

MINIREVIEW

Linking microbial community structure and microbial processes: an empirical and conceptual overview

Raven L. Bier^{1,*}, Emily S. Bernhardt¹, Claudia M. Boot², Emily B. Graham^{1,3,4}, Edward K. Hall², Jay T. Lennon⁵, Diana R. Nemergut^{1,4}, Brooke B. Osborne⁶, Clara Ruiz-González⁷, Joshua P. Schimel⁸, Mark P. Waldrop⁹ and Matthew D. Wallenstein²

¹Department of Biology, Duke University, Durham, NC 27708, USA, ²Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499, USA, ³Environmental Studies Program, University of Colorado, UCB 397, Boulder, CO 80309-0397, USA, ⁴Institute for Arctic and Alpine Research, University of Colorado, UCB 450, Boulder, CO 80309-0450, USA, ⁵Department of Biology, Indiana University, Bloomington, Indiana, 47405, USA, ⁶Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912, USA, ⁷Département des Sciences Biologiques, Université du Québec à Montréal, Montréal, QC, H3C3P8, Canada, ⁸Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA and ⁹US Geological Survey, Menlo Park, CA 94025-3561, USA

*Corresponding author: Department of Biology, Duke University, PO Box 90338, Durham, NC 27708, USA. Tel: +1-919-613-8215; Fax: +1-919-660-7293; E-mail: raven.bier@gmail.com

One sentence summary: Few publications reported statistically testing microbial community structure-process links; 75% of tested links were significant, though had few commonalities in which processes or structures were measured and the techniques used.

Editor: Gerard Muyzer

ABSTRACT

A major goal of microbial ecology is to identify links between microbial community structure and microbial processes. Although this objective seems straightforward, there are conceptual and methodological challenges to designing studies that explicitly evaluate this link. Here, we analyzed literature documenting structure and process responses to manipulations to determine the frequency of structure-process links and whether experimental approaches and techniques influence link detection. We examined nine journals (published 2009–13) and retained 148 experimental studies measuring microbial community structure and processes. Many qualifying papers (112 of 148) documented structure and process responses, but few (38 of 112 papers) reported statistically testing for a link. Of these tested links, 75% were significant and typically used Spearman or Pearson's correlation analysis (68%). No particular approach for characterizing structure or processes was more likely to produce significant links. Process responses were detected earlier on average than responses in structure or both structure and process. Together, our findings suggest that few publications report statistically testing structure-process links. However, when links are tested for they often occur but share few commonalities in the processes or structures that were linked and the techniques used for measuring them.

Keywords: structure–function relationship; microbial community composition; microbially mediated ecosystem function; experimental duration; experimental perturbation; microbial processes

Received: 31 May 2015; Accepted: 9 September 2015

© FEMS 2015. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

INTRODUCTION

Microorganisms dominate Earth's biogeochemistry by virtue of both their numbers and metabolic capabilities (Falkowski, Fenchel and DeLong 2008). Modern-day molecular technologies allow us to identify the myriad microbes that exist, to identify the genes they carry, and even to determine whether those genes are being transcribed and translated into functional proteins. What remains an open question is whether all of this information will enable us to better understand, predict and model the ecosystem processes that microbes perform (e.g. Carney and Matson 2005; van der Heijden, Bardgett and van Straalen 2008; Todd-Brown et al. 2011; Wallenstein and Hall 2011; Petersen et al. 2012; Prosser 2013; Graham et al. 2014).

Microbial and ecosystem ecologists approach the question above with both optimism and caution. On one hand, our capacity to extract, amplify and assess microbial nucleic acids and proteins from environmental samples is staggering and is improving rapidly; we can evaluate the community composition of microbes present within nearly any environmental sample. Yet, this technological progress has repeatedly demonstrated that the phylogenetic identities and metabolic capabilities of microbes within any environmental sample are far more diverse than we had imagined (Prosser 2012) and the variety of metabolic states (growing, active, dormant, deceased) means that 'who is present' is not a proxy for 'who is active' (Jones and Lennon 2010; Lennon and Jones 2011; Blagodatskaya and Kuzyakov 2013; Blazewicz et al. 2013). Perhaps even more importantly, we are increasingly aware that the presence or abundance of particular organisms, genes or gene transcripts may not be well connected to the rates with which the associated biochemical reactions are occurring (Schimel and Schaeffer 2012; Rocca et al. 2015).

Despite these challenges, many recent influential papers and reports have called for incorporating information about microbial communities into assessments of ecosystem functions and improvements of ecosystem models (e.g. Moorhead and Sinsabaugh 2006; Konopka 2009; Allison 2012; Bouskill et al. 2012; Wieder, Bonan and Allison 2013). Some reports distinctly acknowledge that microbial communities temper the influence of natural and anthropogenic disturbances on ecosystem functioning (e.g. Krause et al. 2014). Others suggest that we continue exploring how to use microbes for improving mechanistic predictions of ecosystem processes so that in cases where information about microbial communities 'is' relevant, we will have more accurate predictions (e.g. McGuire and Treseder 2010). A substantial increase in the ease and affordability of acquiring and analyzing microbial community data has spawned significant efforts to study structure-process connections (e.g. Prosser 2012), and we can now examine these connections across spatial, temporal and taxonomic scales (e.g. Walters and Knight 2014). Yet, some researchers continue to challenge the generality of many studies and encourage us to determine where genetic and ecosystem studies overlap (e.g. Fuhrman 2009). Is information resulting from our numerous structure-process studies consistently filling a knowledge gap, or is there little return for our investment (Graham et al. 2014)?

While some studies have identified empirical links between microbial communities and ecosystem processes (Box 1), this body of literature is also replete with studies where structure and process appear uncoupled. Such uncoupling could occur when the ultimate rate-limiting step is abiotic, such as desorption of clay-bound organics or the breakup of aggregates and the release of labile materials; in such cases, the composition of

the decomposer community would be unlikely to visibly influence the rate at which the materials are processed (Schimel and Schaeffer 2012). However, uncoupling of structure and process might occur even when the rate-controlling steps are biotic—in such cases the lack of a relationship might be due to factors including microbial dormancy (Jones and Lennon 2010; Lennon and Jones 2011), horizontal gene transfer (Smets and Barkay 2005), functional redundancy (Allison and Martiny 2008), priority effects (Fukami et al. 2005) and neutral assembly processes (Nemergut et al. 2013; Nemergut, Shade and Violle 2014). When links do occur, the success of identifying them is also likely dependent on the conditions and techniques used in each study (e.g. Shade et al. 2012a), or the time-scale over which measurements occur. Yet, it is not clear how often and with which techniques researchers have identified explicit links between microbial community structure and process, and examining the differences between such studies could guide the direction of future studies.

Box 1 | Term Definitions

Microbial Community Structure: the characteristics of a community of microorganisms including bacteria, archaea and microeukaryotes as measured by any metric of taxa or gene composition, diversity and/or abundance via a range of molecular or cultural techniques.

Microbial Process: microbial activity measured at the community scale, either through direct assessment of enzyme activities that mediate a process (e.g. denitrification enzyme assay), monitoring of end product accumulation over time (e.g. net nitrification) or tracking element cycles through stable isotope tracers (e.g. gross nitrification/denitrification).

