Understanding How Microbiomes Influence The Systems They Inhabit: moving from a correlative to a causal research framework

Authors: Hall, E.K.[1*], Bernhardt, E.S[2], Bier, R.L.[2], Bradford, M.A.[3], Boot, C.M.[1], Cotner, J.B.[4], del Giorgio, P.A.[5], Evans, S.E.[6], Graham E. B [2,7.8], Jones, S.E.[9], Lennon, J.T.[10], Locey, K.J.[10] Nemergut, D.[2], Osborne, B.B.[11], Rocca, J.D.[1,2], Schimel J.S.[12], Waldrop, M.P.[13], Wallenstein, M.W.[1]

Article Type: Perspectives
Words in Abstract: 271
Words in Main Text: 5,167
2 Figures, 0 Tables, 53 References

Running Title: How microbiomes influence ecosystems
Keywords: microbiome, microbial ecology, ecosystem ecology, emergent properties, community aggregated traits

Affiliations: [1]Colorado State University, Fort Collins CO (ed.hall@colostate.edu, claudia.boot@colostate.edu, matt.wallenstein@colostate.edu) [2]Duke University, Durham, N.C. (ebernhar@duke.edu, ebgraham2@colorado.edu, jennifer.rocca@duke.edu) [3]Yale University, New Haven, CT, Dept. of Forestry and Environmental Studies (mark.bradford@yale.edu), [4]University of Minnesota, Saint Paul, MN (cotne002@umn.edu) [5]Université du Québec à Montréal, Montréal, CA (del_giorgio.paul@uqam.ca) [6]Michigan State University, Hickory Corners, MI (evanssar@gmail.com) [7]Institute of Arctic and Alpine Research, University of Colorado at Boulder [8]Pacific Northwest National Laboratory, Richland, WA, USA [9]Notre Dame University, South Bend, IN (sjones20@nd.edu) [10]Indiana University, Bloomington, IN (lennonj@indiana.edu, ken@weecology.org) [11]Brown University, Providence, RI (brooke_osborne@brown.edu) [12]University of California, Santa Barbara, Santa Barbara, CA (josh.schimel@lifesci.ucsb.edu) [13]U.S. Geological Survey, Menlo Park, CA (mwaldrop@usgs.gov) Biological Sciences Division,

* **Corresponding Author**: Dr. Ed Hall, Department of Ecosystem Science and Sustainability, Natural Resource Ecology Laboratory Campus Delivery 1499, Colorado State University, Fort Collins, CO 80523, ed.hall@colostate.edu, 970-491-2162

Contribution: All listed authors have contributed to the conceptualization, writing, and preparation of the current manuscript.

Abstract

Translating the ever-increasing wealth of information on microbiomes (environment, host, or built environment) to advance the understanding of system-level processes is proving to be an exceptional research challenge. One reason for this challenge is that relationships between characteristics of microbiomes and the system-level processes they influence are often evaluated in the absence of a robust conceptual framework and reported without elucidating the underlying causal mechanisms. The reliance on correlative approaches limits the potential to expand the inference of a single relationship to additional systems and advance the field. In this perspective piece we propose that research focused on how microbiomes influence the systems they inhabit should work within a common framework and target known microbial processes that contribute to the system-level process of interest. Here we identify three distinct categories of microbiome characteristics (microbial processes, microbial community properties, and microbial membership) and propose a framework to empirically link each of these categories to each other and the broader system level processes they affect. We posit that it is particularly important to distinguish microbial community properties that can be predicted from constituent taxa (community aggregated traits, CATs) from those properties that are currently unable to be predicted from constituent taxa (emergent properties, EPs). We discuss how a series of existing methods in microbial ecology can be applied to more explicitly elucidate properties within each of these categories and connect these three categories of microbial characteristics with each other. We view this proposed framework, gleaned from a breadth of research on environmental microbiomes and ecosystem processes, as a promising pathway with the potential to advance microbiome science across a broad range of disciplines.

26 ***Current Approaches Linking Microbial Characteristics and Ecosystem***
27 ***Processes***

28 In all ecosystems, virtually all processes are influenced by microorganisms and
29 many processes are carried out primarily/exclusively by microorganisms. This has
30 sometimes led to the assumption that a better description of the microbiome
31 (including its associated transcripts, proteins, and metabolic products) should lead to
32 a better understanding and predictions of system level processes. However, such
33 justifications assume that measureable characteristics of the microbiome (e.g. 16S
34 rRNA gene libraries, metagenomes, enzymatic activities) can inform our ability to
35 better understand and predict system-level processes. Unfortunately, additional
36 information about the microbiome does not always provide a clearer understanding
37 of ecosystem processes beyond what can be predicted by environmental factors
38 alone[1,2] .

39 Two recent meta-analyses[3,4] suggest that research at the intersection of
40 ecosystem science and microbial ecology often rely on assumed correlations
41 between microbiome characteristics and ecosystem processes and less frequently
42 test to see if those correlations are present. The first, an examination of 415 studies,
43 found little evidence that protein-encoding genes (sometimes referred to as
44 “functional genes”) or gene transcripts correlate with associated biogeochemical
45 processes[3]. Although all studies attempted (or presumed) to link microbial genes or
46 transcripts with function, only 14% measured both the copy number of genes or
47 transcripts and the corresponding process (n = 59 studies, comprising 224 individual
48 effects as most studies had measured multiple gene-function relationships). Of the
49 224 effects where both characteristics were measured, only 38% exhibited a positive
50 relationship between molecular and process measurements that many assume to

exist. The effect size for the relationship between gene copy number and process rate had an approximately normal distribution with a mean near zero[3]. This result was consistent whether functional gene or transcript abundance was used as the response variable. The second study, compiled a separate dataset of 148 studies that examined microbial membership and ecosystem processes in response to experimental manipulations[4]. Whereas 40% of included studies reported concomitant changes in microbial membership and an ecosystem process, only one third of those cases reported the relationship between microbial membership and an ecosystem process. Interestingly, of the 53 studies that posed a hypothesis or objective about links between microbial membership and ecosystem processes, more than half (53%) did not test for a statistical structure-function link[4].

Microbiomes are the engines that power system-level processes[5]. However the meta-analyses described above illustrate that the current approach to study the links between microbiome characteristics and ecosystem processes are not well formulated and relationships between microbiome characteristics and system level processes are rarely tested. When linkages are explicitly tested, significant correlations between microbiome characteristics and ecosystem processes are sometimes present, but more frequently not present [3,4]. One reason for the ambiguity between microbiome characteristics and system level processes is that many studies are conducted in the absence of a conceptual framework that illustrates how different measurable microbial characteristics relate to one another and to the system level process of interest.

