Running head: Bacterial diversity and productivity

Microhabitats shape diversity-productivity relationships in freshwater bacterial communities

Marian L. Schmidt[1,#], Bopaiah A. Biddanda[2], Anthony D. Weinke[2], Edna Chiang[1,3], Fallon Januska[2], Ruben Props[1,4] & Vincent J. Denef[1]

[#]Corresponding author: marschmi@umich.edu

[1]Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University Ave., Ann Arbor, MI 48109, USA; [2]Annis Water Resources Institute, Grand Valley State University, 740 West Shoreline Drive, Muskegon, MI 49441, USA; [3]Department of Bacteriology, University of Wisconsin, Madison, WI, USA; [4]Center for Microbial Ecology and Technology (CMET), Department of Biochemical and Microbial Technology, Ghent University, Gent, Belgium

Abstract

Eukaryotic communities commonly display a positive relationship between biodiversity and ecosystem function (BEF). Based on current studies, it remains uncertain to what extent these findings extend to bacterial communities. An extrapolation from eukaryotic relationships would predict there to be no BEF relationships for bacterial communities because they are generally composed of an order of magnitude more taxa than the communities in most eukaryotic BEF studies. Here, we sampled surface water of a freshwater, estuarine lake to evaluate BEF relationships in bacterial communities across a natural productivity gradient. We assessed the impact of habitat heterogeneity - an important factor influencing eukaryotic BEFs - on the relationship between species richness, evenness, phylogenetic diversity, and heterotrophic productivity by sampling co-occurring free-living (more homogenous) and particle-associated (more heterogeneous) bacterial habitats. Diversity measures, and not environmental variables, were the best predictors of particles-associated heterotrophic production. There was a strong, positive, linear relationship between particle-associated bacterial richness and heterotrophic productivity that was strengthened when considering evenness. There were no observable BEF trends in free-living bacterial communities. In contrast, per-capita but not community-wide heterotrophic productivity increased across both habitats as communities were composed of taxa that were more phylogenetically clustered. This association indicates that communities with more phylogenetically related taxa have higher per-capita heterotrophic production than communities of phylogenetically distantly related taxa. Our findings show that lake heterotrophic bacterial productivity can be positively affected by evenness and richness, negatively by phylogenetic diversity, and that BEF relationships are contingent on microhabitats. These results

39 provide a stepping stone to compare biodiversity-productivity theory developed for Eukarya to
40 bacterial ecosystems.

41 **Keywords**: diversity-productivity, biodiversity-ecosystem function, bacterial communities,
42 microhabitats, particle-associated, limnology, heterotrophic productivity

43 **Introduction**

44 Our planet is currently experiencing an extreme species extinction event (Thomas et al.,
45 2004; Wake & Vredenburg, 2008). Concern about such declines in biodiversity has resulted in
46 hundreds of studies evaluating the relationship between biodiversity and ecosystem functions
47 (BEF), with a large focus on terrestrial plant ecosystems. BEF relationships are generally
48 positive and asymptotic and thus biodiversity loss causes a small change in ecosystem function at
49 first and then, at some tipping point, a dramatic decrease in function (Cardinale et al., 2012,
50 2012; Tilman et al., 2014). While the focus of local and global diversity loss is typically on
51 eukaryotic organisms, bacterial biodiversity has also been shown to be decreasing at local scales
52 within the human gut (Blaser, 2014) and terrestrial ecosystems (Singh et al., 2014). Of particular
53 concern is the loss of diversity of bacterial guilds responsible for key geochemical
54 transformations, such as methane oxidation (Levine et al., 2011) that controls rates of methane
55 emissions. Yet, the study of BEF relationships has been more limited for Bacteria and Archaea.
56 Based on the asymptotic BEF relationships observed for eukaryotic communities of up to
57 20 species, the large range of species richness observed in natural bacterial communities
58 (hundreds to thousands) may suggest an absence of bacterial BEF relationships. Several studies
59 have indicated no BEF relationships with broad processes such as heterotrophic respiration or
60 biomass production that are performed by many taxa (see figure 5 in Levine et al., 2011;

Langenheder et al., 2006; Delgado-Baquerizo et al., 2016). Yet, other studies on denitrification (Philippot et al., 2013) and on narrow metabolic processes that are catalyzed by few bacterial taxa, such as methanotrophy (Levine et al., 2011), and the degradation of triclosan and microcystin (Delgado-Baquerizo et al., 2016) found evidence of bacterial BEF relationships.

