Phylogenetic clustering of microbial communities as a biomarker for chemical pollution

Thomas P. Smith [1,*], Rachel Hope[2], Thomas Bell[1]

[1]The Georgina Mace Centre for the Living Planet, Department of Life Sciences, Silwood Park Campus, Imperial College London, Ascot, Berkshire, SL5 7PY, United Kingdom
[2]Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, 1015, Switzerland
*Corresponding author. The Georgina Mace Centre for the Living Planet, Department of Life Sciences, Silwood Park Campus, Imperial College London, Buckhurst Road, Ascot, Berkshire, SL5 7PY, United Kingdom. E-mail: thomas.smith1@imperial.ac.uk
Editor: [Stéphane Vuilleumier]

Abstract

Microbial communities play a critical role in ecosystem functioning and offer promising potential as bioindicators of chemical pollution in aquatic environments. Here we examine the responses of both bacterial isolates and microbial communities to a range of pollutants, focusing on the phylogenetic predictability of their responses. We tested the growth inhibition of environmental bacterial isolates by 168 agricultural pollutants recently shown to have off-purpose antimicrobial activity in human gut bacteria. We also tested the growth responses of whole microbial communities to the same chemical pollutants and quantified changes in the composition of select communities, to link compositional changes to functioning. We found that bacterial isolates exhibited a strong phylogenetic signal in their growth responses, with closely related taxa responding similarly to chemical stress. In microbial communities, pollutants that significantly impacted isolates also reduced community diversity and growth, causing shifts in community structure toward increased phylogenetic clustering, suggesting environmental filtering. The mean phylogenetic distance effectively captured these shifts, indicating its potential as a simple metric for monitoring pollution. Our findings highlight the predictability of microbial responses to pollution and suggest that microbial-based bioindicators, coupled with rapid sequencing technologies, could transform environmental monitoring.

Keywords: microbial communities; composition; ecotoxicology; pollutants; phylogenetics; bioindicators

Introduction

Freshwater ecosystems are functionally diverse environments which play a crucial role in maintaining global biogeochemical cycles, regulating the flow and transformation of key nutrients and greenhouse gases (Sagova-Mareckova et al. 2021, Bani et al. 2022). However, these natural environments are coming under increasing pressure from anthropogenic stressors, such as toxic chemical pollutants (Stehle and Schulz 2015, Thompson et al. 2016). Environmental stressors, including chemical pollutants, are predicted to alter biodiversity-ecosystem functioning relationships across a wide range of ecosystems (De Laender et al. 2016, Rillig et al. 2019, Cadotte and Tatsumi 2025). These impacts can cascade through trophic levels, scaling from populations to communities, leading to diminished ecosystem resilience and disrupted ecosystem services (De Laender et al. 2016, Rillig et al. 2023, van Moorsel et al. 2023).

Microbes are central to the ecological and biogeochemical processes occurring within freshwater environments, driving biogeochemical cycles and providing key ecosystem services (Grossart et al. 2020, Sagova-Mareckova et al. 2021). Microbes are particularly responsive to environmental stressors due to their high metabolic rates and rapid growth (Sagova-Mareckova et al. 2021). Microbes therefore have the potential to be used as rapid 'biosensors' that detect and track chemicals of emerging concern (Smith et al. 2024), particularly because many chemical pollutants

show off-target impacts on microbial populations (Lindell et al. 2024).

The responses of whole communities of microbes to environmental perturbations can be observed through changes in their composition and structure (Astudillo-García et al. 2019). Prior studies of microbial communities have identified individual taxa that can be used as biomarkers for specific contaminants (Sagova-Mareckova et al. 2021, Wijaya et al. 2023). In addition, 'keystone' or 'multitask' taxa have been identified that are hallmarks of multiple environmental conditions and predict general environmental pollution or quality (Wang et al. 2020, Sperlea et al. 2021). A key limitation of indicator taxa is that there are many factors that determine whether a specific taxon may or may not be present in a given environment, and many of those factors are unrelated to environmental stress.

An alternative to indices based on specific taxa is to use taxonomy-free properties of the community for biomonitoring (Cook et al. 2024). One proposed taxonomy-free way to measure disturbed ecosystems is through phylogenetic structure (Helmus et al. 2010). If habitat filtering is the dominant assembly process in disturbed ecosystems, closely related species that share similar traits will be selected, resulting in a loss of genetic diversity (Helmus et al. 2010, Gainsbury and Colli 2019). We can extend this to the effects of chemical pollutants on microbes by considering how species would be filtered by chemi-
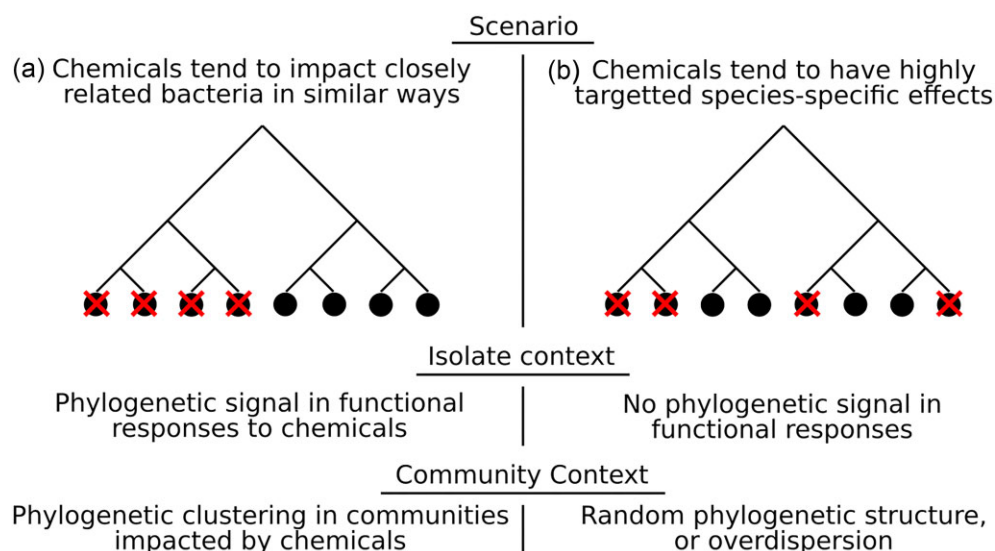
Figure 1. Phylogenetic structure in community responses to chemicals. (a) If chemical stressors negatively affect microbes (indicated by a cross) in a phylogenetically conserved way, i.e. similar taxa are all impacted by a given chemical, then we would expect to observe phylogenetic clustering in communities impacted by the same chemical as whole clades of taxa are filtered out. (b) Conversely, if there is no phylogenetic component to which species are impacted by a chemical, we would not expect to see phylogenetic clustering when compared to community structure in the absence of a chemical stressor.

cals with phylogenetically constrained effects (Fig. 1). If chemicals show fitness-limiting effects on similar microbes in a phylogenetically constrained pattern, then we would expect to observe phylogenetic clustering in communities exposed to these chemicals (Fig. 1a). That is, if clades of closely related species are impacted and filtered out of the community, then the remaining, unimpacted species will also be more closely related than expected (Webb et al. 2002, Cooper et al. 2008). Simulations have shown that microbial community responses to global change are likely to be phylogenetically conserved (Amend et al. 2016), if so, microbial phylogenetic community structure may offer a simple metric to diagnose polluted or perturbed ecosystems.