Microbial characteristics can range in resolution from cellular abundance to the entire genetic potential of a species-rich community (i.e., a metagenome). Because all measureable microbial characteristics do not exist within the same plane of detail,

76 a conceptual structure is required to clarify how various microbial characteristics are
77 related to each other and to the processes they affect. Identifying correlations
78 between a microbial characteristic and an ecosystem process in the absence a
79 broader conceptual underpinning creates the potential for correlations to be
80 mediated by a third (or more) unaccounted variable(s). The absence of an underlying
81 mechanism also limits the potential for each result to be applied to additional
82 systems or to expand to broader spatial and temporal scales, for additional testing,
83 replication, and confirmation.

84 ***Challenges in Linking Microbial Characteristics and Ecosystem Processes***

85 A key challenge in linking microbial information to a system-level process is that
86 conceptual research frameworks often do not effectively align with the methods
87 being applied or the data those methods generate. For example, environmental
88 factors act on the physiology of individual organisms, which alters their competitive
89 ability, abundance, collective physiology, and ultimately their contribution to an
90 ecosystem process. However, designing an observational study or experiment from
91 this framework assumes that environmental and microbial characteristics are
92 measurable across multiple categories of ecological organization (i.e., individuals,
93 populations, and communities) at the temporal and spatial scales at which they
94 influence microorganisms (Figure 1a). In addition, the relationships between
95 environmental variables and microbial characteristics can be decoupled in both time
96 and space[4], and are often non-linear[6]. Recent immigration, phenotypic plasticity,
97 disequilibrium between the environment and the extant microbiome at the time of
98 sampling, functional redundancy, and dormancy can all mask the relationship
99 between measurable microbial characteristics and the processes microorganisms
100 influence (Figure 1b). [4,7,8,9] As micrometer scale characteristics of microbiomes (10^{-6}

101 m) are scaled to the level of ecosystems (m to km), we assume that our conceptual
102 understanding is also scalable. However, each of the aforementioned confounding
103 factors aggregate over multiple orders of magnitude often masking the very
104 relationships we seek to elucidate (Figure 1b). To formalize how measureable
105 microbiome characteristics are linked with system-level processes we have
106 conceptually defined the intersection of microbial and ecosystem ecology and
107 identified three categories of microbial characteristics that illustrate the relationship
108 among each category of microbial characteristics and how they may contribute to an
109 ecosystem process (Figure 2).

110 ***Mapping Ecosystem Processes to Microbial Characteristics***

111 Ecosystem processes are defined as a qualitative change in a pool or a flux from
112 one pool to another (e.g. NH_4^+ to NO_3^-, or dissolved organic matter to CO_2). The first
113 step to understand how the microbiome influences an ecosystem process, is to
114 define the ecosystem process of interest and each sub-process that contributes to it
115 (i.e., the set of constituent reactions that combine to determine net flux). Few, if any
116 ecosystem processes involve a single metabolic pathway, or are carried out by a
117 single organism (e.g., a notable exception being the recent discovery of commamox
118 which can independently carry out nitrification, the conversion of NH_4^+ to NO_3^-)[10,11].
119 Rather, ecosystem processes are composites of complementary or antagonistic sub-
120 processes, carried out by phylogenetically and metabolically diverse
121 microorganisms[12]. For example, net ecosystem productivity (NEP) is the balance
122 between antagonistic processes of C-fixation and C-mineralization. Each sub-
123 process of NEP can be further partitioned into a series of metabolic pathways (e.g.,
124 chemoautotrophic nitrification and photoautotrophic C-fixation or heterotrophic
125 fermentation and aerobic respiration). Partitioning each ecosystem process in this

126 hierarchical manner can continue until the sub-processes maps directly to specific
127 microbial metabolic pathways (e.g., acetoclastic methanogenesis). Subsequently
128 each of these metabolic pathways can be categorized as either phylogenetically
129 broad or narrow[13]. Broad processes are phylogenetically common (i.e., widely
130 distributed among taxa), whereas narrow processes are phylogenetically conserved
131 (i.e., limited to a specific subset of taxa). For example, denitrification and
132 photosynthesis are phylogenetically broad processes, while both methanogenesis
133 and methanotrophy are phylogenetically narrow processes (with at least one notable
134 exception[14]).

135 The second step is to identify, the controls or constraints on each constituent sub-
136 process. For example, the kinetics of a single metabolic pathway in a model
137 organism may help us understand the rate limiting steps of a narrow process, but
138 insights from model organisms are much less likely to capture the full spectrum of
139 responses of a broad process where phenotypic variation among phylogenetically
140 diverse organisms is likely to be much greater[15,16]. Defining the ecosystem process,
141 its critical sub-processes, and the known phylogenetic distribution of the metabolic
142 pathways that drive those sub-processes creates an explicit conceptual pathway that
143 links the ecosystem process to the microorganisms that contribute to it. Once the
144 ecosystem process has been conceptually partitioned into its component parts and
145 their primary controls, a concerted approach can be applied to investigate how
146 characteristics of the microbiome influence the ecosystem process of interest within
147 the complexity of a natural environment.

148 ***Categories of Microbial Characteristics***

149 We propose that attempts to elucidate the microbial contribution to system-level
150 processes needs to explicitly identify three distinct categories of microbial

151 characteristics: 1) microbial processes, 2) microbial community properties, and 3)
152 microbial membership (Figure 2). The contribution of the microbiome to ecosystem
153 processes is exerted through aggregate community properties that are shaped by
154 both microbial membership and environmental factors. This proposed framework
155 allows the researcher to clearly identify how different measurements used to
156 characterize a microbiome interact with each other and identify the potential of each
157 characteristic to elucidate the microbial contribution to the system level process.
158 These categories are hierarchically connected (Fig. 2), but they represent distinct
159 degrees of aggregation that are not simply an additive function of the previous
160 category. Furthermore, each category is potentially subject to different modes of
161 regulation, and, each category has different putative linkages to system-level
162 processes (Fig. 2). All measureable characteristics of microbial communities (e.g.,
163 abundance of cells, sequence of genes, transcripts, or proteins; enzyme expression
164 or activity) can be placed within one of the above categories, but most studies rarely
165 articulate how these measurements differ in their specificity (i.e., the level of
166 phylogenetic resolution at which they are applied), precision (i.e., the ability of the
167 method to repeatedly describe the characteristic of interest), or context (i.e., how a
168 characteristic relates to other characteristics or the ecosystem within which they
169 were measured). This conceptual structure that orientates each microbial category
170 within a broader context creates the opportunity to improve the design of
171 observational and experimental studies in micobiome research.
172 *Microbial Processes* - Microbial processes are the collective metabolism of the
173 microbiome that contribute to changes in pools and fluxes of elements or compounds
174 (i.e., Figure 2, Letter K). This is the level of microbial information that can most
175 readily be incorporated into system-level models because many microbial processes

represent the key sub-processes that contribute to a particular ecosystem process (e.g., methanogenesis + methanotrophy ≈ methane efflux). Commonly measured microbial processes in ecosystem science include nitrogen fixation, denitrification, nitrification, phosphorus uptake and immobilization, carbon fixation, and organic carbon mineralization. The rates of many microbial processes can be approximated through physiological assays (e.g., biological oxygen demand to estimate microbial community respiration), and while they do not open the “black box” of the microbial community, they do directly quantify the microbial contribution (or at least the potential contribution) to changes in resources moving through the box. Microbial processes can be distinguished from other microbial characteristics because they are all rates (i.e., have time in the denominator) and require a bioassay to estimate.

