The Microbiome Stress Project: Toward a Global Meta-Analysis of Environmental Stressors and Their Effects on Microbial Communities

Jennifer D. Rocca[1]*, Marie Simonin[1]*, Joanna R. Blaszczak[1,2], Jessica G. Ernakovich[3], Sean M. Gibbons[4,5,6], Firas S. Midani[7] and Alex D. Washburne[1,8]

[1] Department of Biology, Duke University, Durham, NC, United States, [2] Flathead Lake Biological Station, University of Montana, Polson, MT, United States, [3] Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH, United States, [4] Institute for Systems Biology, Seattle, WA, United States, [5] Molecular and Cellular Biology Program, University of Washington, Seattle, WA, United States, [6] eScience Institute, University of Washington, Seattle, WA, United States, [7] Center for Genomic and Computational Biology, Duke University, Durham, NC, United States, [8] Department of Microbiology and Immunology, Montana State University, Bozeman, MT, United States

Microbial community structure is highly sensitive to natural (e.g., drought, temperature, fire) and anthropogenic (e.g., heavy metal exposure, land-use change) stressors. However, despite an immense amount of data generated, systematic, cross-environment analyses of microbiome responses to multiple disturbances are lacking. Here, we present the Microbiome Stress Project, an open-access database of environmental and host-associated 16S rRNA amplicon sequencing studies collected to facilitate cross-study analyses of microbiome responses to stressors. This database will comprise published and unpublished datasets re-processed from the raw sequences into exact sequence variants using our standardized computational pipeline. Our database will provide insight into general response patterns of microbiome diversity, structure, and stability to environmental stressors. It will also enable the identification of cross-study associations between single or multiple stressors and specific microbial clades. Here, we present a proof-of-concept meta-analysis of 606 microbiomes (from nine studies) to assess microbial community responses to: (1) one stressor in one environment: soil warming across a variety of soil types, (2) a range of stressors in one environment: soil microbiome responses to a comprehensive set of stressors (incl. temperature, diesel, antibiotics, land use change, drought, and heavy metals), (3) one stressor across a range of environments: copper exposure effects on soil, sediment, activated-sludge reactors, and gut environments, and (4) the general trends of microbiome stressor responses. Overall, we found that stressor exposure significantly decreases microbiome alpha diversity and increases beta diversity (community dispersion) across a range of environments and stressor types. We observed a hump-shaped relationship between microbial community resistance to stressors (i.e., the average pairwise similarity score between the control

and stressed communities) and alpha diversity. We used Phylofactor to identify microbial clades and individual taxa as potential bioindicators of copper contamination across different environments. Using standardized computational and statistical methods, the Microbiome Stress Project will leverage thousands of existing datasets to build a general framework for how microbial communities respond to environmental stress.

INTRODUCTION

In the past decade, the advent of high-throughput sequencing technologies has enabled microbial ecologists to characterize microbial community responses to environmental change at an unprecedented pace. Thousands of studies are now available on the impact of natural and anthropogenic stressors in controlled conditions or along environmental gradients spanning a wide range of biomes. Large collaborative endeavors, like the Earth and Human Microbiome Projects, revealed fundamental biogeographic patterns of microbial diversity under "baseline" or "steady state" conditions (Human Microbiome Project Consortium, 2012; Gilbert et al., 2014; Lloyd-Price et al., 2017; Thompson et al., 2017). While this baseline knowledge is crucial, similar large-scale initiatives are necessary for clarifying how microbiomes respond to fluctuating environmental conditions.

Environmental stressors occur over varying magnitudes, frequencies, and durations (Bender et al., 1984), introducing spatiotemporal heterogeneity into the environment. Spatiotemporal heterogeneity is a key driver in both the maintenance and depletion of biodiversity (Connell, 1978; Huston, 1979; Crain et al., 2008; Shade et al., 2012; Piggott et al., 2015). In ecology, *stressor, disturbance, perturbation,* and *threat* are often used interchangeably and refer to a variety of environmental changes (natural, anthropogenic, abiotic or biotic). Here, we use the term stressor to refer to any factor that alters steady-state environmental conditions (biotic or abiotic) and influences the growth or mortality of organisms in a community, resulting in either deterministic or stochastic shifts in stationary relative abundance profiles of microbiomes.

The consequences of environmental stressors on organisms are highly context dependent. Often stressors of increasing intensity/duration induce increasing stress levels but they can also affect organisms non-monotonically. For instance, moderate levels of a stressor, such as exposure to some chemical elements (e.g., micronutrients), can be beneficial to organisms, but extremely high levels can impose adverse effects. This common dose-response pattern is often referred to as hormesis or a subsidy-stress response (Odum et al., 1979; Odum, 1985; Calabrese and Baldwin, 2002). Those direct effects of stressors can propagate through ecological interaction webs causing collateral damage. Thus, stressors have both direct (Schimel et al., 2007) and indirect (Vellend, 2010; Evans and Wallenstein, 2014; Knelman et al., 2014) effects on individual taxa, leading to structural shifts in the microbial communities. Stressor-induced changes in community structure and diversity can, in turn, create

feedbacks that further alter the host state (Reese et al., 2018), environmental conditions (Gibbons et al., 2016; Ratzke and Gore, 2018), or microbial ecosystem function (Bissett et al., 2013; Philippot et al., 2013).

In addition to ecological feedbacks, evolutionary feedbacks can influence how microbial communities respond to a stressor (Sanchez and Gore, 2013). For example, rapid evolution of antibiotic resistance can allow species to expand into environments that would normally be restrictive (Baym et al., 2016). However, despite the rapid expansion of research on environmental and host-associated microbiomes, we still have few generalizable insights for how microorganisms respond to stressors at the individual, population, or community levels (Treseder, 2008; Shade et al., 2012; Holden and Treseder, 2013; Duvallet, 2018). Taken together, the combined ecological data from each independent microbiome study could be a powerful resource for characterizing the processes underlying microbial community assembly and microbial population sensitivity or tolerance to environmental stressors. Pioneering meta-analyses focused on environmental perturbations in soil (Ramirez et al., 2018), human gut ecosystems (Duvallet et al., 2017; Gibbons et al., 2018; Jackson et al., 2018), and even across distinct environments (Shade et al., 2013) now provide valuable frameworks for cross-study analyses of 16S rRNA gene amplicon datasets. However, large-scale disturbance meta-analyses that integrate *both* host-associated and free-living microbiomes are currently lacking.

Here, we present the Microbiome Stress Project, a publicly available database of environmental and host-associated amplicon sequencing studies designed to facilitate our understanding of microbial community responses to disturbed environments through cross-study analyses. More specifically, our goal is to address the following research objectives:

(1) One stressor in one environment: How consistent are microbiome responses to the same stressor within the same environment?
(2) A range of stressors within an environment: Are there common microbiome responses to different stressors within the same environment?
(3) One stressor across environments: Does a given stressor impose consistent effects on microbiomes across multiple environments, including host-associated and free-living systems?
(4) General trends of microbiome stressor responses: Are there general impacts of stressors on microbiomes across all environments?

These objectives lie on four principal axes for how environmental stressors impact microbiomes (**Figure 1A**). Capturing the variability in microbial response to a given stressor in one environment type (e.g., soil) will generate a more detailed understanding of how specific environmental parameters (e.g., soil pH) influence microbiome responses to stressors. We can also determine how the absolute magnitude of stressor treatments may be either dampened through historical exposure or exacerbated by temporal treatment regime (i.e., acute vs. chronic). Examining many stressors within the same environment enables the classification of the stressors themselves in terms of the responses they elicit, which may or may not reflect our assumptions. For instance, two stressors from the same "category" (e.g., two heavy metals) may impose similar impacts on microbiomes, but similar responses may also be possible with two distinct stressors (e.g., water stress and metal contaminations both leading to severe oxidative stress). Conversely, studying just one stressor across many environments allows us to assess the consistencies in the responses among these distinct microbiomes and to identify reliable microbial indicator taxa or clades of specific or multiple disturbances. Finally, by examining all available microbiome datasets, we can identify general effects of stressors on microbial community diversity and structure.

In this article, we introduce the Microbiome Stress Project methodological framework for building and analyzing the amplicon sequencing database (**Figure 1**) and present the results of a proof-of-concept meta-analysis (called pilot study hereafter) using raw sequence data from a subset of the larger database (**Figure 2**). We start by describing the results of a literature search, identifying studies explicitly focused on the impacts of stressors on microbial communities, from which we are building the full Microbiome Stress Project Database. Then, we present the analysis pipeline used to re-process the raw amplicon sequence data in a standardized fashion for the subsequent meta-analyses. The processed data, in the form of exact sequence variant (ESV) tables and corresponding meta-data, is accessible on our website (microbiomestressproject.weebly.com), which hosts the database that will continue to grow as we add studies. The database encompasses both published and unpublished datasets, including *in situ* environmental gradients and controlled experiments, with acute (single) or chronic (repeated) stressor treatments, with either pulse (short-term) or press (long-term) exposure to the stressors. Finally, we present results from our pilot study where we examine the response of soil microbial communities to a wide set of stressors: copper, temperature, antibiotics, oxygen, polycyclic aromatic hydrocarbons, and land use change; and compare the impact of copper contamination on bacterial community structure across a range of environments: soil, gut, sediment, and activated-sludge reactors. We identify community-level responses with alpha and beta diversity, and we highlight individual- and lineage-level responses to stressors using phylofactorization (Washburne et al., 2017).

The Microbiome Stress Project is still growing and is at an early stage. Here, we present our methodological framework to demonstrate how these data can be used to understand how stressors shape microbial ecosystems. We invite the extended microbiome research community to contribute to this collaborative project and help us improve this publicly available database by sharing their data.

METHODS
Literature Search

We performed an extensive literature search to assess the state of current research on the effects of stressors on microbial community structure. Our target questions were: (a) Which microbiome stressors are most commonly studied? (b) In which environments are these studies most commonly performed? (c) Which primers and sequencing platforms are most commonly used? and (d) How much of the microbiome sequencing data is publicly available?

