A global overview of the trophic structure within microbiomes across ecosystems

Wu Xiong a,b, Alexandre Jousset a,*, Rong Li a, Manuel Delgado-Baquerizo c, Mohammad Bahram d,e, Ramiro Logares f, Benjami...
biogeography, particularly of bacteria and fungi (Caporaso et al., 2012; Nilsson et al., 2019; White et al., 2016). Those analyses provided insights on the importance of multiple environmental factors in determining the diversity and community composition of these microbial groups (Delgado-Baquerizo et al., 2018; Tedersoo et al., 2014; Thompson et al., 2017). The diversity and community composition of soil and marine microbial protists is now also described at a global scale (de Vargas et al., 2015; Oliverio et al., 2020). These studies show profound differences in microbiome composition both at the taxonomic and functional level across habitats (Thompson et al., 2017). Most studies have, however, focused almost entirely on single microbial groups. Therefore, it remains unknown if the same abiotic factors structure distinct microbial groups similarly in their taxonomic and functional composition across Earth’s habitats.

Protists (particularly predatory protists) are key determinants of microbial biomass as predators that drive nutrient cycling (coined the microbial loop) in both aquatic (Azam et al., 1983) and terrestrial habitats (Clarholm, 1985). Protistan predators impact bacterial community structures as shown in controlled laboratory and greenhouse experiments (Batani et al., 2016; Bonkowski, 2004). The importance of protistan predation on bacterial communities and direct links between both organism groups is, however, difficult to observe under field conditions. Few studies using correlations between bacterial and protistan community data have provided indirect evidence that both groups are linked across environmental gradients, such as shown by similar diversity patterns of protistan predators and bacterial prey in marine (Steele et al., 2011) and soil habitats (Oliverio et al., 2020; Wilschut et al., 2019). It remains to be explored if correlations between protistan and bacterial communities are equally represented across systems and studies. Protists include also many other functional units (Geisen et al., 2018), among those functional groups are phototrophs (all eukaryotic algae that are a key component of the global biodiversity especially in sunlit waters (de Vargas et al., 2015) and parasites (many animal, plant and algal pathogens such as the malaria causing Plasmodium spp. and oomycte Phytophthora spp.) that can dominate protistan soil communities in the tropics (Mahé et al., 2017). Together, protists compose a major fraction of taxa in Earth’s microbiomes that together with the often interacting bacteria determine major ecosystem functions (Geisen et al., 2018). However, the changes in the structure of the trophic microbiome across habitats from aquatic to terrestrial ecosystems remains largely undetermined. As we did not expect a direct impact of non-predatory protists (such as phototrophs) on bacterial communities, we focused our analyses and hypotheses on links between predatory protists and bacteria.

Overall, communities of protists and bacteria are shaped by different physicochemical processes (Oliverio et al., 2020) and therefore trophic interactions might vary across Earth’s habitats. For example, trophic links might be more important in nutrient-poor systems compared to nutrient-rich systems, as high nutrient availability leads to an increased importance of bottom-up factors that drive microbiome composition (Lenoir et al., 2007; Moore et al., 2003), a link that needs to be confirmed. Compared to aquatic systems, soils are highly heterogeneous and offer more inhabitable niche space that might support a higher microbial diversity (Card et al., 2018; Raynaud and Nunnari, 2014). The micro-scale heterogeneity of soils also implies that bacteria might benefit more than protists as habitable pore space decreases more sharply with size for protists than for bacteria (Rutherford and Juma, 1992). Moreover, anthropogenic factors could also potentially alter the links between bacterial and protistan communities across environmental gradients such as through increases in bottom-up processes. Both climate (Perry et al., 2010; Voigt et al., 2003) and land management (Jennings and Pocock, 2009; Jonsson et al., 2012) are known to change trophic interactions, particularly by predominantly impacting higher rather than lower trophic levels. Profound evidence for this “trophic sensitivity hypothesis” exists for macroscopic animals (Cheng et al., 2017; Voigt et al., 2003). Hardly any information on the existence of the trophic sensitivity hypothesis in microorganisms is available (Thakur et al., 2020).

