Macías-Hernández et al. 2018; Kemp et al. 2019; Siegenthaler et al. 2019). Singleton sequences (MOTUs of abundance = 1 read) were removed after the clustering.

The taxonomic assignment of the representative sequences for each MOTU was performed using ecotag (Boyer et al. 2016) on custom local reference databases, as explained in Wangensteen et al. (2018b). Both databases are publicly available from http://github.com/metabarpark/reference_databases. Ecotag is able to assign sequences without a perfect match using a phylogenetic approach, which selects the best hit in the

reference database and builds a set of reference sequences which are at least as similar to the best hit as the query sequence is. Then, the sequence is assigned to the taxon of the NCBI taxonomy tree including all the reference set sequences. With this procedure, the assigned taxonomic rank varies depending on the similarity of the query sequences and the density of the reference database, so that some sequences can be assigned at the species level, whereas other sequences can be assigned, for example, at the family, order or phylum levels, in case a closer reference sequence is not available for them.

After taxonomic assignment, a blank correction step was performed, following Wangensteen and Turon (2017), where MOTUs with higher than 10% values for the abundance in the blanks to total abundance ratio were removed. The final MOTU datasets were manually checked. Those DNA sequences assigned by ecotag to bacteria or to the root of the tree of life were removed. Other sequences considered as potential contaminants related to human presence or activity (e.g. human DNA and cultivated plants), and sequences from marine organisms (originated by tag switching from the unrelated samples that were sequenced together in the same Illumina run) were also removed. In order to further improve the detection of bacterial sequences produced by unspecific amplifications, an additional refining step was used in the COI pipeline. Sequences of COI MOTUs were queried against a bacterial nucleotide database from Genbank using BLASTn (McGinnis and Madden 2004), and MOTUs which matched a bacterial sequence with an E-value of 10^{-50} or lower were removed. This step allowed to remove an additional 5.9 % of putative bacterial COI MOTUs that had not been assigned as prokaryotic sequences by ecotag. A summary of all software used in the bioinformatics pipelines is available in table S0 (below).

Table S0. Summarized pipelines for the complete metabarcoding procedure including two markers: COI and 18S. Names beginning in "owi_" are custom R scripts available at http://github.com/metabarpark.

Sampling (preservation in ethanol)		
Pre-processing: (homogenization of samples) **DNA extraction: (Norgen Soil DNA Isolation Plus Kit)**		
	PIPELINE FOR COI	PIPELINE FOR 18S
PCR	Tagged Leray primers	Tagged Allshort primers
Library preparation	NEXTflex PCR-free (BIOO)	NEXTflex PCR-free (BIOO)
HT Sequencing	Illumina MiSeq V3 2x250 bp	Illumina MiSeq V3 2x250 bp
Raw sequences QC	fastqc	fastqc
PE alignment	illuminapairedend	illuminapairedend
Demultiplexing	obiannotate/obisplit ngsfilter	obiannotate/obisplit ngsfilter
Length filter	obigrep 300-320 bp	obigrep 75-180 bp
Dereplication	obiuniq	obiuniq
Rename identifiers	obiannotate MBOG1	obiannotate MBOG2
Chimera removal	vsearch uchime_denovo	vsearch uchime_denovo
Clustering	SWARM v2 d=13 obitab owi_recount_swarm delete singletons	SWARM v2 d=1 obitab owi_recount_swarm delete singletons
Taxonomic assignment	ecotag using db_COI_BOLD	ecotag using db_18S
Add higher taxa	owi_add_taxonomy	owi_add_taxonomy
Final refinement	Blank correction Removal of contamination MOTUs Removal of bacterial sequences assigned by ecotag Removal of bacterial sequences using BLASTn	Blank correction Removal of contamination MOTUs Removal of bacterial sequences assigned by ecotag
Community analyses and integration of the results		

References

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) Obitools: a unix-inspired software package for DNA metabarcoding. Mol Ecol Res 16; 176-182

Guardiola M, Wangensteen O, Taberlet P, Coissac E (2016) Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. PeerJ 4: e2807

Geller J, Meyer C, Parker M, Hawk H (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol Ecol Res 13: 851-861