In this paper, we seek to describe the state of recent efforts to characterize microbial community structure and function relationships. We focused this literature synthesis on manipulative experiments because such studies may offer the best opportunity to establish a link between changes in both microbial identity and microbial processes in response to a known (and controlled) experimental driver. We evaluate the frequency with which authors of recent publications (2009–13) have simultaneously investigated microbial community structure and microbial process responses to an experimental manipulation and the time that lapsed before they detected changes in structure, process or both structure and process. To guide our evaluation, we focus on five questions: (1) How frequently do publications report that an experimental manipulation leads to changes in either microbial community structure or microbially mediated ecosystem processes? (2) How often do researchers measure simultaneous changes in both microbial community structure and process? (3) Are particular experimental conditions or techniques more often associated with links between structure and process? (4) Do structure and process respond to disturbance at different rates? (5) How are researchers attempting to evaluate inferential or empirical links between measures of microbial community structure and process?

METHODS

We synthesized recent literature that contained experimental manipulations of environmental factors to induce stress on microbial communities. We excluded field-based observational

studies, such as environmental gradients, out of concern that relationships observed between microbial community structure and ecosystem processes within such studies might result from unobserved drivers (not associated with the gradient of interest) or reverse relationships where ecosystem function affects community composition (as discussed by Krause et al. 2014). By contrast, experimental manipulations allow researchers to determine whether and how structure and process metrics respond to a well-constrained change in the environment.

We used a set of structural terms and a set of process terms to search the ISI Web of Science literature database for papers published between 2009 and 2013 (Fig. S1, Supporting Information). We required that papers include at least one of the process terms and one of the structural terms. To achieve this, we searched for papers containing processes where Topic = 'decomp* OR methan* OR sulfate red* OR denitrif* OR dnf OR nitrif*' and structures where Topic = 'commun* OR gene* OR physiolog*'. The processes indicated by these terms are commonly explored in the ecological literature and the results from this search yielded more papers than a search for 'funct*' alone. These structure-search terms were selected to return experiments involving microbial communities rather than culture isolates. The output from the structure- and process-search terms in the ISI Database yielded 199 749 papers. We refined this search by Topic = 'microb*' to exclude papers focusing exclusively on macroorganisms; this narrowed to the total to 32 386 papers (Fig. S1, Supporting Information). We then restricted these results to 'Environmental Sciences Ecology' as the Research Area and used 'Topic = ecology'. From this output, we selected four of the top five journals with the greatest number of paper results (FEMS Microbiology Ecology, Soil Biology & Biochemistry, The ISME Journal and Microbial Ecology) (we excluded one of the five journals (PLOS One, as a general journal)), two general ecology journals (Ecology and Ecology Letters) and two major full-spectrum journals (Nature and Science). Following a review of this list of journals by experts in the field as part of the Powell Center working group, we added a leading general aquatic science journal (Limnology and Oceanography). Limiting our search to this subset of journals reduced our results to 1189 papers.

We examined the abstract of each paper for the following criteria: (1) the study was experimental, (2) at least one process and one structural metric were measured simultaneously and (3) the study altered at least one chemical or physical condition. We included papers that manipulated biological conditions such as tree girdling if a chemical or physical change to the environment was documented. In total, we obtained 148 papers (12.4% of the original 1189) that comprised the 'full dataset' (Box 2) used in this synthesis.

For each of these 148 papers, we recorded the type of manipulation, the test location (laboratory or field), the duration of the experiment, whether or not immigration or emigration was possible based on open or closed experimental units, and which groups of organisms were examined (microeukaryotes, archaea and/or bacteria). We recorded each experimental treatment-process combination separately, such that each paper could have multiple 'incidences' if multiple processes were measured or if multiple treatments were applied within a single study. For example, a research project that used two treatments (e.g. elevated temperature or a fertilizer addition) and measured both N₂O flux and CO₂ flux in each treatment plot would result in four separate incidences: one for N₂O flux in the elevated temperature plots, one for CO₂ flux in those same elevated temperature plots, and one each for the N₂O flux and CO₂ flux in the

fertilizer plots. Each of these incidences could contain multiple structure metrics if more than one aspect of the community was measured (e.g. 16S rRNA and *nirK* genes). For each process measured, we denoted whether the authors measured ambient or potential (i.e. rate measured with substrate enrichment) microbial processes, as well as the technique used to assess microbial community structure and the type of metric reported: relative abundance, absolute abundance (per gram of soil) or presence-absence.

Box 2 | Dataset Definitions

Full dataset: 148 papers that matched our search terms and contained experiments that measured both a microbial community structure and a microbial process.

Changed dataset: a subset of 112 papers from the full dataset in which both microbial community structure and process changed.

Link-tested dataset: a subset of 128 incidences in 38 papers from the changed dataset where microbial community structure-process incidences were tested for statistical significance.

Linked dataset: a subset of 96 incidences in 28 papers from the link-tested dataset in which structure-process links were found to be statistically significant.

Paper: a single peer-reviewed publication, often containing multiple incidences.

Incidence: the combination of a process and structures between which the authors looked for a link (e.g. an experiment may make one structure measurement and two process measurements resulting in two incidences: 16S rRNA gene with N₂O flux, and 16S rRNA gene with CO₂ flux; or one process and two community measures resulting in one incidence: CO₂ flux with 16S rRNA gene and *nirK*).

To examine the connections between structural and process measures, we investigated whether or not each experimental treatment resulted in (1) no change in either structure or process metrics, (2) a process change only, (3) a structural change only or (4) a change in both structure and process. For those that reported simultaneous change in both structural and process attributes, we further tallied whether or not the authors had statistically tested for a relationship between these attributes and whether a statistical relationship, or link, was found. This statistically tested dataset (referred to as the 'link-tested dataset' hereafter) contained 38 papers (26% of the full dataset) with 96 incidences that found a link and 32 incidences that found no link. For this link-tested dataset, we determined which genes or taxonomic groups of organisms were tested with a process and which metric of community structure had been used to measure them (e.g. qPCR, DGGE and TRFLPs). For detailed information on the generation of the link-tested dataset, see Supplementary Methods.

Because measures of microbial community structure and process were often taken at multiple time points following a disturbance, we examined whether the time since the experimental manipulation affected the likelihood of detecting either a structural or process microbial response to experimental treatments. To investigate this, we examined the duration of experiments in both the full set of experimental papers (148 papers) and the link-tested dataset (38 papers). Using the full set of experimental papers, we compared the duration of incidences in which there was a change in structure only, process

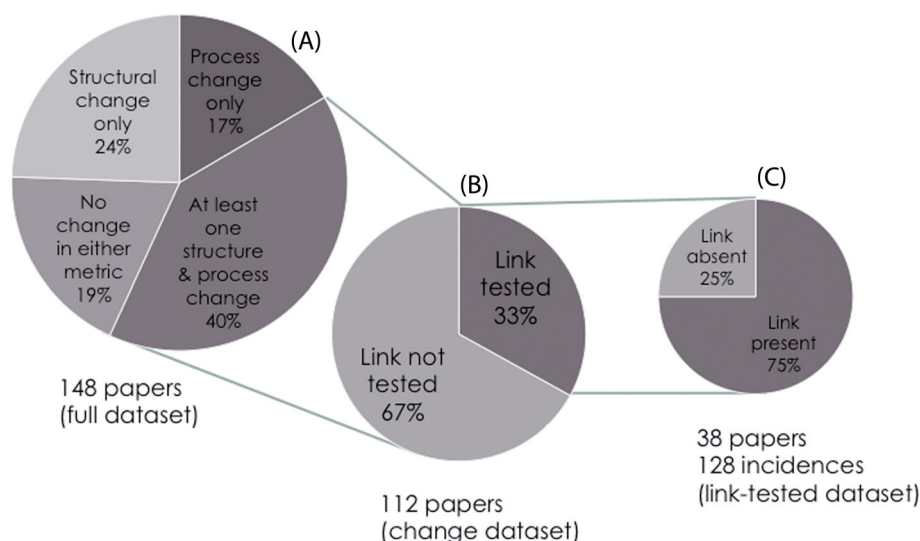


Figure 1. Distribution of the literature synthesis results expressed in proportion of papers (A and B) or incidences (C) derived from the 148 papers that matched selection criteria (see Methods). Papers contained multiple incidences; e.g. the 40% of structure-process change incidences from A occurred within 112 papers.

only, both, or neither. For this analysis, duration was defined as the length of time from the first treatment application to a time at which structure and process were both measured. We included repeated measures of process or structure over the completion of the study. For example, if community structure and a process were measured on the 5th and 30th day of the experiment, a separate incidence was made for each date, so the treatment would have two durations. Using this approach, we were able to capture temporal changes in structure that might occur on a different day from changes in process. Secondly, we used our link-tested dataset to compare the duration of studies in which a statistically significant link was present or absent. Because some studies combined time points in their analysis of a link, the link-tested dataset contained only one time point per incident, the duration recorded for each incident ended the day that both a structural and process change were measured.