Assays used to estimate microbial processes are often logistically challenging, require manipulations that inevitably deviate from the *in situ* conditions, and often depend on the environment from which the microbiome was sampled. For example, the relationship between temperature and microbial processes such as enzyme activity and phosphorus use efficiency (PUE) vary across latitudinal gradients[17] and among seasons[18]. Thus, observations of the effect of temperature on either enzyme activity or PUE depend on where (e.g., at what latitude) and when (e.g., during which season) they were measured. In the absence of an understanding of the underlying physiological mechanism (e.g., the physiological change that allows a community to perform differently at different temperatures), the relationship between and environmental driver (e.g., temperature) and a microbial process must be measured through a direct assay at each location and at each time. This limits the inference possible from relying only on measurements of microbial processes alone.

200 *Microbial Community Properties* - Microbial community properties include a broad
201 set of microbial characteristics such as community biomass or biomass elemental
202 ratios (e.g., biomass C:N or C:P ratios) and the majority of phylogenetically
203 undifferentiated aggregate sequence based measurements (e.g., gene abundance,
204 metagenomes, transcriptomes). Microbial community properties (Figure 2) represent
205 an integrated characteristic of the microbiome that has the potential to predict or at
206 least constrain the estimates of microbial processes. For example, microbial
207 community biomass C:N (a community property) indicates a microbiome's potential
208 to mineralize or immobilize N.[19] Community biomass stoichiometry has been shown
209 to be a useful predictor of nutrient immobilization or mineralization during litter
210 decomposition[19], and in soils can predict both respiration and N-mineralization better
211 than microbial biomass alone[20]. The power of biomass elemental ratios to explain
212 nutrient cycling has also been shown in freshwater[21] and marine ecosystems[22]
213 including the seminal paper that demonstrated the similarities between the
214 stoichiometry of marine algal biomass and that of the dissolved fraction of nutrients
215 in the ocean[23].
216 Microbial community properties can be separated into two categories, emergent
217 properties (EPs) and community aggregated traits (CATs). Emergent properties have
218 been used to refer to a variety of phenomena in ecology, however here we use
219 emergent properties as it has been defined by Salt (1979)[24]: "An emergent property
220 of an ecological unit is one which is wholly unpredictable from observation of the
221 components of that unit", which is consistent with its contemporary use in microbial
222 ecology.[25] For example, the potential importance of emergent properties to influence
223 ecosystem processes has been demonstrated in a series of experimental flow-
224 through flumes that mimicked development and metabolism of stream biofilms[26].

225 Transient storage (i.e., an increase in residence time of the water and its solutes
226 near the biofilm relative to the flow around it) increased as the microbial biofilm
227 density increased[26]. Microbial biofilm formation is an EP[27] that affects the important
228 ecosystem process of hydrological transient storage[26]. Another example of an
229 emergent property is the relative abundance of a certain traits within a microbiome.
230 Trait based approaches have a rich history in ecology and have been increasingly
231 applied to address questions in multiple areas of microbial ecology.[28] For example,
232 specific functions (i.e. uptake of an individual organic substrate) are associated with
233 traits which can involve multiple genes, among different taxa, all capable of
234 performing the function albeit with differences in the underlying physiology
235 and efficiency. The distribution and expression of these functional gene variants
236 generates a trait structure among microbiomes, which determines the overall
237 performance of the microbiome for that given function (i.e. uptake of a given organic
238 substrate), but which cannot be predicted simply from the presence of the taxa that
239 carry the genes conferring that trait.[29] While characterization of EPs may improve the
240 understanding of microbial processes (Figure 2, Letter G) they cannot, in principle,
241 be estimated or predicted on the basis of the constituent taxa (i.e. membership)
242 alone (Figure 2, Letter F), and thus must remain as an intermediary between
243 environmental drivers (Figure 2, Letter C) and microbial processes (Figure 2, Letter
244 G).
245 Unlike EPs, CATs can potentially be estimated from characteristics of their
246 constituents and provide a pathway to link microbial community membership to the
247 community properties that drive microbial processes (Figure 2, Letter E)[30]. For
248 example, CATs may include commonly measured community properties such as
249 functional gene abundance as estimated from qPCR (e.g., pmoA which encodes a

250 subunit of the enzyme involved in methane oxidation, can be used to estimate
251 potential for methanotrophy and as a phylogenetic marker for methanotrophs)[31]. A
252 recent perspective article discussed the role of CATs in microbial ecology and noted
253 a series of additional CATs (e.g., maximum growth rate, dormancy, osmoregulation)
254 that could be inferred from metagenomic data of the extant community[30].
255 Understanding when, and which, community properties that shape microbial
256 processes can be predicted by membership is a critical research question, and an
257 important step in understanding how the microbiome contributes to system level
258 processes. Whether or not a community property is an EP or a CAT is an exciting
259 area of research and provides an important framework to advance research at the
260 microbial-ecosystem nexus. New approaches, like studying higher-level interactions
261 in ecological communities could help understand how a microbiome's constituents
262 interact to from emergent properties.[32] This is not a trivial task, yet a suite of existing
263 methods, discussed below, already provides the ability to directly pursue this
264 challenge.
265 *Microbial Community Membership* **-** Although the now commonplace analysis of
266 community membership by sequencing phylogenetic marker genes (e.g., regions of
267 the 16S, 18S, or ITS genes) or suites of phylogenetically conserved protein
268 sequences identifies constituent microbial taxa, the direct coupling of microbial
269 phylogeny to physiology and ecology remains elusive (Figure 2, Letter H).[33,34,35] In
270 general the paucity of associated physiological data or information on population
271 phenotypes that accompany phylogenetic sequence data limits the system-level
272 inference that is possible from analyses of community membership. This constrains
273 our ability to attribute microbial processes to community membership of even
274 relatively simple environmental consortia. Whereas it is clear that microbial

275 populations are not randomly distributed in space and time[35], and that some
276 microbial traits are conserved at coarse taxonomic scales [28,36,37], the physiological
277 mechanisms underlying non-random distributions of microbial taxa across
278 environmental gradients is often unknown. The limited understanding of the
279 metabolism of most bacterial phyla limits an explicit understanding between the
280 organism's abundance and its role in the microbial process that contributes to an
281 ecosystem process.

