Osburn *et al. Environmental Microbiome* (2025) 20:131
https://doi.org/10.1186/s40793-025-00790-7

Environmental Microbiome

RESEARCH **Open Access**

Relative abundances of bacterial phyla are strong indicators of community-scale microbial growth rates in soil

Ernest D. Osburn[1*], J. L. Weissman[2,3], Michael S. Strickland[4,5], Mo Bahram[6,7,8], Bram W. Stone[9] and Steven G. McBride[10]

Abstract

Background To improve our understanding of microbial systems, it is essential to refine the conceptual frameworks that connect microorganisms to their ecological functions. While trait-based approaches can provide nuanced perspectives on how microorganisms influence ecosystem processes, there is ongoing debate over the link between microbial taxonomic classifications and life history traits. Here, we integrate genomic, metagenomic, amplicon sequencing, and experimental (stable isotope probing) data to investigate the scaling of bacterial growth traits from individual taxa to complex assemblages and to identify specific taxonomic groups of soil bacteria that can be used as indicators of community-scale microbial growth.

Results Our results revealed broadly different distributions of growth rates among bacterial phyla, including significantly different mean and median rates. This, in turn, manifested in strong relationships between relative abundances of some phyla and community-scale growth rates in soil. Specifically, we calculated community weighted mean growth rates using measured growth rates of constituent taxa and found that the fast-growing taxa that had sufficient abundance and ubiquity across samples to contribute to variation in community-average growth were mostly lineages of Proteobacteria (e.g., *Sphingomonas*). As a result, the relative abundance of phylum Proteobacteria was the single strongest taxonomic predictor of community-average growth, explaining up to ~ 60% of the variation in growth rates across communities. In contrast, Verrucomicrobia were consistent indicators of slower community-average growth. These patterns were especially strong when using taxon-level growth rates measured following carbon and nitrogen additions to soil.

Conclusions Our results demonstrate that phylum relative abundances can be strong indicators of community-level bacterial growth despite the wide variation in growth rates observed within phyla. The stronger phylum-growth relationships for whole assemblages than are apparent for individual taxa are due to relative abundance-weighted trait averaging in complex assemblages, i.e., at the community scale, broad differences in growth traits among phyla become more important than variation within phyla. Overall, our results provide clarity regarding the use of bacterial taxonomic information for inferring traits, demonstrating that high taxonomic ranks can be valid indicators of microbial traits in soil provided that inferences are drawn at the appropriate scale.

---

*Correspondence:
Ernest D. Osburn
e.osburn@uky.edu

Full list of author information is available at the end of the article

Keywords Growth rate, Bacteria, Soil, Genome, Metagenome, 16S, Amplicon sequencing, Stable isotope probing, Microbial community

Background

The advent of microbiome sequencing techniques has greatly advanced our knowledge of the taxonomic composition of microbial communities, but there remains substantial debate over whether and how microbial traits and functions can be inferred from taxonomic information. On the one hand, microbiome studies are often conducted under the premise that they will improve our understanding of microbial ecosystem functions [1] and many tools have been developed to translate microbial taxonomy/phylogeny into function [2–5]. In contrast, other studies are highly critical of inferring traits and functions from microbial taxonomy, claiming that problems such as horizontal gene transfer and poor phylogenetic conservation of traits will obscure relationships [5–9]. As microbiome analyses become increasingly integrated into environmental research, we need clear guidance on how microbial datasets can be used to glean information on microbe-mediated ecosystem function. Here, we address the use of bacterial taxonomy for inferring a key microbial functional trait: growth rate.

Growth rate is a fundamental bacterial life history trait related to important ecosystem processes such as the decomposition of organic matter, nutrient turnover, and soil carbon (C) storage [10–14]. The importance of microbial growth is highlighted by its centrality within popular life history concepts in microbial ecology, including the copiotroph–oligotroph dichotomy [15, 16] and the growth yield–resource acquisition–stress tolerance (Y–A–S) framework [10]. The importance of microbial growth is further underscored by emerging efforts to incorporate microbial traits, including growth rates, into process-based ecosystem models [17–20]. Recently, research on microbial growth has been facilitated by the emergence of high-resolution methods for estimating microbial growth rates, including (meta)genomic methods that estimate growth potential [15, 21] and stable isotope probing (SIP) methods that measure in situ growth [22]. However, while powerful, these methods are challenging and resource-intensive and may not be feasible for all studies to implement. Thus, it is reasonable to question whether reliable indicators of microbial growth could be inferred from taxonomic identity or community-level taxonomic composition. Some prior work has suggested that such inferences may indeed be possible, demonstrating that high bacterial taxonomic ranks (e.g., phyla) have broadly different ecological traits, including growth rates [15, 23], though the degree of phylogenetic conservation varies depending on the specific trait in question [24, 25]. That conclusion was supported by experimental work demonstrating distinct growth responses of bacterial phyla to sucrose additions [26]. In subsequent years, many studies have used these phylum-level trait categorizations to draw inferences regarding microbial growth traits in soils [27–30]. The idea that high taxonomic ranks could be used as microbial growth indicators is attractive given that complex environmental microbiomes, e.g., soil communities, are often poorly characterized at lower taxonomic levels but have reliable high-level taxonomic information.

Though attractive, the aforementioned trait categorizations have also been criticized in recent years, with several studies claiming that life history categorizations of bacterial phyla are inaccurate and/or do not reflect the substantial variation in traits within phyla [7, 9, 31]. These criticisms are supported by recent SIP studies, which demonstrated that phylum affiliation is a poor predictor of the growth rates of individual bacterial taxa in soil [16, 32]. These new studies have led to uncertainty over whether and how taxonomic information can be used as indicators of microbial ecological traits. However, while convincing, the SIP studies also introduce a new conceptual difficulty: inferences in the SIP studies are made at the level of individual taxa, whereas many studies that use life history designations of high taxonomic ranks are doing so to gain insight into the functioning of whole microbial communities [27–29]. This mismatch of ecological levels of organization highlights the need to explore the scaling of growth rates from individual taxa to complex assemblages. A first step at accomplishing this would be to combine taxon-level growth data with information about the occurrence and relative abundance patterns of those taxa across distributed natural assemblages. This would facilitate identification of taxonomic groups that are likely to influence variation in community-scale microbial growth, thus yielding important insights regarding the patterns and drivers of microbial growth at varying ecological scales.

We hypothesize that growth trait designations of bacterial phyla can be valid at the community level even when they are not reliable for individual taxa. Justification of our hypothesis partly comes from the taxon-level growth data itself, e.g., the observation that fast-growing taxa are only observed in phyla historically considered copiotrophic (e.g., Proteobacteria, Firmicutes) and are sometimes highly abundant [16]. Further justification of our hypothesis comes from prior studies that reported broad differences in growth rate distributions among phyla [15, 33]. This observation suggests that even if there is high variation in growth rates within phyla, trait averaging across taxa in diverse assemblages could lead to predictable relationships between growth rate and high taxonomic ranks in whole communities [34]. Though investigation of scaling effects and emergent properties in microbial systems is a growing area of research [1, 34–36], no studies to our knowledge have assessed how ecological scale influences the relationships between bacterial taxonomy and growth rates. To address these questions, we integrate a collection of datasets on bacterial growth rates to gain insight into the taxonomic groups that could be used as reliable indicators of microbial growth in soil. Specifically, we extend prior analyses to explore how bacterial taxonomy-growth relationships vary at different ecological scales, i.e., taxon versus community levels. By considering the role of ecological scale in microbial trait designations, our work is an important addition to microbial life history theory, contributing to our goal of providing clearer guidance on relating microbial taxonomic profiles to the important ecological traits of those communities.