RESULTS

How frequently do publications report that an experimental manipulation led to changes in microbial community structure or microbially mediated ecosystem processes? And how often do these changes cooccur?

The set of 1189 papers matching our search terms comprised less than 3% of the total number of papers published in the targeted journals between 2009 and 2013 (Table S1, Supporting Information). The majority of these papers were published in *Soil Biology & Biochemistry*, *FEMS Microbiology Ecology*, *Microbial Ecology* and *The ISME Journal*, respectively (Fig. S2, Supporting Information). Moreover, only 12.4% ($n = 148$) of these papers contained experiments that measured both microbial community structure and microbial process in response to an environmental manipulation (Fig. S1, Supporting Information). For 19% of incidences (from 52 papers, 236 of 1082 incidences), authors concluded no change in either a structure or process metric, while 24% (68 papers, 219 incidences) reported only structural shifts, and 17% (46 papers, 139 incidences) reported only process changes. In the remaining 40% of incidences (112 papers, 488 incidences), authors detected both a process change

and a structural change in response to an experimental manipulation (Fig. 1). This subset was our 'changed dataset'. Within each paper, we examined whether the authors did or did not test for statistical links between microbial structural and process responses to experimental manipulations. We found that only 38 of the 112 papers from the changed dataset were included in the link-tested dataset because they specifically tested for a statistical link between structure and process metrics (Fig. 1). Many of these papers measured multiple structure-process incidences so that our dataset included a total of 128 tested links. We found that in 75% of incidences (from 28 papers, 96 incidences) the authors identified a statistically significant relationship between structural and process responses to their experimental manipulation and in 25% of incidences (from 16 papers, 32 incidences) the tested link was not statistically significant. The set of statistically significant linked incidences comprises our linked dataset.

Do particular experimental conditions or techniques more often associate with observed links between community structure and microbial processes?

Experimental design

Although studies manipulated different variables, statistical links between structure and process were most commonly tested in studies that manipulated fertilizers and climate change drivers (e.g. temperature and CO_2) (Table S2, Supporting Information). Among the incidences associated with fertilization treatments ($n = 36$), such as the addition of ammonium nitrate or urea, 78% resulted in a significant link between structure and process. In climate change manipulations ($n = 27$), which comprised the next most common type of manipulation, 63% (17 incidences) showed a significant statistical relationship. Treatments that exhibited significant relationships between structure and process included warming (53% of incidences), elevated CO_2 (35% of incidences) or a combination of both elevated temperature and CO_2 (12% of incidences).

The proportion of significant links varied depending on the techniques used for measuring community structure (Fig. 2A), but not according to the group of organisms targeted (Fig. 2B), the location of the experiment (laboratory or field, Fig. 2C) or

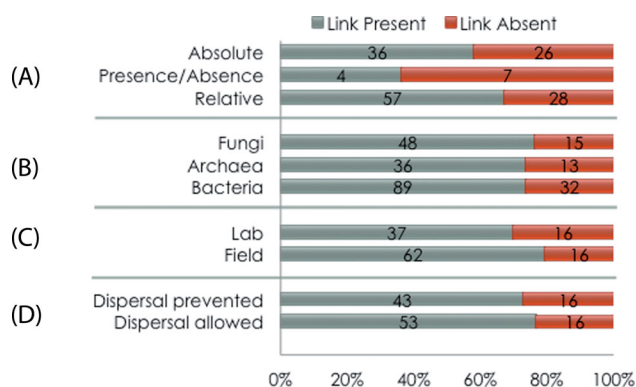


Figure 2. Distribution of incidences among different experimental attributes from the 38 papers that tested for a link between microbial community structure and process. The number of incidences is indicated within each bar. (A) Type of quantification of community structure; (B) major taxonomic groups targeted; (C) laboratory versus field experiments; (D) allowance of microbial dispersal during the experiment.

whether experimental design prevented dispersal into or out of experimental units (Fig. 2D). Moreover, those experiments using presence-absence measurements of community membership reported a smaller proportion of significant links between structure and process (36%) in comparison to those experiments assessing relative (67%) or absolute abundances (58%) of the present taxa. However, the number of incidences reporting presence-absence values was small ($n = 11$) compared to either of the other two categories ($n = 62$ and 85 for absolute and relative abundances, respectively) and included ordination techniques, measures of diversity and specific taxonomic groups.

When examining our full dataset (148 papers), we found that the duration of the experiment significantly affected some of the observed responses (Fig. 3A, Kruskal-Wallis test, $df = 3$, $P < 0.01$). Using incidence medians, changes in process alone occurred after shorter periods (27 days) than changes in structure alone (61 days), or than concurrent changes in structure and process (56 days) (Mann Whitney pairwise comparisons with Bonferroni corrections, $P < 0.02$). However, there was no difference in the duration of experiments that produced structure changes or concurrent structure-process changes ($P = 1.0$). When we compared durations with and without a statistically significant link using the link-tested dataset, though, there was no significant

difference between the duration of linked and unlinked experiments (Mann Whitney U test, $df = 1$, $Z = -0.33$, $P = 0.74$, Fig. 3B). Using the entire datasets, experimental duration for studies in the link-tested dataset was longer than the mean experimental duration of the full dataset (Mann Whitney U test, $df = 1$, $Z = 3.23$, $P < 0.01$).

Ecosystem processes and community structure in the link-tested dataset

Within the linked dataset, CO_2 fluxes and nitrification were the most frequently measured ecosystem processes (Fig. 4, Table S3, Supporting Information). These processes were each significantly linked to a microbial community structure attribute in ca. 80% of incidences where a statistical test was performed. Nitrification was the most commonly measured process (18 incidences, 10 papers), followed by CO_2 flux (17 incidences, 9 papers), N_2O flux (17 incidences, 10 papers) and CH_4 flux (15 incidences, 5 papers). A link between community structure and microbial process was present in 100% of experiments that attempted to link CH_4 flux to a microbial community attribute. The same was true for ammonia oxidation, though we came across only four such incidences, all within a single paper. On the other end of the spectrum, there were no significant links with community structure in experiments that measured organic nitrogen decomposition, ammonification and total activity from a suite of nine enzymes, though these conclusions were supported by only two or three incidences for each process.