A literature search was performed on Web of Science (Core Collection) in April 2018 using a comprehensive set of basic microbiome keywords (full list of keywords in Supplementary Material **Table S1**). We constrained the search to studies published between 2010 and 2018, which is the period of time when high-throughput sequencing methods are standardized and commonly used to measure microbial community structure and composition. We narrowed our initial search results to identify potential studies for the Microbiome Stress Project Database as those which had the basic microbiome keywords, in addition to at least one of our specific subcategory stressor keywords. We used these keywords to identify which categories of stressors ($n = 16$, e.g., heavy metals, fire, pathogens, etc.) were studied and in which environment ($n = 11$, e.g., soil, gut, reactor, etc.). These categories enabled us to address the first (a) and second (b) questions about the research landscape and determine whether or not the number of studies available is sufficient to perform robust and balanced meta-analyses.

We performed a bibliometric network analysis on the articles identified as potential candidates for the database using the unified VOS mapping technique (VOS viewer software; Van Eck and Waltman, 2010) and clustering (Waltman et al., 2010). In these networks, we investigated where the microbiome stressor studies were published and the co-citations between these journals to determine the disciplines covered by our literature search and their interactions. We also created a network based on the co-occurrence of stressor terms that were most frequently mentioned in the microbiome studies.

We investigated primer and sequencing platform usage and data availability (questions c and d) using a random subset of 150 studies, as each study needed to be manually screened for the necessary information since Web of Science keyword searches do not extract this information (i.e., this methodological information is not commonly included in abstracts or keywords). The information gleaned from question (c) was essential for standardizing the database and for developing our analytical workflow.

Selection of Studies for the Pilot Study

We selected nine studies (606 samples; 1 independent sample = 1 microbiome) with publicly available Illumina MiSeq V4 hypervariable region 16S rRNA gene sequencing data (**Figure 2**; Fuentes et al., 2016; Lin et al., 2016; Nunes et al., 2016; Sun et al.,
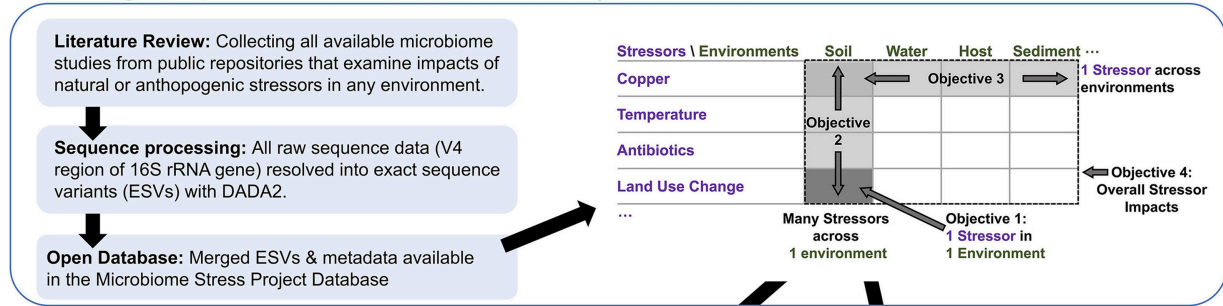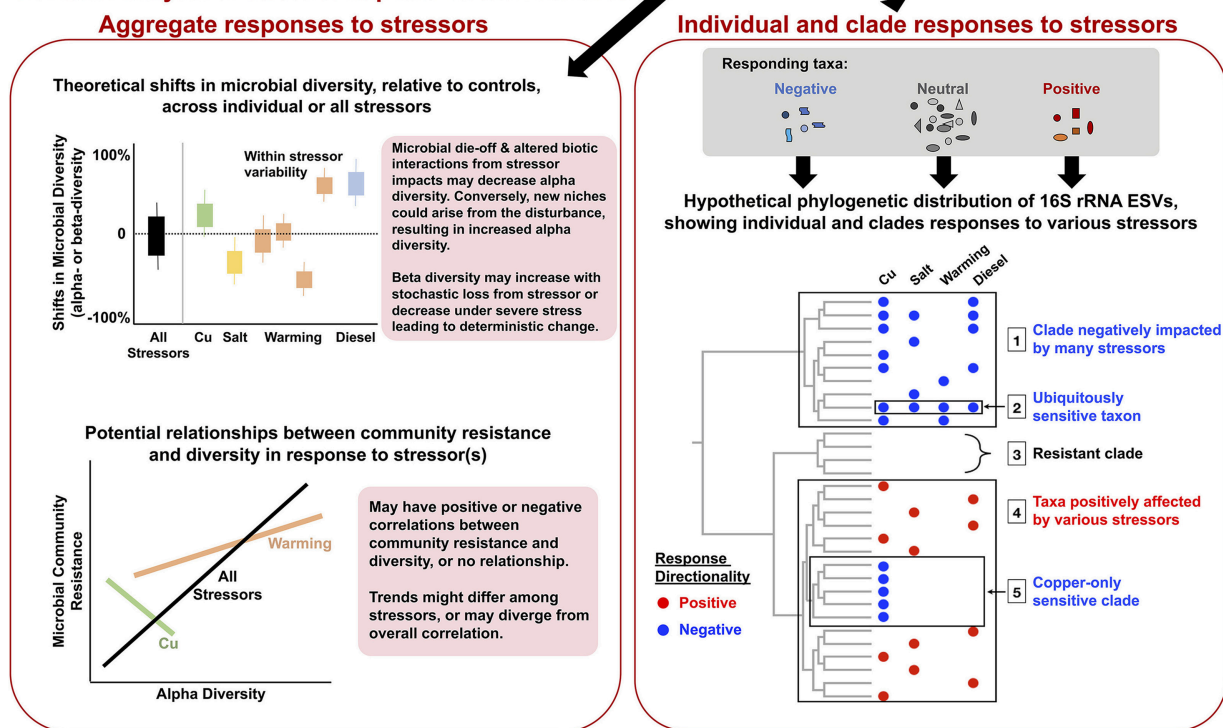
FIGURE 1 | Conceptual overview of the methodological framework of the Microbiome Stress Project: from **(A)** the development of a publicly available database comprised of standardized microbial community datasets, to **(B)** the examination of wholesale impacts of stressors on microbiome diversity, and the more targeted identification of putative indicator microbial taxa or clades to a comprehensive range of stressors across multiple environments. *"Cu" stands for copper*.

2016; Zhang et al., 2016; Ernakovich et al., 2017; Jurburg et al., 2017; Zhai et al., 2017), which were all controlled experiments that included control and treated microbiomes, enabling the calculation of proportional shifts in microbial diversity (see below section Downstream Analyses: Finding Ecological Patterns Hidden in ESVs Across Studies). These nine studies span the four axes presented in **Figure 1A**, targeting our primary objectives:

(1) One stressor in one environment: studies restricted to soil microbiomes exposed to warming (3 studies, 79 treated microbiomes).

(2) Multiple stressors within an environment: soil microbiomes exposed to 12 stressors, including exposure to: copper, temperature alteration, oxygen, antibiotics, polycyclic aromatic hydrocarbons, precipitation change, and over-grazing (6 studies, 325 treated microbiomes).

(3) One stressor across environments: microbial communities exposed to copper across four environments: soil, sludge, mice gut, and sediment (4 studies, 102 treated microbiomes). The Nunes et al. (2016) study fell within objectives 2 and 3, which investigated copper exposure on soil microbiomes.
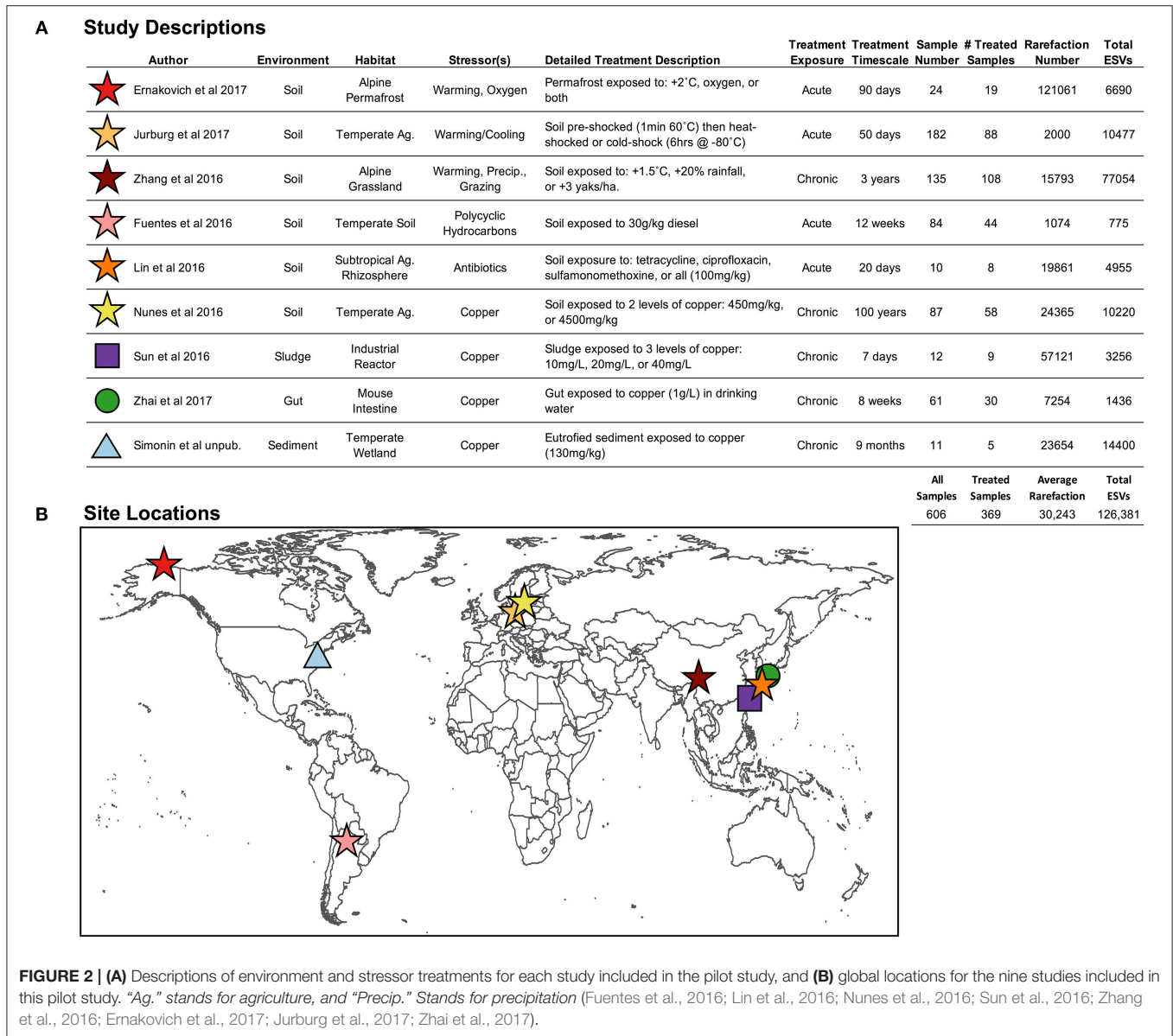
A **Study Descriptions**