Here, we aimed to obtain a cumulative global-scale overview of the structure of the trophic microbiomes of bacterial and protistan communities across major terrestrial and aquatic habitats by combining and re-analysing datasets that simultaneously examined bacterial and protistan communities by high-throughput sequencing of 16S and 18S rRNA gene markers. We hypothesized that trophic composition of microbiomes will drastically change across habitats, but that the linkages between bacteria and protists in terms of community structures are consistent across freshwater, marine and terrestrial ecosystems. Due to fundamental differences in the habitats as visualized by rather open planktonic compared to matrix-composed soils, we anticipated differences in the ratios of larger-sized total protist and predatory protists to bacteria richness as well as the potential interrelations between them. In addition, we investigated if the trophic structure is affected by anthropogenically managed agricultural systems compared with natural grassland/forest soils.

2. Methods

2.1. Dataset collection

We performed a literature review to select papers that used high-throughput sequencing to analyze both protistan (via 18S rRNA genes) and bacterial (via 16S rRNA genes) communities. To this end, we performed a literature screening using Web of Science (https://webofknowl edge.com/) with the search terms “TS = 18S AND TS = (Protist OR Protozoa) AND TS = 16S AND TS = (Sequencing OR Illumina OR Miseq OR HiSeq OR High-throughput OR Highthroughput OR HTS)” and using Google Scholar (https://scholar.google.com/) with the search term “16S rRNA” AND “18S rRNA” AND (HTS OR “meta-barcoding” OR “high-throughput sequencing” OR “next-generation sequencing”) AND bacteria AND (protist OR protozoa) AND (DNA OR RNA) AND PCR AND Illumina” in March 2020. We removed papers/datasets based on the following criteria: 1) single-end or merged 18S rRNA reads which did not cover the V4 (<300 bp) or V9 (<100 bp) region; 2) 18S rRNA reads produced by using a specific primer set that has been shown to cover a only a minor part of the diversity of protists (Geisen et al., 2019); 3) single-end or merged 16S rRNA reads which do not cover the V4 region (<200 bp); 4) raw sequencing data were not publicly available or not provided by the authors upon request. To reduce likely introduced biases from other sequencing platforms (e.g., Roche 454 or Ion Torrent Personal Genome Machine), we extracted only datasets obtained by Illumina sequencing platforms that provide comparable results (Caporaso et al., 2012). By contacting other researchers, we obtained 8 additional unpublished datasets including one dataset from Lake Malawi, one dataset with samples taken from the air and one greenhouse experiment dataset from the Netherlands, as well as six datasets of different agricultural systems including cropping rotation, continuous cropping and different fertilizer experiments in China. Based on these selection criteria, a total of 38 datasets were selected with both 18S rRNA and 16S rRNA gene sequences for further analyses (Table S1). Only samples recovered from DNA templates were included as only a few analyses used RNA templates.

In total, we used 1,899 samples targeting the V4 region of 18S rRNA gene and 1,359 samples targeting the V9 region of this marker. These samples simultaneously contained bacterial 16S rRNA gene sequences, in which 80 samples including both the V4 and V9 regions of 18S rRNA gene (Table S2). The global distribution of all sampling points was visualized using the “ggmap” package (Kahle and Wickham, 2013) in R (version 3.5.3) as shown in Fig. 1a. We obtained climatic data including mean annual temperature (MAT) and mean annual precipitation (MAP) from the WorldClim database (https://www.worldclim.org/) based on latitude and longitude information for all collected samples with the “raster” package (Hijmans and van Etten, 2016) in R.
2.2. Bioinformatic re-analyses of protistan and bacterial community data

To avoid differences in bioinformatic pipelines (or clustering methods), we re-analyzed raw sequencing data of all samples using previously established protocols (Xiong et al., 2020, 2018) with the following modifications:

For 18S rRNA gene sequences, pair-end reads were merged with USEARCH v11 (Edgar, 2010). Single and merged sequences for each sample with expected errors > 1.0 or a length shorter than 300 bp (V4 region) or 100 bp (V9 region) were removed. We extracted the V4 region of 18S rRNA gene sequences with the primer sets: 616*f (TTAAARVGYTCGTAGTYG) and TAReukREV3 (ACTITGGTCITTGATYRA) or the V9 region of the 18S rRNA gene with the primer sets 1391F (GTACACCCGCGGT) and EukBr (TGATCCTTCTGCAGGTTCACCTAC) by “search_pcr2” command in USEARCH (Edgar, 2010) or with “Pcr.seqs” command via Mothur (Schloss et al., 2009) if one primer region was not included or removed. In order to obtain an equivalent sequencing depth for later analyses, we rarefied each sample to 10,000 reads (but kept the samples if they did not contain 10,000 reads at this stage). After trimming the V4 region of 18S rRNA reads to 300 bp (V9 of 18S rRNA reads to 100 bp), unique sequences were identified by VSEARCH (Rognes et al., 2016). We then identified amplicon sequence variants (ASV) or so-called “zOTU” with UNOISE3 (Edgar, 2010) for the V4 and V9 regions separately, with simultaneous removal of chimeras. We aimed at focusing on globally abundant taxa that should reduce study-specific differences in bioinformatic pipelines (or clustering methods). To avoid differences in bioinformatic pipelines (or clustering methods), we re-analyzed raw sequencing data of all samples using previously established protocols (Xiong et al., 2020, 2018) with the following modifications:

For 16S rRNA gene analyses, single or merged sequences for each sample with expected errors > 1.0 or a length shorter than 200 bp were removed. We extracted the V4 region of the 16S rRNA with the primer sets: 520F (AYTGGGYDTAAAGNG) and 802R (TACNVGGGTATCC), and rarefied each sample to 10,000 reads. After we trimmed the merged 16S rRNA reads to 200 bp, unique sequences were identified by VSEARCH. ASVs were generated with simultaneously removing chimeras. We further removed the 16S ASVs that contained fewer than 100 reads across all the samples. Finally, the 16S ASV representative sequences were matched against the RDP database (Cole et al., 2009; Wang et al., 2007). To focus on bacteria, we removed the reads assigned as chloroplast, mitochondria, archaea and eukaryotes.

Both bacterial and protistan ASV tables were further rarefied to 1,000 sequences for each sample to calculate diversity and community structure of bacteria and protists. We kept the unrarefied samples to calculate relative abundances of bacterial and protistan taxa. Analysis of similarities (ANOSIM) was applied to assess the statistical significance of different habitats/systems on the bacterial and protistan communities (Hellingen transformed) based on Bray–Curtis distance metrics with the “anosim” function in the “vegan” package (Oksanen et al., 2019) in R (https://www.r-project.org/) with 999 permutations. As controlled conditions (lab/greenhouse experiment) and plant rhizospheres significantly determined (ANOSIM, P < 0.001) both soil bacterial and protistan communities (Table S4), we removed soil samples from controlled conditions and from plant rhizospheres from further analyses. We finally obtained 4 main habitats with 19 independent studies, in which city water contained 1 independent study with 60 retained samples, lake water contained 1 independent study with 26 retained samples, marine habitats contained 8 independent studies with 264 retained samples, soils contained 8 independent studies and included 4 natural vegetation types of shrubland (retained sample size n = 23), moss (n = 29), forest (n = 59) and grassland (n = 391) as well as heavily human disturbed agricultural soils (n = 323). We assigned one lightly-fertilizer treated grassland from Tibetan alpine meadow (Soil No.009) as natural to distinguish from heavily managed agricultural soils (Table S1).
2.3. Statistical analyses