Overall, the greatest number of incidences that tested for a significant link with process targeted the 16S rRNA gene (51 incidences, 12 papers), the denitrification gene *nosZ* (39 incidences, 12 papers) or the bacterial ammonia monooxygenase gene *amoA* (37 incidences, 14 papers) (Fig. 5, Table S4, Supporting Information). The proportion of incidences that were linked varied by the metric of community structure that was applied and included both phylogenetically specific and universal genes or phylogenetically broad markers. All of the incidences in which community structure measurements were defined using most probable number of methanogens ($n = 8$) and methanotrophs ($n = 8$), 16S rRNA genes (archaea) ($n = 2$) or *cbh1* (cellobiohydrolase gene) ($n = 6$) resulted in a statistical link with process, although these incidences were drawn from only one or two papers. Nitrogen cycling genes *nosZ* and *nirS* as well as the methane monooxygenase gene *pmoA* had the next highest percentage of links

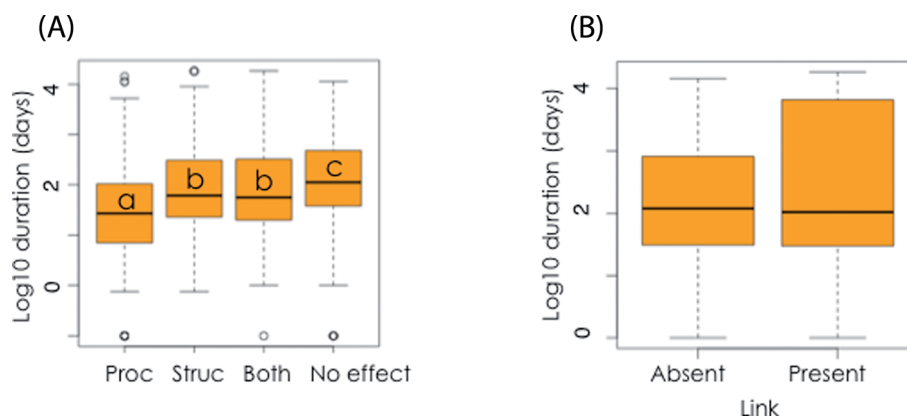


Figure 3. (A) Duration of experiments from the 148 papers in which process changes [Proc] ($n = 140$), structural changes [Struc] ($n = 221$), simultaneous structure-process changes [Both] ($n = 488$), or no changes [No effect] ($n = 237$) were reported. Letters reflect Kruskal-Wallis rank sum test ($df = 3$, $P < 0.01$) followed by Mann Whitney pairwise comparisons with Bonferroni corrections. Different letters indicate significant differences ($P < 0.02$) between categories. (B) Duration of experiments from 38 papers where both structure and process changed and a link was statistically tested (Mann Whitney U test, $df = 1$, $Z = -0.33$, $P = 0.74$).

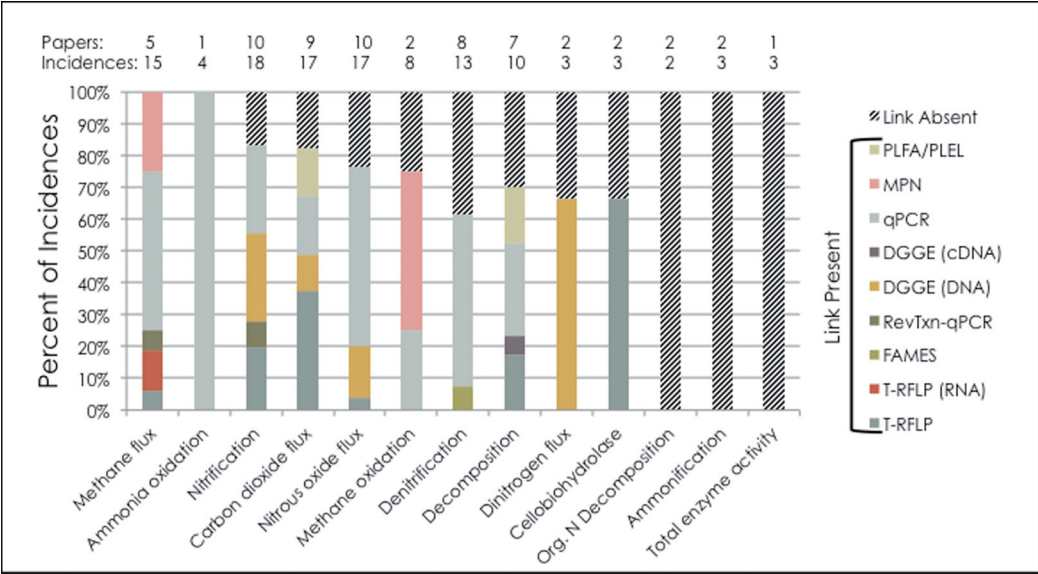


Figure 4. Distribution of incidences among microbial processes from the 38 papers that tested for statistical links between structure and process (link-tested dataset). Different colors indicate the relative proportion of linked incidences associated with the different process metrics used. The numbers above each bar indicate the number of papers from which the incidences were extracted and the total number of incidences. See Table S3 (Supporting Information) for additional processes not included in figure due to fewer than two incidences.

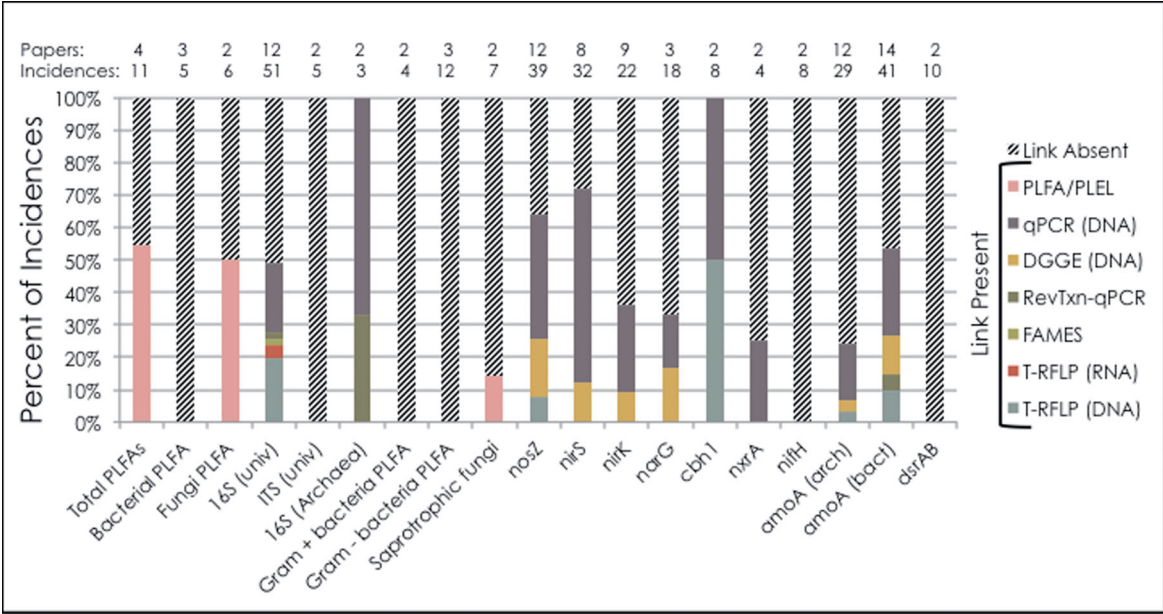


Figure 5. Distribution of incidences among structural measurements from the 38 papers that tested for a statistical link between microbial community structure and microbial process (link-tested dataset). Different colors indicate the relative proportion of positive incidences associated with the different compositional metrics used. The numbers above each bar indicate the number of papers from which the incidences were extracted and the total number of incidences. Additional metrics with fewer than two papers found in Table S4 (Supporting Information).

with 64–70% of incidences linked ($n = 39, 30$ and 6 total incidences from $12, 8$ and 1 paper(s), respectively). When tested for, the remaining genes and organisms were linked with a process in $<55\%$ of the incidences. Categories in which no link with process was detected also included phylogenetically broad and narrow groups. Broad categories included Gram-positive and -negative bacteria ($n = 4$ and 12), often assessed using phospholipid fatty acid (PLFA) techniques, as well as genes in fungi associated with the internal transcribed spacer region ($n = 5$). More specific genes not linked to any process included the nitrogen

cycling *nifH* gene and the sulfate-reducing *dsrAB* genes ($n = 8$ and 10).