Recent work has shown that microbial community assembly is highly predictable and repeatable, resulting in set community structures (Pascual-Garcia et al. 2025). While communities may differ in their composition from site to site due to environmental differences or other factors (e.g. priority effects), we would nonetheless expect that phylogenetically related taxa will respond to pollutants in a structured, repeatable way, driven by their shared evolutionary history. In previous work, we demonstrated that bacterial isolates vary widely in their responses to chemical mixtures, and that a synthetic community composed of these isolates was generally more resilient than individual strains grown in monoculture (Smith et al. 2024). This raises the important question of whether ecotoxicological tests based on bacterial isolates can meaningfully predict how natural, more complex communities will respond to chemical stress. We build upon that work here by testing whether bacterial responses to individual pollutants scale to whole-community changes in composition and structure. Specifically, we tested (i) whether there was a phylogenetic signal in the responses of bacterial isolates to chemical pollutants; (ii) whether chemical pollutants altered the composition and phylogenetic structure of microbial communities; and (iii) whether the responses of monocultured taxa are also observed in the community context.

Methods

Bacterial isolation and culture

In these experiments, we assayed 26 strains of bacteria and six whole microbial communities from our −80°C stored culture library. Twenty strains of bacteria were isolated from pristine freshwater stream sediments from Iceland; the isolation and identification methods have been described previously (Smith et al. 2024). The six whole microbial communities were obtained from the original sediments that the Icelandic strains were isolated from. Note that the communities are numbered based on their original stream locations, and thus their numeric identifier is not representative of their position in this experiment. The remaining six bacterial isolates were obtained from soils from Silwood Park, Berkshire, UK, as described previously (Mombrikotb et al. 2022). We selected these strains as representing a wide diversity of taxa in our isolate library for the robustness of our phylogenetic analyses (see Fig. 2a), but also, crucially, for their capacity to be reliably revived in a rich culture media and grown to carrying capacity within a short time period (72 h). Bacterial strains and communities were revived from frozen glycerol stocks and grown to carrying capacity in Luria–Bertani media (Sigma-Aldrich) before chemical toxicity testing, as ascertained through absorbance readings at 600 nm (A_{600}).

Chemical toxicity experiments

A previous study tested off-target toxicity of 1076 industrial and agricultural chemicals on human gut bacteria, and found that 168 were active against at least one bacterial strain (Lindell et al. 2024). In this study, we tested the toxicity of those 168 chemicals on our environmental bacteria and communities. Stock chemicals were prepared in dimethyl sulfoxide (DMSO) following the same methodology as previously (Lindell et al. 2024). Bacterial cultures were grown to carrying capacity over 4 days, and then inoculated into clear 96-well assay plates via a 1:100 dilution into fresh media adjusted with the chemicals at a concentration of 20 μM (1% DMSO). Control wells were pre-
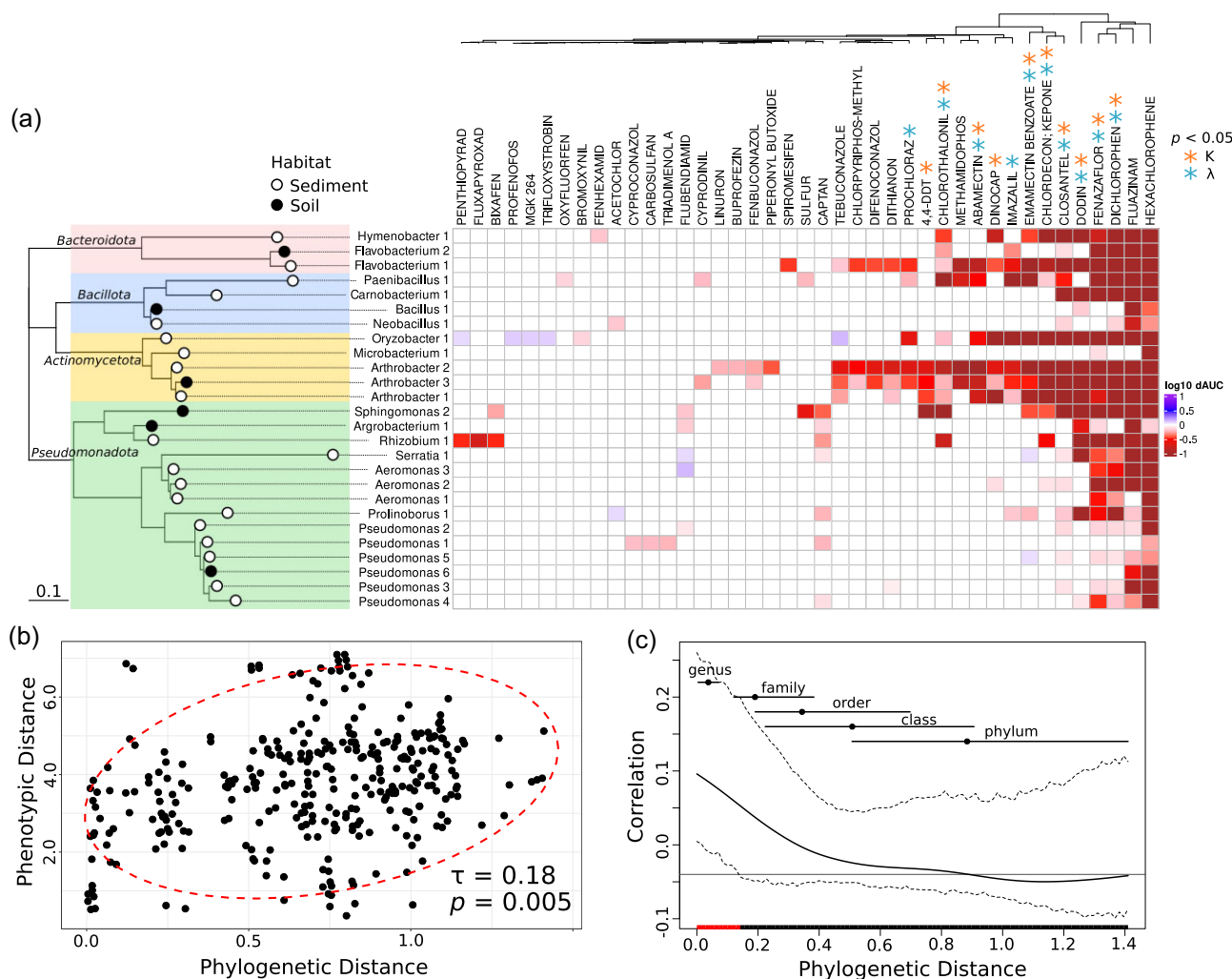
Figure 2. Phylogenetic signal in isolate responses to chemical pollutants. a. Heatmap of chemical impacts on isolate growth. Colours show the impact on growth on a $\log_{10}$ scale, i.e. -1 is 10x less growth (red), 0 is the same as control growth (white), $+1$ is 10x more growth (blue). Dendrogram above shows hierarchical clustering of the chemical responses based on Euclidean distance. Rows are ordered by the phylogeny of the isolates built from their 16S sequences, shown to the left. Different phyla are highlighted by colour, scale bar shows phylogenetic distance. Only the chemicals with significant impacts on the growth of at least one isolate are shown. Orange and blue asterisks show chemicals where significant phylogenetic signal (P < .05, adjusted for multiple testing) was found via Blomberg's K or Pagel's λ, respectively. (b) Pairwise phylogenetic distance plotted against pairwise phenotypic distance (from hierarchical clustering), with the Mantel test statistic (Kendall's τ) used to compare the matrices. Ellipse shows the 95% prediction region for the data distribution. (c) Phylogenetic correlogram, representing the phylogenetic distances at which the correlation between traits is observed. Solid black line shows the Moran's I index and dashed lines show the 95% confidence envelope. The horizontal black line indicates the expected value of Moran's I with no phylogenetic correlation (null hypothesis). Red highlight indicates the range of phylogenetic distances with significant trait correlation (i.e. where the null hypothesis value of Moran's I is outside of the 95% confidence interval). Points and horizontal lines represent the mean and range of phylogenetic distances between taxa in different levels of taxonomic organization, respectively. For genus this is the mean pairwise distance between taxa of different genera within the same family, for family this is the mean pairwise distance between taxa of different families within the same order, and so on.