Kemp J, López-Baucells A, Rocha R, Wangensteen OS, Andriatafika Z, Nair A, Cabeza M (2019) Bats as potential suppressors of multiple agricultural pests: A case study from Madagascar. Agric Ecosyst Environ 269: 88-96

Macías-Hernández N, Athey K, Tonzo V, Wangensteen OS, Arnedo M, Harwood JD (2018) Molecular gut content analysis of different spider body parts. PLoS ONE 13: e0196589

McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32(S2): W20-W25

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4: e2584

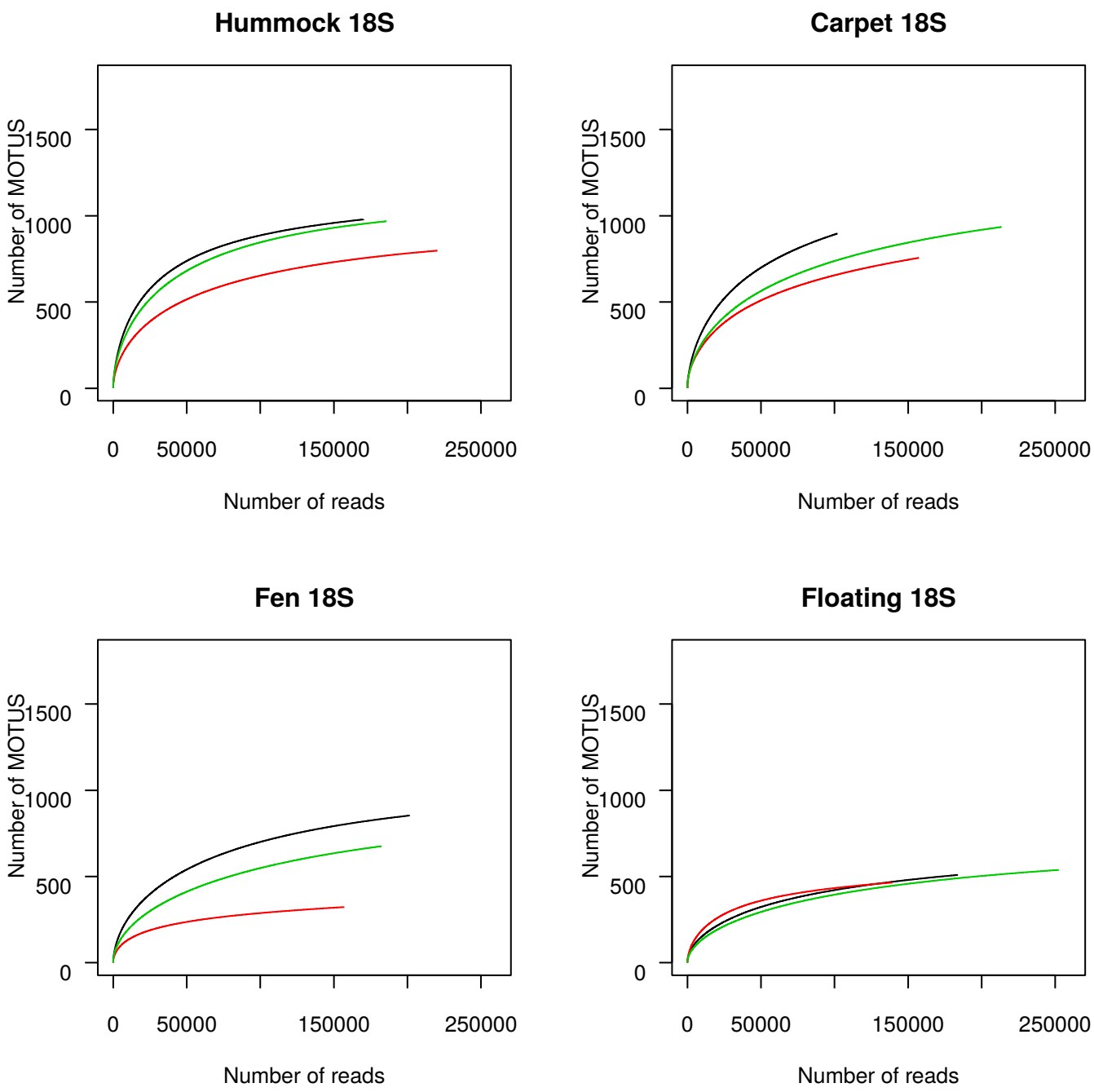
Siegenthaler A, Wangensteen OS, Benvenuto C, Campos J, Mariani S (2919) DNA metabarcoding unveils multiscale trophic variation in a widespread coastal opportunist. Mol Ecol 28: 232-249