Researchers attempted to link non-specialized processes with genes that were narrowly distributed and vice versa. Occasionally, these genes were also not directly related to the process. To explore whether there were differences in the degree of metabolic specialization of processes tested with a universal or specific gene, we examined incidences measuring either the universal 16S rRNA gene or bacterial *amoA* gene (Fig. S4, Supporting Information). These genes were both common in our

link-tested dataset. Multiple studies attempted to link bacterial 16S rRNA genes with both CO₂ flux, a broad process performed by all heterotrophs and with N₂O flux, a more narrowly distributed process derived from both nitrification and denitrification. Community structure assessed by the 16S rRNA gene was linked to CO₂ flux in 83% of the experiments where it was tested ($n = 5$ of 6) but was never statistically associated with N₂O flux ($n = 9$). Attempts to link specific functional genes were consistently well correlated with the associated process, e.g. nitrification rates were linked with bacterial *amoA* gene frequencies in 100% of tests ($n = 14$). Conversely, specific genes were not significantly related to the broad processes with which researchers attempted to link them, e.g. *amoA* with CO₂ flux ($n = 3$).

Molecular techniques used in the link-tested dataset

Community structure metrics used in the link-tested dataset (38 papers) were dominated by four techniques: qPCR, T-RFLPs, DGGE and PLFA. Approaches using DNA resulted in the highest percent of linked structure-process incidences. Quantitative PCR (qPCR) of functional genes was used in 8 out of the 10 most commonly linked processes and was the only technique that had a higher occurrence of links present (50 incidences) than absent (33 incidences). Primarily, these analyses used relative abundances, (copy number ng DNA⁻¹) (66%); though absolute abundances (copy number g soil⁻¹) were associated with about one third of the incidences (33%). Roughly two-thirds (68%) of structure-process pairs using relative abundance with qPCR were statistically linked. This was a less common occurrence for links tested using absolute abundance (56% of pairs linked). The next most commonly used techniques were DNA-based Terminal Restriction Fragment Length Polymorphism (T-RFLP) and DNA-based Denaturing Gel Gradient Electrophoresis (DGGE) which both were linked in 43% of incidences and used for the 16S rRNA gene and functional genes (Fig. 5). DNA-based T-RFLP was more commonly used (24 incidences linked: 20 relative abundance, 4 presence-absence) than DNA-based DGGE (15 incidences linked: 11 relative abundance, 4 presence-absence). Methane flux was statistically linked to community structure in 100% of tested incidences in which structure was characterized using five different techniques from five different papers, while the other 100% linked process (ammonia oxidation) used only qPCR and was drawn from a single paper. RNA-based techniques (reverse transcription qPCR and TRFLPs or DGGE using cDNA) were used solely for exploring links with CH₄ flux or oxidation, denitrification, nitrification and decomposition. These were associated with a small percentage of the linked incidences in CH₄ flux ($n = 3$), nitrification ($n = 2$) and decomposition ($n = 1$) (Fig. 4). There was no apparent connection between the number of different techniques used and the likelihood of detecting a community structure-process link (Fig. S3, Supporting Information).

How are researchers attempting to identify links between measures of microbial community structure and process?

Abundance, diversity and presence-absence were all used to measure community structure, though abundance was used most frequently and contributed to more links than presence-absence or diversity measures. Three of the four community structure metrics that yielded 100% of linked incidences were made from copy numbers or organism counts of methanogens, methanotrophs or 16S rRNA genes (archaea) (Fig. 5, Table S4, Supporting Information). Microorganisms with the *cbh1* (cellulohydrolase) gene were other group with 100% of linked in-

cidences and had incidences evenly split between qPCR abundance ($n = 4$) and T-RFLP-based diversity indices ($n = 4$). Each of these fully linked metrics of community structure, however, relied on results from only one or two papers. Nitrogen cycling genes *nosZ* and *nirS* (12 and 8 papers, respectively) as well as the methane monooxygenase gene *pmoA* (1 paper) had the next highest percentage of links present (64–70%). The majority of these links were also obtained using abundance instead of diversity metrics. With the *nosZ* gene, 15 out of the 25 linked incidences (60%) used abundance data while the remaining 40% used diversity metrics. In *nirS*, 19 incidences used abundance (83%) and four used diversity (17%). Using *pmoA*, links with diversity and abundance were evenly split with two of four incidences in each category. The most commonly used metric of community structure targeted the universal segment of the 16S rRNA gene (51 incidences) and yielded links nearly evenly split between abundance (12 incidences) and diversity metrics (13 incidences). The second most common metric, the bacterial ammonia monooxygenase gene *amoA* (41 incidences), had nine incidences linked through diversity measures and 13 linked through abundance.

The majority of structure-function links were tested using correlation analysis. Likely reflecting the prevalence of abundance metrics based on qPCR, the techniques used to test for links were dominated by Spearman or Pearson's correlation analyses (68% of incidences, details not shown). Roughly 77% of incidences that were tested using correlation analysis yielded links, which mirrored the percentage of links present in the total dataset (75% of total incidences had a link present, Fig. 1). Canonical correspondence analysis was the second most frequently used technique and represented 11% of tests, but only 44% of those yielded a link. Incidences based on redundancy or co-inertia analysis had 100% of incidences linked to process. Regression analysis was associated with only 5% of incidences, 75% of which had a link present.

DISCUSSION

The literature synthesis presented here revealed that researchers explicitly tested for a statistical link between microbial community structure and process in only one third of incidences from the experimental studies detecting structure and process rate changes in response to experimental manipulations. Yet, when authors reported testing for links they were commonly found; three-quarters of tested incidences were statistically linked. Ideally, theories involving structure-process links would generate publications with a stated hypothesis that was statistically tested and fed back to theory development (Fig. 6). However, this flowpath occurred in only 17% of papers. This suggests that many datasets may be available to explore for structure-process linkages or may support hypothesis generation. It is possible that due to publication bias toward statistically significant results, these datasets were tested previously and only significant results were reported. Regardless of whether or not there was a greater number of unlinked structure-process pairs than we report here, our analyses identify many challenges to consider when designing and conducting experiments to investigate microbial community structure and process links.