A principal coordinate analysis (PCoA) based on Bray–Curtis distance metrics was performed to explore differences in bacterial and protistan communities (Hellinger transformed) at both ASV and phylum/supergroup levels across different habitats/systems. We further used variation partitioning analysis to determine the contributions of habitats, vegetation and management types as well as individual data sets and their interactions to the variation of bacterial and protistan communities (Hellinger transformed) through the “varpart” function in the “vegan” package in R. At the ASV level, we also used Bray–Curtis distance to evaluate the community dissimilarity in different systems for both bacteria and protists. Mantel test was performed to determine the correlations between the distances of bacterial and protistan communities with “vegan” package in R. We defined the core/dominant bacterial or protistan ASVs as those with an average relative abundance over 0.5% across samples and which were observed in more than 50% of samples in each of the main habitats (Table S6). As diversity indices turned out to be highly correlated between numbers of observed ASVs, Chao1 richness, ACE richness, Shannon diversity and Shannon evenness of both bacteria and protists (Fig. S2), we selected richness (number of observed ASV) as a commonly used biodiversity metric to evaluate the diversity of bacteria and protists. It should be noted that the richness index is not meant as a complete inventory of existing bacterial and protistan diversity but to compare the relation of higher to lower trophic diversity across habitats. To focus on predatory protistan richness, we extracted predatory protist ASVs from the rarefied protistan table (not including the often mixotrophic and difficult to functionally place Dinoflagellata). Furthermore, we performed a test to evaluate the robustness of our approach by correlating rarefied predatory protistan richness (rarefied each sample to 100 sequences) and un-rarefied predatory protistan richness. Both highly correlated ($R^2 = 0.78$, $P < 0.001$) across the samples (data not shown here). We used the correlations between predatory protists (both richness and relative abundance) and bacterial richness as a proxy to study the interrelationships between microbial predators and prey across systems. For those correlation-based analyses, we focus on marine and soils as highly repeated habitats (both containing 8 individual datasets) including three main soils of forest, grassland and agricultural soils (sample size $> 50$). We tested relationships (linear, quadratic and cubic regressions) between predatory protist richness and the relative abundance of predatory protists with bacterial richness by the “lm” function in R. We identified the best model for the regression by the lowest Akaike Information Criteria (Burnham et al., 2011; Peruggia, 2003). We calculated Spearman’s rank correlations between the relative abundance of predatory protists and main bacterial taxa with the “corr.test” function in the package “psych” (Revelle, 2020) in R, the $P$ values were adjusted by the false discovery rate method (Benjamini and Hochberg, 1995). Further zooming into soils, we tested the Spearman’s rank correlations between the MAT, MAP and absolute latitude with the main microbial parameters. To relax the non-homogeneously distributed data, we used “Welch one-way test” for multiple comparisons, followed by “pairwise.t.test” with $P$ values being adjusted by the false discovery rate method. Student’s $t$-test was used to compare significance of the difference between the means of two treatments.
3. Results

3.1. Patterns of bacterial and protistan communities across Earth’s habitats

Phylum/supergroup level analyses revealed clear differences in the taxonomic composition of the major bacterial and protistan lineages across the main habitats (Fig. 2 and Table S5). Proteobacteria dominated bacterial communities across all habitats. Among Proteobacteria, Alphaproteobacteria was the most abundant taxon in city water, marine and soils, while Betaproteobacteria was most abundant in lake habitats (Table S5). Soils showed higher proportions (pairwise T-tests, \( P < 0.05 \)) of Acidobacteria and Actinobacteria compared to aquatic habitats (city water, lake and marine samples). The proportions of Cyanobacteria (phototrophic bacteria) was higher (pairwise T-tests, \( P < 0.05 \)) in marine samples, than in lake, city water and soils (with lowest level in soils). Among the dominant protistan taxa, Alveolata showed the highest proportion in marine samples (mostly Dinoflagellata; Table S6), Alveolata and Stramenopiles in lake, Stramenopiles in city water and Rhizaria in soils (mostly Cercozoa; Table S6). The most abundant protistan functional groups across all habitats were phototrophic and predatory protists (Fig. 2 and Table S5). Phototrophs were more abundant (pairwise T-tests, \( P < 0.05 \)) in aquatic habitats, while predatory protists dominated (pairwise T-tests, \( P < 0.05 \)) in soils. Protistan pathogens of plants were higher (pairwise T-tests, \( P < 0.05 \)) in soils compared with aquatic habitats. The relative abundance of parasitic protists was higher

---

**Fig. 3.** Mantel tests of correlation between the distance of protistan communities and bacterial communities in the main habitats (a, b, c and d) and soils (vegetation and management types) (e, f and g). Solid line presents a linear correlation.
(pairwise T-tests, \( P < 0.05 \)) in soils than in lake, city water and marine samples. Specific analyses in soils revealed differences in the taxonomic and functional composition of bacteria and protists between forest, grassland and agricultural soils (Fig. 2 and Table S5). For example, agricultural soils showed higher (pairwise T-tests, \( P < 0.05 \)) proportions of Cyanobacteria, Proteobacteria and Archaeplastida (and in line with the latter all combined phototrophic protists), but lower (pairwise T-tests, \( P < 0.05 \)) proportions of Acidobacteria, Alveolata, Rhizaria and parasitic protists compared with natural forest and grassland soils.