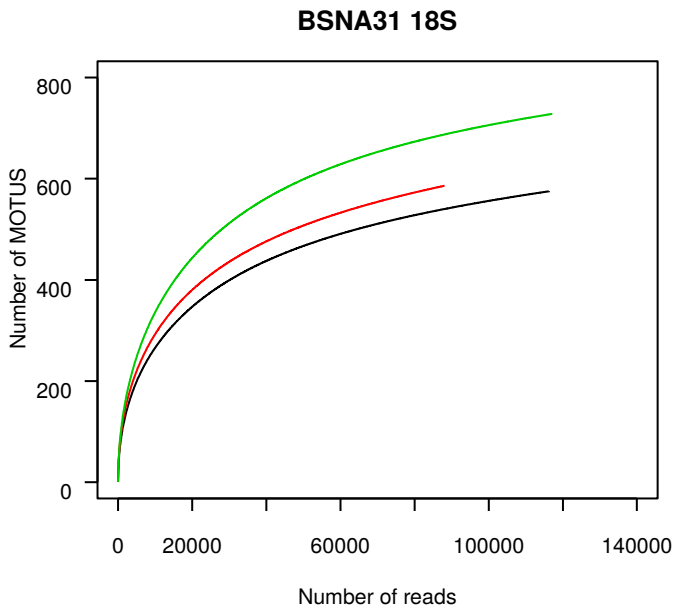
Wangensteen OS, Turon X (2017) Metabarcoding techniques for assessing biodiversity of marine animal forests In: Rossi S, Bramanti L, Gori A, Orejas C (eds) Marine Animal Forests. The Ecology of Benthic Biodiversity Hotspots, Springer International Publishing, Switzerland, ISBN 978-3-319-21011-7 pp. 445-473

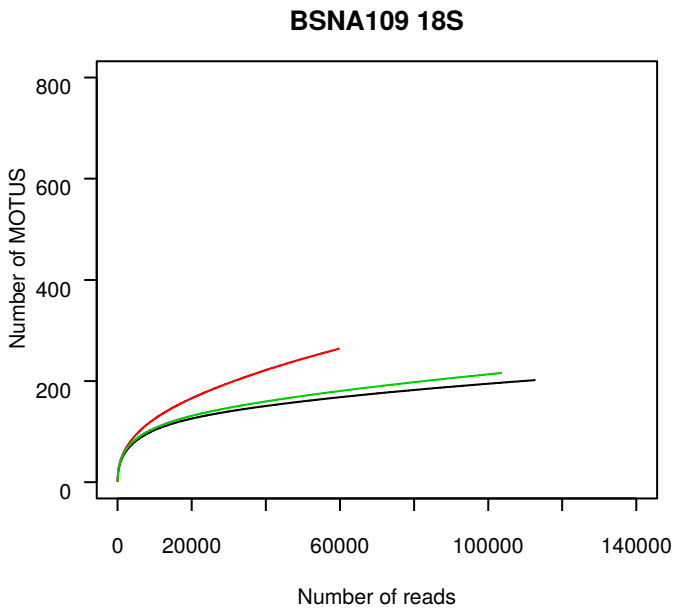
Wangensteen OS, Cebrian E, Palacín C, Turon X (2018a). Under the canopy: community-wide effects of invasive algae in marine protected areas revealed by metabarcoding. Mar Poll Bull 127: 54-66