282 ***A Path Forward***

283 We suggest that a challenging but necessary step for microbiome science is to
284 move away from identifying correlative relationships between characteristics of the
285 microbiome and system level processes, and towards identifying more causative and
286 mechanistic relationships. The conceptual diagram (Figure 2) is a road map to
287 organize and link the diverse suite of measureable microbial characteristics that are
288 currently available to researchers. Figure 2 does not represent how these
289 components necessarily interact in the environment; rather it is a map that identifies
290 potential links between measureable microbial and system-level characteristics that
291 can help structure our exploration of how microorganisms influence the systems they
292 inhabit. Ecosystem ecology has traditionally been confined to interactions between
293 environmental parameters and ecosystem processes (depicted within the horizontal
294 arrow, Figure 2). Similarly, microbial ecology (depicted within the vertical arrow,
295 Figure 2) has historically focused on phylogenetically undefined aspects of microbial
296 communities (e.g., bacterial abundance) and microbial processes (e.g., bacterial
297 production) or on the physiology of microbial isolates (e.g., sulfate reducing bacteria)
298 or the collective physiology of highly reduced communities with known membership
299 (e.g., waste water treatment microbiome). The routine inclusion of sequence-based

300 approaches in studies of environmental microorganisms has lead to an increasingly
301 detailed description of the world's microbiomes and an increasing interest in how
302 constituents of those communities interact to influence the system as a whole.
303 The drive to include microbial characteristics into system-level science has led to a
304 range of approaches for linking characteristics of the microbiome to ecosystem
305 processes. Direct connections between microbial membership and ecosystem
306 processes (Figure 2, Letter I), or community properties and ecosystem processes
307 (Figure 2, Letter J), have proven consistently difficult to establish[3,4]. We propose 1)
308 identifying which microbial processes are likely to contribute to ecosystem-level
309 pools and fluxes *a priori* (Figure 2, Letter K), 2) determining which microbial
310 community properties best describe and predict these microbial processes (Figure 2,
311 Letter G), and 3) identifying whether the community properties that best describe
312 each process are a CAT or an EP (Community Properties, Figure 2). If the
313 community property is a CAT then exploring the link between microbial membership
314 and community properties may lead to further understanding and perhaps an
315 enhanced predictive power (Figure 2, Letter E). However, if the community property
316 is an EP elucidating the microbial membership that contributes to the EP is unlikely
317 to improve understanding of the drivers of that community property (Figure 2, Letter
318 F) and understanding how environmental drivers structure the EP will be more
319 insightful. Formalizing microbiome research into a structured, conceptual framework
320 will allow the research community to better focus on potential links between
321 microbiome characteristics and system-level processes that are most likely to be
322 detected empirically. This approach will also allow researchers working in different
323 systems to test the same pathways among defined microbiome characteristics and
324 thus increase the possibility of understanding the casual mechanism (or absence of

325 causality) for observed correlations. Equally as important, we suggest that
326 attempting to link community membership (Figure 2, Letter I) or community
327 properties (Figure 2, Letter J) directly to ecosystem processes is by definition
328 correlative and therefore a less powerful approach to integrating microbiome
329 characteristics into system-level science. Thus future research endeavors will be
330 most powerful if they focus on elucidating connections through the complete path of
331 microbial ecology (Figure 2, blue arrow, Letters E, F, and G) and not direct
332 connections between microbial membership or community properties and ecosystem
333 processes (Figure 2, Letter I and J).

334 ***Applying and Testing the Proposed Framework***

335 Applying and testing the proposed framework will depend on the ability to more
336 robustly evaluate each category of microbial characteristics and to directly measure
337 the arrows that connect each category (Figure 2). Both labeling/sorting approaches
338 and phenotypic description of isolates provide an opportunity to better understand
339 how microbial membership contributes to community properties (Figure 2, Letter E or
340 F). Labeling and cell sorting approaches (e.g., fluorescent in situ hybridization (FISH)
341 coupled with flow cytometry cell sorting[38], or immunocapture such as with
342 bromodeosyuridine, BrdU)[39] provide powerful tools to constrain the complexity of the
343 microbiome and directly test hypotheses that link membership to community
344 properties or microbial processes. Labeling and sorting techniques allow the cells
345 that can be targeted with a stain or other label to be separated from the broader
346 community and then assayed for membership or phenotypes such as activity or
347 biomass composition. For example, a study of an Arctic Ocean bacterial community
348 labeled the actively growing component of the community using BrdU and then
349 separated those populations from the rest of the community using an immunocapture

350 technique to better understand the portion of the microbiome that was driving
351 community dynamics[39]. By simplifying the community, researchers were able to link
352 membership to secondary production (a microbial process, Figure 2, Letter H) and
353 begin to better understand which members of the community were contributing to
354 secondary production.
355 In addition, physiological studies of isolates from a broader distribution of
356 representative phyla are key to advancing our understanding of how membership
357 contributes to community properties (Figure 2, Letter E). Finding isolates that are
358 representative of important community properties has the potential to better
359 understand phenotypic plasticity and how constituent populations do (CATs) or do
360 not (EPs) contribute to a community property[15]. For example, work on the marine
361 bacterioplankton SAR11 has led to an increased understanding of how this
362 ubiquitous member of the marine microbiome interacts with elemental cycles in the
363 open ocean[40]. Similarly, a rich body of work on multiple isolates of the comparably
364 ubiquitous photoautotroph *Prochlorococcus* has advanced our understanding of the
365 ecology and physiology of one of the most abundant phototrophs on the planet[41].
366 Detailed studies of isolates of common environmental OTUs have clearly
367 demonstrated immense variation within a given OTU (i.e., “microdiversity”) that in
368 part explains the challenge of linking membership to a community property[16]. For
369 example, work on *Prochlorococcus has led to* a better understanding of how
370 ecotypes within a single taxonomic unit (OTU) can lead to specialization in
371 temperature and substrate affinity[41]. OTUs that form a substantial portion of the
372 microbiome’s sequence abundance provide potential candidates for further
373 investigation of possible phenotypic plasticity and or microdiversity[16]. For example, a
374 single phylotype of the class *Spartobacteria* within the phyla *Verrucomicrobia* was