One major challenge in identifying and examining relationships between microbial community composition and process responses is that we have little understanding of the temporal scales at which changes in community structure and related functional attributes occur. Microbial enzymes can be modified by chemistry and biology of the surrounding

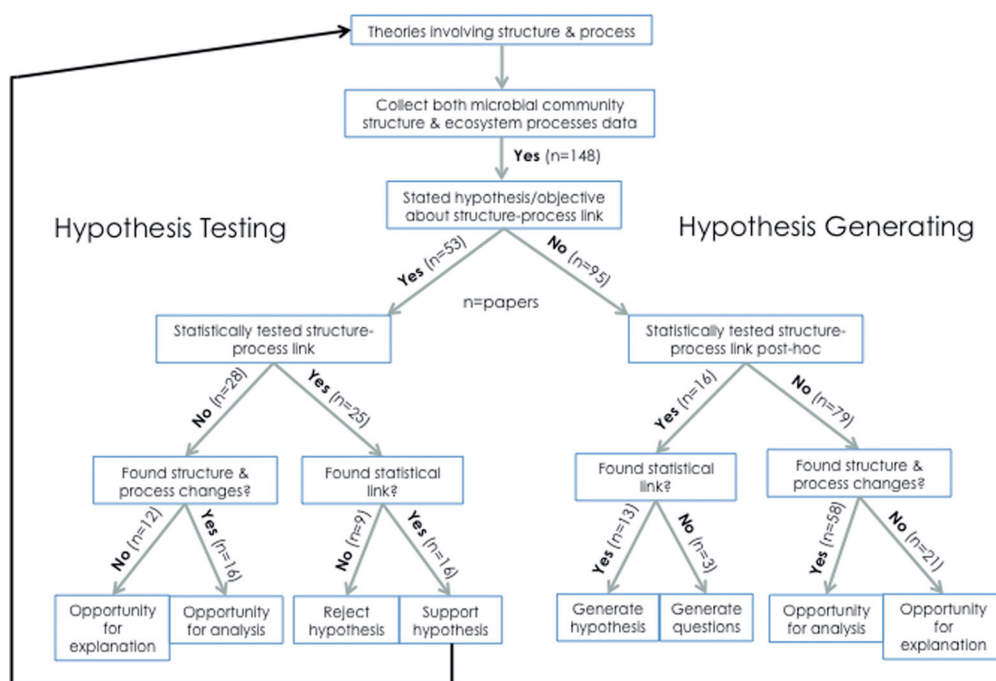


Figure 6. Flowchart of guidelines for research involving microbial community structure and ecosystem processes overlain with 'yes/no' data from this literature synthesis (n = number of papers). Research decisions lead to hypothesis testing or hypothesis generating paths.

environment before they are relevant for an ecosystem process, thus creating a temporal disconnect between structure and process. For example, phenotypic plasticity can lead to differences in activity over short periods without an apparent change in taxonomic composition. This is illustrated by experiments incorporating single-cell techniques such as microautoradiography combined with fluorescence in situ hybridization which have shown that bacterial activity can change greatly without an apparent change in community membership (e.g. Ruiz-González et al. 2012). This suggests that transformations in both microbial structure and processes sometimes can be decoupled in time, potentially misleading interpretations about links between them. Within our link-tested dataset of incidences reporting that both structure and process had changed, there was no difference between the median duration of experiments where significant links were detected and those where there were none. This indicates that when experimental data captures structure and process responses, the timescale of the experiment does not affect the likelihood of detecting a link. In our full dataset (148 papers), however, the median duration in which structure changed (61 days) or both structure and process changed (56 days) was approximately twice as long as the mean duration of studies reporting a process change alone (27 days) (Fig. 3). This supports the idea that physiological responses precede, and perhaps do not even require, community shifts (Comte, Fauteux and del Giorgio 2013). Therefore, if processes are changing prior to structure, researchers making early or infrequent measurements may not capture data necessary to support a connection between these two parameters. For longer experiments, there is evidence that researchers sample less frequently: in their literature exploration of microbial responses to disturbance, Shade et al. (2012a) identified a negative relationship between sampling frequency and experiment duration. Thus, changes in process from our full dataset (148 papers) may reflect higher frequency sampling whereas changes in structure could result from less frequent sampling over a longer duration. Because this elevates

the difficulty of identifying a connection, explicit consideration of temporal factors in study design may decrease these discrepancies.

Considerable experimental and empirical evidence has shown that alteration of environmental variables such as temperature, salinity, pH and nutrient concentrations often coincide with shifts in structure and/or processes of microbial communities across a variety of ecosystems (e.g. Lozupone and Knight 2007; Braker, Schwarz and Conrad 2010; Vishnivetskaya et al. 2011; Herold, Baggs and Daniell 2012; Reed and Martiny 2012; Shade et al. 2012b; Wertz, Leigh and Grayston 2012). In our compilation of experiments within single study systems, most often either only one or neither attribute responds to environmental disturbance. Many studies have shown that pH can drive multiple types of compositional and process changes (e.g. Liu et al. 2010; Rousk et al. 2010; Meron et al. 2012; Cheng et al. 2013). Therefore, it is surprising that the studies in our link-tested dataset rarely manipulated pH directly (two incidences), though other manipulations such as N additions often indirectly alter pH. The addition of fertilizers was among the most commonly used disturbance. Our finding that fertilization treatments such as urea or ammonium nitrate additions were most likely to yield a link between structure and process may have been a consequence of fertilizer serving as a microbial resource, increasing plant-derived carbon availability, and altering pH (Pierre 1928; Geisseler and Scow 2014). Alternatively, nitrogen could inhibit microbial growth and activity. Meta-analyses of nitrogen enrichment studies found that under elevated nitrogen, microbial biomass and CO₂ flux may decline (Treseder 2008) and organic matter decomposition may be impeded (Janssens et al. 2010) or the recalcitrant soil carbon pool may be less effectively decomposed (Ramirez, Craine and Fierer 2010).

Of all the techniques used to characterize microbial community structure, links with microbial processes were most commonly detected with qPCR. This suggests that a microbial process is more likely to coincide with the relative or absolute

membership of the responsible organisms instead of indirect metrics such as composition or diversity of the community. Diversity metrics may be more representative of within community dynamics than functional potential, yet in processes catalyzed by organisms with specialized metabolisms, diversity can also adequately predict process stability and magnitude (Levine et al. 2011). Often our 'snapshot' analysis of microbial communities also attempts to link structure to ecosystem processes without identifying the influence of underlying ecological conditions such as competition, assembly, tradeoffs and feedbacks (Prosser et al. 2007; Prosser 2012), but arguably, this is a challenging task for any field of ecology, not just microbial ecology.

In our link-tested dataset, both relative and absolute abundance yielded a similar percentage of links. While this result suggests that neither approach had an advantage in terms of uncovering links, the contributing studies may have had little variation in total biomass between the control and treatment samples, resulting in similar relative and actual population sizes. Presence-absence characterizations of communities, however, provided less frequent links to process, likely because presence-absence data provides little information about the dominant organisms within a community that might be responsible for the process. This highlights that commonness or rarity attributes of microbial communities may be important for understanding function (Aanderud et al. 2015).

Given the large information output, decreasing costs of sequencing and suggested ecological applications (Poisot, Péquin and Gravel 2013) we anticipated that next generation sequencing (NGS) would be a frequently used technique, but in fact, none of the studies used NGS as a method for examining links between structure and process. This may be due to the use of NGS in many observational studies instead of experiments, which would have excluded them from our datasets or because of prohibitive costs which increase quickly when striving to fulfill replication requirements (Prosser 2010). While these costs continue to decrease, the time span we used for our literature search may have captured studies that were completed while NGS costs were still high. The use of NGS data enhances our ability to characterize the diversity and composition of microbial communities at different levels of resolution, which could influence the detectability of relationships between structure and process. For instance, if two taxa within a family respond oppositely to an environmental perturbation, coarser-scale taxonomic resolution such as order could obscure the actual change in structure. This inconsistency in the response of closely related taxa has been documented across complex environmental gradients (e.g. Bier, Voss and Bernhardt 2015), but may be more the exception than the rule (Philippot et al. 2010; Lennon et al. 2012). A further consideration for NGS information is that it provides relative abundance data, and can result in substantial data mining and type II error. Thus, it is critical for experiments investigating structure-process relationships with NGS to be hypothesis motivated.