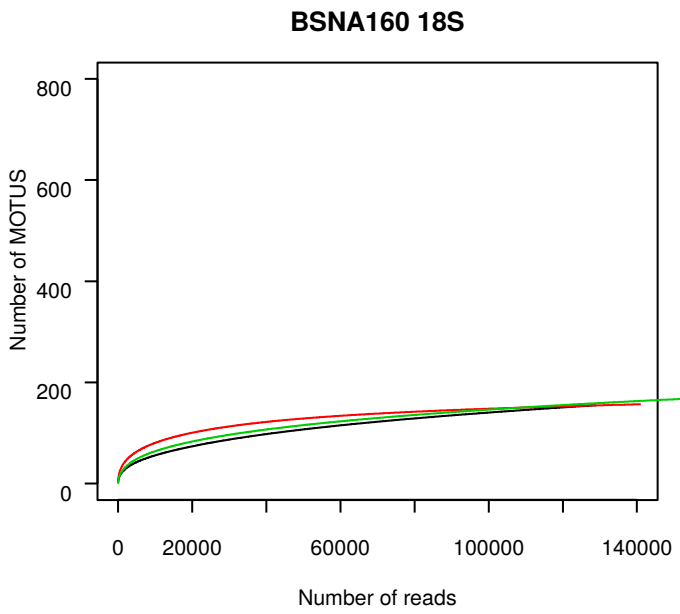
Wangensteen O S, Palacín C, Guardiola M, Turon X (2018b) DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ 6: e4705.

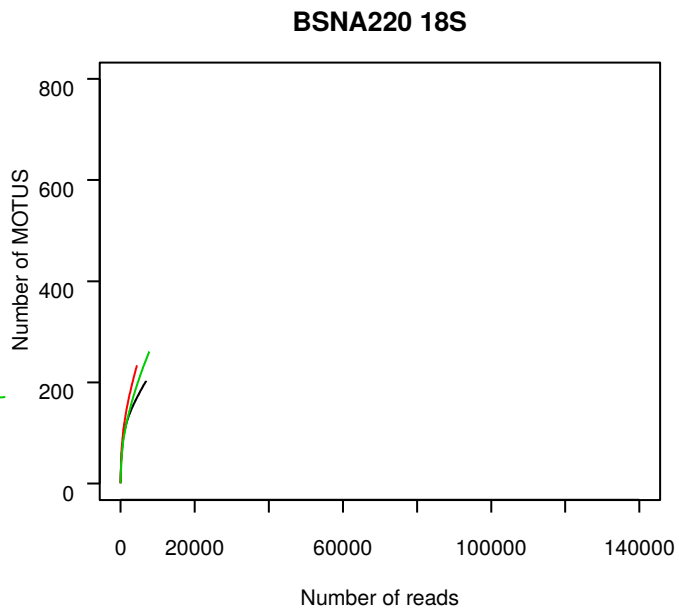
Supplementary file S2. Rarefaction curves for individual samples using 18S rRNA V7 and COI Leray-XT markers in modern and sedimentary communities (excluding singletons).