375 found to be present in a broad range of soil ecosystems and comprised as much as
376 31% of all 16S sequences returned from prairie soils[42], making it an excellent
377 candidate for targeted isolation and physiological studies. Whereas it is challenging
378 to isolate and culture many microorganisms from the environment, existing
379 approaches to isolation have been led to successful isolation of both abundant and
380 rare members of environmental microbiomes. A recent study isolated members of an
381 apple orchard soil microbiome where most isolates were from the least abundant
382 members of the community[43]. Previous studies have had success isolating members
383 of the pelagic marine microbiome by using filter-sterilized seawater with a dilution to
384 extinction approach[44]. Thus there is a potential to target both abundant[40,41] and rare[43]
385 members of diverse microbiomes to learn more about their influence on community
386 level properties. Studies of environmental isolates are essential in building a broader
387 understanding of how community membership does or does not contribute to
388 community properties (H2, Letter E and F).
389 The most commonly measured microbial characteristics can be associated with
390 one of the three categories of microbial characteristics defined here. However, the
391 key to moving from a correlative and descriptive approach to a causative and
392 mechanistic approach comes in measuring the arrows represented by letters in
393 Figure 2. There is a suite of powerful methods already being employed in microbial
394 ecology that can actively measure many of the arrows illustrated in Figure 2. These
395 include: stable isotope probing of mixed communities[45], single cell methods that can
396 assay cells in the physiological state they occur in in the environment[46], and labeling
397 individual cells with stable isotopes for single cell analyses[46]. Studies that use stable
398 isotope probing or any form of tracking stable isotopes into a population have been
399 successful in linking microbial membership to microbial processes (Figure 2, Letter

400 H). For example, a study of sulfate reduction in a Scottish peatland using SIP
401 revealed that a single species of *Desulfosporosinus* was most likely responsible for
402 the totality of sulfate reduction within the peatland even though it only comprised
403 0.006% of the retrieved 16S rRNA gene sequences[47]. In this case, the
404 *Desulfosporosinus* species represented the only known sulfate reducer within the
405 community and thus the kinetics of this organism seemingly defined the kinetics of
406 sulfate reduction for the entire system. Whereas this is a single example of using SIP
407 to link microbial membership to microbial processes (Figure 2, Letter H), there is a
408 suite of culture-free techniques (such as Raman microspectroscopy (MS),
409 NanoSIMS, or X-ray microanalysis, XRMA) that complement sequence-based
410 microbiome analyses by reporting on the physiological and phenotypic
411 characteristics of individual cells *in situ* [46,48,49]. For example, Raman MS has the
412 ability to elucidate the macromolecular composition of uncultured individual cells *in*
413 *situ*[46]. Incorporation of stable isotopes into a cell's macromolecules can be visualized
414 as a shift in the Raman spectra. This provides information not only on which cells are
415 incorporating the substrate but what macromolecular pool those substrates are being
416 allocated to. Similarly, NanoSims allows for visualization of isotopes within a cell,
417 and while it cannot be used to identify which macromolecule pool an isotopic label
418 went into, it can visualize which cells are actively incorporating isotopically labeled
419 substrates. Both Raman and NanoSims can be coupled with a range of *in situ*
420 hybridization techniques (e.g., fluorescent in situ hybridization, FISH) to identify
421 which populations are contributing to community properties (Figure 2, Letter E) or
422 microbial processes (Figure 2, Letter H). For example, the study of a microbial
423 consortia from the Sippewissett Salt Marsh on the coast of Massachusetts, USA
424 used a combination of FISH and NanoSIMs to confirm a syntrophic association

425 between a population of autotrophic purple sulfur bacteria and heterotrophic sulfate
426 reducing bacteria (SRB)[50]. Whereas several pieces of evidence pointed to a
427 syntrophic association, the authors confirmed the association by first using FISH to
428 visualize the physical association of each population. Further analysis with
429 NanoSIMS after incubation with ^{34}S enriched sulfate and ^{14}C enriched bicarbonate
430 confirmed the presence of carbon fixation by the purple photosynthetic bacteria with
431 sulfide as the electron donor was coupled to the reduction of sulfate by the SRB.
432 Using existing methods of confirmatory ecophysiology allows for direct
433 measurements of the arrows connecting membership with microbial processes in a
434 stable microbial consortia (in this case both carbon fixation and sulfate reduction,
435 Figure 2, Letter H). These approaches applied in concert with sequence-based
436 analyses have the potential to empirically link the categories of microbial information
437 defined here (Figure 2), moving microbiome science from a descriptive and
438 correlative approach to a mechanistic and causative approach.
439 These culture free approaches also create the potential to begin to determine
440 which community properties are EPs, and which are CATs. For example, microbial
441 community biomass stoichiometry (e.g., biomass C:N or C:P) cannot currently be
442 predicted (or even constrained) from a list of it's constituent taxa (Figure 2, Letter F).
443 However, microbial biomass stoichiometry is a community property with power to
444 predict the microbial contribution to nutrient cycling (Figure 2, Letter G).[19-22]
445 Independently, the biomass stoichiometry of microbial isolates grown on the same
446 media has been shown to differ among different taxa suggesting a relationship
447 between an organisms' identity and the elemental composition of its biomass.[51,52]
448 However, there is an abundance of evidence that suggests that the biomass
449 stoichiometry of many isolates is a function of the stoichiometry of the media they

450 were cultured on[52]. Electron dispersal spectroscopy (EDS) has the power to
451 measure the C:N:P of individual bacterial cells growing *in situ* (i.e. not in culture)[49].
452 The potential to couple EDS analysis with a phylogenetic label presents the
453 opportunity to assay mixed microbial communities and assess the link between
454 phylogenetic identity and biomass stoichiometry under natural conditions[53]. Thus,
455 community biomass stoichiometry can potentially be deconstructed into the biomass
456 stoichiometry of its constituent taxa[53]. This approach would provide a direct link
457 between community membership and a community property (e.g., biomass C:N,
458 Figure 2, Letter E), that influences an important microbial process (i.e. nutrient
459 recycling).