	Author	Environment	Habitat	Stressor(s)	Detailed Treatment Description	Treatment Exposure	Treatment Timescale	Sample Number	# Treated Samples	Rarefaction Number	Total ESVs
★	Ernakovich et al 2017	Soil	Alpine Permafrost	Warming, Oxygen	Permafrost exposed to: +2˚C, oxygen, or both	Acute	90 days	24	19	121061	6690
★	Jurburg et al 2017	Soil	Temperate Ag.	Warming/Cooling	Soil pre-shocked (1min 60˚C) then heat-shocked or cold-shock (6hrs @ -80˚C)	Acute	50 days	182	88	2000	10477
★	Zhang et al 2016	Soil	Alpine Grassland	Warming, Precip., Grazing	Soil exposed to: +1.5˚C, +20% rainfall, or +3 yaks/ha.	Chronic	3 years	135	108	15793	77054
★	Fuentes et al 2016	Soil	Temperate Soil	Polycyclic Hydrocarbons	Soil exposed to 30g/kg diesel	Acute	12 weeks	84	44	1074	775
★	Lin et al 2016	Soil	Subtropical Ag. Rhizosphere	Antibiotics	Soil exposure to: tetracycline, ciprofloxacin, sulfamonomethoxine, or all (100mg/kg)	Acute	20 days	10	8	19861	4955
★	Nunes et al 2016	Soil	Temperate Ag.	Copper	Soil exposed to 2 levels of copper: 450mg/kg, or 4500mg/kg	Chronic	100 years	87	58	24365	10220
■	Sun et al 2016	Sludge	Industrial Reactor	Copper	Sludge exposed to 3 levels of copper: 10mg/L, 20mg/L, or 40mg/L	Chronic	7 days	12	9	57121	3256
●	Zhai et al 2017	Gut	Mouse Intestine	Copper	Gut exposed to copper (1g/L) in drinking water	Chronic	8 weeks	61	30	7254	1436
▲	Simonin et al unpub.	Sediment	Temperate Wetland	Copper	Eutrofied sediment exposed to copper (130mg/kg)	Chronic	9 months	11	5	23654	14400

								All Samples	Treated Samples	Average Rarefaction	Total ESVs
								606	369	30,243	126,381

B **Site Locations**

FIGURE 2 | (A) Descriptions of environment and stressor treatments for each study included in the pilot study, and **(B)** global locations for the nine studies included in this pilot study. *"Ag." stands for agriculture, and "Precip." Stands for precipitation* (Fuentes et al., 2016; Lin et al., 2016; Nunes et al., 2016; Sun et al., 2016; Zhang et al., 2016; Ernakovich et al., 2017; Jurburg et al., 2017; Zhai et al., 2017).

(4) Overall impact of stressors on microbiomes: all microbial communities from any environment and exposed to any stressor (9 studies, 369 treated microbiomes).

16S rRNA Gene Sequence Processing Pipeline

Raw fastq sequence data from the 606 samples were downloaded from NCBI's Sequence Read Archive (SRA) and from an unpublished dataset, and corresponding metadata was extracted for each study via SRA file names, direct correspondence, and the published manuscript itself to identify treatment levels and treatment duration for each sample.

We chose to define phylotypes at the level of exact sequence variants (ESVs; Callahan et al., 2017). The ESV methodology identifies error-adjusted single nucleotide differences in sequence datasets to generate standardized phylotypes that can be directly compared across independently processed datasets. ESVs can be determined individually by sample or study and subsequently merged into the same feature table, allowing thousands of microbiome studies to be rapidly and efficiently combined. The ESV method provides the finest-scale taxonomic partitioning of microbial phylotypes possible and permits targeted identification of consistently-defined indicator species across datasets.

To prepare the datasets for ESV delineation, we trimmed the 16S rRNA gene sequences to the same 250 bp V4 hypervariable region within the Earth Microbiome Project primer set (515F/806R) (Gilbert et al., 2014). We used DADA2 version 1.6.0 (Callahan et al., 2017) in R version 3.5.1 (R Core Team, 2018) to resolve ESVs in each study dataset separately. DADA2 is optimized for running on separate forward and reverse fastq files. However, most of the sequence data we downloaded from SRA is only available as processed "contigs" (i.e., forward/reverse

paired-ends already merged), so we proceeded with DADA2 quality control removal and dereplication on merged paired-end sequences. We acknowledge that using different algorithms to join paired ends [i.e., Qiime2 (Caporaso et al., 2010), PEAR (Kozich et al., 2013), PANDAseq (Masella et al., 2012), and UPARSE (Edgar, 2013)] may generate slightly different sequences during the joining of the forward and reverse Illumina reads due to differences in allowable mismatches and other default settings. Putative chimeras were removed with the "consensus" method using *removeBimeraDenovo*. We then merged the resulting feature tables from each study and merged the representative ESVs.

Taxonomic assignment was performed with a naive Bayesian classifier against the SILVA ribosomal RNA gene database v128 (Desantis et al., 2006; Quast et al., 2013) using *assignTaxonomy* in DADA2, and species assignment was performed by an exact string-matching algorithm using *assignSpecies* in DADA2. Chloroplast, mitochondrial, and archaeal sequences were removed from the ESV tables and representative ESVs using these taxonomic assignments (Note: several archaeal sequences were retained in the representative sequences as outgroups for the alignment and phylogeny). The final dataset contained 126,381 unique ESVs and a total sequence count of 31,512,311. An alignment and phylogeny of the representative ESVs were simultaneously estimated using the Practical Alignment method in SATé and TRAnsitivity (PASTA; Mirarab et al., 2015) against the SILVA v128 reference 16S rRNA gene sequence alignment and tree.

Our pipeline for standardized processing of the V4 hypervariable 16S rRNA Illumina raw data into ESV table, representative ESVs, and taxonomy are detailed here: https://github.com/MicrobiomeStressProject/Frontiers_Microbiology_2018/blob/master/main.md.

Downstream Analyses: Finding Ecological Patterns Hidden in ESVs Across Studies

For the downstream analyses, instead of rarefying to the lowest suitable sequencing depth across the entire merged dataset, which would have been 1,074 sequences, we opted for per-study rarefactions (levels reported in **Figure 2**) in order to retain as much microbiome data as possible, ensuring maximal quality of the subsequent analyses (alpha diversity, beta diversity, resistance, and phylofactorization).

For each study treatment, we used the *core-metrics-phylogenetic* command in QIIME2 (Caporaso et al., 2010) to estimate three indices of alpha diversity (observed ESVs, Shannon-Weiner, and Pielou's evenness) and one of beta-diversity (Bray-Curtis dissimilarity). We used the *betadisper* command in the Vegan package version 2.5-2 in R (Oksanen et al., 2018) to calculate the distance to group centroid from the Bray-Curtis distance matrix for each study treatment. To identify significant study treatment effects on alpha diversity and beta diversity, we ran within-study Kruskal-Wallis tests for differences between the treated microbiomes relative to undisturbed control samples.

To conduct the meta-analysis, we calculated the proportional changes in treated samples relative to control samples for each study treatment (e.g., 25% decrease in observed ESVs or 15% increase in Pielou's evenness after diesel exposure in soil). Study-by-study treatment effect sizes were combined for each diversity metric and used to examine impacts of stressors to address our four objectives: (1) impact of warming on soil microbiomes ($n = 79$ samples), (2) the impact of a range of stressors in a soil environment ($n = 325$ samples), (3) the impact of copper in a variety of environments ($n = 102$ samples), and (4) overall impacts of stressors on microbiomes ($n = 369$ exposed microbiomes). We used a one-sample t-test of the distribution of effect sizes with $\mu_o = 0$ to test for significance. The mean and confidence intervals for each categorical set of effect sizes were displayed as forest plots, along with the individual by study treatment results displayed in boxplots, each with corresponding symbols of significance ($p < 0.05$). For each study treatment, we estimated community resistance as the average pairwise similarity [1- Bray-Curtis dissimilarity] between the control microbiomes and the treated microbiomes (De Vries and Shade, 2013). Significant positive or negative correlations were identified on individual study treatments by examining the relationship between resistance and each alpha diversity index using linear models, and non-linear model fits for the overall relationship across all studies.