Materials and methods

Bacterial growth rates at the taxon level

Our first goal was to revisit patterns in taxon-level growth traits of soil bacteria derived from both genomic and experimental (i.e., SIP) estimates of growth. For these analyses, we considered organisms within six bacterial phyla (Acidobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria, Verrucomicrobia), as these groups typically comprise the majority of soil bacterial communities and were present in all datasets we used. Furthermore, these are the major groups that have previously been categorized as fast versus slow-growing [26, 31] and have therefore been the focus of more recent work investigating the growth traits of individual bacterial lineages [16]. For the genomic analysis, we used growth rates of soil bacteria found in the EGGO database, a compilation of minimum doubling time predictions for publicly available genome assemblies using gRodon [15]. The gRodon model is based on patterns of codon usage bias, the strongest genomic indicator of bacterial growth potential [21]. The updated predictions we used also included a correction for optimal growth temperature [33]. Details of the gRodon model mechanics are provided elsewhere [15, 33, 37]. We selected soil bacteria from the database by filtering the genomes only to include those representing soil organisms as indicated by the Joint Genome Institute's GOLD database [38]. For intercomparison with SIP data that were generated by 16S rRNA gene sequencing (below), and for which taxonomic assignments are unreliable below the genus level, we aggregated the gRodon estimates at the genus level (207 genera from 10,096 genomes). For consistency with the SIP data, we calculated maximum growth rates from the gRodon minimum doubling times (in h) using an exponential growth equation, where growth rate (h^{-1}) = ln(2)/h.

We compared the genomic predictions of growth rates to those empirically measured in soil bacteria via quantitative stable isotope probing (qSIP) [22, 39]. The qSIP dataset we used has been described previously [32, 40, 41]. Briefly, soils for the qSIP experiment were collected from four ecosystems in northern Arizona and incubated under three different experimental treatments: C addition (1 mg glucose-C g soil^{-1}), C + N addition (glucose + 0.1 mg $[NH_4]_2SO_4$-N g soil^{-1}), and no additions (control). All treatments received ^{18}O water to enable isotopic labelling of microbial DNA. After the incubation, DNA was extracted, subjected to CsCl density gradient centrifugation, and then split into ~ 20 density fractions of ~ 150 µL. 16S rRNA genes (515F/806R primers) were quantified in the fractions using qPCR and sequenced on an Illumina MiSeq (raw sequences: PRJNA521534). Raw sequences were processed using DADA2 [42] and taxonomy was assigned to the amplicon sequence variants (ASVs) using a naïve-bayes classifier [43] trained on the SILVA database (version 132) [44]. The molecular weight of the DNA of each ASV was calculated using the qSIP equations [22] and the DNA molecular weights were used to estimate growth rates of each taxon using an exponential growth model [39, 45]. This ultimately yielded 2277 ASVs with growth rate estimates. To facilitate comparison with the genomic data, we calculated the average growth rates of the 279 genera present. We did this for each experimental treatment separately (control, C, C + N) and for all treatments averaged together. Full details are provided in the Supplementary Methods section.

Bacterial growth rates at the community level

Our next goal was to extend the analysis to the community level to investigate the scaling of microbial growth and to identify potential indicator taxa of community-aggregated growth. We used two complementary approaches to addressing this question. First, as a community-level analogue of the genomic growth rate estimates, we used a previously described global dataset of soil metagenomes [46] (raw sequences: PRJEB18701) to estimate the growth potential of soil bacterial communities. As described previously, we predicted community average maximum growth rates for the metagenomes using gRodon2 [14], which considers codon usage bias in highly expressed ribosomal protein genes and the relative coverages of genes in the community [15, 37]. Full details are provided in the Supplementary Methods. To relate the growth rates to the taxonomic composition of the communities, we used a dataset of 16S rRNA gene amplicons generated from the same soils [46]. The amplicons were produced using the 515F/806R primer set and sequenced on an Illumina MiSeq (raw sequences: PRJEB19856). The raw 16S sequences were processed and classified as described for the qSIP data above.

Our second approach to the community-level analysis was to identify the qSIP ASVs in external community-level data: the Earth Microbiome Project (EMP) 16S rRNA gene amplicon dataset [47] (https://ftp.microbio.me/emp/release1/otu_tables/deblur/). The rationale for using the EMP dataset instead of the original qSIP dataset for this analysis was that the qSIP study featured few whole communities, i.e., the original experimental replicates. The EMP dataset provided a large number of distinct communities, which allowed us to draw on broader patterns of bacterial taxonomic composition. This approach was intended to add ecological context to the qSIP data by simultaneously considering both the measured growth rates of the taxa as well as the patterns of occurrence and relative abundance of the taxa across distributed soil assemblages. We accomplished this by calculating community weighted mean growth rates of the EMP assemblages using the taxon-level qSIP growth rates, thus allowing us to identify specific taxonomic groups that strongly contributed to (and are therefore indicators of) variation in community growth across samples.

To identify the qSIP ASVs in the EMP communities, we first subset the EMP data to only include soil communities and then used VSEARCH [48] to identify 100% identity alignment matches between the EMP ASVs and the qSIP ASVs. We used the longest available EMP ASVs (the 150 bp release) to perform the alignments. Of the 2277 qSIP ASVs, 1574 had exact sequence matches in the EMP soil communities. We interpret these to be the same (or very similar) taxa as those present in the qSIP dataset that should have similar growth traits. We then subset the EMP communities to only include the qSIP ASVs and removed samples with fewer than 5000 sequences remaining. Relative abundances of the ASVs in the community subsets were then calculated as the sequence count of each ASV divided by the rarefied sequence depth (5000). The resulting community subsets had microbial richness values ranging from 120 to 600 ASVs per sample. Next, following prior approaches of scaling up taxon-level processes [49, 50], we calculated a community weighted mean growth rate for each EMP community subset by multiplying the relative abundance of each ASV by the qSIP-estimated growth rate of that ASV and then summing across all of the ASVs in each community subset. We did this separately using growth rates from the substrate/nutrient conditions of the three qSIP experimental treatments (control, C, and C + N). Prior to further analysis, we removed one EMP study from the dataset (study ID '864') as this study contained unusual bacterial communities with many extreme outliers in terms of community average growth rates (Supplementary Fig. S1). The final datasets analyzed for community-average growth included 701, 698, and 636 EMP samples with 1208, 1158, and 867 qSIP ASVs for the control, C, and C + N, treatments, respectively. These differences are due to differences in the specific ASVs present in the different qSIP treatments, which, in turn, influenced the number of EMP samples meeting the sequence depth threshold in each case.

A limitation of our qSIP-EMP approach is that the qSIP ASVs comprise only ~24% of the EMP sequence reads and ~2% of the total ASVs within the EMP soil samples. Therefore, our calculated community growth rates should not be interpreted as representing the original EMP communities, since EMP ASVs without growth rate information were removed prior to calculation. However, while the community subsets are not fully representative of the original EMP communities at the ASV level, we did observe strong correlations in the relative abundances of higher taxonomic groups between the community subsets and the original full EMP communities (Supplementary Fig. S2). This indicates that the community subsets are at least representative of the original EMP communities at higher taxonomic levels. In any case, the community subsets could alternatively be viewed as somewhat hypothetical assemblages that are distinct from the original EMP communities but that still exhibit wide natural variation in the relative abundances of taxa for which growth has been measured. As such, analysis of the community subsets is appropriate for accomplishing the central goal of our study, i.e., exploring the patterns by which growth rates of individual taxa might scale up to complex soil assemblages.