460 ***Designing microbiome research***

461 It is critical that we recognize the influence of the taxonomic and functional
462 composition of the microbiome is exerted through multiple pathways, some that are
463 direct and can be readily identified, some that are indirect, and mediated by complex
464 interactions at the community level. We must also recognize that the influence of
465 microbiomes will vary depending on the system-level process in question, because
466 analysis of microbial characteristics may simply not improve the environmentally-
467 based prediction of certain processes (Figure 2, Letter A), whereas other system-
468 level processes may indeed benefit from the inclusion of microbial characteristics
469 including membership. For the latter, the challenge then is to determine which
470 microbial category is the most relevant predictor of the system level process of
471 interest: microbial processes, community properties, or microbial membership.
472 Establishing the links between these microbial dimensions (Figure 2, Letters E, F,
473 and G) further contributes to our understanding of the mechanistic underpinnings

474 that affect system level processes and thus will have greater explanatory power in a
475 broader range of systems.
476 The framework presented here provides one approach to formalize inquiry across
477 microbiome science and encourages empirical linkages between the presence of
478 organisms in a system and the processes that characterize that system. Whereas we
479 draw examples from environmental microbiomes and the ecosystems they inhabit,
480 this structured approach has the potential to benefit the analysis of microbiomes
481 associated with other systems such as host organisms and those of the built
482 environment. As important as establishing causal links among microbial
483 membership, community properties, microbial processes, and ecosystem processes,
484 is determining when these links are unlikely to be present. Research that
485 indiscriminately seeks to identify correlations, which does not recognize the
486 hierarchy of effects, and that places all metrics on an equal plane are susceptible to
487 confirmation bias and will continue to yield conflicting and ambiguous results that not
488 only fail to provide new insight into ecosystem processes, but also blurs the
489 connections that do exist. We suggest that rather than looking for linkages among
490 microbiome membership and system-level processes in every study, research efforts
491 would benefit from strategically targeting the linkages and processes for which an *a*
492 *priori* understanding of microbial physiology should allow us to improve our
493 understanding of the ecosystem process.

Acknowledgements

This work is a product of the Next Generation of Ecosystem Indicators Working Group, supported by the USGS John Wesley Powell Center for Synthesis and Analysis. Preparation of this manuscript was supported by NSF DEB IOS #1456959 awarded to EKH. Chuck Pepe-Ranney and Ariane Peralta provided valuable feedback on previous versions of this manuscript. This paper is dedicated to Diana Nemergut, an integral part of our working group who passed away during the preparation of this manuscript. She is one of a kind.

Figures

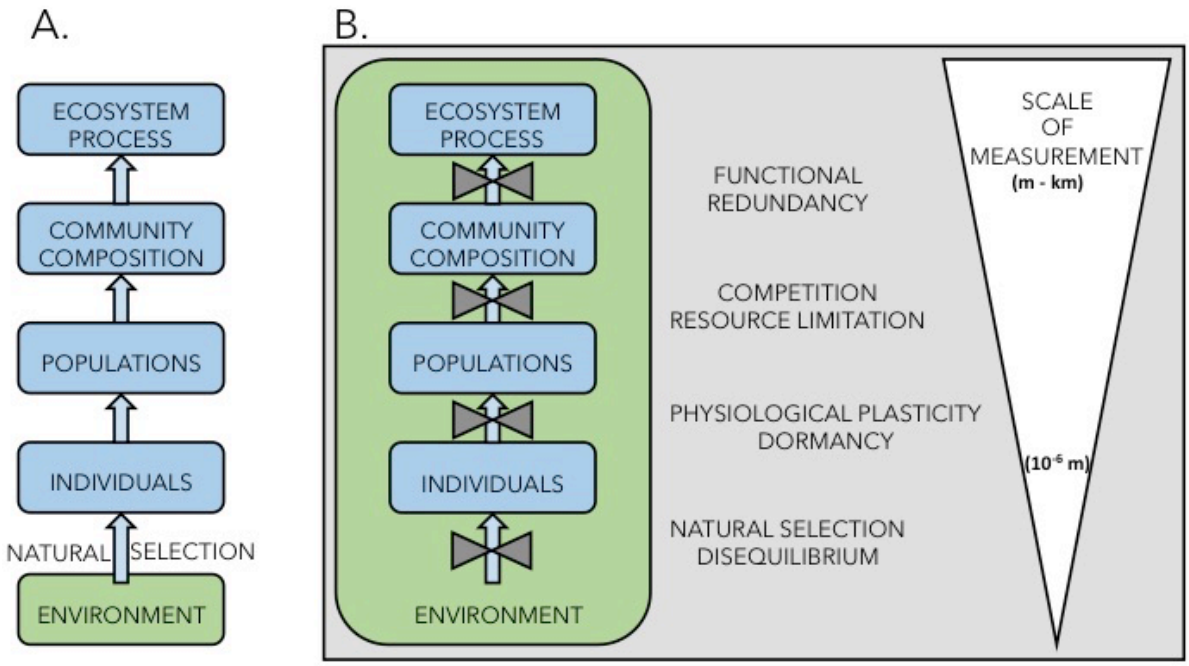

Figure 1 Diagram of microbial-ecosystem linkages A) how linkages are commonly conceptualized across levels of ecological organization and B) the series of ecological phenomena that create challenges when attempting to link metrics from one level of ecological organization to the other.