Finally, to identify ESVs and lineages responding to stressors, we performed phylofactorization (Phylofactor; Washburne et al., 2017) on just the subset of studies examining copper contamination (Objective 3). We acknowledge that there are a plethora of algorithms enabling differential abundance analysis for identifying taxa and/or clade responses to disturbance (Gloor et al., 2017; Weiss et al., 2017). We performed ANCOM analysis (Mandal et al., 2015) on each individual study treatment, along with Phylofactor, but decided to present just the Phylofactor results here, as our datasets spanned a wide range of environments (**Figure 2**), with very little overlap in ESVs among them. A central advantage of phylofactorization is the ability to detect similar shifts in microbial clades among studies when mapped to the same phylogeny, even with minimal ESV overlap. For example, the microbiomes from two environments may have no overlap in ESVs yet may still show overlap in clade membership (e.g., different microbial taxa identified as Actinobacteria). Even in the absence of ESV overlap among studies, if all members of a clade respond negatively to a stressor, this response will be apparent in the phylofactorization result. For computational ease and simplicity, we focused phylofactorization on the most prevalent members of microbial communities, defined as ESVs present in at least 20% of the microbiomes in a given study. Because the studies varied in both absolute copper treatment and actual copper exposure (i.e., difference in environmental retention of copper and bioavailability), certain studies with larger effect sizes would washout any phylofactor patterns of studies with more subtle responses. Consequently, to control for different effect sizes across studies, we inferred the phylogenetic factors, hereafter "phylofactor(s)," within each individual study and then mapped these phylofactors to a common phylogeny to

identify overlap in clades among the copper stressor studies. Additionally, phylofactor is suitable for our dataset because it has no requirements for choosing a reference frame (i.e., non-responding microorganisms) to determine changes in microbial abundance in response to a disturbance. Phylogenetic factors were chosen as edges maximizing the F-statistic from least squares regression predicting isometric log-ratio (ILR) abundances with treatment. The number of significant factors was determined using Holm's sequentially-rejective multiple test procedure (Holm, 1979) with a 10% family-wise error rate cutoff.

RESULTS AND DISCUSSION

Results of the Literature Search Reveals Thousands of Studies on Biotic and Abiotic Stressors in Host-Associated and Free-Living Environments

Our keyword search (restricted to dates between 2010 and April 2018) identified 12,687 published microbiome studies with 43% of the publications referring to at least one stressor keyword (n = 5,480, i.e., potential studies for the Microbiome Stress Project Database; **Table S2**). The bibliometric network of the journals that published these microbiome stress studies shows that this collective research spans a wide variety of disciplines, including medicine, ecology, ecotoxicology, food science, soil science, and plant sciences (**Figure S1**).

When examining the stressors keywords cited in these 5,480 papers (no environment keywords included in this search), more than half of the studies (55%) referred to keywords categorized as "biotic" stressors (n = 3,306; e.g., pathogens, invasive species, herbivory, predation; **Table S2**). Of the remaining stressor categories, "contaminants" had the largest share (e.g., inorganic and organic pollutants; 27%) and "physical" had the smallest (e.g., fire, UV radiation; 2%). When examined within a co-occurrence network, stressor terms within studies clustered primarily into two separate groups—one dominated by global change (e.g., climate change) or ecotoxicology stressors (e.g., contamination), and the other by human health (e.g., disease, obesity; **Figure S2**). Pathogens occupied a central position within the network, highlighting the common interest shared by environmental and medical fields for these stressors. When microbiome studies were separated by environment (n = 4,293 when all environmental, stressor, and microbiome keywords combined; **Figure 3**; **Table S2**), the top three categories were: animal-associated (including human microbiomes, 38%), soil (34%), and aquatic ecosystems (18%). The three least referenced environment categories were: sediment, aerial, and plant-associated microbiomes, which were only referenced in a total of 5% of the studies.

Of the subset of 150 studies randomly screened from the 5,480 full set of studies in our literature search, 96% of the screened studies performed amplicon sequencing, using primarily Illumina (59%) and 454 pyrosequencing (31%) platforms and only 4% performed shotgun metagenomic sequencing. 89% of the studies sequenced the 16S rRNA gene to study the diversity of archaea and bacteria. Of the 16S rRNA

gene studies, the V4 hypervariable region was the most frequently sequenced (76% of studies using primer sets: V3-V4, V4, V3-V5, & V4-V5). The next most common region was V3 (39% of studies using primer sets V1-V3, V2-V3, V3, V3-V4, V3-V5). Hence, the most common combination of sequencing platform and hypervariable region of the 16S rRNA gene studied was Illumina and the V4 region, respectively, which represented 57% of the 150 studies. Independent of those two classifications across all sequencing platforms and targeted genes, only 48 studies (64% of 150) had raw sequencing data available in public repositories (e.g., Sequence Read Archive) or via direct contact. Of the studies targeting the V4 hypervariable region of the 16S rRNA gene with Illumina sequencing, 69% had raw sequence data available in public repositories. Based on this literature search, for now we focus the Microbiome Stress Project Database and pilot study on the most common primer and sequencing platform combination (V4 region of 16S rRNA gene; Illumina MiSeq), with the pilot study results presented in section Results of the Pilot Study.

Moreover, while screening each individual study, we encountered many issues associated with unclear or incorrect meta-data, making these datasets unusable. The accessibility of both raw data and meta-data is absolutely crucial for integrating studies into larger meta-analyses like the Microbiome Stress Project. Our results clearly indicate that the microbiome research community still needs to commit to generating data that is open and accessible for reproducibility and synthesis.

In preparation for the full Microbiome Stress Project Database, we are also targeting unpublished microbiome studies to minimize publication biases, where published studies are more likely to have significant results. Including unpublished datasets, in addition to the published set, is crucial for minimizing publication bias for our meta-analyses. We have launched an online campaign using social media (Twitter: @MicrobialStress) and our website (microbiomestressproject.weebly.com) to solicit authors to contribute their unpublished datasets to our effort. In addition to contributing to the growing database, researchers will also be able to see how their microbial studies fit alongside comparable studies, or among studies of the same environment or stressor.

Results of the Pilot Study

Among the 606 microbiomes, the average read depth was 39,761 sequences per sample ($\pm$38,902), resulting in 126,381 total merged ESVs (**Figure 2**), with only 2.2% (2882 ESVs) overlap among studies. The average rarefaction was 30,243 ESVs, and the actual rarefaction values for each study are presented in **Figure 2**. Despite large differences in per-study sequencing depth and in the number of ESVs observed, the two values were not correlated ($r^2 = 0.018$).

Stressors Impact on Community Diversity and Resistance of Microbiomes

Our first objective was to assess the specific impacts of warming on soil microbiomes. We found that warming treatments altered alpha diversity through decreased observed ESV richness (-21%, $P < 0.0001$) and Shannon-Weiner index (-7.2%, $P =$
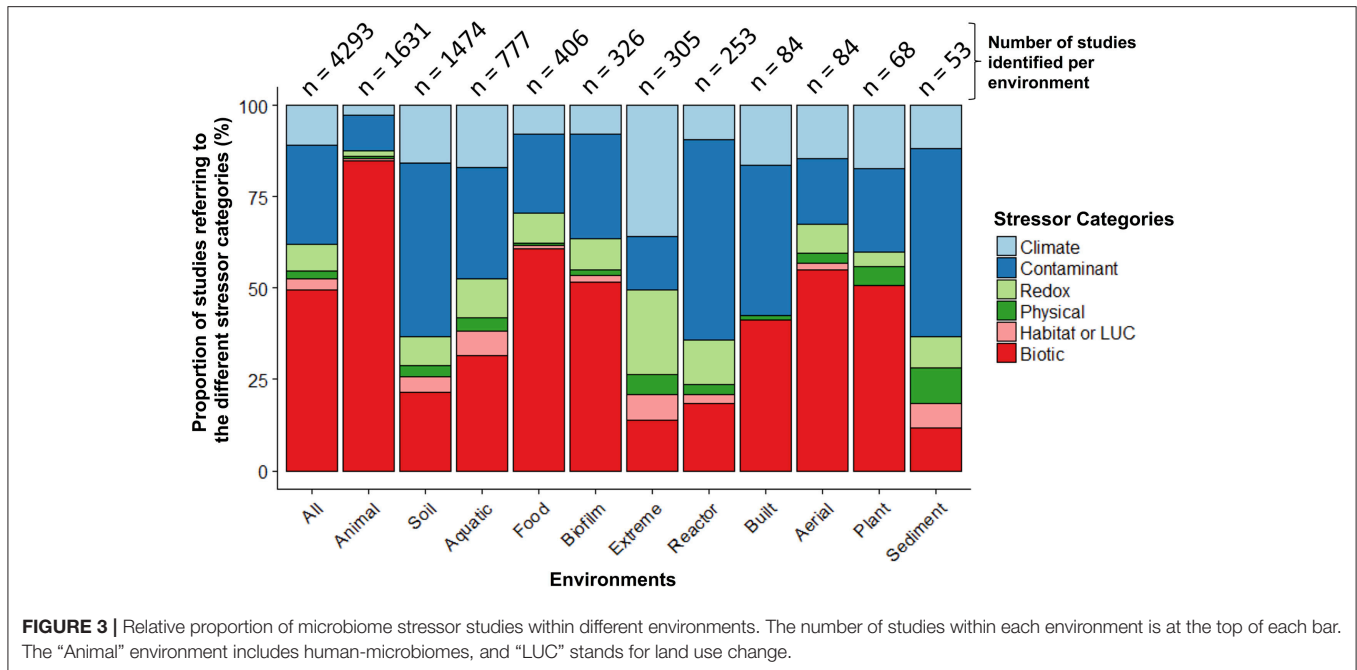
FIGURE 3 | Relative proportion of microbiome stressor studies within different environments. The number of studies within each environment is at the top of each bar. The "Animal" environment includes human-microbiomes, and "LUC" stands for land use change.