pared with media containing 1% DMSO but no chemical pollutant. Further, wells with no bacterial inoculant were also prepared as controls for contamination. Growth was assayed by measuring A_{600} once per hour, for 72 h using an automated plate reader (Synergy 2, Agilent BioTek) and stacker (Biostack 4, Agilent BioTek). Plates were briefly shaken before reading to homogenize samples and disrupt biofilm formation. Three independent biological replicates of each bacterial-chemical combination were carried out on different days and in different plate layouts of the chemicals to prevent edge effects. For community assays, plates were subsequently stored at $-80°C$ for DNA extraction.

Quantification of chemical impacts on growth

A chemical may impact different or multiple metrics of bacterial growth (e.g. the lag time, maximum growth rate, carrying capacity). Bacterial growth curves are also likely to deviate from classic sigmoidal shapes and display nonstandard dynamics under stressful conditions, leading to difficulties in growth model fitting. Therefore, rather than focus on specific growth parameters, we calculated the area under the growth curve (AUC) as an overall metric of fitness, as previously used in high-throughput microbial toxicity screens (Maier et al. 2018, Lindell et al. 2024, Smith et al. 2024). AUC was calculated by fitting a spline curve to each of the growth curves, then integrating over the 72 h growth period. To

determine whether chemicals had a significant impact on AUC for each bacterial culture tested, we performed a Dunnett's test against the control (DMSO only) wells. For downstream analysis, we calculated the ratio of growth in the presence of a chemical pollutant to the growth in the presence of DMSO (control) to give the relative growth (dAUC),

$$dAUC_i = \frac{\overline{AUC_i}}{\overline{AUC}_{DMSO}}$$

where i is one of the 168 chemicals tested and the bar represents the mean of biological replicates. We calculated a dAUC value for each bacterial × chemical combination in this manner.

All growth calculations were performed in R, version 4.3.2 (R core Team 2022).

Nanopore sequencing of 'community 9'

Based on the dAUC results, we selected 23 chemical treatments to sequence the postgrowth assay communities for 'community 9', as well as the DMSO control (see Supplementary Table S1 for details). We sequenced two replicates of each treatment, for a total of 48 samples sequenced from this community. DNA extraction was performed using a Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research, Irvine, CA). The Nanopore 16S barcoding kit-24 (SQK-16S024) was used to prepare 16S amplicon libraries for sequencing on a MinION Mk1B device (Oxford Nanopore Technologies, UK), following the manufacturer's protocol. Barcoded polymerase chain reaction (PCR) products were quantified using a Qubit fluorometer and adjusted to 10 ng/µl with Tris-HCl for each sample. Barcoded samples were then pooled into two libraries of 24 samples, loaded onto a MinION flow cell (R 9.4.1), and read for 12 h. The flow cell was washed between libraries using the Nanopore flow cell wash kit (EXP-WSH004).

The Nanopore runs of the two libraries generated a total of 3.24 million and 2.68 million reads respectively (>100 k reads per sample), with an N50 of 1.65 kb. Basecalling, barcode trimming, adapter trimming, and read splitting were performed in Guppy version 6.4.6, using the high accuracy model for basecalling. Reads were then size-filtered to retain sequences of 1300–1950 bp in length using SeqKit version 2.30., following (Matsuo et al. 2021). Species level taxonomic abundances were estimated using Emu, version 3.4.4 (Curry et al. 2022), with reads mapped to the prebuilt 16S database (a combination of rrnDB v5.6 and NCBI 16S RefSeq from 17–09–2020).

Phylogenetic reconstruction

We constructed a phylogeny for the bacterial isolates in our experiments using their 16S sequences. These are the 16S sequences directly collected from those strains as part of the process to characterize them. Following previous methodology (Smith et al. 2022), these sequences were aligned to the SILVA 16S reference database using the SILVA Incremental Aligner (Pruesse et al. 2012). Based on this alignment, we inferred 100 phylogenetic trees in RAxML (v8.1.1) using a general time reversible substitution model with gamma distributed rates, and selected the tree with the highest log-likelihood as the consensus phylogeny (Fig. 2a). For the sequenced 'community 9' samples, we extracted the 16S sequences from the Emu database, matching every unique taxon identified across the sequencing runs. We then performed the same process in SILVA and RAxML to produce a phylogeny representing the community (Supplementary Fig. S3).

Statistical analysis

To test for phylogenetic signal in the responses of bacterial isolates to the suite of chemical pollutants, we used the dAUC as the phenotype for each bacteria-chemical combination. We calculated the pairwise Euclidean distance of phenotypic responses between the different strains and compared the pairwise distance matrix to the pairwise distances extracted from the phylogeny using a Mantel test (Fig. 2b). As previously, we used Kendall's rank correlation τ as the test statistic for the Mantel test due to the nonparametric distribution of the pair-wise phenotypic distances (Smith et al. 2024). The Mantel test was performed using the '*Vegan*' R package (v2.6–8) (Oksanen et al. 2022).

We also tested for phylogenetic signal in the responses of the bacteria to each chemical individually by quantifying Blomberg's K and Pagel's λ, which quantify the tendency of closely related species to exhibit similar trait values (dAUC) under a Brownian motion model of trait evolution. These metrics and tests for their significance were calculated in the '*Phytools*' R package (Revell 2012). *P*-values from these tests were adjusted to account for multiple testing using a Benjamini-Hochberg false discovery rate correction.

To analyse microbial community dissimilarities, we calculated the Bray-Curtis dissimilarities based on the taxonomic abundance profiles in '*Vegan*'. We used the previously constructed community phylogeny to calculate UniFrac distances weighted by the taxon abundances, using the '*phyloseq*' R package (v1.46) (McMurdie and Holmes 2013). We visualized these as ordinations using principal co-ordinate analysis (PCoA) implemented in the '*ape*' R package (v5.8) (Paradis and Schliep 2019). For the community data, we also quantified measures of community diversity—species richness and Shannon diversity—using the '*Vegan*' package. We also calculated the standard effect size (SES) of mean phylogenetic distance (MPD) between taxa within each sequenced community sample using the '*pez*' R package (v1.2–4) (Pearse et al. 2015). The SES of MPD quantifies how much the observed MPD deviates from the expectation under a random community assembly null model.