Guided by earlier hypotheses (Schimel 1995; Schimel, Bennett and Fierer 2005), we expected that a confined guild of microbes would contain a more similar overall genetic makeup and would be more likely to respond to a perturbation using similar mechanisms, whereas a guild containing a much more diverse genetic toolbox would respond to the same perturbation in a greater variety of ways. Thus, there would be less variation in process output from the confined guild and hence a greater likelihood of a statistical link between guild structure and the process measured. However, we found that links were not only identified for processes governed by microbes with nar-

row phylogenetic distributions such as methane oxidizers, methanogens and sulfate reducers, but also for processes performed by a wide diversity of taxa, for example, carbon substrate utilization (Martiny, Treseder and Pusch 2013) (Fig. S4, Supporting Information). These findings support the idea that for some processes, such as those related to soil moisture adaptation, organisms at coarse taxonomic levels have ecological coherence (Philippot et al. 2010; Lennon et al. 2012).

An important consideration in 'linkage studies' should be whether the linkage is direct and causal, or incidental being driven by a master variable. This synthesis led us to reflect on the use of universal genes for linking with specific, phylogenetically narrow processes. For instance, the 16S rRNA gene was used in correlation with methane flux and nitrification (Fig. S4, Supporting Information). While employing the 16S rRNA gene for sequencing would allow researchers to reduce costs for testing connections between structure and multiple processes, this broad approach may be more appropriate for hypothesis generation or incidental associations than targeted research questions (see Prosser 2013). Moreover, statistical links were tested between indirectly related structure-process pairs. For instance, archaeal ammonia oxidation genes were linked with denitrification processes (Fig. S4, Supporting Information). This may stem from a temptation to test every structure-process pair merely because the data are available. Given that the probability of finding a correlation increases with the number of variables, studies with a small sample size run the risk of coincidental discoveries. In this synthesis, though, the number of different structure or process metrics used per study did not exceed 10 and had no influence on the likelihood of detecting a structure-process link (Fig. S3, Supporting Information).

Although for this synthesis we used literature reporting experimental manipulations of environmental variables, linkage studies can also result from another class of experiments that directly manipulate microbial community structure. This structure-manipulation approach would potentially allow one to assess the strength of structure-process links in a highly controlled environment. For example, altering decomposer richness may increase rates of CO₂ mineralization (Bell et al. 2005), while microbial evenness can affect responses to salt stress (Wittebolle et al. 2009). These studies unequivocally demonstrate that composition affects function under some conditions and can complement environmental manipulations to aid our understanding of microbial process responses.

We found that Spearman and Pearson's correlations dominated the statistical techniques used to assess structure-process links, but these types of assessments may not be the most effective when considering structure changes in a community of microbes. Correlations are useful for linear quantitative analysis with specific genes, but do not establish causality. Microbes both influence and respond to variations in the environment, and these are difficult to separate with correlations. Further, when assessing the community as a whole, correlation analyses may obscure microbial interactions and non-linear responses to manipulations if researchers have not designed a study with the intent of exploring non-linearities. For community analyses, multivariate statistics may yield more appropriate approaches to testing these connections, but the multitude of options can be difficult to assess. In response to this challenge, some researchers have attempted to make the appropriate techniques more accessible. For example, the GUiDe to STatistical Analysis in Microbial Ecology (GUSTA ME) is a web resource that guides users through new and accepted methods of analysis (Buttigieg and Ramette 2014). Structural equation modeling is another

important statistical technique that moves beyond the univariate analysis of microbial communities and examines relationships among all the interacting biotic and abiotic variables within an ecosystem (Grace et al. 2010). This technique has been used successfully to determine, for example, that *amoA* abundance information is important for understanding N cycling rates in soils (Petersen et al. 2012).

CONCLUSIONS

By compiling recent literature where environmental manipulations were conducted, we show that 36% of the papers collecting microbial community structure and ecosystem process data specified objectives or hypotheses regarding structure-process links, yet less than half of these papers specifically reported checking for the presence of a direct link between the two properties. And 17% of papers without structure-process hypotheses tested for a link *post hoc*. Certainly, there are different objectives specific to each study, but over and above the biological complexity of these studies, our conclusions are likely complicated by the addition of biases toward reporting positive results in academic culture. Given this bias, the low frequency with which links were statistically explored and the even lower frequency with which they were reported should encourage us to think critically about the contribution of our data to hypothesis testing as we reflect on the return for our investment. While hypothesis generation is not without merit, there was a nearly equal contribution of papers supporting hypotheses ($n = 16$) as those generating hypotheses ($n = 13$) (Fig. 6). As the structure-process knowledge base develops, we look forward to a greater portion of studies testing hypotheses that may contribute to theories involving structure and process links.

In addition, we had anticipated more links associated with phylogenetically narrow groups, yet the prevalence of detected structure-process links did not seem to follow any particular type of perturbation, experimental design or analytical metric aside from qPCR that would target these groups, thus limiting our ability to elucidate the strength and ubiquity of connections between microbial community structure and microbially mediated processes. Our potential for identifying connections is improving as our fields implement standard methods (e.g. Earth Microbiome Project) (<http://www.earthmicrobiome.org/>) and metadata requirements (e.g. for uploading data to the Joint Genome Institute and Metagenomic Rapid Annotations using Subsystems Technology server). As we move forward, building on collaborative, targeted efforts that yield experimental designs with empirical associations among communities and ecosystems may aid in our discovery of broader conclusions about the relationships between structure and microbial processes.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers and additional members of the Next Generation of Ecological Indicators working group for their contributions and the John Wesley Powell Center for Analysis and Synthesis for its support.

FUNDING

This project was supported by the United States Geological Survey, Powell Center Working Group titled Next Generation of Eco-

logical Indicators: Defining Which Microbial Properties Matter Most to Ecosystem Function and How to Measure Them.

Conflict of interest. None declared.

REFERENCES

- Aanderud Z, Jones S, Fierer N, et al. Resuscitation of the rare biosphere contributes to pulses of ecosystem activity. *Front Microbiol* 2015;6:1–11.
- Allison SD. A trait-based approach for modelling microbial litter decomposition. *Ecol Lett* 2012;15:1058–70.
- Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. *P Natl Acad Sci USA* 2008;105:11512.
- Bell T, Newman JA, Silverman BW, et al. The contribution of species richness and composition to bacterial services. *Nature* 2005;436:1157–60.
- Bier RL, Voss KA, Bernhardt ES. Bacterial community responses to a gradient of alkaline mountaintop mine drainage in Central Appalachian streams. *ISME J* 2015;9:1378–90.
- Blagodatskaya E, Kuzyakov Y. Active microorganisms in soil: critical review of estimation criteria and approaches. *Soil Biol Biochem* 2013;67:192–211.
- Blazewicz SJ, Barnard RL, Daly RA, et al. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J* 2013;7:2061–8.
- Bouskill NJ, Tang J, Riley WJ, et al. Trait-based representation of biological nitrification: model development, testing, and predicted community composition. *Front Microbiol* 2012;3:364.
- Braker G, Schwarz J, Conrad R. Influence of temperature on the composition and activity of denitrifying soil communities. *FEMS Microbiol Ecol* 2010;73:134–48.
- Buttigieg PL, Ramette A. A Guide to Statistical Analysis in Microbial Ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiol Ecol* 2014;90:543–50.
- Carney KM, Matson PA. Plant communities, soil microorganisms, and soil carbon cycling: does altering the world below-ground matter to ecosystem functioning? *Ecosystems* 2005;8:928–40.
- Cheng Y, Wang J, Mary B, et al. Soil pH has contrasting effects on gross and net nitrogen mineralizations in adjacent forest and grassland soils in central Alberta, Canada. *Soil Biol Biochem* 2013;57:848–57.
- Comte J, Fauteux L, del Giorgio PA. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front Microbiol* 2013;4:1–11.
- Falkowski PG, Fenchel T, DeLong EF. The microbial engines that drive Earth's biogeochemical cycles. *Science* 2008;320:1034–9.
- Fuhrman JA. Microbial community structure and its functional implications. *Nature* 2009;459:193–9.
- Fukami T, Bezemer TM, Mortimer SR, et al. Species divergence and trait convergence in experimental plant community assembly. *Ecol Lett* 2005;8:1283–90.
- Geisseler D, Scow KM. Long-term effects of mineral fertilizers on soil microorganisms—a review. *Soil Biol Biochem* 2014;75:54–63.
- Grace JB, Anderson TM, Olff H, et al. On the specification of structural equation models for ecological systems. *Ecol Monogr* 2010;80:67–87.
- Graham EB, Wieder WR, Leff JW, et al. Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes. *Soil Biol Biochem* 2014;68:279–82.