0.0002), and no significant impacts on microbiome evenness ($P < 0.0657$) (**Figure 4A**). The overall decreased alpha diversity was primarily driven by one study focused on heat shock impacts in temperate agricultural soils that simulated extreme temperature conditions (+40°C, Jurburg et al., 2017). We acknowledge that this set of results is based on just three distinct studies (79 treated microbiomes), with substantial variation in the site-level characteristics as well as differences in the warming treatments, and as a proof-of-concept analysis, these trends should not be generalized. In trying to explain differential responses of diversity to warming, prior work in macro-ecosystems suggests a variety of diversity responses to disturbance, including non-monotonic responses (Mackey and Currie, 2001). Species often live in environmental optima, suggesting that moving in either direction from that optimum will result in a decrease in that species abundance (Holt, 2009). These unimodal patterns in niche space likely also give rise to unimodal diversity-disturbance relationships, if a large enough range of an environmental gradient is explored. This might explain the contrasted response of permafrost soil microbiomes to elevated temperature (+2C; Ernakovich et al., 2017), with increased alpha diversity across all indices (Observed ESVs: +34.5%, Shannon-Weiner: +21%, Pielou's Evenness: +16.2%, $n = 7$) (**Figure 4B**). Permafrost is frozen soil, so even slight increases in temperature results in drastic changes to the environment conditions. Thawing permafrost could potentially explain the increased alpha diversity, by "awakening" many microorganisms that are encysted, or otherwise metabolically inactive, in the frozen conditions (Makhalanyane et al., 2016; Wurzbacher et al., 2017) Alternatively, increased alpha diversity can also result from the growth of rare microbial taxa (below detection limit in controls)

that rapidly respond to the warming disturbance (Shade et al., 2012; Coveley et al., 2015).

While warming decreased alpha diversity overall, community dispersion increased by 7% relative to control community dispersion ($P = 0.0010$), driven primarily by one significant increase from the same study examining heat-shocking temperate agriculture soils. Increases in microbial community dispersion have been observed in prior diversity-disturbance experiments (Gibbons et al., 2016), and may reflect a flattening of ecological selection as the environment becomes more heterogeneous. Additionally, this result could be explained by the fact that warming, and disturbances more generally, have stochastic rather than deterministic effects on microbiome composition leading to an increased dispersion compared to controls (i.e., the *Anna Karenina principle*, Zaneveld et al., 2017). Despite significant overall impacts of warming on soil microbial diversity, there was high variability in responses for each study treatment, which highlights the need for integrating more independent studies into the Microbiome Stress Project Database. This will increase the statistical power in these objectives, thereby generating reliable insights on how soil microbiomes respond to specific stressors like warming.

In our second objective, we expanded the analysis of soil microbiome responses to a wide range of stressors, which also resulted in decreased alpha diversity and increased community dispersion. Observed ESVs decreased by 17.7% ($P < 0.0001$), Shannon-Weiner by 9.1% ($P < 0.0001$), and Pielou's evenness decreased by 5.3% ($P < 0.0001$) across all soil microbiomes studies (**Figure 4A**), while community dispersion increased by 5.4% relative to control microbiomes ($P = 0.0006$).
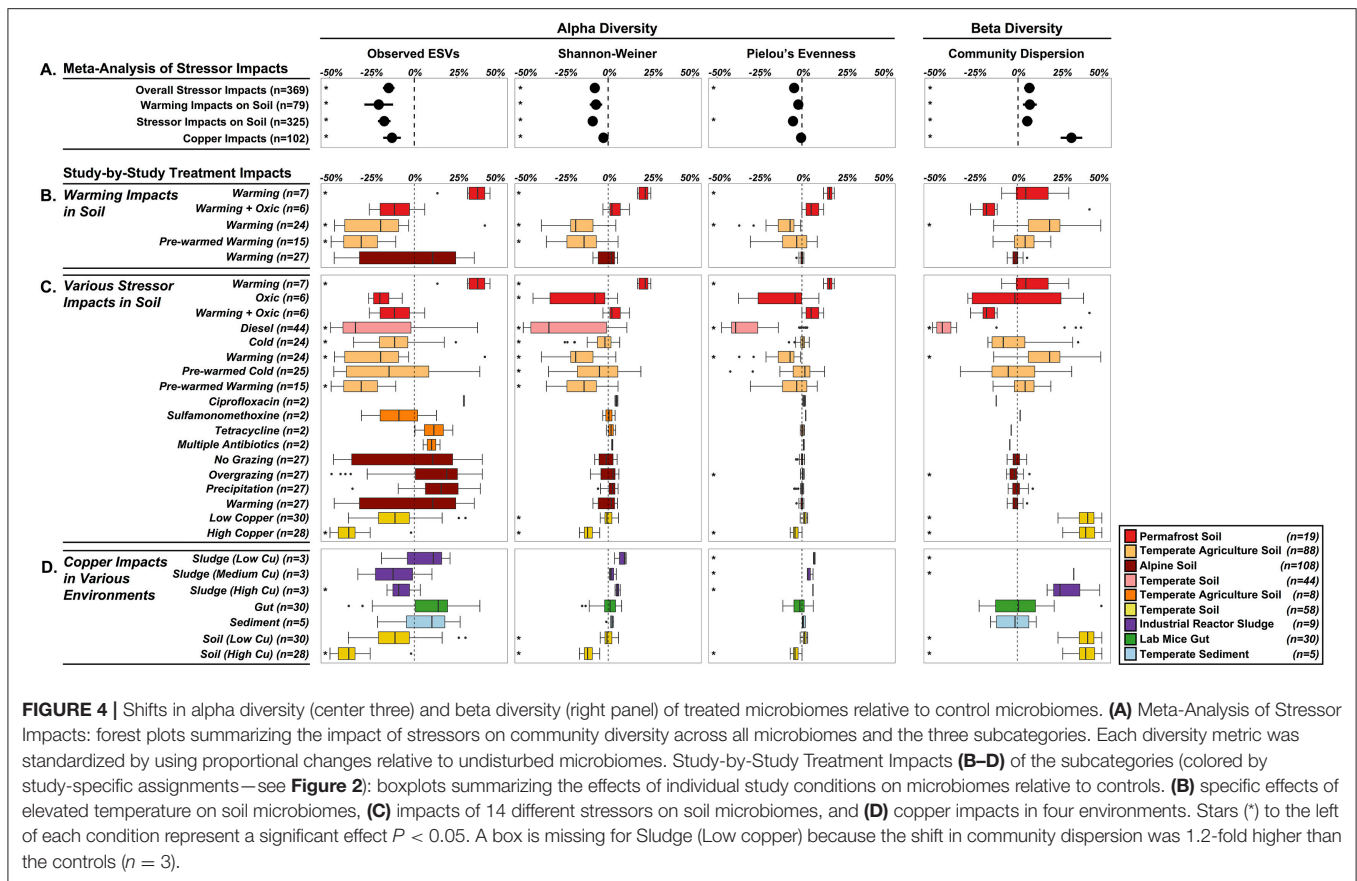
FIGURE 4 | Shifts in alpha diversity (center three) and beta diversity (right panel) of treated microbiomes relative to control microbiomes. **(A)** Meta-Analysis of Stressor Impacts: forest plots summarizing the impact of stressors on community diversity across all microbiomes and the three subcategories. Each diversity metric was standardized by using proportional changes relative to undisturbed microbiomes. Study-by-Study Treatment Impacts **(B–D)** of the subcategories (colored by study-specific assignments—see **Figure 2**): boxplots summarizing the effects of individual study conditions on microbiomes relative to controls. **(B)** specific effects of elevated temperature on soil microbiomes, **(C)** impacts of 14 different stressors on soil microbiomes, and **(D)** copper impacts in four environments. Stars (*) to the left of each condition represent a significant effect $P < 0.05$. A box is missing for Sludge (Low copper) because the shift in community dispersion was 1.2-fold higher than the controls ($n = 3$).

Similar to the study specific impacts of warming, there was high variability in the soil microbiome responses to the range of stressors as well (**Figure 4C**). The studies testing the effects of warming, diesel, and copper primarily decreased across all alpha diversity metrics. Warming and copper imposed significant increases in community dispersion, consistent with the meta-analysis trend. In contrast, the diesel and overgrazing stressors resulted in significant decreases in community dispersion, which suggests that these stressors may strengthen ecological selection and lead to more deterministic shifts in composition (Zaneveld et al., 2017). Here, we only selected 12 stressors for this pilot study, but we already have 27 different stressors in the database, with multiple studies available per stressor. In future analyses, covering more soil systems and stressor exposures will allow us to validate the effects observed on alpha and beta-diversity in the pilot study and uncover the abiotic and biotic factors driving soil microbiome responses to disturbed conditions.

Our third objective was to evaluate the effects of an individual stressor across multiple environments. We examined copper contamination on 102 samples across the four environments (soil, sludge, gut, and sediment), and found similar decreases in alpha diversity: a 13.1% decrease in observed ESVs ($P < 0.0001$) and a 2.8% decrease in Shannon-Weiner ($P = 0.0002$). Like the other objectives, there was substantial among-study variability in alpha diversity in response to copper (**Figure 4D**). The overall negative impacts were primarily driven by high