Beyond the impact of the number of species, phylogenetic relatedness is predicted to influence BEF relationships based on the phylogenetic limiting similarity hypothesis. The phylogenetic limiting similarity hypothesis posits that distantly related organisms will have more dissimilar niches and therefore reduced competition and a higher likelihood of coexistence (Violle et al., 2011). Therefore, it predicts that communities will have high phylogenetic diversity due to competitive exclusion of closely related species. Indeed, some papers show relationships across different ecosystems between phylogenetic diversity and ecosystem functions (Cadotte et al., 2008; Jiang et al., 2010; Violle et al., 2011). However, studies with freshwater green algae (Fritschie et al., 2014; Venail et al., 2014) did not find this relationship. A recent study found the opposite result by showing that closely related green algal species had weaker competition and more facilitation than distantly related species (Narwani et al., 2017). While relationships between phylogenetic relatedness among community members and ecosystem function have been assessed in bacterial systems (Tan et al., 2012; Galand et al., 2015; Roger et al., 2016), most work has focused on low-diversity, experimentally-assembled communities with bacteria that can be grown in culture. We need to expand these findings to communities with richness levels typically found in natural communities.

The nature of BEF relationships and the mechanism(s) that underpins them may depend on habitat structure or heterogeneity. Increasing habitat heterogeneity has been found to enhance the strength of BEF relationships (Tylianakis et al. 2008), presumably due to a greater role for

84 niche complementarity effects in heterogeneous environments (Cardinale 2011). While habitat
85 heterogeneity contributes to increased diversity within bacterial populations and communities
86 (Zhou et al., 2008; Shade et al., 2008), the influence of habitat heterogeneity on BEF
87 relationships remains unknown for bacterial systems.

88 In this study, we hypothesized that bacterial diversity would be positively correlated with
89 bacterial heterotrophic production, and that this relationship would be stronger in more
90 heterogeneous environments. We simultaneously surveyed free-living and particle-associated
91 surface water bacterial communities. Particulate matter comprises a variety of types and sizes of
92 particles with each particle also harboring physicochemical gradients (Simon et al., 2002), and
93 hence represents a more heterogeneous habitat than the surrounding water. We tested BEF
94 relationships using a variety of diversity metrics including observed richness, species dominance,
95 and phylogenetic diversity. We focused on heterotrophic bacterial production as our measure of
96 ecosystem function, as it is a key process affecting freshwater bacterial growth that in turn fuels
97 the macroscopic food web through their recycling of nutrients bound in organic matter (Cotner &
98 Biddanda, 2002).

99 **Methods**

100 ***Lake sampling and sample processing***

101 Surface water samples were collected at 1 meter depth from 4 long-term sampling stations
102 (Steinman et al., 2008) in mesotrophic Muskegon Lake (**Figure S1**), which is a freshwater
103 estuarine lake connecting the Muskegon River and Lake Michigan. These stations included the
104 mouth of the Muskegon River (43.250133,-86.2557), the channel to Bear Lake (43.238717,-

86.299283; a hypereutrophic lake), channel to Lake Michigan (43.2333,-86.3229; oligotrophic lake), and the deepest basin of Muskegon Lake (43.223917,-86.2972; max depth = 24 m).