Statistical analyses

Statistical analyses were conducted in R [51]. To test the hypothesis that the phyla exhibit distinct distributions of growth rates, we used multi-sample Anderson–Darling tests [52]. To compare the phyla in terms of their median and mean growth rates, we used Mood's median test and a generalized linear model (gamma distribution, log-link function) [53], respectively. To analyze relationships between variables, we used Pearson correlation analysis. In cases where multiple correlation coefficients were calculated for an analysis, we used the Benjamini–Hochberg method to correct the p values for false discovery rates [54]. To identify the specific taxonomic groups that most strongly contributed to variation in the calculated community-average growth values across samples, we used random forest regression [55]. For the regression models, we considered taxon relative abundances as candidate predictors of community-average growth. We quantified the "importance" of taxa in contributing to variation in community growth by determining the increase in model mean squared error when each taxon was randomly shuffled across the dataset. Note that in the correlation analyses relating to microbial growth and in the regression models, the x and y variables are not fully independent since the taxon relative abundances (predictors) and the community growth rates (response) were both determined using the original ASV sequences. Because of this, the statistical results we present are interpreted only for the purpose of revealing the specific taxonomic groups that were the strongest contributors to calculated community growth. All data and analysis scripts are provided in Figshare: https://doi.org/10.6084/m9.figshare.28264322.v1 [56].

Results and discussion

Taxon-level patterns in bacterial growth traits

The genomic and qSIP methods produced similar patterns of growth rates among the six bacterial phyla (Fig. 1A, B). However, the magnitude of the values was very different between the methods (Fig. 1C). Not surprisingly, the maximum rates predicted from genomes were much higher than the qSIP-measured rates (Fig. 1C). This reflects the difference between growth potential and actual growth and also likely reflects a bias towards higher growth in cultured isolates with sequenced genomes [15]. Similar to prior studies [16], both methods revealed wide variation in growth rates within most bacterial phyla (Fig. 1A, B). These patterns held when examining the average growth rate of genera across all qSIP treatments (Fig. 1B) and for the three qSIP treatments analyzed separately (Supplementary Figs. S3, S4, and S5).

Despite the wide variation in growth rates observed within phyla, we also observed differences in the growth rate distributions among phyla (Fig. 1A, B) ($A^2 = 16.3$ and 18.8 for gRodon and qSIP, respectively, both $P < 0.001$). Those differences included different mean rates ($\chi^2 = 37.1$ and 29.9 for gRodon and qSIP, respectively, both GLM

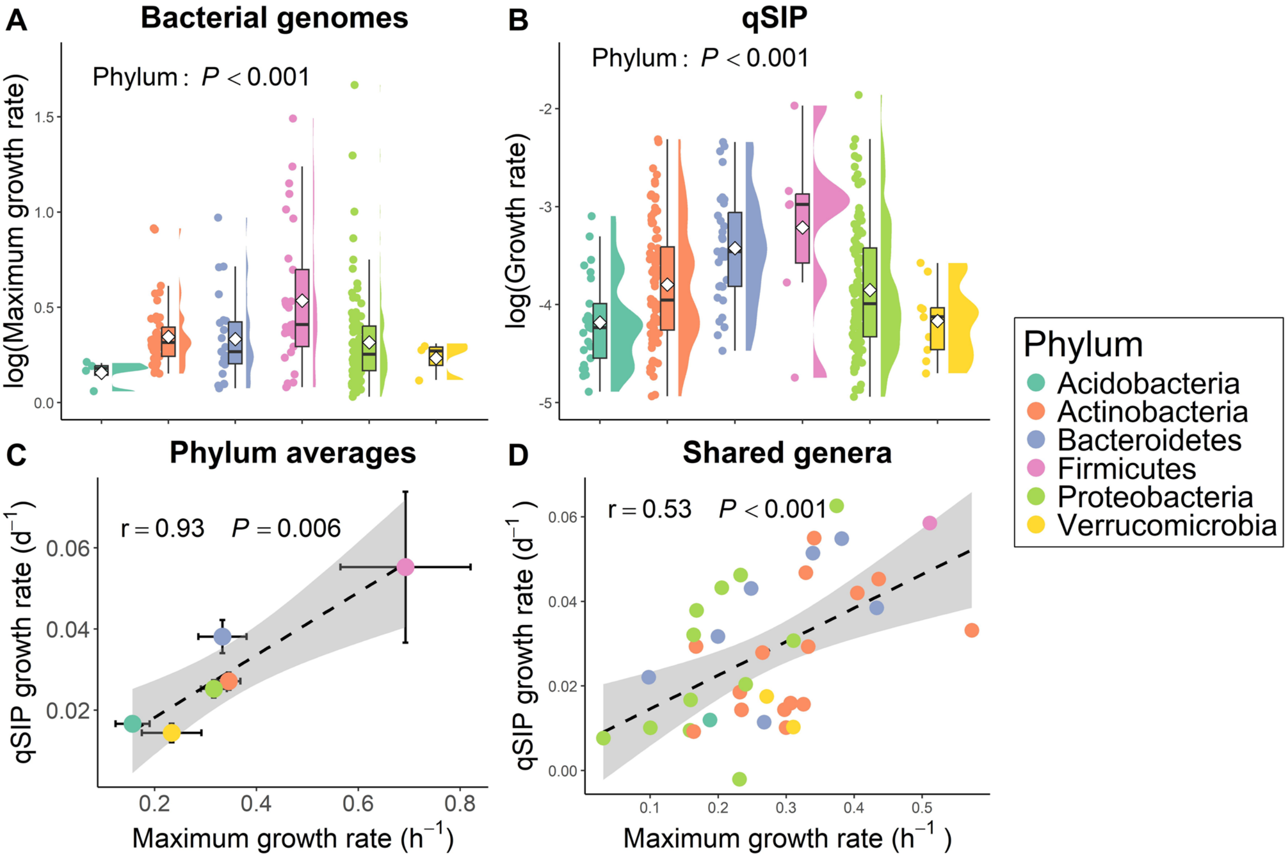

Fig. 1 Growth rates within and among the major soil bacterial phyla. Shown are distributions of maximum growth rates of 207 genera of soil bacteria from the gRodon model (**A**) and measured growth rates of 276 bacterial genera from quantitative stable isotope probing (qSIP) (**B**). Panel **B** shows the overall average growth rates across all qSIP experimental treatments. The growth distributions separated by qSIP treatment are provided on Supplementary Figs. S3, S4, and S5. *P* values on panels (**A**) and (**B**) indicate statistically different mean growth rates among phyla (diamond symbols on the plots) and are from generalized linear models. Panel **C** shows the correlation between the two growth rate methods in terms of their phylum-level mean growth rates. Error bars on (**C**) indicate one standard error of the mean and the correlation coefficient (r) and *P* value provided are from Pearson correlation analysis. Panel **D** shows the correlation of the growth rates of the genera shared between the two datasets and the correlation coefficient (r) and *P* value provided are from Pearson correlation analysis