copper exposure in soil and sludge, while low copper treatments in soil and sludge and the copper treatments in gut and sediment were not significantly impacted by copper addition (**Figure 2B**). There was no significant overall impact of copper on community evenness ($P = 0.2248$). Community dispersion was 31.5% greater in copper treated microbiomes than control microbiomes ($P < 0.0001$), however this was primarily driven by the two copper treatments in soil (Nunes et al., 2016). The high variability in response of alpha diversity and beta diversity (**Figure 5**) may be due in part to the differential intrinsic resistance capacity of the microbiomes (e.g., initial alpha diversity or community network structure before exposure). Additionally, the magnitude of the treatment disturbance between studies (**Figure 2**) and the ecological memory of antecedent exposure to copper or other stressors may explain this range of copper effects observed on diversity (Griffiths and Philippot, 2013; Azarbad et al., 2016). The concentration and mode of exposure to copper were very contrasted between the four different studies which may explain a large part of the differences among studies (**Figure 2**). However, a very high chronic copper exposure in a mouse gut (1 g/L) did not cause any change in diversity (**Figure 4**), which suggests that the absolute magnitude of treatment may still not reflect actual exposure level for the microbiome, as in this case copper may be absorbed or transformed by the host before gut microbiome exposure (Cholewinska et al., 2018). Additionally, different redox potentials between environments may also determine actual
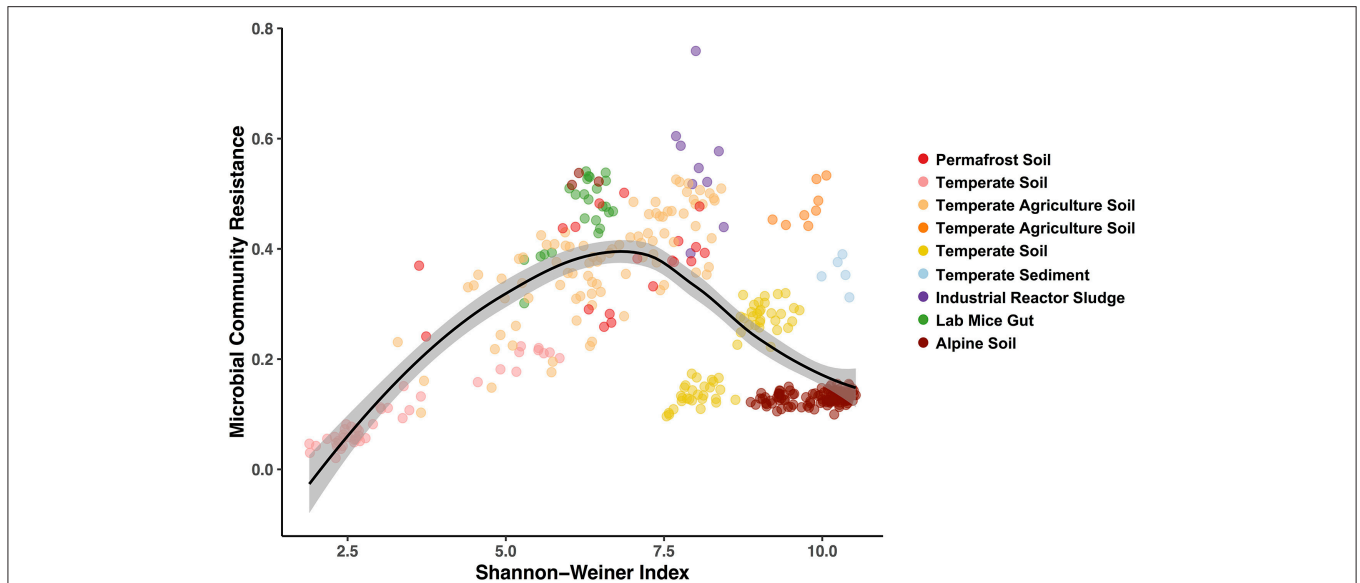
FIGURE 5 | Hump-shaped relationship between microbiome community resistance to stressors and alpha diversity (Shannon-Weiner) of all treated microbiomes included in the pilot study. Data points represent a stressor treated microbiome ($n = 369$) and are colored by each study environment.

exposure and speciation of copper, influencing the toxicity and bioavailability of this metal for microorganisms (Fleming and Trevors, 1989; Giller et al., 1998). Stochastic death throughout the treated microbiomes might also explain these patterns, with many microbial taxa driven to local extinction under high copper, resulting in distinct community structure and hence high beta diversity. While these results are interesting, we acknowledge again here that this pilot study does not have the statistical power to reliably draw any generalizable insights of copper impacts across multiple environments. With multiple studies for each environment-copper comparison (similar to the warming in soils) in the full dataset, we will be able to repeat the analysis with more power, accounting for copper bioavailability and historical exposures among studies.

Our fourth objective to analyze all 369 treated microbiomes in the pilot study revealed an average reduction in alpha diversity and increase in beta diversity. Similar to the results presented on the three other objectives of the meta-analysis, the exposure to all stressors led to significant decreases in the alpha diversity: observed ESVs decreased by 15% ($P < 0.0001$), Shannon-Weiner alpha diversity decreased by 7.8% ($P < 0.0001$), and evenness decreased by 4.6% ($P < 0.0001$) (**Figure 4A**). Beta diversity was also significantly affected by stressors, with community dispersion increasing by 6.7% ($P < 0.0001$) across all studies. These findings show that the diversity patterns observed for soil warming, a range of stressors in soil and copper across multiple environments still hold when performing a global meta-analysis of microbiome responses. These first results suggest that very contrasted stressors in diverse ecosystems lead to similar consequences for microbial community diversity. If these trends are substantiated with the same analysis at a larger scale, more research will be needed to understand how these shifts in diversity affect host health and ecosystem
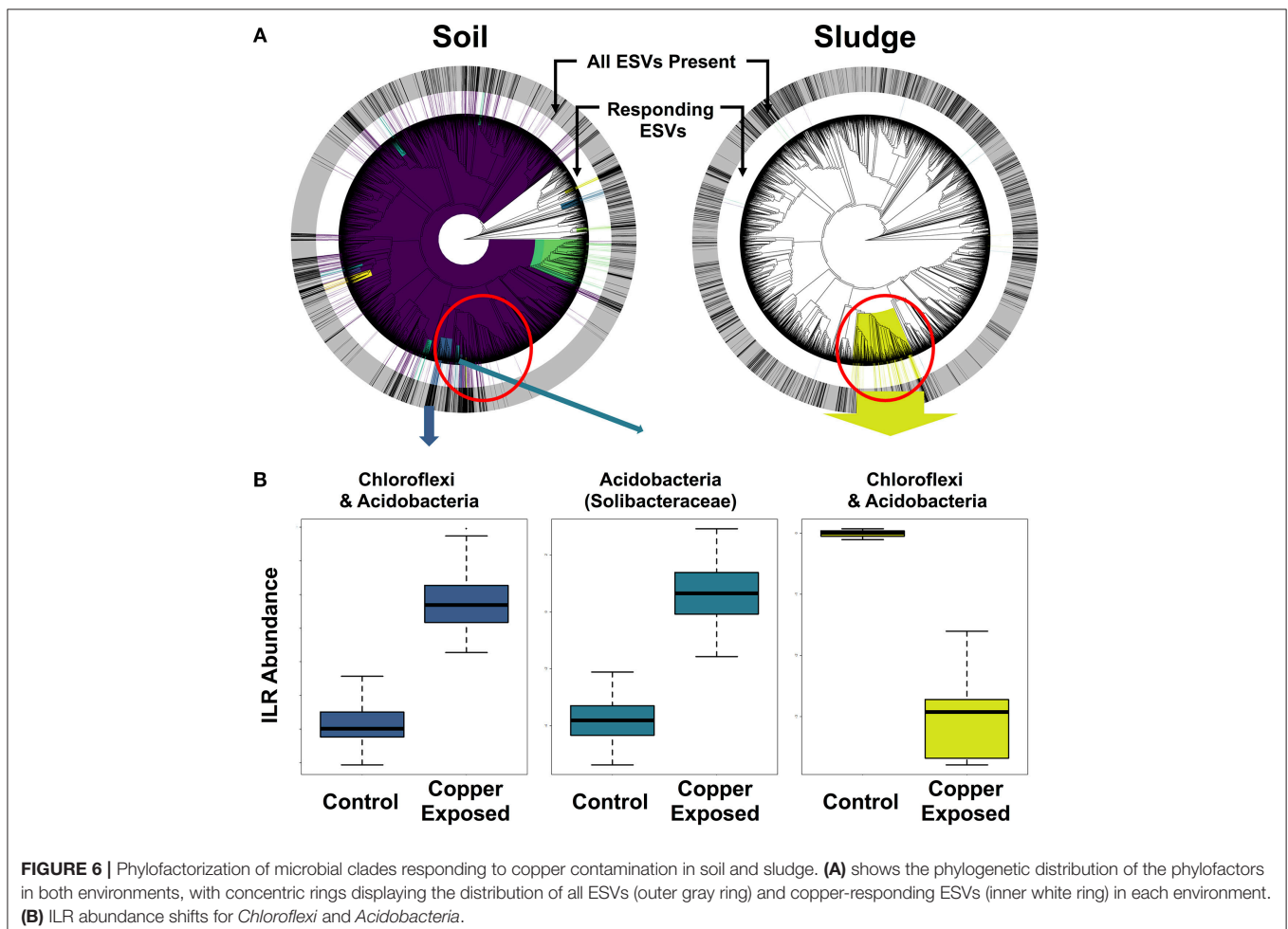
function. Moreover, using the growing Microbiome Stress Project database, researchers will be able to identify knowledge gaps and target research efforts for specific stressors or environments.

Finally, we used the Bray-Curtis similarities to estimate community resistance to stressors and explored potential correlations with the three alpha diversity indices (**Figure 5**; **Table S3**). We expected to see a positive correlation between alpha diversity and the microbiomes capacity to resist change (Tilman and Downing, 1994) and to remain stable when exposed to a stressor (i.e., the *Insurance Hypothesis*, Yachi and Loreau, 1999). However, when including all the exposed microbiomes, we primarily observed hump-shaped relationships between microbial community resistance and alpha diversity (**Figure 5**; relationship with Shannon-Weiner index). These results indicate that at low diversity levels (2 to 7.5 Shannon-Wiener index values, **Figure 5**), microbiome resistance is positively correlated to diversity but that this relationship becomes negative at higher diversity levels (7.5 to 10, **Figure 5**). These results suggest an alpha diversity optimum at which the microbiome resistance is maximal, which has been shown to exist in other microbial systems (Gibbons et al., 2016; Locey and Lennon, 2016). Elevated diversity can also be an indicator of a community in flux, where high dispersion rates in the environment increase the incidence of new microbial colonizers (Cadotte, 2006; Evans et al., 2017). In these systems, microbial populations may present a low residence time and a high community composition turnover (Locey and Lennon, 2016), which may drive in part the lower resistance observed at high diversity here. These preliminary results on diversity-stability relationships emphasize the need for a greater depth of studies for each stressor—environment combination to determine if these patterns are ubiquitous or occur under specific environmental conditions.