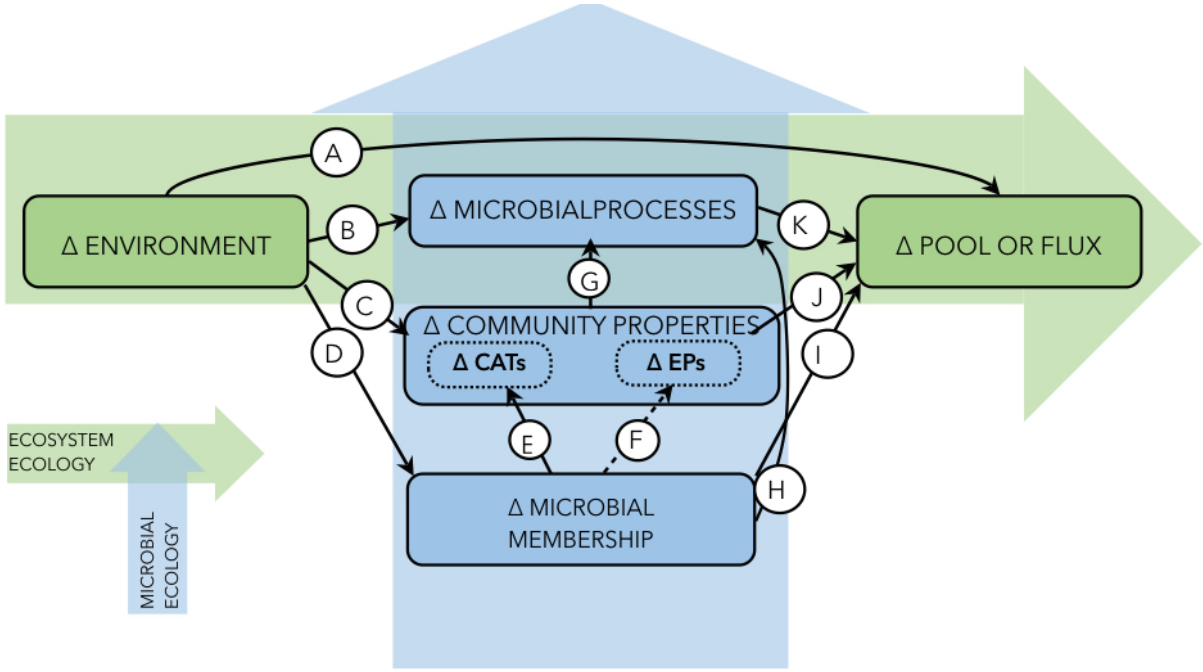

Figure 2 Shown is a conceptual map of the intersection between microbial (vertical) and ecosystem (horizontal) ecology with each of the three categories of microbial characteristics (microbial processes, community properties, and microbial membership) as defined in the text. We argue for an increased focus on studies that elucidate pathways E, F, and G. In addition, we note that pathways I and J are less likely to effectively incorporate microbiome characteristics into system-level science. The delta symbol in each category indicates an emphasis on how changes within a category may lead to a change in a connected category. The dotted arrow for letter F denotes that many emergent properties cannot currently be linked to membership and is an important area for active research.

References

1.Graham E. B., Wieder, W.R., Leff, J.R., Weintraub, S. W., Townsend, A.R., Cleveland, C.C., Philippot L., Nemergut D.R. (2014) Do we need to understand microbial communities to predict ecosystem function? A comparison of statistical models of nitrogen cycling processes Soil Biology and Biochemistry 68, 279–282

2.Graham E. B. et al. (2016) Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? Frontiers in Microbiology 7: 214 doi:10.3389/fmicb.2016.0021

3.Rocca J.D., Hall E.K., Lennon J.T., Evans S.E., Waldrop M.P., Cotner J.B., Nemergut D.R., Graham E.B., Wallenstein M.D. (2015) Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed ISME J. 9: 1693–1699

4.Bier R.L., E.S. Bernhardt, CM Boot, EB Graham, EK Hall, JT Lennon, D Nemergut, B Osborne, C Ruiz-González, JP Schimel, MP Waldrop, MD. Wallenstein (2015) How are we forging conceptual, analytical, and mechanistic links between microbial community structure and ecosystem process? FEMS microbiology ecology FEMS microbiology ecology 91(10): fiv113

5. Falkowski, P.G., Fenchel, T., and DeLong, E.F. (2008) The microbial engines that drive earth's biogeochemical cycles Science 320: 1034-1038

6.Felip M., Pace M.L., Cole, J.J. (1996) Regulation of planktonic bacterial growth rates: The effects of temperature and resources. Microb. Ecol. 31(1): 15-28.

7. Lennon, J. T. and S. E. Jones (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. Nature Microbiology Reviews 9: 119-130.

8. Comte, J., Fauteux, L. and del Giorgio, P.A. (2013) Links between Metabolic Plasticity and Functional Redundancy in Freshwater Bacterioplankton Communities. Frontiers in Microbiology 4: 112. doi: 10.3389/fmicb.2013.00112

9. Adams, H. E., Crump, B.C. and Kling, G.W. (2014) Metacommunity Dynamics of Bacteria in an Arctic Lake: The Impact of Species Sorting and Mass Effects on Bacterial Production and Biogeography. Frontiers in Microbiology 5 (2014): 82. 10.3389/fmicb.2014.00082

10. Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierhellig, J., Bulaev, A., Kirkegaard, R. H., von Bergen M., Rattei, T., Bendinger, B., Nielsen, P.H., Wagner, M. (2015) Complete nitrification by Nitrospira bacteria Nature 528: 504-509 doi: 10.1038/nature16461

11. van Kessel, Jaartje A.H.J., Speth, D.R., Albertsen, M., Nielsen, P.H., Op den Camp, H.J.M., Kartal, B.,Jetten, M.S.M., Lueker, S. (2015) Complete nitrification by a single microorganism 528: 555-559 doi: 10.1038/nature16549

12. Schimel, J.P., J. Bennett, and N. Fierer. (2005) Microbial community composition and soil N cycling: is there really a connection? In: Biological diversity and function in soils. Bardgett, R.D., D.W. Hopkins, and M.B. Usher (Eds.) Cambridge University Press. pp. 171-188

13. Schimel, J.P. (1995) Ecosystem consequences of microbial diversity and community structure in Arctic and Alpine Biodiversity: Patterns, Causes and Ecosystem Consequences Ecological Studies Volume 113, pp 239-254

14. Lenhart K., M. Bunge, S. Ratering, T. R. Neu, I. Schüttmann, M. Greule, C. Kammann, S. Schnell, C. Müller, H. Zorn & F. Keppler. (2012) Evidence for methane production by saprotrophic fungi Nature Comm. 3:1046

15. Hagstroem, A., Azam, F., Berg, C., Zweifel U.L. (2017) Isolates as models to study bacterial ecophysiology and biogeochemistry Aquat. Microb. Ecol. 80:15-27
16. Larkin, A.A, and Martiny, A. (2017) Microdiversity shapes the traits, niche space and biogeography of microbial taxa Environ. Microbiol. Rep. 9(2): 55-70 doi: 10.111/1758-2229
17. German D.P., Marcelo, K.R.B., Stone, M.M and Allison, S.D. (2012) The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study Global Change Biology 18: 1468-1479
18. Hall, E. K., A. R. Dzialowski, S. M. Stoxen, and J. B. Cotner. (2009) The effect of temperature on the coupling between phosphorus and growth in lacustrine bacterioplankton communities. Limnol. and Oceanogr. 54: 880-889.
19. Manzoni, S., R. B. Jackson, J. A. Trofymow, and A. Porporato. (2008) The global stoichiometry of litter nitrogen mineralization. Science 321: 684-686.
20. Buchkowski, R.W., Schmitz, O.J., Bradford, M.A. (2015) Microbial stoichiometry overrides biomass as a regulator of soil carbon and nitrogen cycling. Ecology 96: 1139-1149.
21. Elser, J.J., Chrzanowski, T.H., Sterner, R.W., Schampel, J.H., Foster D.K. (1995) Elemental ratios and the uptake and release of nutrients by phytoplankton and bacteria in three lakes of the canadian shield Microbial Ecology 29: 145-162
22. Goldman, J.C., Caron, D.A., and Dennett, M.R. (1987) Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C:N ratio Limnology and Oceanography 32 (6): 1239-1252
23. Redfield, A.C (1958) The biological control of chemical factors in the environment American Scientist 46 (3): 205-221
24. Salt, G.W. (1979) A comment on the use of the term emergent properties American Naturalist 113: 145-148
25. Konopka, A. (2009) What is microbial community ecology? ISME J., 3 (11): 1223-12230
26. Battin, T., L. A. Kaplan, L. Newbold, X. Cheng, and C. Hansen. (2003) Effects of current velocity on the nascent architecture of stream microbial biofilms. Appl. and Env. Microbiol. 69: 5443-5452.
27. Flemming, H-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S.A., Kjelleberg, S. (2016) Biofilms: an emergent form of bacterial life Nature Reviews Microbiology 14: 563-575
28. Martiny JBH, Jones SE, Lennon JT, Martiny AC (2015) Microbiomes in light of traits: a phylogenetic perspective. Science 350: doi: 10.1126/science.aac9323
29. Ruiz-González, C., Niño-García, J. P., Lapierre, J. F., & del Giorgio, P. A. (2015) The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. Global Ecology and Biogeography *24*(12): 1487-1498.
30. Fierer, N., A. Barberán, Laughlin, D. (2014) Seeing the forest for the genes: Using metagenomics to infer the aggregated traits of microbial communities. Frontiers in Microbiology 5: 614
31. Judd, C.R., Koyama, A., Simmons, M.P., Brewer, P., and von Fischer, J.C. (2016) Co-variation in methanotroph community composition and activity in three temperate grassland soils Soil Biol. and Biochem. 95:78-86
32. Grilli, J., Barabas G., Michalska-Smith, M.J., and Allesina, S. (2017) Higher-order interactions stabilize dynamics in competitive network models Nature. 2017 548: 210–213