The taxonomic differences between habitats led to expected differences in community structures of both bacteria and protists at the ASV and phylum/supergroup levels (Fig. S1). We found that different habitats, vegetation and management types as well as individual datasets all significantly determined (Adonis, \( P < 0.001 \)) structures of bacteria and protists. Variation partitioning analysis revealed that individual datasets most strongly influenced structures for bacteria and protists (Fig. S1). Interestingly, ASV level-targeted analyses showed that protists exhibited higher (student’s t-test, \( P < 0.001 \)) community dissimilarities than bacteria in all aquatic habitats (Fig. S3a), while the opposite pattern was evident across soils (Fig. S3a and Fig. S3b). soils displayed higher (pairwise T-tests, \( P < 0.05 \)) heterogeneity of both bacterial and protistan communities than aquatic habitats (Fig. S3a).

Diversity of both bacteria and protists, as determined by richness, was highest (pairwise T-tests, \( P < 0.05 \)) in soils compared with aquatic habitats (Fig. S3c) with lowest bacterial richness in marine samples. In addition, soils hosted higher (pairwise T-tests, \( P < 0.05 \)) diversities of predatory protists than aquatic habitats (Fig. S4a). Across soils, grassland showed higher (pairwise T-tests, \( P < 0.05 \)) richness of bacteria, protists and predatory protists than forest and agricultural soils (Fig. S3d and Fig. S4b).

3.2. Links between bacterial and protistan communities

Mantel tests showed that the distance of protistan communities significantly (Mantel test: \( P < 0.001 \)) correlated with the distance of bacterial communities across all habitats and within distinct soils (forest, grassland and agricultural soils) (Fig. 3). Among the 4 main habitats, marine samples showed the highest (pairwise T-tests, \( P < 0.05 \)) ratio of total protists to bacteria richness (compared to lake, soils and city water samples) (Fig. 4a). This pattern was driven to a large extent by the lowest bacterial richness in marine samples (Fig. S3c). Soil habitats showed a higher (pairwise T-tests, \( P < 0.05 \)) richness ratio of predatory protists to bacteria compared to aquatic habitats (Fig. 4c). Across soils, ratios of total protists to bacteria richness and predatory protists to bacteria richness in grassland was higher (pairwise T-tests, \( P < 0.05 \)) than in forest and agricultural soils (Fig. 4b and Fig. 4d).

We further tested whether bacterial diversity correlates with the richness and abundance of predatory protists. Both predatory protist richness (Cubic regression: \( R^2 = 0.278, P < 0.001 \), Linear regression: \( R^2 = 0.100, P < 0.001 \)) and the relative abundance of predatory protists (Cubic regression: \( R^2 = 0.200, P < 0.001 \), Linear regression: \( R^2 = 0.139, P < 0.001 \)) positively correlated with bacterial richness in marine samples (Fig. 5a and Fig. 5f). Predatory protist richness showed a concave-shaped (Cubic regression: \( R^2 = 0.038, P < 0.001 \)) relationship with bacterial richness in soils (Fig. 5b). The relative abundance of predatory protists, however, exhibited a hump-shaped (Quadratic regression: \( R^2 = 0.059, P < 0.001 \)) relationship with bacterial richness (Fig. 5g). In forest soils, the richness of predatory protists to bacteria positively correlated (Linear regression: \( R^2 = 0.261, P < 0.001 \)) (Fig. 5c). In grassland soils, the relative abundance of predatory protists positively correlated with bacterial richness (Linear regression: \( R^2 = 0.086, P < 0.001 \)) (Fig. 5i). In contrast to those natural soils, in

\[
\begin{align*}
\text{Main habitats} & : \quad \text{Citywater} \quad \text{Lake} \quad \text{Marine} \quad \text{Soils} \\
\text{Forest} & \quad \text{Grassland} \quad \text{Agriculture}
\end{align*}
\]