$P < 0.001$) and different median rates among phyla (both Mood's median $P < 0.001$) (Fig. 1A, B). Further, the general pattern of mean growth rates among phyla was consistent between the qSIP and genomic methods (Fig. 1C). For both methods, Firmicutes genera exhibited the highest growth rates on average, Verrucomicrobia and Acidobacteria exhibited the lowest growth rates, and the other phyla were intermediate (Fig. 1C). Specifically, Firmicutes genera had ~ fourfold and ~ threefold higher growth rates than Acidobacteria genera on average and ~ threefold and fourfold higher growth rates than Verrucomicrobia genera on average for the genomic and qSIP methods, respectively. Similar patterns among phyla have been reported in studies focused on marine and gut bacteria [15, 33]. These results also support prior work identifying Firmicutes as the most potentially copiotrophic bacterial phylum in soil [16]. It is worth noting, however, that some phyla were lacking in representative genera—very few soil Acidobacteria and Verrucomicrobia genera were within the EGGO database and few Firmicutes genera were in the qSIP dataset. However, the strong agreement between the two methods across all phyla shows broadly conserved growth patterns and is a promising demonstration of our ability to connect in situ growth rates with intrinsic genomic traits. Furthermore, while only 40 genera overlapped between the two datasets, the genomic and qSIP growth rates of those genera were also significantly positively correlated (Fig. 1D). We conclude from these results that the patterns among taxa support the premise of our hypothesis, i.e., that the major phyla of soil bacteria are broadly different in terms of observed and potential growth rates.

This conclusion may seem unsurprising since it has long been known that certain phyla have many fast-growing members, particularly common laboratory organisms, e.g., *Bacillus* (Firmicutes), *Escherichia* (Proteobacteria), etc. However, recent work has also shown that many environmental strains within those phyla do not conform to the trait designations that have been historically assumed for the phyla [16]. Those new studies emphasize high within-phylum variation in growth rates [16] and demonstrate that phylum affiliation is a poor predictor of traits of individual taxa [16, 32]. Here, we re-emphasize the broad among-phylum differences in growth rates to explore the possibility that among-phylum differences take precedence over within-phylum variation when growth rates are scaled to the community level.

Community-aggregated growth rates of soil bacteria

Since the community level is the ecological scale at which life history designations of soil bacterial phyla have often been interpreted [26–29], we sought to explore the scaling of bacterial growth from individual taxa to diverse assemblages. Growth at the community scale is important since investigators interested in ecosystem processes such as nutrient turnover or C sequestration may be more interested in the aggregate growth of entire bacterial communities rather than individual taxa [1, 11]. Specifically, our goal was to identify taxa that might be used as reliable indicators of community-average growth. To investigate the scaling of bacterial growth, we applied qSIP-measured growth rates of bacterial taxa to those same organisms in EMP soil communities and calculated community weighted mean growth rates for the EMP assemblages. The specific lineages across all taxonomic levels that most strongly contributed to variation in community average growth are shown in Supplementary Table S1. In general, we found that the weighted average growth of communities was consistently positively associated with the relative abundance of Proteobacteria in the assemblages and negatively associated with Verrucomicrobia (Fig. 2A). These are among the relationships that would be expected since Verrucomicrobia genera had among the lowest growth rates on average while Proteobacteria genera had intermediate growth rates on average but with some very high growth potential taxa (Fig. 1). When using growth rates measured under C and C + N addition, the relationships became much stronger (Fig. 2B, C). The strength of those relationships is striking and suggests that some phyla can be strong indicators of the growth status of complex soil communities, particularly under substrate and nutrient addition conditions. These results add important ecological context to the taxon-level growth rate data, revealing that the taxa that are sufficiently ubiquitous and abundant enough to drive variation in community-average growth across distributed soil samples tend to be members of the Proteobacteria and Verrucomicrobia phyla.

Community growth estimates also showed that the relative abundances of other phyla did not consistently contribute to variation in community-average growth (Fig. 2B–D). For example, the generally slow-growing Acidobacteria (Fig. 1) were not consistently associated with slower community growth (Fig. 2). Similarly, contrary to what might be expected, fast-growing Firmicutes (Fig. 1) did not consistently contribute to faster community growth (Fig. 2). That said, we did detect a significant positive association between Firmicutes and community growth under C addition (Fig. 2B), which was attributed to a small cluster of samples that exhibited unusually high relative abundance of Firmicutes and also had high community-average growth. This relationship is expected from the genus-level growth distributions and suggests that Firmicutes can be functionally important in some communities. This conclusion is supported by revisiting the EMP study we initially removed from the analysis: study 864. This study was characterized by