Samples were collected during the morning to early afternoon of 3 days in 2015 (May 12, July 21, & September 30) aboard the R/V W.G. *Jackson*. All water samples were collected with vertical Van Dorn samplers. Additionally, a vertical profile of temperature (T), pH, specific conductivity (SPC), oxidation-reduction potential (ORP), chlorophyll (Chl*a*), total dissolved solids (TDS), and dissolved oxygen (DO) was constructed at each station to characterize the water column using a calibrated YSI 6600 V2-4 multiparameter water quality sonde (Yellow Springs Instruments Inc.). Total Kjeldahl nitrogen (TKN), ammonia (NH_3), total phosphorus (TP), and alkalinity (Alk) were processed from whole water while nitrate (NO_3), phosphate (PO_4), and chloride (Cl^-) were hand filtered using a 60 mL syringe fitted with Sweeny filter holder with a 13 mm diameter 0.45 µm pore size nitrocellulose filters (Millipore) and were determined by standard wet chemistry methods in the laboratory (EPA, 1993).

Bacterial abundance by epifluorescence microscopy

Lake surface water samples were processed within 2-6 hours of their collection for determination of heterotrophic bacterial abundance. Samples (5 mL) were preserved with 2% formalin and 1 mL subsamples were stained with acridine orange stain and filtered onto black 25 mm 0.2 µm pore size polycarbonate filters (Millipore) at a maximum pressure of 0.1 Bar or 1.5 PSI. Prepared slides were stored frozen until enumeration by standard epifluorescence microscopy at 1000x magnification under blue light excitation (Hobbie et al. 1977). Bacteria within the field of view (100 µm x 100 µm) that were not associated with any particles were counted as free-living bacteria, whereas bacteria that were on particles were counted as particle-associated. Sample filtration may bias counts due to free-living or particle-associated cells being hidden on the

128 underside of particles, , free-living bacteria settling on top of particles, or particle-associated
129 cells dislodging. In the absence of any quantitative studies that have rigorously addressed this
130 issue, we have assumed the net effect of these opposing methodological biases to be negligible in
131 the present study.

132 ***Heterotrophic bacterial production measurements***

133 Community-wide heterotrophic bacterial production was measured using [^{3}H] leucine
134 incorporation into bacterial protein in the dark (Kirchman et al. 1985; Simon and Azam, 1989).
135 At the end of the incubation with [^{3}H]-leucine, cold trichloroacetic acid-extracted samples were
136 filtered onto 3 μm filters that represented the leucine incorporation by particle-associated
137 bacteria (>3.0 μm). Each filtrate was collected and filtered onto 0.2 μm filters and the activity
138 therein represented incorporation of leucine by free-living bacteria (>0.2 μm-<3 μm). Measured
139 leucine incorporation during the incubation was converted to bacterial carbon production rate
140 using a standard theoretical conversion factor of 2.3 kg C per mole of leucine (Simon and Azam,
141 1989). Per-capita heterotrophic production was estimated by dividing heterotrophic production
142 by the cell counts measured in each fraction.

143 ***Preservation of bacterial filters in the field***

144 Microbial biomass for the particle-associated (> 3 μm) fraction and the free-living (3–0.22 μm
145 fraction) bacterial fraction was collected by sequential in-line filtration on 3 μm isopore
146 polycarbonate (TSTP, 47 mm diameter, Millipore, Billerica, MA, USA) and 0.22 μm Express
147 Plus polyethersulfone membrane filters (47 mm diameter, Millipore, MA, USA). We used 47
148 mm polycarbonate in-line filter holders (Pall Corporation, Ann Arbor, MI, USA)and an E/S
149 portable peristaltic pump with an easy-load L/S pump head (Masterflex®, Cole Parmer

Instrument Company, Vernon Hills, IL, USA). The total volume filtered varied from 0.8–2.2 L with a maximum filtration time of 16 minutes per sample. Filters were submerged in RNAlater (Ambion) in 2 mL cryovials, frozen in liquid nitrogen and transferred to a −80°C freezer until DNA extraction.