Phylofactorization of Microbiomes From Four Environments Exposed to Copper

In order to isolate responding microbial ESVs and identify overlap in whole clades responding to copper contamination, we implemented Phylofactor. The four distinct environments exposed to copper had markedly different numbers of significant phylofactors (i.e., clades or terminal nodes responsive relative to the rest of the microbiome). The soil dataset had 780 phylofactors, covering 80% of the ESVs in the community, highlighting an extremely high community turnover in response to copper exposure (**Figure 6A**). In contrast, the gut microbiome had 20 phylofactors significantly responding to copper, the sludge had 17 phylofactors (**Figure 6B**), and the sediment had only five phylofactors containing 7 ESVs covering <0.2% of the dataset. The higher number of phylofactors in soil and sludge are consistent with the large shifts in alpha and beta diversity observed in these two studies (**Figure 4**), and the large effect size in the soil dataset yielded a phylofactorization implicating nearly the entire phylogeny as responsive to copper treatment. One particularly large phylofactor in soil identified a lineage of 688 ESVs accounting for 47% of the community. While, there were no common ESVs among the copper responders across the environments, there was some overlap in responding clades found to be affected by copper exposure. In particular, a clade of 116 ESVs in the sludge dataset was close to four phylofactors in the soil dataset containing 20, 62, 10, and 9 ESVs, respectively (red circle in **Figure 6**). These phylofactors primarily contained uncultivated lineages annotated as members of *Acidobacteria*, *Solibacteraceae*, but also other *Acidobacteria* and Chloroflexi. These lineages are phylogenetically close, and in some cases nested clades but they interestingly presented opposite responses to copper treatment in the two environments (**Figure 6**). In the soil study, the responding community subsets of *Chloroflexi* and *Acidobacteria* were increased in the copper treated microbiomes, while they decreased in the sludge study (**Figure 6**). Other studies have demonstrated the sensitivity of both *Acidobacteria* and *Chloroflexi*, while Actinobacteria are generally tolerant to copper exposure in soils (Li et al., 2015) and in sludge (Sun et al., 2016). Due to the compositional nature of sequence-count data, it is impossible to determine how much of the effect is due to the populations of *Solibacteraceae*, for example, having opposite responses in the two environments and how much is due to the remaining community whose abundance shifts can confound directional inferences in compositional data.

FIGURE 6 | Phylofactorization of microbial clades responding to copper contamination in soil and sludge. **(A)** shows the phylogenetic distribution of the phylofactors in both environments, with concentric rings displaying the distribution of all ESVs (outer gray ring) and copper-responding ESVs (inner white ring) in each environment. **(B)** ILR abundance shifts for *Chloroflexi* and *Acidobacteria*.

Rather than yielding clear inferences about potential bioindicator lineages in the current state, phylofactorization yielded more insight about future research on the difficulty of comparing results across microbiome studies, and the opportunities of datasets like the Microbiome Stress Project may yield to test and improve statistical tools. In addition to improving computational tools, it is important for researchers to improve quantification of absolute abundances by running qPCR or by including DNA spike-ins prior to sequencing to limit our reliance to relative abundance data. Concomitantly, future experimental work could directly target the microbial clades identified as responders to better understand their responses to individual stressor–environment scenarios and validate the inferences made in the amplicon sequencing meta-analyses.

Concluding Remarks

In this article, we presented the objectives, methodological framework and bibliographic landscape covered by the Microbiome Stress Project and used our pilot study as a proof-of-concept meta-analysis to demonstrate the potential of this database. With thousands of studies available on the impact of environmental stressors on microbial communities, the Microbiome Stress Project Database is well-poised to allow researchers to tackle key ecological issues regarding the resistance, resilience, and response of microbial communities following exposure to environmental stressors. This project will generate a wealth of information on the natural history of microbial taxa, especially on their stressor tolerance and important life-history strategies for surviving in fluctuating environments. The full database will enable the identification of indicator taxa and clades to specific or multiple stressors that could eventually be used for monitoring or microbiome engineering. The Microbiome Stress Project provides an ideal resource for developing and testing new statistical methods with comprehensive amplicon sequencing datasets. We anticipate that research into common responses of microbial communities to stressors will lead to better microbiome diagnostics, allowing researchers to make inferences that are robust to both subtle and large-scale changes in species composition across independent studies.

The aim of this project is to identify global patterns in microbiome responses to stressors, but we encourage the scientific community to use the database to examine their own research questions. Adding your studies to the Microbiome Stress Project Database is a valuable contribution to the research community and also provides valuable context for your own work. Finally, building the database will reveal knowledge gaps in terms of missing or low replication studies of specific stressors or particular environments. We hope that the Microbiome Stress Project will forge new interdisciplinary collaborations leading to important breakthroughs in our understanding of microbial communities' responses to environmental change and for improving our ability to engineer microbiomes for improved human and environmental health.

AUTHOR CONTRIBUTIONS

JR, MS, JB, and FM conceived the project. MS and JB collected the raw sequences meta-data and conducted the literature search. FM and JR processed the sequence data. FM, AW, and SG developed the downstream analyses. MS, JR, and JE developed the conceptual framework and background synthesis. JR, MS, JB, JE, SG, FM, and AW contributed equally to writing the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.03272/full#supplementary-material

REFERENCES

Azarbad, H., van Gestel, C. A. M., Niklinska, M., Laskowski, R., Roling, W. F. M., and van Straalen, N. M. (2016). Resilience of soil microbial communities to metals and additional stressors: DNA-based approaches for assessing "stress-on-stress" responses. *Int. J. Mol. Sci.* 17, 933. doi: 10.3390/ijms17060933

Baym, M., Lieberman, T. D., Kelsic, E. D., Chait, R., Gross, R., Yelin, I., et al. (2016). Spatiotemporal microbial evolution on antibiotic landscapes. *Science* 353, 1147–1152. doi: 10.1126/science.aag0822

Bender, E. A., Case, T. J., and Gilpin, M. (1984). Perturbation experiments in community ecology: theory and practice. *Ecology* 65, 1–13. doi: 10.2307/1939452

Bissett, A., Brown, M. V., Siciliano, S. D., and Thrall, P. H. (2013). Microbial community responses to anthropogenically induced environmental change: towards a systems approach. *Ecol. Lett.* 16, 128–139. doi: 10.1111/ele.12109

Cadotte, M. (2006). Metacommunity influences on community richness at multiple spatial scales: a microcosm experiment. *Am. Nat.* 87, 1008–1016. doi: 10.1890/0012-9658(2006)87[1008:MIOCRA]2.0.CO;2

Calabrese, E. J., and Baldwin, L. A. (2002). Defining hormesis. *Hum. Exp. Toxicol.* 21, 91–97. doi: 10.1191/0960327102ht217oa

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303.QIIME

Cholewinska, E., Ognik, K., Fotschki, B., Zdunczyk, Z., and Juskiewicz, J. (2018). Comparison of the effect of dietary copper nanoparticles and one copper (II) salt on the copper biodistribution and gastrointestinal and hepatic morphology and function in a rat model. *PLoS ONE* 13:e0197083. doi: 10.1371/journal.pone.0197083

Connell, J. H. (1978). Diversity in tropical rain forests and coral reefs. *Science* 199, 1302–1310. doi: 10.1126/science.199.4335.1302

Coveley, S., Elshahed, M. S., and Youssef, N. H. (2015). Response of the rare biosphere to environmental stressors in a highly diverse ecosystem (Zodletone spring, OK, USA). *PeerJ* 3:e1182. doi: 10.7717/peerj.1182

Crain, C. M., Kroeker, K., and Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecol. Lett.* 11, 1304–1315. doi: 10.1111/j.1461-0248.2008.01253.x

De Vries, F. T., and Shade, A. (2013). Controls on soil microbial community stability under climate change. *Front. Microbiol.* 4:265. doi: 10.3389/fmicb.2013.00265

Desantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S RRNA gene database and workbench compatible with ARB. *Appl. Env. Microbiol.* 72, 5069–5072. doi: 10.1128/AEM.03006-05

Duvallet, C. (2018). Meta-analysis generates and prioritizes hypotheses for translational microbiome research. *Microb. Biotechnol.* 11, 273–276. doi: 10.1111/1751-7915.13047

Duvallet, C., Gibbons, S. B., Gurry, T., Irizarry, R. A., and Alm, E. J. (2017). Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat. Commun.* 8, 1784. doi: 10.1038/s41467-017-01973-8

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Ernakovich, J. G., Lynch, L. M., Brewer, P. E., Calderon, F. J., and Wallenstein, M. D. (2017). Redox and temperature-sensitive changes in microbial communities and soil chemistry dictate greenhouse gas loss from thawed permafrost. *Biogeochemistry* 134, 183–200. doi: 10.1007/s10533-017-0354-5

Evans, S. E., Martiny, J. B., and Allison, S. D. (2017). Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J.* 11, 176–185. doi: 10.1038/ismej.2016.96

Evans, S. E., and Wallenstein, M. D. (2014). Climate change alters ecological strategies of soil bacteria. *Ecol. Lett.* 17, 155–164. doi: 10.1111/ele.12206

Fleming, C. A., and Trevors, J. T. (1989). Copper toxicity and chemistry in the environment: a review. *Water Air Soil Pollut.* 44, 143–158.