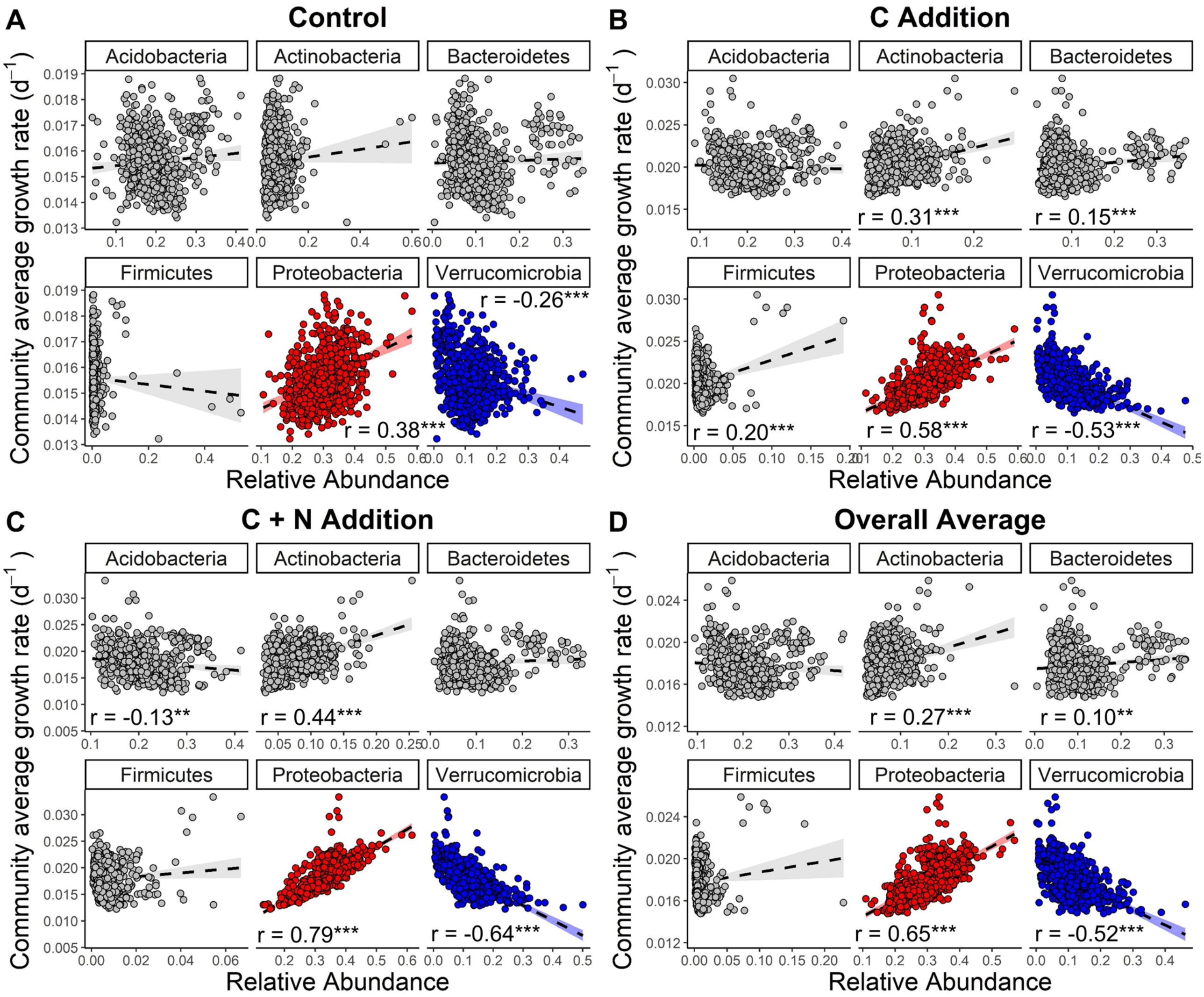

Fig. 2 Relationships between phylum relative abundances and calculated community-average growth rates for the subset Earth Microbiome Project (EMP) soil communities. EMP communities were filtered to only include the qSIP ASVs prior to calculation of the community-average rates. Shown are calculated community growth rates for the control (no addition) (**A**), the glucose-C addition treatment (**B**), the C + N addition treatment (**C**) and the overall average growth across all experimental treatments (**D**). Correlation coefficients are from Pearson correlation analysis. Asterisks indicate significant correlations at the following levels: **$P < 0.01$, ***$P < 0.001$. Colors indicate consistent contributors to fast growth (red) and slow growth (blue), i.e., Proteobacteria and Verrucomicrobia, respectively

atypical soil communities that were often dominated by Firmicutes (sometimes > 80% of sequences), with correspondingly high estimated growth rates in those cases (Supplementary Fig. S6). Within that study, the relative abundance of Firmicutes was strongly positively associated with growth (Supplementary Fig. S6). We conclude from this that Firmicutes should be considered important community members in soils where they are found to be abundant (e.g., > 10% relative abundance). However, they may not be consistently abundant enough to be a reliable indicator of community-level growth. Proteobacteria and Verrucomicrobia, on the other hand, do appear to be consistent indicators of community growth, lending support to prior work that has interpreted those phyla as such [28–30]. This conclusion is supported by the dataset of maximum growth rates predicted from global soil metagenomes using the gRodon2 model [14], which revealed community-aggregated growth to be correlated only with the relative abundances of Proteobacteria and Verrucomicrobia (Supplementary Fig. S7). However, the correlations in the metagenome data were weaker (Supplementary Fig. S7). This may be because, unlike the qSIP study, those samples were not enriched in rapidly growing taxa due to C and N addition, i.e., the conditions under which the influence of high growth potential community members becomes particularly evident [41] (Fig. 2). Therefore, variation in maximum growth might have been somewhat constrained in the metagenome dataset, potentially preventing us from observing stronger relationships. Future experimental studies that involve soil C and N additions could also incorporate the metagenomic maximum growth estimates to resolve this question.

We sought additional support for these results by repeating the analysis using phylum relative abundances from the original full EMP communities to determine whether the observed taxon-growth correlations held for the original EMP communities and not just for the subsets comprised only of the qSIP ASVs. Phylum relative abundances in the full EMP communities had essentially the same relationships with growth compared with the subsets (Supplementary Fig. S8), supporting our finding that the community subsets are representative of the full EMP communities at high taxonomic levels (Supplementary Fig. S2). We also observed essentially the same relationships after correcting the ASV counts for 16S rRNA gene copy number (Supplementary Fig. S9), demonstrating that variation in 16S copy number was not a confounding factor in our analysis and providing additional support for our findings.

Though we observed consistent patterns across several iterations of our analysis, we do suggest some caution in interpreting phylum relative abundances as indicators of bacterial growth given that the community-average rates we report here are only calculated and not actual measured growth rates for those communities. Actual rates could differ from our calculated rates for various reasons, e.g., varying responses of taxa to differences in soil abiotic factors or ecological factors such as density dependence, interspecific interactions, etc. [39]. Although the qSIP growth rates applied to the EMP organisms may not represent their actual growth in all cases, the qSIP rates were generated across climatically divergent sites (desert grassland, arid shrubland, dry pine forest, and temperate mixed conifer forest) and so should be applicable to many systems [32]. In addition, the qSIP growth rates appear to have a genomic basis (Fig. 1), so we would expect strong phylogenetic conservation of growth traits across environments [57], especially when averaging across hundreds of ASVs in a community. This should be particularly true with C + N additions, which would eliminate some of the site-specific growth constraints and where the rates are likely more reflective of genomic growth potential. For all these reasons, it is likely that the qSIP rates provide reasonable estimates of the growth potential of the EMP soil communities.

An additional caveat is that our study does not capture the highly dynamic nature of microbial growth in soils. Because the growth estimates we used are based on inherent genomic traits or on qSIP measurements at a single time point, temporal patterns in soil microbial growth are not reflected in our results. For example, the qSIP measurements do not consider temporal changes in environmental conditions (e.g., soil moisture, resource inputs), which can cause rapid shifts in taxonomic composition and activity/dormancy patterns among taxa in soil communities [58–60], all of which would strongly influence microbial growth rates. Another limitation is that the qSIP approach likely yields growth rate estimates that are varyingly accurate among microbial life strategies, e.g., qSIP likely underestimates the growth rates of boom-bust adapted organisms with exceptionally high growth potential that turn over before the end of the 1-week incubation (e.g., some Firmicutes). Due to these sampling and methodological limitations, our community-average growth estimates (and the original individual estimates) need to be interpreted with caution. Indeed, while our work demonstrates that the estimated growth of a microbial community at a snapshot in time is strongly related to broad patterns in the taxonomic composition of that community, we emphasize that our results should not be interpreted as static properties of those soils under all conditions. We encourage future studies on microbial growth to include sampling at varying time scales to improve our understanding of the temporal dynamics of microbial growth at individual and community levels.

It is also worth revisiting here the issue of non-independence of the calculated community growth rates and the taxon relative abundances. Because both variables were determined using the original ASV sequences, they are inherently mathematically related. Indeed, in some ways the results of our analyses are self-evident, i.e., communities with more fast-growing constituent taxa will obviously have faster community-average growth. However, it is also important to note that taxonomic information per se was not used in the calculation of community-average growth. Thus, it is still useful to investigate the taxonomic affiliations of the organisms that were important in contributing to variation in calculated community growth. Therefore, while the relationships need to be interpreted carefully, they are ecologically relevant in that they reveal the taxonomic groups that are likely to influence the aggregate growth of whole bacterial assemblages, owing to the measured growth rates of those taxa in combination with the observed occurrence and relative abundance patterns of those taxa across actual communities.