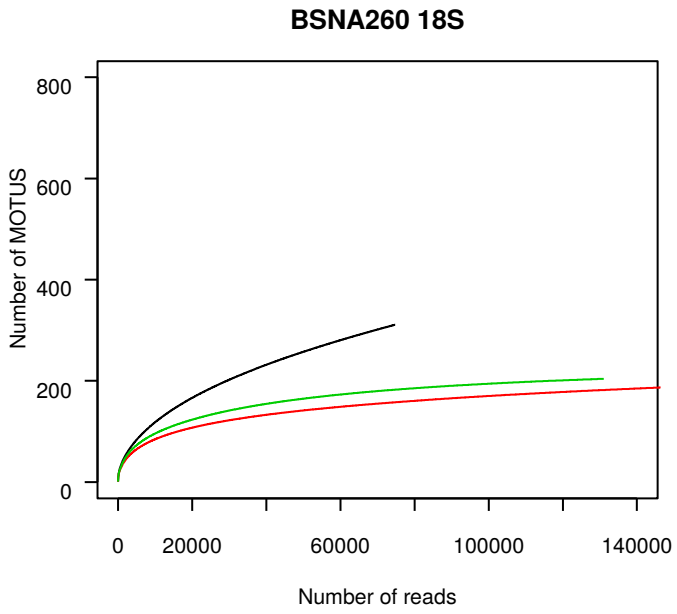


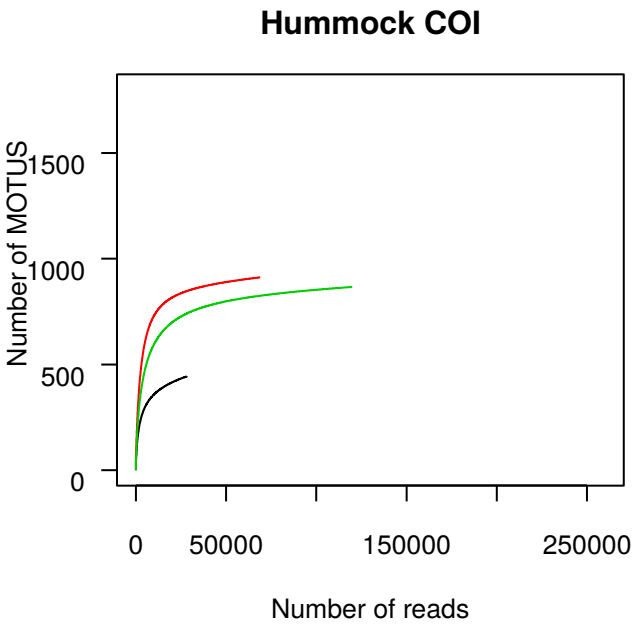

BSNA31 18S
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads

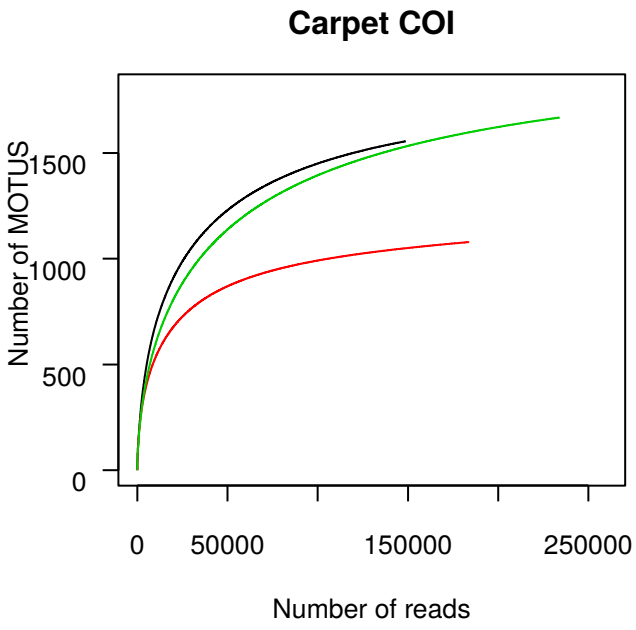

BSNA109 18S
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads

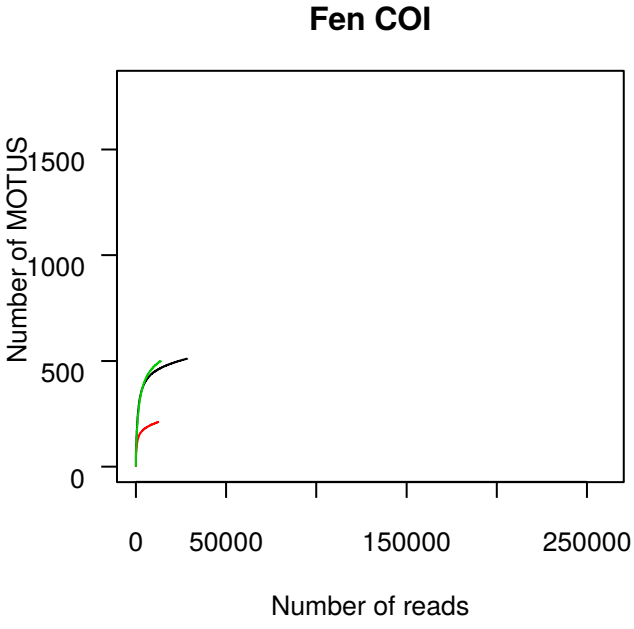

BSNA160 18S
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads

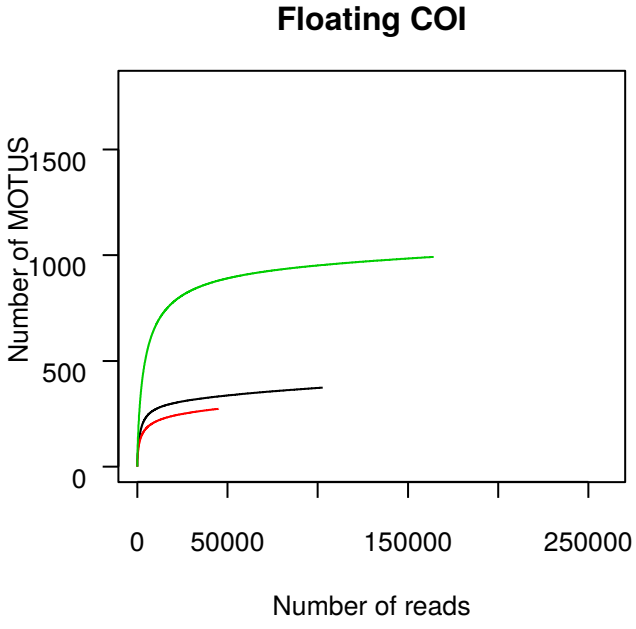

BSNA220 18S
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads

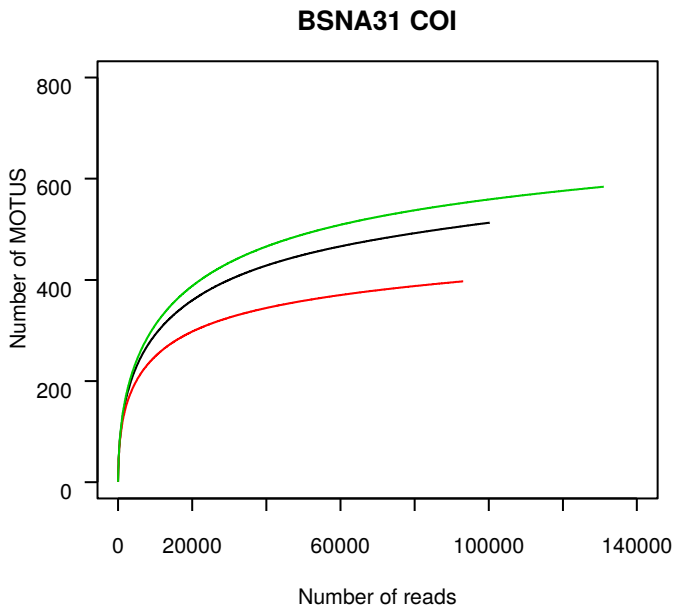

BSNA260 18S
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads

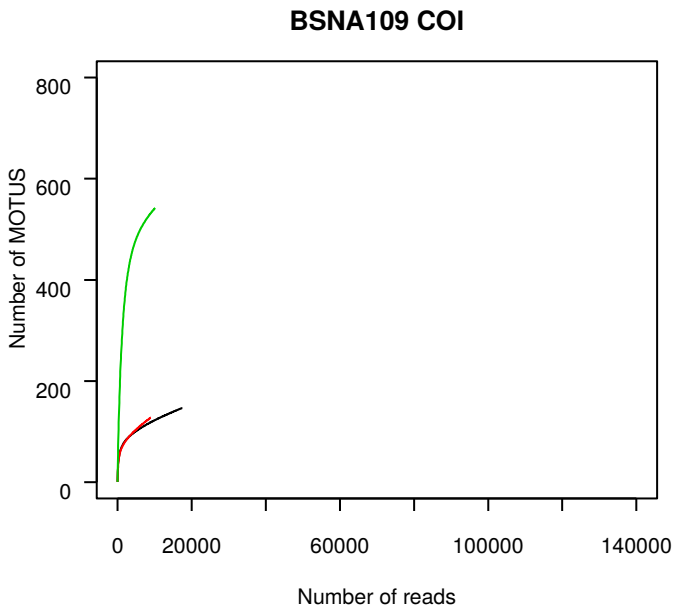

Hummock COI
Number of MOTUS
0
500
1000
1500
0
50000
150000
250000
Number of reads