Results

We assayed the impact of 168 chemical pollutants on 26 strains of bacteria in our culture library, isolated from two different habitats. These habitats were stream sediments in Iceland, see Smith et al. (2024), and soils from Silwood Park, Berkshire, see Mombrikotb et al. (2022). The chemical pollutants were chosen based on a previous analysis of off-target effects of industrial and agricultural chemicals on human gut bacteria (Lindell et al. 2024). We quantified the growth of bacterial isolates by measuring the optical density of liquid cultures every hour for 72 h (3 days) using an automated plate reader. We quantified the AUC as an overall metric of growth. We calculated the ratio of AUC in the presence of a chemical pollutant to the AUC in the presence of DMSO (control) to find the relative growth (dAUC) in the presence of a chemical. Of the potential 4368 bacteria-chemical responses (26 strains × 168 chemicals), 242 were significantly different from the controls, i.e. the chemical tested had a significant impact on growth of the isolate (Dunnett's test, Supplementary Fig. S1). These significant responses involved only 41 of the 168 chemicals tested. Of these significant responses, most were negative (232/242, 95.9% of significant bacteria-chemical responses dAUC <1), i.e. growth was significantly reduced in the presence of the chemical (Fig. 2a). The responses also displayed a nested pattern. That is, some more susceptible isolates were impacted by many chemicals, whereas
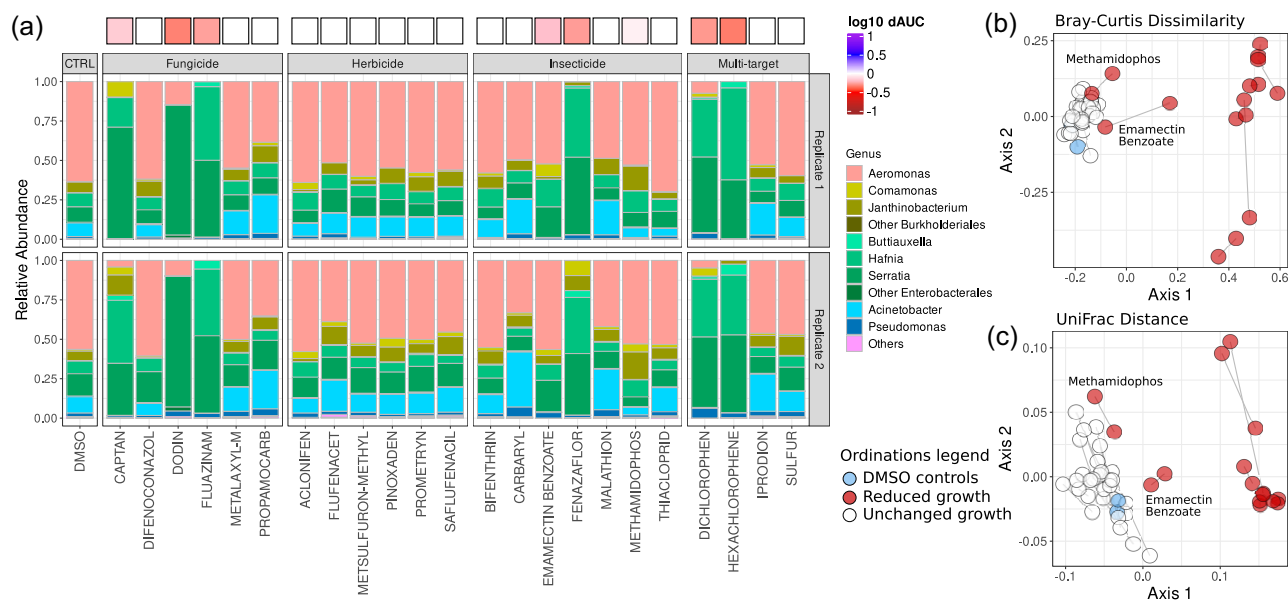
Figure 3. Community composition profiles. (a) Relative abundance of taxa from Community 9 grown under differing chemical conditions. Growth with DMSO (left panel) is the control condition (CTRL). Relative abundances are aggregated to the genus level, with the most abundant genera shown. Top and bottom panels are sequencing profiles from two replicates for each condition. The colour boxes above the bars show the relative growth (dAUC) on the same $\log_{10}$ colour scale as Fig. 2a. Chemicals are grouped by broad target classes, with insecticides including acaricides. Multitarget chemicals span more than one of these classifications. (b) Principal co-ordinates analysis (PCoA) based on Bray-Curtis dissimilarity. The chemicals which don't impact community functioning (growth) cluster on with the DMSO controls (blue) and represent the chemical treatments which have largely unaffected the community composition. The compositions in the presence of chemicals which impacted community growth (captan, dodin, dichlorophen, emmamectin benzoate, fenazaflor, fluazinam, hexachlorophene and methamidophos; coloured red) are shifted in the ordination space by differing degrees. (c) PCoA based on weighted UniFrac distance, a measure of phylogenetic dissimilarity of communities. The chemicals which caused reductions in community growth also caused those communities to be more phylogenetically dissimilar to the control condition. In both ordinations (b and c), lines join the two replicate points for each condition. Methamidaphos and emamectin benzoate are labelled to show that these communities cluster near the controls despite showing (moderate) responses in their growth.

some more resilient isolates were impacted by only a few chemicals, but these few were a subset of those impacting the susceptible isolates (Fig. 2a).

We assayed bacterial isolates from a wide range of phylogenetic groups, covering four bacterial phyla, obtained from stream sediments or soils (Fig. 2a). Similar taxa clustered together in their responses to the chemicals but were not perfectly grouped by taxonomy (Supplementary Fig. S1). We tested for phylogenetic signal in the responses of the isolates to each chemical using both Blomberg's K and Pagel's λ. We found significant phylogenetic signal from at least one of these metrics in the responses to 12 of the chemicals, within which eight showed significant phylogenetic signal by both metrics (Fig. 2a, Supplementary Table S2). We also tested phylogenetic signal across the combined suite of chemical responses via a Mantel test (using Kendall's τ) comparing the pair-wise distance matrix computed from the chemical responses to the distance matrix of the phylogeny (Fig. 2b). We found a significant correlation between the phenotypic distance (responses to the chemicals) and the phylogenetic distance (Mantel statistic $\tau = 0.1829$, $P = .0054$). This means that related bacteria tend to respond to the chemical treatments in the same way, despite being obtained from two disparate environments. We located where this phylogenetic signal occurred within the taxonomy, using a phylogenetic correlogram (Fig. 2c). A phylogenetic correlogram shows the correlation of a trait across species as a function of their phylogenetic distance, illustrating how trait similarity changes with increasing evolutionary separation (Keck et al. 2016).We find that the correlation occurs only at relatively short phylogenetic distances (<0.14), which at most equates to the dis-

tance between taxa in different genera or different families in our dataset (Fig. 2c).