Overall, our results suggest that some phyla may be informative of the growth rates of whole bacterial communities. It is possible that this is also true for other aspects of bacterial life history, e.g., resource acquisition or stress tolerance [10], which should be evaluated by future studies. While prior studies have criticized the interpretation of phylum relative abundances in microbial ecology research [7, 9, 31], we offer a clarified position by emphasizing ecological scale as a critical consideration

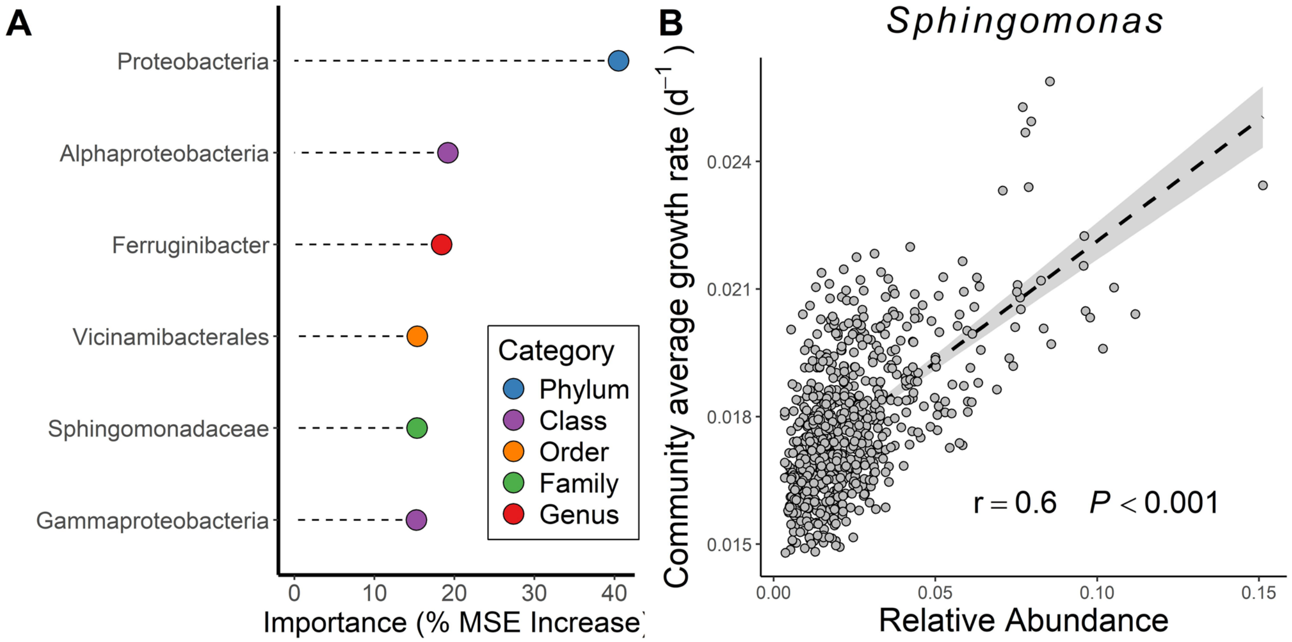

Fig. 3 Regression models showing the strongest taxonomic indicators of community average bacterial growth. Panel **A** shows the top importance individual taxa when scaled relative abundances of all taxa were combined into a single random forest regression model. 'Importance' in (**A**) was calculated as the increase in model mean squared error (MSE) when each respective predictor was randomly shuffled across the dataset. Panel **B** shows the correlation between community average growth and the relative abundance of *Sphingomonas* from the Alpharoteobacteria lineage Sphingomonadales, the most consistently important lower taxon across all models (Supplementary Table S1). The correlation coefficient and *P* value shown are from Pearson correlation analysis

when applying phylum-level trait assignments. Phylum-level taxonomy should not be used to draw firm conclusions about the traits of any particular taxon, given the high variation in trait values within phyla (Fig. 1) [16]. However, at the community level, not every organism belonging to a particular phylum needs to exhibit the traits assigned to the phylum for the relative abundance of that phylum to reliably indicate community function. As long as the traits in question (e.g., rapid growth) only appear within abundant and ubiquitous members of specific phyla, it seems reasonable to assume that the relative abundances of those phyla will accurately correlate with trait values across large collections of communities. Our analysis suggests that this assumption holds for two phyla, Proteobacteria and Verrucomicrobia (Fig. 2), and that the role of those focal phyla should be more thoroughly evaluated by future work. This conclusion could find practical use in the field of ecosystem modelling, where there is great interest in incorporating microbial traits into model structures [17–20]. While it may not be feasible to incorporate trait information of hundreds of constituent taxa into models, it may be tractable to incorporate one or a few phyla that reliably correlate with community-level trait values.

Indicator taxa of community-aggregated growth

In addition to correlating community-level growth with phylum relative abundances, we sought to quantitatively assess the contributions of taxa to the observed variation in community growth rates across samples. To do this, we used random forest regression, where models were constructed with different levels of taxonomic resolution and the relative abundances of the taxa were used as the predictor features in the models. For this analysis, we used the average growth across all qSIP treatments (shown in Fig. 2D), as this dataset contains the largest number of samples and taxa and encompasses all the experimental growth conditions assessed. As stated previously, since the response and predictors in these models are not fully independent, we only use the models to aid in identification of the specific taxa that were the strongest contributors to calculated community growth. We first evaluated an overall model using re-scaled relative abundances of all taxa at all taxonomic levels as candidate predictors. In this model, phylum Proteobacteria emerged as the single strongest predictor of community growth, with more than two-fold higher importance than any other individual taxon across all taxonomic levels (Fig. 3A). This reinforces our conclusion that high taxonomic groups can be strong indicators of the growth of complex bacterial communities. Indeed, comparison of the regression models for different taxonomic levels revealed that additional taxonomic information beyond the class level had negligible additional benefit in accounting for variation in community growth (Supplementary Table S1). This indicates that when considering whole microbial communities, the class level may represent a useful compromise between the improved trait information of finer taxonomic levels versus the simplicity and scaling benefits of broader taxonomic levels. In all, these results further underscore the potential utility of high taxonomic levels in the context of microbial community growth.

Taken together, these results suggest that high taxonomic resolution may not be important for predicting community-level growth across large and distributed microbial datasets. In contrast, prior findings from the same qSIP experiment demonstrated that predicting growth rates of individual taxa requires greater taxonomic resolution than phylum-level affiliations [16, 32]. These differences between taxon vs. community-level patterns emerge as a product of the broad among-phylum differences in growth (Fig. 1) in combination with variation in the relative abundances of taxa in complex assemblages. In other words, relative abundance-weighted trait averaging results in emergent predictability of growth [34], where phylum-growth relationships become stronger for assemblages than they are for individual taxa. Put simply, when scaling growth to the community level, broad among-phylum differences become more important than within-phylum variation. Indeed, while it is important to investigate the complex variation in the growth traits of individual taxa, when one additionally considers the observed occurrence and relative abundance patterns of those taxa across natural assemblages, the importance of broad microbial taxonomic groups (i.e., Proteobacteria) to community-level growth becomes clear. This is despite the fact that many individual Proteobacteria actually grow slowly and the fact that taxa in other phyla (e.g., Firmicutes) generally grow faster on an individual basis (Fig. 1). We expect that this observation will be particularly useful for outside soil microbiome datasets where many constituent taxa will be unclassified at fine taxonomic levels or members of poorly characterized lower taxonomic groups with no a priori knowledge of their growth potential.