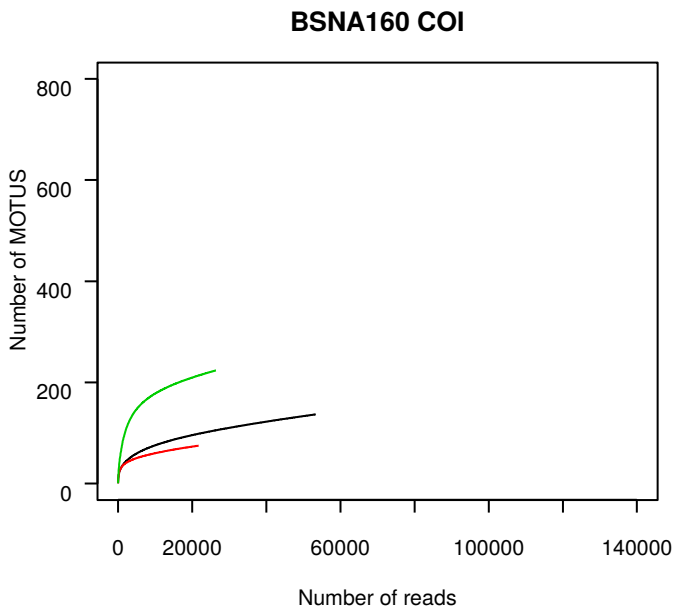

Carpet COI
Number of MOTUS
0
500
1000
1500
0
50000
150000
250000
Number of reads

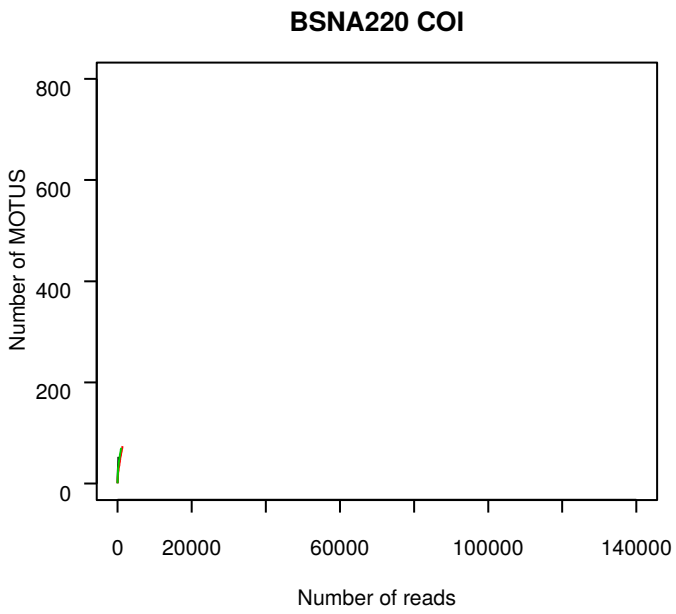

Fen COI
Number of MOTUS
0
500
1000
1500
0
50000
150000
250000
Number of reads

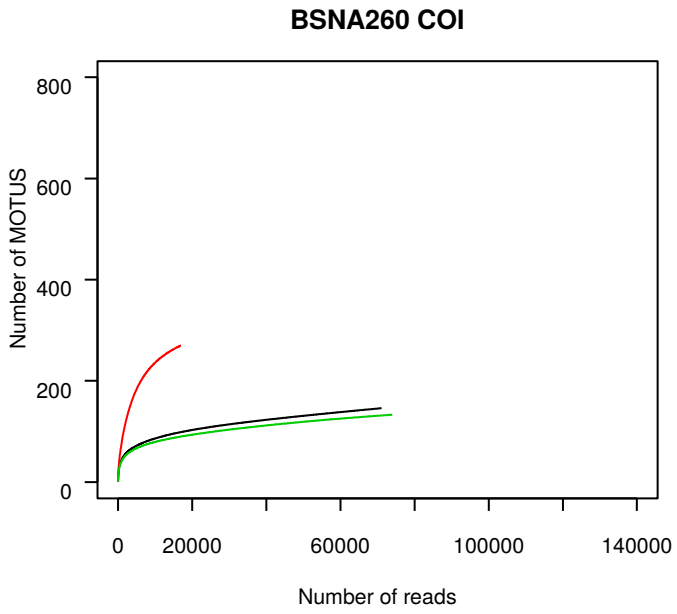

Floating COI
Number of MOTUS
0
500
1000
1500
0
50000
150000
250000
Number of reads


BSNA31 COI
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads


BSNA109 COI
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads


BSNA160 COI
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads


BSNA220 COI
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads


BSNA260 COI
Number of MOTUS
0
200
400
600
800
0
20000
60000
100000
140000
Number of reads