DNA extraction, sequencing and processing

DNA extractions were performed using an optimized method based on the AllPrep DNA/RNA/miRNA Universal kit (Qiagen; McCarthy et al., 2015; details in supplementary methods). Extracted DNA was sequenced using Illumina MiSeq V2 chemistry 2 × 250 (500 cycles) of dual index-labelled primers that targeted the V4 hypervariable region of the 16S rRNA gene (515F/806R) (Caporaso et al., 2012; Kozich et al., 2013) at the Microbial Systems Laboratories at the University of Michigan Medical School in July 2016. RTA V1.17.28 and MCS V2.2.0 software were used to generate data. Fastq files were submitted to NCBI sequence read archive under BioProject accession number PRJNA412984. We analyzed the sequence data using MOTHUR V.1.38.0 (seed = 777; Schloss et al., 2009) based on the MiSeq standard operating procedure accessed on 3 November 2015 and modified with time (see data accessibility and supplemental methods). For classification of operational taxonomic units (OTUs), a combination of the Silva Database (release 123; Quast et al., 2013) and the freshwater TaxAss 16S rRNA database and pipeline (Rohwer et al., 2017, accessed August 18, 2016). All non-bacterial and chloroplast sequences were pruned out of the dataset and replicate samples were merged by summing sample sequencing read counts using the *merge_samples* function (phyloseq). A batch script for our protocol can be found in this project's GitHub page at

https://github.com/DenefLab/Diversity_Productivity/blob/master/data/mothur/mothur.batch.taxass.

Estimating Diversity

To get the best estimate of each diversity metric, each sample was subsampled to 6,664 sequences (the smallest library size) with replacement and were averaged over 100 trials. Observed richness, Shannon entropy, and inverse Simpson's index were calculated using the *diversity* function within the vegan (Oksanen et al., 2013) R package via the *estimate_richness* function in the phyloseq (McMurdie and Holmes, 2013) R package. Simpson's Evenness was calculated by dividing the inverse Simpson's index by the observed richness (Magurran, 2004). To calculate phylogenetic diversity, we first removed OTUs that had a count of 2 sequences or less throughout the entire dataset, as these are more prone to be artefacts originating from sequencing errors or the OTU clustering algorithm. Representative sequences of each of the 1,891 remaining OTUs were collected from the aligned fasta file produced within mothur, and header names in the mothur output fasta file were modified using bbmap (Bushnell, 2016) to only include the OTU name. A phylogenetic tree was created with FastTree using the GTR+CAT (general time reversible) model (Price et al., 2010). Mismatches between the species community data matrix and the phylogenetic tree were checked with the *match.phylo.comm* command (picante). Finally, both abundance-unweighted and -weighted phylogenetic diversity was estimated using specifications described in the next paragraph with the picante R package.

The most common phylogenetic diversity (PD) measure is Faith's PD (Faith, 1992), however, this metric is very strongly correlated with species richness (**Figure S2**). Instead, the mean pairwise phylogenetic distance (or MPD) was calculated (*ses.mpd* function in the *Picante* R package (Kembell et al., 2010), null.model = "independentswap"). The MPD measures the

194 average phylogenetic distance between all combinations of two taxa pulled from the observed
195 community and compares it with a null community of equal richness pulled from the gamma
196 diversity of all the samples (*see supplemental methods for more details*). Values higher than zero
197 indicate phylogenetic evenness or overdispersion (higher phylogenetic diversity) while values
198 less than zero indicate phylogenetic clustering (lower diversity) or that species are more closely
199 related than expected according to the null community (Kembel, 2009). Thus, this phylogenetic
200 metric is relative. From here, we will refer to the SES_{MPD} as the "phylogenetic diversity" for
201 simplicity and clarity.

202 ***Statistical analysis***

203 Further analysis of sequence data was performed in R version 3.4.2 (R Core Team 2017;
204 *see supplemental methods for more details*). To test which variable(s) were the best predictors of
205 community and per-capita heterotrophic production, we performed variable selection via a lasso
206 regression (using the *glmnet* R package, alpha = 1, and lambda.1se as the tuning parameter
207 (Friedman et al., 2010)) on all of the environmental, biodiversity, and principal component
208 variables. To further validate the lasso regression results, we performed ordinary least squares
209 (OLS) regressions on all variables, including the principal components (PCA) of the euclidean
210 distances of the environmental data. We used the Akaike information criterion (AIC) (accessed
211 with the *broom::glance()* command) to select the best performing OLS regression model.