We also assayed the growth responses of six whole bacterial communities obtained from the same stream sediments that the sediment isolates were cultured from to the same set of chemical stressors. The communities were generally resilient in their growth responses, clustering together and with the more resilient *Pseudomonas* isolates (Supplementary Fig. S1). Only 10 of the chemicals had impacts on the overall growth of any of the communities, and none of the chemicals completely prevented community growth (Supplementary Fig. S1). To understand how the community growth responses related to differences in community composition, we sequenced one of the communities ('community 9') after growth in 23 of the chemicals, and the DMSO control. We selected a range of chemicals (see Supplementary Table S1), including those most impacting the communities' growth, to understand whether impacts on growth were due to changes in community composition. We also selected some chemicals which elicited no significant impact on growth, to investigate functional redundancy—whether communities may have lost taxa but showed no change in functional response. We found remarkable consistency in the abundance profiles between replicates impacted by the same chemicals (Fig. 3a, Supplementary Fig. S2). We also observed that the greatest changes in abundance profiles were in the presence of chemicals which also caused a change in community growth (Fig. 3a). These responses were very similar when aggregated to the genus-level, and exemplified by a loss of the *Aeromonas* taxa, with the exception of emmamectin benzoate and methamidophos (Fig. 3a). However, we note that at the
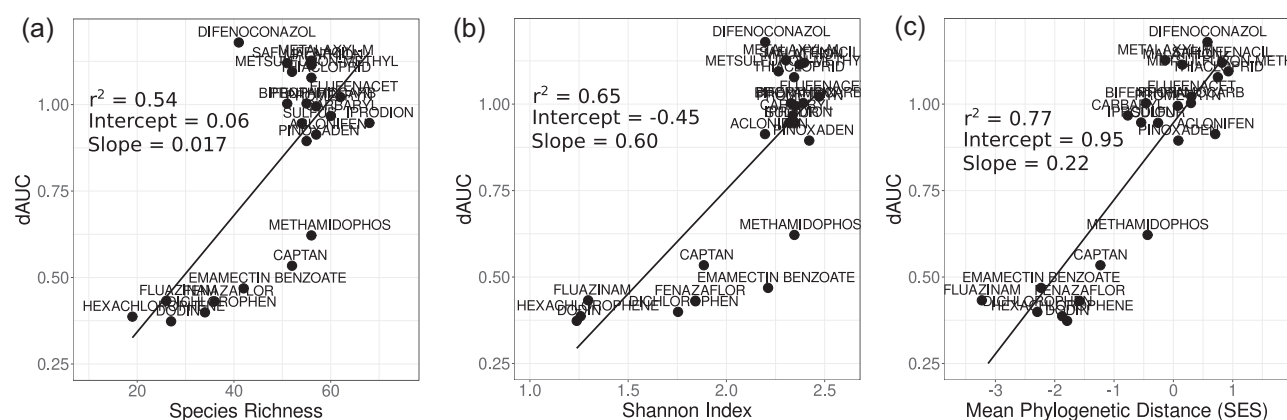
Figure 4. Community diversity, structure, and functioning. Linear regressions testing the impact of community diversity (species richness, a; Shannon index, b; standard effect size of mean phylogenetic distance, c) on functioning (growth). We find that while all three metrics are significant predictors of differences in community growth, the phylogenetic distance (SES.MPD) is the strongest predictor ($r^2 = 0.77$).

species-level, distinct taxonomic profiles arise in response to each of the chemicals which induce a growth response (Supplementary Figs S2 and S3).

We further examined differences in community composition via PCoA, using the Bray-Curtis dissimilarity metric (Fig. 3b) and using weighted UniFrac distance, a measure of phylogenetic dissimilarity (Fig. 3c). In both cases, we found that replicates generally clustered together, showing repeatability of community assembly in response to the chemical treatments. We found that communities grown in the presence of chemicals which did not affect growth clustered together with the DMSO controls in both ordinations, implying no difference in composition compared to the control condition. We found that the compositions in the presence of chemicals which impacted growth were shifted away from the DMSO controls, in a similar direction in the ordination space, implying a constrained community response to chemical pollutants (Fig. 3b and c). As observed in their taxonomic profiles (Fig. 3a) emamectin benzoate and methamidophos did not induce the same compositional changes as the other impactful chemicals, clustering more with the controls and unimpacted communities than the communities with reduced growth (Fig. 3b and c).

To further test how community composition was impacted by the chemical treatments and to link this to the functional measurements, we quantified three measures of community diversity—species richness, the Shannon index, and the standard effect size of mean phylogenetic distance (SES.MPD). The species richness is simply the count of unique species in the communities. The Shannon index measures the diversity based on the species richness but also accounts for the evenness of the community (how evenly the individuals are distributed among species). SES.MPD quantifies the degree to which species in a community are phylogenetically related, with negative values indicating that species are more closely related than expected by chance. We took the mean of the two replicates when calculating each of these metrics. Each of these metrics explained the differences in growth of the communities compared to the control to different degrees, with SES.MPD providing the best model fit (linear models: species richness $r^2 = 54\%$, Shannon index $r^2 = 65\%$, SES.MPD $r^2 = 77\%$; Fig. 4). These results show that communities with lower growth compared to the DMSO controls had lower diversity as quantified by the species richness and Shannon index (Fig. 4). They also show that communities with lowered growth were more phylogenetically clustered (negative SES.MPD), i.e. species in these

communities were more closely related than in the communities showing no growth response.

The responses of single isolates do not account for biotic interactions between taxa which occur in communities of microbes. We therefore compared the effects of chemicals on specific taxa within a community context to their responses to those chemicals in monoculture. Between our monocultured isolates and sequenced community, there were five shared genera: *Aeromonas*, *Carnobacterium*, *Flavobacterium*, *Pseudomonas*, and *Serratia*. We tested whether the functional responses of members of a genus in monoculture were correlated with changes in the abundance of the same genus in a community by comparing the mean functional response (dAUC) of that genus to its aggregated fold change in relative abundance within the community in response to each of the 23 chemical pollutants (Fig. 5). We do not suggest direct causation or dependence between these variables nor assume that one predicts the other, but quantified the strength of relationship between them. We quantified both Pearson and Spearman's rank correlations (Table 1), capturing linear and monotonic trends respectively. We found significant positive relationships between these chemical response measures for three of our genera: *Aeromonas*, *Carnobacterium*, and *Flavobacterium*. Conversely, we found a negative relationship for *Serratia*, i.e. *Serratia* relative abundance tended to be increased in the presence of chemicals that caused a decrease in their monoculture growth. However, this negative relationship was not significant by Spearman's rank after accounting for multiple testing (Table 1). We found no significant relationship between these measures for *Pseudomonas*.