33. Newton, R. J, S. E. Jones, A. Eiler, K. D McMahon, Bertilsson, S. (2011) A guide to the natural history of freshwater lake bacteria. Microbiol. and Molec. Biol. Rev. 75:1, doi:10.1128/MMBR.00028
34. Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D.A., Finstad, K.M., Amundson, R., Thomas, B.C., Banfield, J.F. (2016) 1: 16048 doi: 10.1038/nmicrobiol.2016.48
35. Martiny, J.B.H., B. Bohannan, J. Brown, R. Colwell, J. Fuhrman, J. Green, M.C. Horner-Devine, M. Kane, J. Krumins, C. Kuske, P. Morin, S. Naeem, L. Ovreas, A.-L. Reysenbach, V. Smith, J. Staley. (2006) Microbial biogeography: Putting microorganisms on the map. Nature Reviews Microbiology 4: 102-112
36. Lennon JT, Aanderud ZA, Lehmkuhl BK, Schoolmaster DR (2012) Mapping the niche space of soil microorganisms using taxonomy and traits. Ecology 93: 1867-1879
37. Treseder KK, Lennon JT (2015) Fungal traits that drive ecosystem dynamics. Microbiology and Molecular Biology Reviews 79: 243-262
38. Czechowska, K., Johnson, D.R., van der Meer, J.R. (2008) Use of flow cytometric methods for single-cell analysis in environmental microbiology Curr. Opin. In Microbiol. 11:3, 205-212
39. Galand P.E., L. Alonso-Sáez, S. Bertilsson, C. Lovejoy, and E. O. Casamayor (2013) Contrasting activity patterns determined by BrdU incorporation in bacterial ribotypes from the Arctic Ocean in winter Front Microbiol. (2013); 4: 118
40. Giovannoni S.J. 2017 SAR11 Bacteria: The most abundant plankton in the oceans. Annual Review Marine Science 9: 231-255
41. Biller, S.J., Berube, P.M., Lindell, D., Chisholm, S.W. (2015) *Procholorcoccus:* The structure and function of collective diversity Nature Reviews Microbiology 13:13-27
42. Brewer, T.E., Handley, K.M., Carini, P., Gilbert, J.A., Fierer, N. (2016) Genome reduction in an abundant and ubiquitous soil bacterium ‘Candidatus Udaeobacter copiosus’ Nature Microbiology 2: 16198
43. Shade, A., Hogan, C.S., Klimowicz, A.K., Linske, M., McManus, P.S., and Handelsman, J. (2012) Culturing captures members of the soil rare bioshphere Environmental Microbiology 14(9): 2247-2252
44. Schut, F., De Vries, E.J., Gottschal, J.C., Robertson, B.R., Harder, W., Prins, R.A., Button, D.K. (1993) Isolation of typical marine bacteria by dilution culture: Growth, maintenance and characteristics of isolates under laboratory conditions Applied and Environmental Microbiology 59(7): 2150-2160
45. Neufeld J.D., Vohra J., Dumont M.G., Lueders T., Manefield M., Friedrich M.W., Murrell J.C. (2007) DNA stable-isotope probing. Nat Protoc. 2:4, 860-866
46. Wagner, M. (2009) Single cell ecophysiology of microbes as revealed by Raman microspectroscopy or secondary ion mass spectrometry imaging. Annu Rev Microbiol 63:411-429.
47. Pester, M., Bittner, N., Deevong, P., Wagner, M., and Loy, A. (2010) A 'rare biosphere' microorganism drives sulfate reduction in a peatland. ISME J. 4:1591–1602.
48. Behrens, S., Kappler, A., Obst M. (2012) Linking environmental processes to the in situ functioning of microorganisms by high-resolution secondary ion mass spectrometry (NanoSIMS) and scanning transmission X-ray microscopy (STXM) Env. Microbiol. 14:11, 2851-69

49. Norland S, Fagerbakke K, Heldel M. (1995) Light element analysis of individual bacteria by X-ray microanalysis. Appl. Environ. Microbiol. 61:1357–62.
50. Wilbanks, E.G., Jaekel, U., Salman, V., Humphrey, P.T., Eisen, J.A., Faciotti, M.T., Buckley, D.H., Zinder, S.H., Druschel, G.K., Fike, D.A., Orphan,V.J., (2014) Microscale sulfur cyclingin the photophrophic pink berry consortia of the Sippewissett Salt Marsh Environ. Micro 16: 3398-3415
51. Mouginot, C., Kawamura, R., Matulich, K.L., Berlemont, R., Allison, S.D., Amend, A.S. and Martiny, A.C. (2014) Elemental stoichiometry of fungi and bacteria strains from grassland leaf litter Soil Biol. & Biochem. 76, 278-285
52. Godwin, C.M. and Cotner, J.B. (2015) Stoichiometric Flexibility in Diverse Aquatic Heterotrophic Bacteria Is Coupled to Differences in Cellular Phosphorus Quotas. Frontiers in Microbiology 6: doi:10.3389/fmicb
53. Hall, E.K., Maixner F., Franklin O., Daims, H., Richter, A. and Battin, T. (2011) Linking microbial and ecosystem ecology using ecological stoichiometry: A synthesis of conceptual and empirical approaches Ecosystems 14: 261-273