Fig. 4. Ratio of total protists to bacteria richness in the main habitats (a) and soils (vegetation and management types) (b); ratio of predatory protists to bacteria richness in the main habitats (c) and soils (vegetation and management types) (d). Different letters indicate a significant difference at \( P < 0.05 \) according to “pairwise.
test”. Red dot indicates the median value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
agricultural soils richness of predatory protists and bacteria were linked in a concave shape (Quadratic regression: $R^2 = 0.165, P < 0.001$), while the relative abundance of predatory protists showed hump-shaped (Cubic regression: $R^2 = 0.275, P < 0.001$) relationship with bacterial richness (Fig. 5e and Fig. 5j). In addition, we found that correlations between the relative abundance of predatory protists and dominant bacterial taxa differed across systems (Fig. S5). For example, predatory protists positively (Spearman’s rank correlation, $P < 0.01$) correlated with Cyanobacteria, Planctomycetes and Verrucomicrobia in marine samples, while predatory protists negatively (Spearman’s rank correlation, $P < 0.001$) correlated with diverse bacterial phyla including Cyano- bacteria, Planctomycetes and Verrucomicrobia in soils. Predatory protists positively (Spearman’s rank correlation, $P < 0.01$) correlated with Proteobacteria (Betaproteobacteria and Gammaproteobacteria) in grassland, but negatively (Spearman’s rank correlation, $P < 0.001$) in agricultural soils.

Targeted analysis in soils revealed that the diversity and composition of both bacteria and protists are correlated with climate factors and location (Fig. 6). We showed that richness of bacteria, protists and predatory protists negatively (Spearman’s rank correlation, $P < 0.001$) correlated with mean annual temperature. Protistan and predatory protistan richness positively (Spearman’s rank correlation, $P < 0.001$) correlated with mean annual precipitation, but this pattern was opposite (Spearman’s rank correlation, $P < 0.001$) for bacterial richness. Ratios of total protists to bacteria richness and predatory protists to bacteria richness positively (Spearman’s rank correlation, $P < 0.001$) correlated with mean annual precipitation and absolute latitude, but negatively (Spearman’s rank correlation, $P < 0.001$) correlated with mean annual temperature. Phototrophic protists positively (Spearman’s rank correlation, $P < 0.001$) correlated with mean annual temperature, while parasites and plant pathogens showed the opposite trends (Spearman’s rank correlation, $P < 0.001$).

4. Discussion

Our global synthesis provides a catalogue of the trophic structure of the microbiome across the major habitats of the planet. We also found predictable associations between bacteria and protists across and within contrasting habitats. Moreover, our study reveals that terrestrial ecosystems dominated by predatory protists are far more heterogeneous in their microbiomes than aquatic environments, which are dominated by phototrophic organisms.

Our data supports previously reported biogeographic patterns of bacterial and protistan communities. For example, we confirm the reported dominance of Proteobacteria (Alphaproteobacteria) in upper ocean (Sunagawa et al., 2015) and soil habitats (Delgado-Baquerizo et al., 2018), as well as of protistan Dinoflagellates in marine habitats (Bescot et al., 2016) and of Cercozoa in soils (Oliverio et al., 2020). Our results also mirrored known patterns at the protistan functional level, such as a dominance of phototrophic protists (Decelle et al., 2015; Sanders, 2011) as well as predatory protists in marine plankton (de Vargas et al., 2015) and of predatory protists in soils (Oliverio et al., 2020).

Unlike most previous studies, our focus was on the direct comparison of bacterial and protistan communities. We found that the distance of total protistan community structures (protistan beta-diversity) were more dissimilar than bacteria in aquatic habitats, as suggested before in marine habitats (Logares et al., 2020), while the opposite pattern was found in soils. The increased dominance of bacteria in highly heterogeneous soils can be attributed to their ability to inhabit smaller soil pores than their protistan predators (Rutherford and Juma, 1992). We provide evidence that soil ecosystems are far more heterogeneous across all the habitats in terms of bacterial and protists communities. Overall, soils host most diversity and have a higher dissimilarity of both bacteria and protistan community structures as well as the most diverse microbiomes, which can be attributed to their higher spatial heterogeneity compared to aquatic systems (Fierer, 2017; Walters and Martiny, 2020). Moreover, both richness and relative abundance of predatory protists positively correlated with bacterial richness in marine samples, which supports the importance of protistan predation on bacteria in marine food webs (Azam et al., 1983; de Vargas et al., 2015; Sherr and Sherr, 2002).