Discussion

Our study subjected bacterial isolates and whole bacterial communities to a large collection of environmental pollutants and showed the predictability of microbial responses to those chemical pollutants. We found phylogenetic signal in the growth responses of bacterial isolates to chemicals, i.e., evolutionarily similar taxa responded similarly to the same chemical treatments. This means that we can use the responses of target taxa to infer the responses of other, similar species. We also found nestedness in the responses of isolates, with some chemicals impacting all or most of the isolates, but others impacting smaller subsets of the bacteria. This shows that within our dataset there were resilient taxa which were impacted by only a small number of chemicals,
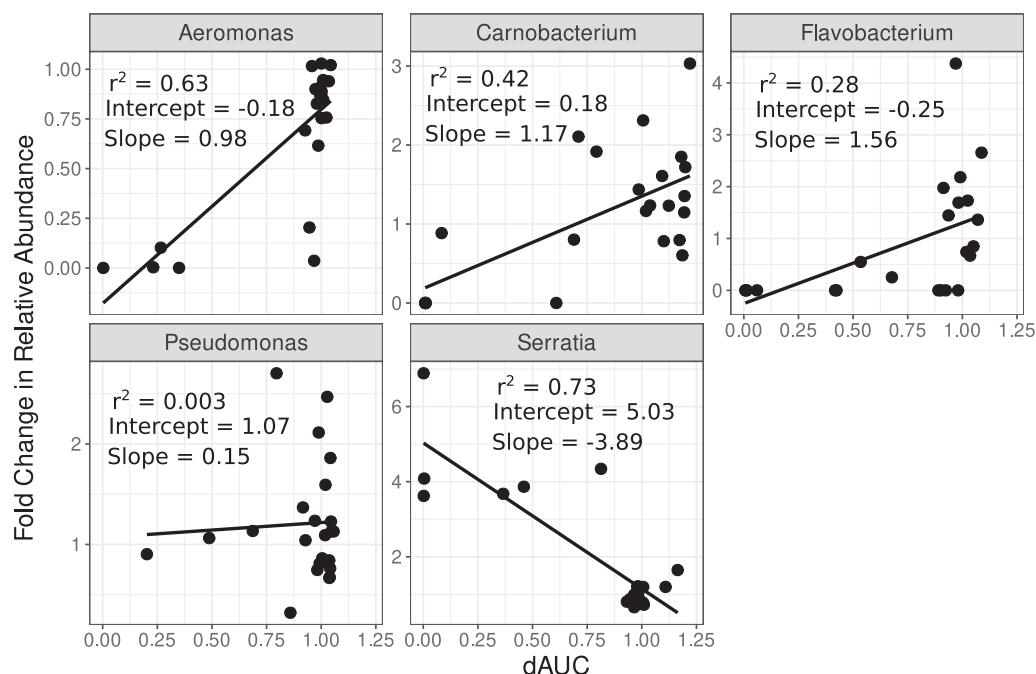
Figure 5. Comparison of shared genera in monoculture and in a community. The mean change in growth compared to control growth (dAUC) for five focal genera in the presence of each chemical pollutant is plotted against the fold change in abundance of those genera in the community data. We also provide the statistical outputs from linear regressions, for comparison to the correlation tests in Table 1.

Table 1. Correlations between monoculture and community responses.

Genus	Pearson's r	P-value (Pearson)	Adjusted P (Pearson)	Spearman's ρ	P-value (Spearman)	Adjusted P (Spearman)
Aeromonas	0.79	>.001*	**>.001***	0.68	>.001*	**.0012***
Carnobacterium	0.65	>.001*	**.0013***	0.52	.010*	**.017***
Flavobacterium	0.53	.0098*	**.012***	0.70	>.001*	**>.001***
Pseudomonas	0.05	.81	.81	-0.03	.90	.90
Serratia	-0.86	>.001*	**>.001***	-0.42	.044*	.056

*$P < .005$.
Pearson's and Spearman's rank test statistics and P-values from correlation tests of isolate growth responses (change in dAUC) against community growth responses (fold change in relative abundance) in the presence of the chemical pollutants (see Fig. 5). The adjusted P-values are after application of a Benjamini-Hochberg false discovery rate (FDR) adjustment to account for multiple testing. Bold font indicates significant P-values after FDR adjustment.

and more sensitive taxa that were impacted by many more of the chemicals. Many of the broad-spectrum chemicals here also showed broad-spectrum activity in screens of human gut bacteria, further highlighting the generalizability of their off-target activity (Lindell et al. 2024). Phylogenetic signal was primarily observed in the responses to chemicals with intermediate levels of impact, i.e. those that affected a moderate number of strains. This is consistent with expectations, as signal detection requires variation in trait values: if all strains respond similarly (either all sensitive or all resistant), there is little structure to explain.

The single strain results were reflected in the community analysis. The chemicals which most impacted the isolates also most impacted the community that we sequenced, reducing both the community's growth and diversity. A key finding is that the phylogenetic structure of the community was also altered in the microcosms with lower growth. We quantified the SES.MPD, which gives an overall view of the community's phylogenetic structure, providing insight into broad ecological and evolutionary processes, such as large-scale environmental filtering (Webb et al. 2002). Positive values of SES.MPD imply phylogenetic overdispersion (species in a community are more distantly related than expected by

chance), whereas negative values imply phylogenetic clustering (species are more closely related than expected) (Webb et al. 2002, Cooper et al. 2008). Alternatively, community assembly may not be strongly influenced by phylogeny (Cooper et al. 2008), with values close to zero implying stochastic community assembly. We can therefore use SES.MPD to detect patterns where environmental conditions select for closely related species, leading to phylogenetic clustering (Webb et al. 2002, Cavender-Bares et al. 2009). Our results indicate a shift in community structure from primarily stochastic community assembly under control conditions to strong environmental selection in the communities with lowered growth. This is expected if clades of closely related taxa are lost from the community, as predicted from the phylogenetic signal in the responses of isolates. However, the processes driving phylogenetic structure in communities may be more complex than environmental filtering alone, and clustering may also result from biotic interactions (Mayfield and Levine 2010, Kraft et al. 2015, Cadotte and Tucker 2017). In our study system, it is possible that both abiotic filtering (via direct chemical stress) and biotic processes (changes in the outcomes of competitive exclusion) contributed to the observed patterns of community assembly. We

also quantified UniFrac distance, the pairwise phylogenetic dissimilarities between the sequenced communities. We found that the communities showing no change in growth tended to cluster together and with the controls in ordination space based on UniFrac distance, while the communities whose growth was impacted clustered away from the controls. This further highlights the use of phylogenetic measures for diagnosing disturbed microbial communities.

None of the isolates tested showed a response to just one chemical, more often they showed similar responses to many of the chemical pollutants. Similarly, few chemicals impacted only one of the isolates tested. From a biosensors perspective, it may therefore be difficult to pinpoint contamination by a specific chemical pollutant based on the response of single isolate 'biosensors'. Indeed, previous work has identified groups of taxa which may be indicative of multiple environmental variables (Sperlea et al. 2021). We also found a lack of sensitivity in our community responses— many chemicals which impacted the communities caused similar shifts in the community composition. This suggests the need for a broader approach when using microbial communities as bioindicators. Our findings of sets of sensitive and resilient taxa in both our isolate and community work suggest that loss of sensitive taxa from a community may be an early indication of environmental perturbation. However, we also found that the sensitivities of certain taxa in a community context couldn't always be inferred from their monoculture responses. Indeed, we found an increased relative abundance of *Serratia* in communities dosed with the same chemicals that elicited a negative response in our *Serratia* isolate when grown in monoculture. This may be due to the compositional nature of microbial community datasets, which are difficult to interpret in terms of changes in abundances of certain taxa in a community, when we can only obtain specimens of a certain read depth (Weiss et al. 2017). Furthermore, the biological responses of a taxon may simply be different in a community context where dynamics are influenced by ecological interactions not reflected in monoculture growth (Gralka et al. 2020). This also raises questions, as in previous work (Smith et al. 2024), about the usage of model organisms in ecotoxicology testing, given the difficulties in generalizing monoculture responses to broader contexts. Our phylogenetics results show that instead of focusing on the direct responses of indicator taxa, we can infer the loss of sensitive taxa from increased phylogenetic clustering of microbial communities, suggesting that a simple metric (SES.MPD) could be used to move forward with microbial-based bio-monitoring.