Nevertheless, it is still useful to identify the individual orders, families, or genera that may also be strong indicators of community growth. The most consistently important lower taxa across the different random forest models were members of the Alphaproteobacterial lineage Sphingomonadales (Supplementary Table S1), which positively contributed to community growth (Fig. 3B). The highest importance Verrucomicrobia genus was *Candidatus* Udaeobacter, which was negatively associated with community average growth (Supplementary Fig. S10). It is not surprising to find *Candidatus* Udaeobacter to be associated with slow growth given its auxotrophic lifestyle [61]. Other important taxa included the Bacteroidetes genus *Ferruginibacter* (Fig. 3A), which was positively associated with community growth (Supplementary Fig. S11) and the Acidobacteria order Vicinamibacterales (Fig. 3C), which was negatively associated with community growth (Supplementary Fig. S12). Interestingly, however, none of these taxa exhibited spectacularly high or low growth rates themselves. For example, under C + N addition, *Sphingomonas* was at the 73rd percentile for growth rate while *Candidatus* Udaeobacter was at the 35th percentile. The contributions of these taxa to community growth, therefore, are due not only to their traits per se but also to their ubiquity and high variation in relative abundance across communities. In contrast, organisms in the extreme upper and lower tails of the growth distributions generally had very low relative abundance, and thus did not significantly contribute to community-level growth. This indicates that extremely high and low growth potential may represent “fringe” ecological strategies in soil that rarely confer high abundance and may not substantially influence broad processes such as overall community growth. In any case, it is clear from our results that even trait values only slightly different from the average can manifest changes in broad community-level functions provided that the taxa possessing those traits are highly abundant. As such, future studies on microbial growth should focus on taxa that are common and abundant across communities [62] but do not currently have growth rate estimates. Indeed, though our analysis included many widely distributed and abundant taxa, only a minority of the organisms present in the EMP soils had measured growth rates in the qSIP growth rate data we used. Therefore, we expect that additional taxonomic groups that are important in contributing to community-scale growth will be revealed as research on individual microbial growth rates continues to advance.

Conclusions

To improve our understanding of microbe-mediated ecosystem functioning, it is imperative to determine how key microbial traits such as growth rate scale from individuals to assemblages. By evaluating how taxonomy-growth relationships vary across ecological scales, our study sought to identify specific taxa that can be used as indicators of the growth of complex microbial assemblages. Our work highlights the importance of ecological scale when inferring life history traits of bacterial taxa. While high-level bacterial taxonomic ranks (e.g., phyla) do not provide precise information about the growth traits of individual taxa, we demonstrate that high ranks can be strong indicators of the growth of complex bacterial assemblages. Specifically, we found that Proteobacteria were consistently positively associated with community-average growth rates whereas Verrucomicrobia were consistently negatively associated with community-average growth. These patterns manifested because the most abundant and ubiquitous members of those phyla in soils exemplify the respective phylum-level trait designations, e.g., fast-growing *Sphingomonas* for Proteobacteria and slow-growing *Candidatus* Udaeobacter for Verrucomicrobia. These patterns were strongest under conditions of substrate and nutrient additions, indicating that our findings may be particularly useful in studies involving resource gradients and/or manipulations. Overall, our results demonstrate that high taxonomic ranks can be accurate indicators of bacterial life history traits and that the ecological relevance of those categories only emerges at the community level.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40793-025-00790-7.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
EDO conducted initial data analyses, wrote the first draft of the manuscript and prepared the figures. JLW, BWS, and SGM provided guidance and suggestions on additional data analyses. All authors contributed to reviewing and editing the manuscript.

Funding
MSS was supported by National Science Foundation award 2131837. SGM was supported by the United States Department of Agriculture award 2020-67034-41310. BWS is grateful for support from the Linus Pauling Distinguished Postdoctoral Fellowship program through Pacific Northwest National Laboratory.

Data availability
All data and analysis scripts are provided in Figshare: https://doi.org/10.6084/m9.figshare.28264322.v1. The raw sequence data used in this study can be found at the following accession numbers: PRJEB19856, PRJEB18701, PRJNA521534. The Earth Microbiome Project data used in this manuscript can be found here: https://ftp.microbio.me/emp/release1/otu_tables/deblur/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Plant and Soil Sciences, University of Kentucky, Lexington, KY, USA
[2]Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, USA
[3]Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY, USA
[4]Department of Soil and Water Systems, University of Idaho, Moscow, ID, USA
[5]Deep Soil Ecotron, University of Idaho, Moscow, ID, USA
[6]Department of Agroecology, Aarhus University, Slagelse, Denmark
[7]Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia
[8]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden
[9]Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA
[10]Department of Biology, Radford University, Radford, VA, USA

Received: 24 June 2025 / Accepted: 16 September 2025

References

1. Hall EK, Bernhardt ES, Bier RL, Bradford MA, Boot CM, Cotner JB, et al. Understanding how microbiomes influence the systems they inhabit. Nat Microbiol. 2018;3:977–82.
2. Louca S, Parfrey LW, Doebeli M. Decoupling function and taxonomy in the global ocean microbiome. Science. 2016;353:1272–7.
3. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol. 2020;38:685–8.
4. Donhauser J, Doménech-Pascual A, Han X, Jordaan K, Ramond J-B, Frossard A, et al. Modelling soil prokaryotic traits across environments with the trait sequence database *ampliconTraits* and the R package *MicEnvMod*. Ecol Inform. 2024;83:102817.
5. Djemiel C, Maron P-A, Terrat S, Dequiedt S, Cottin A, Ranjard L. Inferring microbiota functions from taxonomic genes: a review. Gigascience. 2022;11:giab090.
6. Prosser JI. Ecosystem processes and interactions in a morass of diversity. FEMS Microbiol Ecol. 2012;81:507–19.
7. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. Nat Rev Microbiol. 2017;15:579–90.
8. Nannipieri P, Ascher-jenull J, Ceccherini MT, Pietramellara G, Renella G, Schloter M. Beyond microbial diversity for predicting soil functions: a mini review. Pedosphere. 2020;30:5–17.
9. Alteio LV, Séneca J, Canarini A, Angel R, Jansa J, Guseva K, et al. A critical perspective on interpreting amplicon sequencing data in soil ecological research. Soil Biol Biochem. 2021;160:108357.
10. Malik AA, Martiny JBH, Brodie EL, Martiny AC, Treseder KK, Allison SD. Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. ISME J. 2020;14:1–9.
11. Foley MM, Stone BWG, Caro TA, Sokol NW, Koch BJ, Blazewicz SJ, et al. Growth rate as a link between microbial diversity and soil biogeochemistry. Nat Ecol Evol. 2024. https://doi.org/10.1038/s41559-024-02520-7.
12. Cotrufo MF, Wallenstein MD, Boot CM, Denef K, Paul E. The microbial efficiency-matrix stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter? Glob Change Biol. 2013;19:988–95.
13. Malik AA, Puissant J, Goodall T, Allison SD, Griffiths RI. Soil microbial communities with greater investment in resource acquisition have lower growth yield. Soil Biol Biochem. 2019;132:36–9.
14. Osburn ED, McBride SG, Bahram M, Strickland MS. Global patterns in the growth potential of soil bacterial communities. Nat Commun. 2024;15:6881.
15. Weissman JL, Hou S, Fuhrman JA. Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. Proc Natl Acad Sci USA. 2021;118:e2016810118.
16. Stone BWG, Dijkstra P, Finley BK, Fitzpatrick R, Foley MM, Hayer M, et al. Life history strategies among soil bacteria—dichotomy for few, continuum for many. ISME J. 2023;17:1–9.
17. Morrissey EM, Kane J, Tripathi BM, Rion MSI, Hungate BA, Franklin R, et al. Carbon acquisition ecological strategies to connect soil microbial biodiversity and carbon cycling. Soil Biol Biochem. 2023;177:108893.
18. Treseder KK, Balser TC, Bradford MA, Brodie EL, Dubinsky EA, Eviner VT, et al. Integrating microbial ecology into ecosystem models: challenges and priorities. Biogeochemistry. 2012;109:7–18.
19. Wallenstein MD, Hall EK. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. Biogeochemistry. 2012;109:35–47.