Besides the cross-habitat differences in microbial compositions, we also found that anthropogenic factors might induce differences on...
microbiomes with differences between trophic level groups. The lower soil biodiversity including bacteria and protists (including predatory protists) in agricultural ecosystems compared with natural grassland soils, supports findings on bacteria under long-term agricultural practices (Fierer et al., 2013) and protists in respect to increased fertilizer inputs (Zhao et al., 2019). The relative abundances of Cyanobacteria, Archaeplastida and photrotrophic protists were enriched in agricultural soils compared with natural forest and grassland soils, suggesting that anthropogenic factors including nutrient input and increased light availability due to the removal of plants might promote phototrophic algae as shown in aquatic systems (Davidson et al., 2014). In line with Schulz et al. (2019) who showed substantial reduction of animal and plant parasites with increasing land use intensity, we also found lower parasites in agricultural versus natural forest and grassland soils, which might be associated with a loss in the diversity of other macroscopic organisms including soil fauna (Tsiafouli et al., 2015) and aboveground insects (Seibold et al., 2019).

Differences between soils were more profound at higher trophic level protists than at bacteria such as shown for the decreased ratios of both total protists and predatory protists to bacteria richness in agricultural than in natural grassland soils. This confirms our hypothesis that protists are more severely impacted than bacteria by anthropogenic effects. This provides further support that higher trophic level protists are more sensitive to disturbance than their bacterial prey (Thakur et al., 2020) and therefore follow the trophic sensitivity hypothesis claimed for animals (Cheng et al., 2017; Voigt et al., 2003). These effects can be attributed to direct negative effects of management on protists and by indirect effects through increased nutrient inputs leading to a dominance of bottom-up control by nutrients (Moore et al., 2003; Scheu and Schaefer, 1998). This suggestion is supported by positive correlations between predatory protists and bacterial richness in natural forest and grasslands, but not in agricultural soils.

We acknowledge several potential biases inherent with our approach. For example, there was an expected influence of individual datasets (Ramirez et al., 2018) and biases through differences in sampling and DNA extraction, primer-use and PCR regime (Djurhuus et al., 2017; Pawluczynk et al., 2015; Penton et al., 2016). Yet, our combined re-analysis of multiple studies and the focus on highly repeated habitats likely reveals highly reliable trophic microbiome profiles in marine systems and soils.

5. Conclusions

Our study shows that the associations between communities of bacteria and protists are consistent across global habitats, despite that contrasting community compositions and specific links within dominate these systems. Microbiome diversity and heterogeneity was highest in soils compared with aquatic environments. Taken together, our study advances our knowledge on the trophic structure of the Earth microbiome, and alert that anthropogenic factors such as land management might affect potential trophic structure and ultimately the ecosystem functions that are driven by them.

6. Authors’ contributions

S.G., A.J. and W.X. conceived and designed this study. W.X. performed the bioinformatics processing and data analysis. S.G., A.J., W.X., M.D-B. and Ramiro L. interpreted the results. W.X. and S.G. wrote the manuscript with substantial inputs and final approval from all authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Prof. Dr. Noah Fierer from University of Colorado for constructive comments on the manuscript. We also thank Angela M. Oliverio, Walter Traunspurger and Juan Li for providing part of the datasets. W.X. and A.J. were supported by NWO from the Netherlands Organisation for Scientific Research (ALWGR.2017.016). S.G. was supported by an NWO-VENI grant from the Netherlands Organisation for Scientific Research (016.Veni.181.078). M.D-B. is supported by a Ramón y Cajal grant from the Spanish Ministry of Science and Innovation (RYC2018-025483-I).
Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2021.106438.

References