A key challenge in using microbial community data for routine water quality monitoring is the time and cost involved with sequencing approaches. Traditional sequencing technologies, such as Illumina, are resource-intensive and can take days to generate results. However, recent advancements in sequencing technology, particularly the use of portable Nanopore sequencing devices, offer a promising alternative. Indeed, studies have already demonstrated the feasibility of using Nanopore technology in the field for rapid sequencing analyses (Menegon et al. 2017, Parker et al. 2017, Tyler et al. 2023). The ability to monitor ecosystems in real time could revolutionize bioindication practices, providing a more immediate understanding of pollutant impacts. We highlight here that we generated the sequencing data for the current study using a portable Nanopore MinION device, further showing the utility of this technology.

Despite the potential for biomonitoring, it is important to note that our work was conducted in a controlled lab environment, growing both isolates and microbial communities in culture media. Therefore, while our findings show promise for future applications, they need to be validated in the field where microbial communities are influenced by a range of variables not present in the lab, such as natural variations in nutrient levels, temperature, and flow conditions. Furthermore, real-world environments often involve exposure to mixtures of pollutants rather than single compounds, which can interact in nonadditive ways (Orr et al. 2020, Morris et al. 2022, Turschwell et al. 2022, Smith et al. 2024). In our previous work, we found that bacterial responses to mixtures of chemicals were not easily predicted from their responses to individual stressors, nor from phylogenetic relatedness (Smith et al. 2024). However, a synthetic community showed greater resilience to chemical mixtures, highlighting the buffering effects of biotic interactions in multi-species communities. Future research should investigate how these two levels of complexity—chemical mixtures and microbial community interactions—impact each other, and whether patterns of nestedness and phylogenetic signal persist in more realistic exposure scenarios.

More generally, our results have importance for understanding the responses of microbial communities to perturbations. Our finding of broad repeatability in the community composition in response to different chemicals, while useful for biomonitoring processes, is also informative for understanding microbial community assembly. We find that community assembly may have different end-points depending upon the chemical perturbation, but that replicates converge on the same composition. This is in agreement with recent work showing that microbial community assembly converges on a small number of set structural compositions (Pascual-garcía et al. 2025). Furthermore, our data suggest that microbial communities exhibit some degree of resilience, with communities showing greater capacity to tolerate chemical stress than individual bacterial isolates. For example, while hexachlorophene severely inhibited the growth of most isolates, all tested communities maintained some growth in its presence, highlighting the importance of biotic interactions and functional redundancy in determining ecosystem resilience (Biggs et al. 2020). However, this resilience was not without limits. In cases where chemicals led to a clear loss in community diversity, we also observed a corresponding decrease in growth, indicating a potential lack of functional redundancy. This suggests that while communities may be more robust than individual isolates, there is a tipping point where loss of diversity directly translates to loss of ecosystem function. These results highlight the importance of protecting microbial diversity in natural ecosystems, as it underpins both their ecological resilience and biogeochemical functions (Cavicchioli et al. 2019).

In this study we move away from a focus on specific indicator taxa, and suggest a broader, phylogenetically informed approach to microbial bioindication. Our findings demonstrate that microbial responses to pollutants are, in many cases, phylogenetically conserved and repeatable, and that these patterns scale from individual isolates to whole-community changes in composition and structure. Importantly, we show that simple community phylogenetic metrics such as SES.MPD can capture these responses, offering a promising route for the development of generalizable and scalable biomonitoring techniques. By pairing this framework with portable, *in situ* sequencing technologies, there is clear potential to transform environmental monitoring, enabling rapid, real-time water quality assessments. Future work should focus on refining these tools for field applications, particularly by testing microbial community responses to a broader range of ecosystems and pollutant mixtures directly under natural environmental variability. Ultimately, our results show that incorporating microbial phylogenetic structure into routine biomonitoring could

help to detect environmental degradation, guiding more timely and effective ecosystem management interventions.

Author contributions

Thomas P. Smith (Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Supervision, Visualization, Writing—original draft, Writing—review & editing), Rachel Hope (Data curation, Formal Analysis, Investigation, Visualization, Writing—review & editing), and Thomas Bell (Conceptualization, Methodology, Project administration, Supervision, Writing—review & editing)

Supplementary data

Supplementary data is available at *FEMSEC* Journal online.

Conflict of interest : None declared.

Data availability

Supplementary data files provided include the phenotypic chemical response data, extended details of the chemical library, extended details of the isolates screened, and the species abundance tables from the community analysis.

The raw growth curve data are made available on figshare: https://doi.org/10.6084/m9.figshare.28684982

The raw Nanopore sequencing dataset for the community analysis has been made available on the NCBI sequence read archive under the BioProject ID PRJNA1222838.

References

Amend AS, Martiny AC, Allison SD *et al.* Microbial response to simulated global change is phylogenetically conserved and linked with functional potential. *ISME J* 2016;**10**:109–18. https://doi.org/10.1038/ismej.2015.96.

Astudillo-García C, Hermans SM, Stevenson B *et al.* Microbial assemblages and bioindicators as proxies for ecosystem health status: potential and limitations. *Appl Microbiol Biotechnol* 2019;**103**:6407–21. https://doi.org/10.1007/s00253-019-09963-0.

Bani A, Randall KC, Clark DR *et al.* Mind the gaps: What do we know about how multiple chemical stressors impact freshwater aquatic microbiomes? In *Functional Microbiomes*, Vol. **67**, 1st edn, London, United Kingdom: Elsevier Ltd. 2022. https://doi.org/10.1016/bs.aecr.2022.09.003.

Biggs CR, Yeager LA, Bolser DG *et al.* Does functional redundancy affect ecological stability and resilience? A review and meta-analysis. *Ecosphere* 2020;**11**:e03184. https://doi.org/10.1002/ecs2.3184.

Cadotte MW, Tatsumi S. Uncovering the mechanisms underpinning divergent environmental change impacts on biodiversity and ecosystem functioning. *Ecology* 2025;**106**:1–13. https://doi.org/10.1002/ecy.70040.

Cadotte MW, Tucker CM. Should environmental filtering be abandoned? *Trends Ecol Evol* 2017;**32**:429–37. https://doi.org/10.1016/j.tree.2017.03.004.

Cavender-Bares J, Kozak KH, Fine PVA *et al.* The merging of community ecology and phylogenetic biology. *Ecol Lett* 2009;**12**:693–715. https://doi.org/10.1111/j.1461-0248.2009.01314.x.

Cavicchioli R, Ripple WJ, Timmis KN *et al.* Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Micro* 2019;**17**:569–86. https://doi.org/10.1038/s41579-019-0222-5.

Cook LSJ, Briscoe AG, Fonseca VG *et al.* Microbial, holobiont, and Tree of Life eDNA /eRNA for enhanced ecological assessment. *Trends Microbiol* 2025;**33**:48–65. https://doi.org/10.1016/j.tim.2024.07.003.

Cooper N, Rodríguez J, Purvis A. A common tendency for phylogenetic overdispersion in mammalian assemblages. *Proc R Soc B* 2008;**275**:2031–7. https://doi.org/10.1098/rspb.2008.0420.

Curry KD, Wang Q, Nute MG *et al.* Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nat Methods* 2022;**19**:845–53. https://doi.org/10.1038/s41592-022-01520-4.

De Laender F, Rohr JR, Ashauer R *et al.* Reintroducing environmental change drivers in biodiversity–ecosystem functioning research. *Trends Ecol Evol* 2016;**31**:905–15. https://doi.org/10.1016/j.tree.2016.09.007.