Fuentes, S., Barra, B., Caporaso, J. G., and Seeger, M. (2016). From rare to dominant: a fine-tuned soil bacterial bloom during. *Appl. Environ. Microbiol.* 82, 888–896. doi: 10.1128/AEM.02625-15

Gibbons, S. M., Duvallet, C., and Alm, E. J. (2018). Correcting for batch effects in case-control microbiome studies. *PLoS Comput. Biol.* 14:e1006102. doi: 10.1371/journal.pcbi.1006102

Gibbons, S. M., Scholz, M., Hutchinson, A. L., Dinner, A. R., Gilbert, J. A., and Coleman, M. L. (2016). Disturbance regimes predictably alter diversity in an ecologically complex bacterial system. *mBio* 7:e01372–e01316. doi: 10.1128/mBio.01372-16

Gilbert, J. A., Jansson, J. K., and Knight, R. (2014). The earth microbiome project: successes and aspirations. *BMC Biol.* 12:69. doi: 10.1186/s12915-014-0069-1

Giller, K. E., Witter, E., and McGrath, S. P. (1998). Toxicity of heavy metals to microorganisms and microbial processes in agricultural soils: a review. *Soil Biol. Biogeochem.* 30, 1389–1414. doi: 10.1016/S0038-0717(97)00270-8

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 8:2224. doi: 10.3389/fmicb.2017.02224

Griffiths, S. B., and Philippot, L. (2013). Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiol. Rev.* 37, 112–129. doi: 10.1111/j.1574-6976.2012.00343

Holden, S. R., and Treseder, K. K. (2013). A meta-analysis of soil microbial biomass responses to forest disturbances. *Front. Microbiol.* 4:163. doi: 10.3389/fmicb.2013.00163

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70. doi: 10.2307/4615733

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19659–19665. doi: 10.1073/pnas.0905137106

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Huston, M. (1979). A general hypothesis of species diversity. *Am. Nat.* 113, 81–101. doi: 10.1086/283366

Jackson, M. A., Verdi, S., Maxan, M., Shin, C. M., Zierer, J., Bowyer, R. C. E., et al. (2018). Gut microbiota associations with common diseases and prescription medications in a population-based cohort. *Nat. Commun.* 9:2655. doi: 10.1038/s41467-018-05184-7

Jurburg, S. D., Nunes, I., Brejnrod, A., Jacquiod, S., and Priemé, A., Søren, et al. (2017). Legacy effects on the recovery of soil bacterial communities from extreme temperature perturbation. *Front. Microbiol.* 8:1832. doi: 10.3389/fmicb.2017.01832

Knelman, J. E., Nemergut, D. R., Klotz, M. G., Blackwood, C., and Rodrigues, J. L. M. (2014). Changes in community assembly may shift the relationship between biodiversity and ecosystem function. *Front. Microbiol.* 5:424. doi: 10.3389/fmicb.2014.00424

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/AEM.01043-13

Li, J., Ma, Y. B., Hu, H. W., Wang, J. T., Liu, Y. R., and He, J. Z. (2015). Field-based evidence for consistent responses of bacterial communities to copper. *Front Microbiol.* 6:31. doi: 10.3389/fmicb.2015.00031

Lin, H., Jin, D., Freitag, T. E., Sun, W., Yu, O., Fu, J., et al. (2016). A Compositional shift in the soil microbiome induced by tetracycline, sulfamonomethoxine and ciprofloxacin entering a plant-soil system. *Environ. Pollut.* 212, 440–448. doi: 10.1016/j.envpol.2016.02.043

Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A. B., et al. (2017). Strains, functions and dynamics in the expanded human microbiome project. *Nature.* 550, 61–66. doi: 10.1038/nature23889

Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U.S.A.* 113, 5970–5975. doi: 10.1073/pnas.1521291113

Mackey, R. L., and Currie, D. J. (2001). The diversity-disturbance relationship: is it generally strong and peaked? *Ecology* 82, 3479–3492. doi: 10.2307/2680166

Makhalanyane, T. P., Van Goethem, M. W., and Cowan, D. A. (2016). Microbial diversity and functional capacity in polar soils. *Curr. Opin. Biotechnol.* 38, 159–166. doi: 10.1016/j.copbio.2016.01.011

Mandal, S., van Treuren, W., White, R. A., Eggesbo, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* 26, 27663. doi: 10.3402/mehd.v26.27663

Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G., and Neufeld, J. D. (2012). PANDAseq: paired-end assembler for illumine sequences. *BMC Bioinformatics* 13:31. doi: 10.1186/1471-2105-13-31

Mirarab, S., Nguyen, N., Guo, S., Wang, L., Kim, J., and Warnow, T. (2015). PASTA: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *J. Comput. Biol.* 22, 377–386. doi: 10.1089/cmb.2014.0156

Nunes, I., Jacquiod, S., Brejnrod, A., Holm, P. E., Johansen, A., Brandt, K. K., et al. (2016). Coping with copper: legacy effect of copper on potential activity of soil bacteria following a century of exposure. *FEMS Microbiol. Ecol.* 92:fiw175. doi: 10.1093/femsec/fiw175

Odum, E. P. (1985). Trends expected in stressed ecosystems. *Bioscience* 35, 419–422.

Odum, E. P., Finn, J. T., and Franz, E. H. (1979). Perturbation theory and the subsidy-stress gradient. *Bioscience* 29, 349–352.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2018). *R Package 'Vegan'* (Vienna).

Philippot, L., Spor, A., Henault, C., Bru, D., Bizouard, F., Jones, C. M., et al. (2013). Loss in microbial diversity affects nitrogen cycling in soils. *ISME J.* 7, 1609–1619. doi: 10.1038/ismej.2013.34

Piggott, J. J., Townsend, C. R., and Matthaei, C. D. (2015). Climate warming and agricultural stressors interact to determine stream macroinvertebrate community dynamics. *Glob. Change Biol.* 21, 1887–1906. doi: 10.1111/gcb.12861

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Ramirez, K. S., Knight, C. G., De Hollander, M., Brearley, F. Q., Constantinides, B., Cotton, A., et al. (2018). Detecting macroecological patterns in bacterial communities across independent studies of global soils. *Nat. Microbiol.* 3, 189–196. doi: 10.1038/s41564-017-0062-x

Ratzke, C., and Gore, J. (2018). Modifying and reacting to the environmental PH can drive bacterial interactions. *PLoS Biol.* 16:e2004248. doi: 10.1371/journal.pbio.2004248

Reese, A. T., Cho, E. H., Klitzman, B., Nichols, S. P., Wisniewski, N. A., Villa, M. M., et al. (2018). Antibiotic-induced changes in the microbiota disrupt redox dynamics in the gut. *Elife* 7:e35987. doi: 10.7554/eLife.35987

Sanchez, A., and Gore, J. (2013). Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biol.* 11:e1001547. doi: 10.1371/journal.pbio.1001547

Schimel, J., Balser, T. C., and Wallenstein, M. D. (2007). Microbial stress-response physiology and its implications for ecosystem function. *Ecology* 88, 1386–1394. doi: 10.1890/06-0219

Shade, A., Caporaso, J. G., Handelsman, J., Knight, R., and Fierer, N. (2013). A meta-analysis of changes in bacterial and archaeal communities with time. *ISME J.* 7, 1493–1506. doi: 10.1038/ismej.2013.54

Shade, A., Peter, H., Allison, S. D., Baho, D. L., Berga, M., Burgmann, H., et al. (2012). Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* 3:417. doi: 10.3389/fmicb.2012.00417

Sun, F. L., Fan, L. L., and Xie, G. J. (2016). Effect of copper on the performance and bacterial communities of activated sludge using illumina miseq platforms. *Chemosphere* 156, 212–219. doi: 10.1016/j.chemosphere.2016.04.117

Thompson, L. R., Sanders, J. G., Mcdonald, D., Amir, A., Ladau, J., Locey, K. J., et al. (2017). A communal catalogue reveals earth's multiscale microbial diversity. *Nature* 551, 457. doi: 10.1038/nature24621

Tilman, D., and Downing, J. A. (1994). Biodiversity and stability in grasslands. *Nature* 367, 363–365. doi: 10.1038/367363a0

Treseder, K. K. (2008). Nitrogen additions and microbial biomass: a meta-analysis of ecosystem studies. *Ecol. Lett.* 11, 1111–1120. doi: 10.1111/j.1461-0248.2008.01230.x

Van Eck, N. J., and Waltman, L. (2010). VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538. doi: 10.1007/s11192-009-0146-3

Vellend, M. (2010). Conceptual synthesis in community ecology. *Q. Rev. Biol.* 85, 183–206. doi: 10.1086/652373

Waltman, L., Van Eck, N. J., and Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *J. Informetr.* 4, 629–635. doi: 10.1016/j.joi.2010.07.002

Washburne, A. D., Silverman, J. D., Leff, J. W., Bennett, D. J., Darcy, J. L., Mukherjee, S., et al. (2017). Phylogenetic factorization of compositional data yields lineage-level associations in microbiome datasets. *PeerJ* 5:e2969. doi: 10.7717/peerj.2969

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzales, A., Lozupone, C., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5:27. doi: 10.1186/s40168-017-0237-y

Wurzbacher, C., Attermeyer, K., Kettner, M. T., Flintrop, C., Warthmann, N., Hilt, S., et al. (2017). DNA metabarcoding of unfractioned water samples relates phyto-, zoo- and bacterioplankton dynamics and reveals a single-taxon bacterial bloom. *Environ. Microbial. Rep.* 9, 383–388. doi: 10.1111/1758-2229.12540

Yachi, S., and Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1463–1468. doi: 10.1073/pnas.96.4.1463

Zaneveld, J. R., McMinds, R., and Thurber, R. V. (2017). Stress and stability: applying the anna karenina principle to animal microbiomes. *Nat. Microbiol.* 2, 17121. doi: 10.1038/nmicrobiol.2017.121

Zhai, Q., Li, T., Yu, L., Xiao, Y., Feng, S., Wu, J., et al. (2017). Effects of subchronic oral toxic metal exposure on the intestinal microbiota of mice. *Sci. Bull.* 62, 1–10. doi: 10.1016/j.scib.2017.01.031

Zhang, Y., Dong, S., Gao, Q., Liu, S., Zhou, H., Ganjurjav, H., et al. (2016). Climate change and human activities altered the diversity and composition of soil microbial community in alpine grasslands of the qinghai-tibetan plateau. *Sci. Tot. Environ.* 562, 353–363. doi: 10.1016/j.scitotenv.2016.03.221

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.