- Herold MB, Baggs EM, Daniell TJ. Fungal and bacterial denitrification are differently affected by long-term pH amendment and cultivation of arable soil. *Soil Biol Biochem* 2012;**54**:25–35.
- Janssens IA, Dieleman W, Luyssaert S, et al. Reduction of forest soil respiration in response to nitrogen deposition. *Nat Geosci* 2010;**3**:315–22.
- Jones SE, Lennon JT. Dormancy contributes to the maintenance of microbial diversity. *P Natl Acad Sci USA* 2010;**107**:5881–6.
- Konopka A. What is microbial community ecology? *ISME J* 2009;**3**:1223–30.
- Krause S, Le Roux X, Niklaus PA, et al. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front Microbiol* 2014;**5**:251.
- Lennon JT, Aanderud ZT, Lehmkuhl BK, et al. Mapping the niche space of soil microorganisms using taxonomy and traits. *Ecology* 2012;**93**:1867–79.
- Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 2011;**9**:119–30.
- Levine UY, Teal TK, Robertson GP, et al. Agriculture's impact on microbial diversity and associated fluxes of carbon dioxide and methane. *ISME J* 2011;**5**:1683–91.
- Liu B, Mørkved PT, Frostegård A, et al. Denitrification gene pools, transcription and kinetics of NO, N₂O and N₂ production as affected by soil pH. *FEMS Microbiol Ecol* 2010;**72**:407–17.
- Lozupone CA, Knight R. Global patterns in bacterial diversity. *P Natl Acad Sci USA* 2007;**104**:11436–40.
- McGuire KL, Treseder KK. Microbial communities and their relevance for ecosystem models: decomposition as a case study. *Soil Biol Biochem* 2010;**42**:529–35.
- Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 2013;**7**:830–8.
- Meron D, Rodolfo-Metalpa R, Cunnig R, et al. Changes in coral microbial communities in response to a natural pH gradient. *ISME J* 2012;**6**:1775–85.
- Moorhead DL, Sinsabaugh RL. A theoretical model of litter decay and microbial interaction. *Ecol Monogr* 2006;**76**:151–74.
- Nemergut DR, Schmidt SK, Fukami T, et al. Patterns and processes of microbial community assembly. *Microbiol Mol Biol R* 2013;**77**:342–56.
- Nemergut DR, Shade A, Violle C. When, where and how does microbial community composition matter? *Front Microbiol* 2014;**5**:2012–4.
- Petersen DG, Blazewicz SJ, Firestone M, et al. Abundance of microbial genes associated with nitrogen cycling as indices of biogeochemical process rates across a vegetation gradient in Alaska. *Environ Microbiol* 2012;**14**:993–1008.
- Philippot L, Andersson SGE, Battin TJ, et al. The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Microbiol* 2010;**8**:523–9.
- Pierre WH. Nitrogenous fertilizers and soil acidity: I. Effect of various nitrogenous fertilizers on soil reaction. *J Am Soc Agron* 1928;**20**:254–69.
- Poisot T, Péquin B, Gravel D. High-throughput sequencing: a roadmap toward community ecology. *Ecol Evol* 2013;**3**:1125–39.
- Prosser JL. Replicate or lie. *Environ Microbiol* 2010;**12**:1806–10.
- Prosser JL. Ecosystem processes and interactions in a morass of diversity. *FEMS Microbiol Ecol* 2012;**81**:507–19.
- Prosser JL. Think before you sequence. *Nat Rev Microbiol* 2013;**494**:41.
- Prosser JL, Bohannan BJM, Curtis TP, et al. The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 2007;**5**:384–92.
- Ramirez KS, Craine JM, Fierer N. Nitrogen fertilization inhibits soil microbial respiration regardless of the form of nitrogen applied. *Soil Biol Biochem* 2010;**42**:2336–8.
- Reed HE, Martiny JB. Microbial composition affects the functioning of estuarine sediments. *ISME J* 2012;**7**:868–79.
- Rocca JD, Hall EK, Lennon JT, et al. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J* 2015;**9**:1693–9.
- Rousk J, Bååth E, Brookes PC, et al. Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J* 2010;**4**:1340–51.
- Ruiz-González C, Lefort T, Massana R, et al. Diel changes in bulk and single-cell bacterial heterotrophic activity in winter surface waters of the northwestern Mediterranean Sea. *Limnol Oceanogr* 2012;**57**:29–42.
- Schimel J. Ecosystem consequences of microbial diversity and community structure. In: Chapin FS, Körner C (eds). *Arctic and Alpine Biodiversity: Patterns, Causes, and Ecosystem Consequences*. Berlin: Springer, 1995, 239–54.
- Schimel JP, Bennett J, Fierer N. Microbial community composition and soil N cycling: is there really a connection? In: Bardgett RD, Hopkins DW, Usher MB (eds). *Biological Diversity and Function in Soils*. New York: Cambridge University Press, 2005, 171–88.
- Schimel JP, Schaeffer SM. Microbial control over carbon cycling in soil. *Front Microbiol* 2012;**3**:348.
- Shade A, Peter H, Allison SD, et al. Fundamentals of microbial community resistance and resilience. *Front Microbiol* 2012a;**3**:417.
- Shade A, Read JS, Youngblut ND, et al. Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME J* 2012b;**6**:2153–67.
- Smets BF, Barkay T. Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. *Nat Rev Microbiol* 2005;**3**:675–8.
- Todd-Brown KEO, Hopkins FM, Kivlin SN, et al. A framework for representing microbial decomposition in coupled climate models. *Biogeochemistry* 2011;**109**:19–33.
- Treseder KK. Nitrogen additions and microbial biomass: a meta-analysis of ecosystem studies. *Ecol Lett* 2008;**11**:1111–20.
- van der Heijden MGA, Bardgett RD, van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* 2008;**11**:296–310.
- Vishnivetskaya TA, Mosher JJ, Palumbo A V, et al. Mercury and other heavy metals influence bacterial community structure in contaminated Tennessee streams. *Appl Environ Microb* 2011;**77**:302–11.
- Wallenstein MD, Hall EK. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. *Biogeochemistry* 2011;**109**:35–47.
- Walters WA, Knight R. Technology and techniques for microbial ecology via DNA sequencing. *Ann Am Thorac Soc* 2014;**11**:S16–20.
- Wertz S, Leigh AKK, Grayston SJ. Effects of long-term fertilization of forest soils on potential nitrification and on the abundance and community structure of ammonia oxidizers and nitrite oxidizers. *FEMS Microbiol Ecol* 2012;**79**:142–54.
- Wieder WR, Bonan GB, Allison SD. Global soil carbon projections are improved by modelling microbial processes. *Nat Clim Change* 2013;**3**:909–12.
- Wittebolle L, Marzorati M, Clement L, et al. Initial community evenness favours functionality under selective stress. *Nature* 2009;**458**:623–6.