Gainsbury AM, Colli GR. Phylogenetic community structure as an ecological indicator of anthropogenic disturbance for endemic lizards in a biodiversity hotspot. *Ecol Indic* 2019;**103**:766–73. https://doi.org/10.1016/j.ecolind.2019.03.008.

Gralka M, Szabo R, Stocker R *et al.* Trophic interactions and the drivers of microbial community assembly. *Curr Biol* 2020;**30**:R1176–88. https://doi.org/10.1016/j.cub.2020.08.007.

Grossart HP, Massana R, McMahon KD *et al.* Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. *Limnol Oceanograph* 2020;**65**:S2–S20. https://doi.org/10.1002/lno.11382.

Helmus MR, Keller W, Paterson MJ *et al.* Communities contain closely related species during ecosystem disturbance. *Ecol Lett* 2010;**13**:162–74. https://doi.org/10.1111/j.1461-0248.2009.01411.x.

Keck F, Rimet F, Bouchez A *et al.* Phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol Evol* 2016;**6**:2774–80. https://doi.org/10.1002/ece3.2051.

Kraft NJB, Adler PB, Godoy O *et al.* Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* 2015;**29**:592–9. https://doi.org/10.1111/1365-2435.12345.

Lindell AE, Kamrad S, Roux I *et al.* Off-purpose activity of industrial and agricultural chemicals against human gut bacteria. *Biorxiv* 2024;1–26. preprint: 2024.09.05.610817.

Maier L, Pruteanu M, Kuhn M *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* 2018;**555**:623–8. https://doi.org/10.1038/nature25979.

Matsuo Y, Komiya S, Yasumizu Y *et al.* Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. *BMC Microbiol* 2021;**21**:1–13. https://doi.org/10.1186/s12866-021-02094-5.

Mayfield MM, Levine JM. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol Lett* 2010;**13**:1085–93. https://doi.org/10.1111/j.1461-0248.2010.01509.x.

McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;**8**:e61217. https://doi.org/10.1371/journal.pone.0061217.

Menegon M, Cantaloni C, Rodriguez-Prieto A *et al.* On site DNA barcoding by nanopore sequencing. *PLoS One* 2017;**12**:1–18. https://doi.org/10.1371/journal.pone.0184741.

Mombrikotb SB, Van Agtmaal M, Johnstone E *et al.* The interactions and hierarchical effects of long-term agricultural stressors on soil bacterial communities. *Environ Microbiol Rep* 2022;**14**:711–8. https://doi.org/10.1111/1758-2229.13106.

Morris OF, Loewen CJG, Woodward G *et al.* Local stressors mask the effects of warming in freshwater ecosystems. *Ecol Lett* 2022;**25**:2540–51, https://doi.org/10.1111/ele.14108.

Oksanen J, Simpson GL, Blanchet FG *et al.* Vegan: community Ecology Package (R package version 2.6-4) 2022. https://cran.r-project.org/package=vegan.

Orr JA, Vinebrooke RD, Jackson MC *et al.* Towards a unified study of multiple stressors: divisions and common goals across research disciplines. *Proc R Soc B* 2020;**287**:20200421. https://doi.org/10.1098/rspb.2020.0421.

Paradis E, Schliep K. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019;**35**:526–8. https://doi.org/10.1093/bioinformatics/bty633.

Parker J, Helmstetter AJ, Devey D *et al.* Field-based species identification of closely-related plants using real-time nanopore sequencing. *Sci Rep* 2017;**7**:1–8. https://doi.org/10.1038/s41598-017-08461-5.

Pascual-garcía A, Rivett D, Jones ML *et al.* Replicating community dynamics reveals how initial composition shapes the functional outcomes of bacterial communities. *Nature Communications* 2025;**16**:3002.

Pearse WD, Cadotte MW, Cavender-Bares J *et al.* Pez: phylogenetics for the environmental sciences. *Bioinformatics* 2015;**31**:2888–90. https://doi.org/10.1093/bioinformatics/btv277.

Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;**28**:1823–9. https://doi.org/10.1093/bioinformatics/bts252.

R core Team. *R: a language and environment for statistical computing (4.2.1)* Vienna, Austria: R Foundation for Statistical Computing, 2022. https://www.r-project.org/.

Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 2012;**3**:217–23. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

Rillig MC, Ryo M, Lehmann A. *et al.* The role of multiple global change factors in driving soil functions and microbial biodiversity. *Science* 2019;**366**:886–90. https://doi.org/10.1126/science.aay2832.

Rillig MC, van der Heijden MGA, Berdugo M *et al.* Increasing the number of stressors reduces soil ecosystem services worldwide. *Nat Clim Chang* 2023;**13**:478–83. https://doi.org/10.1038/s41558-023-01627-2.

Sagova-Mareckova M, Boenigk J, Bouchez A *et al.* Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Res* 2021;**191**:116767. https://doi.org/10.1016/j.watres.2020.116767.

Smith TP, Clegg T, Ransome E *et al.* High-throughput characterization of bacterial responses to complex mixtures of chemical pollutants. *Nat Microbiol* 2024;**9**:938–48, https://doi.org/10.1038/s41564-024-01626-9.

Smith TP, Mombrikotb S, Ransome E *et al.* Latent functional diversity may accelerate microbial community responses to temperature fluctuations. *eLife* 2022;**11**:e80867. https://doi.org/10.7554/eLife.80867.

Sperlea T, Kreuder N, Beisser D *et al.* Quantification of the covariation of lake microbiomes and environmental variables using a machine learning-based framework. *Mol Ecol* 2021;**30**:2131–44. https://doi.org/10.1111/mec.15872.

Stehle S, Schulz R. Agricultural insecticides threaten surface waters at the global scale. *Proc Natl Acad Sci USA* 2015;**112**:5750–5. https://doi.org/10.1073/pnas.1500232112.

Thompson MSA, Bankier C, Bell T *et al.* Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biol* 2016;**61**:2037–50. https://doi.org/10.1111/fwb.12676.

Turschwell MP, Connolly SR, Schafer RB *et al.* Interactive effects of multiple stressors vary with consumer interactions, stressor dynamics and magnitude. *Ecol Lett* 2022;**25**:1483–96, https://doi.org/10.1111/ele.14013.

Tyler AD, McAllister J, Stapleton H *et al.* Field-based detection of bacteria using nanopore sequencing: Method evaluation for biothreat detection in complex samples. *PLoS One* 2023;**18**:1–8. https://doi.org/10.1371/journal.pone.0295028.

van Moorsel SJ, Thébault E, Radchuk V *et al.* Predicting effects of multiple interacting global change drivers across trophic levels. *Global Change Biol* 2023;**29**:1223–38. https://doi.org/10.1111/gcb.16548.

Wang B, Zheng X, Zhang H *et al.* Keystone taxa of water microbiome respond to environmental quality and predict water contamination. *Environ Res* 2020;**187**:109666. https://doi.org/10.1016/j.envres.2020.109666.

Webb CO, Ackerly DD, McPeek MA *et al.* Phylogenies and community ecology. *Annu Rev Ecol Syst* 2002;**33**:475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448.

Weiss S, Xu ZZ, Peddada S *et al.* Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 2017;**5**:1–18. https://doi.org/10.1186/s40168-017-0237-y.

Wijaya J, Byeon H, Jung W *et al.* Machine learning modeling using microbiome data reveal microbial indicator for oil-contaminated groundwater. *J Water Process Eng* 2023;**53**:103610. https://doi.org/10.1016/j.jwpe.2023.103610.