20. Wan J, Crowther TW. Uniting the scales of microbial biogeochemistry with trait-based modelling. Funct Ecol. 2022;36:1457–72.
21. Vieira-Silva S, Rocha EPC. The systemic imprint of growth and its uses in ecological (meta)genomics. PLoS Genet. 2010;6:e1000808.
22. Hungate BA, Mau RL, Schwartz E, Caporaso JG, Dijkstra P, van Gestel N, et al. Quantitative microbial ecology through stable isotope probing. Appl Environ Microbiol. 2015;81:7570–81.
23. Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB, et al. The ecological coherence of high bacterial taxonomic ranks. Nat Rev Microbiol. 2010;8:523–9.
24. Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. ISME J. 2013;7:830–8.
25. Goberna M, Verdú M. Predicting microbial traits with phylogenies. ISME J. 2016;10:959–67.
26. Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. Ecology. 2007;88:1354–64.
27. Zhou Z, Wang C, Luo Y. Effects of forest degradation on microbial communities and soil carbon cycling: a global meta-analysis. Glob Ecol Biogeogr. 2018;27:110–24.
28. Osburn ED, Badgley BD, Strahm BD, Aylward FO, Barrett JE. Emergent properties of microbial communities drive accelerated biogeochemical cycling in disturbed temperate forests. Ecology. 2021;102:e03553.
29. Stephenson T, Hudiburg T, Mathias JM, Jones M, Lynch LM. Do Tasmanian devil declines impact ecosystem function? Glob Change Biol. 2024;30:e17413.
30. Osburn ED, Badgley BD, Aylward FO, Barrett JE. Historical forest disturbance mediates soil microbial community responses to drought. Environ Microbiol. 2021;23:6405–19.
31. Ho A, Lonardo D, Paolo D, Bodelier PLE. Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiol Ecol. 2017 [cited 2018 Mar 16], p 93. Available from: https://academic.oup.com/femsec/article/93/3/fix006/2937747.
32. Morrissey EM, Mau RL, Hayer M, Liu X-JA, Schwartz E, Dijkstra P, et al. Evolutionary history constrains microbial traits across environmental variation. Nat Ecol Evol. 2019;3:1064–9.
33. Xu L, Zakem E, Weissman JL. Improved maximum growth rate prediction from microbial genomes by integrating phylogenetic information. bioRxiv; 2024 [cited 2024 Oct 11], p 2024.10.03.616540. Available from: https://doi.org/10.1101/2024.10.03.616540v1.
34. Moran J, Tikhonov M. Emergent predictability in microbial ecosystems [Internet]. bioRxiv; 2024 [cited 2024 Dec 3], p 2024.03.26.586882. Available from: https://doi.org/10.1101/2024.03.26.586882v2
35. van den Berg NI, Machado D, Santos S, Rocha I, Chacón J, Harcombe W, et al. Ecological modelling approaches for predicting emergent properties in microbial communities. Nat Ecol Evol. 2022;6:855–65.
36. Averill C, Werbin ZR, Atherton KF, Bhatnagar JM, Dietze MC. Soil microbiome predictability increases with spatial and taxonomic scale. Nat Ecol Evol. 2021;5:747–56.
37. Weissman JL, Peras M, Barnum TP, Fuhrman JA. Benchmarking community-wide estimates of growth potential from metagenomes using codon usage statistics. mSystems. 2022;7:e00745–22.
38. Mukherjee S, Stamatis D, Li CT, Ovchinnikova G, Bertsch J, Sundaramurthi JC, et al. Twenty-five years of Genomes OnLine Database (GOLD): data updates and new features in vol 9. Nucleic Acids Res. 2023;51:D957–63.
39. Stone BW, Blazewicz SJ, Koch BJ, Dijkstra P, Hayer M, Hofmockel KS, et al. Nutrients strengthen density dependence of per-capita growth and mortality rates in the soil bacterial community. Oecologia. 2023;201:771–82.
40. Morrissey EM, Mau RL, Schwartz E, McHugh TA, Dijkstra P, Koch BJ, et al. Bacterial carbon use plasticity, phylogenetic diversity and the priming of soil organic matter. ISME J. 2017;11:1890–9.
41. Li J, Mau RL, Dijkstra P, Koch BJ, Schwartz E, Liu X-JA, et al. Predictive genomic traits for bacterial growth in culture versus actual growth in soil. ISME J. 2019;13:2162–72.
42. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.
43. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
44. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.
45. Koch BJ, McHugh TA, Hayer M, Schwartz E, Blazewicz SJ, Dijkstra P, et al. Estimating taxon-specific population dynamics in diverse microbial communities. Ecosphere. 2018;9:e02090.
46. Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, et al. Structure and function of the global topsoil microbiome. Nature. 2018;560:233–7.
47. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature. 2017;551:457–63.
48. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584.
49. Piñeiro J, Dang C, Walkup JGV, Kuzniar T, Winslett R, Blazewicz SJ, et al. Shifts in bacterial traits under chronic nitrogen deposition align with soil processes in arbuscular, but not ectomycorrhizal-associated trees. Glob Change Biol. 2024;30:e17030.
50. Walkup JG, Morrissey EM. Scaling up taxon-specific microbial traits to predict community-level microbial activity in agricultural systems. Soil Biol Biochem. 2024;200:109622.
51. R Core Development Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2019. Available from: http://www.R-project.org.
52. Scholz FW, Stephens MA. K-sample Anderson-Darling tests. J Am Stat Assoc. 1987;82:918–24.
53. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol. 2009;24:127–35.
54. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
55. Cutler F original by LB and A, Wiener R port by AL and M. randomForest: Breiman and Cutler's random forests for classification and regression [Internet]. 2022 [cited 2022 Jun 23]. Available from: https://CRAN.R-project.org/package=randomForest.
56. Osburn ED. Data and analysis code for the community-aggregated bacterial growth study [Internet]. figshare; 2025 [cited 2025 Jan 23]. Available from: https://figshare.com/articles/dataset/Data_and_analysis_code_for_the_community-aggregated_bacterial_growth_study/28264322/1.
57. Walkup J, Dang C, Mau RL, Hayer M, Schwartz E, Stone BW, et al. The predictive power of phylogeny on growth rates in soil bacterial communities. ISME Commun. 2023;3:73.
58. Placella SA, Brodie EL, Firestone MK. Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. Proc Natl Acad Sci USA. 2012;109:10931–6.
59. Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol. 2011;9:119–30.
60. Evans SE, Wallenstein MD. Climate change alters ecological strategies of soil bacteria. Ecol Lett. 2014;17:155–64.
61. Brewer TE, Handley KM, Carini P, Gilbert JA, Fierer N. Genome reduction in an abundant and ubiquitous soil bacterium '*Candidatus* Udaeobacter copiosus'. Nat Microbiol. 2016;2:1–7.
62. Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, et al. A global atlas of the dominant bacteria found in soil. Science. 2018;359:320–5.

Publisher's note