212 ***Data and code availability***

213 Original fastq files can be found on the NCBI sequence read archive under BioProject accession
214 number PRJNA412984. Processed data and code can be found on the GitHub page for this

project at https://deneflab.github.io/Diversity_Productivity/ with the main analysis at https://deneflab.github.io/Diversity_Productivity/Final_Analysis.html.

Results

Free-living communities had more cells and higher community-wide heterotrophic production but particle-associated communities had higher per-capita heterotrophic production

We observed an order of magnitude more cells per milliliter ($p = 1 \times 10^{-6}$, Figure 1A) and ~2.5 times more community-wide heterotrophic production in the free-living fraction ($p = 0.024$, Figure 1B). However, when calculated per-capita, particle-associated bacteria were on average an order of magnitude more productive than free-living bacteria ($p = 7 \times 10^{-5}$, Figure 1C). Particle-associated and free-living cell abundances in samples taken from the same water sample did not correlate (Figure S3A). Heterotrophic production between corresponding free-living and particle-associated fractions from the same water sample were positively correlated for both community (Adjusted $R^2 = 0.40$, $p = 0.017$; Figure S3B) and per-capita production rates (Adjusted $R^2 = 0.60$, $p = 0.003$; Figure S3C).

Particle-associated communities are more diverse in terms of observed richness and Shannon Entropy while free-living communities are more phylogenetically diverse

Across all samples, particle-associated bacterial communities were more diverse than free-living communities when considering richness and Shannon entropy (Figures 2A & S4A), but similar in the inverse Simpson's index and Simpson's evenness (Figure 2B & S4B).

Particle-associated bacterial community richness was always higher than in free-living communities and was maintained across the four sampling stations in the lake (Figure S5A).

Particle-associated samples at the river and Bear lake stations were on average more OTU-rich than the outlet to Lake Michigan and the Deep stations. Additionally, the river station had almost twice the inverse Simpson's value as compared with all other lake stations (Mean inverse Simpson Indices: Outlet = 23.6; Deep = 23.7; Bear = 35.3; River = 59.1; Figure S5A).

Particle-associated communities were more phylogenetically clustered than free-living communities based on unweighted phylogenetic diversity ($p = 0.01$, Figure 3A). Compared to other particle-associated samples, the outlet station that connects to oligotrophic Lake Michigan had a much larger unweighted phylogenetic diversity, indicating phylogenetic overdispersion (Figure S5A). Nevertheless, no sample across the entire dataset differed significantly from the null model with a significance threshold p-value of 0.05. There was no difference between weighted phylogenetic diversity in particle-associated versus free-living communities (Figure S5A).

Diversity-Productivity relationships are only observed in particle-associated communities

There was a strong, positive, linear BEF relationship between community-wide (Figures 2C-D & S4C-D) and per-capita (Figures 2E-F & S4E-F) heterotrophic productivity and all richness and evenness diversity metrics in the particle-associated communities, while no BEF relationships were observed for the free-living communities. The inverse Simpson's index explained the most amount of variation in community-wide (Figure 2D; Adjusted $R^2 = 0.69$, $p = 5 \times 10^{-4}$) and per-capita (Figure 2F; Adjusted $R^2 = 0.69$, $p = 0.001$) heterotrophic production. These results are robust across a range of minimum OTU abundance filtering thresholds (see *Sensitivity Analysis of Rare Taxa* in the supplemental methods and Figure S6) and hold up for all threshold levels in Inverse Simpson and for richness until removal of 25 counts (community-wide heterotrophic production) and 15 counts (per-capita heterotrophic production). When the

particle-associated and free-living samples were combined together into one linear model to test an overall relationship between diversity and productivity, there was no relationship (richness: $p = 0.86$; Shannon: $p = 0.99$; Inverse Simpson: $p = 0.36$), with the exception of a weak correlation for Simpson's Evenness (Adjusted $R^2 = 0.12$, $p = 0.054$). However, when particle-associated and free-living samples were combined together into one linear model to test an overall relationship between diversity and productivity, there was a strong relationship with observed richness (Adjusted $R^2 = 0.63$, $p = 3 \times 10^{-6}$), which broke down as evenness was weighed more (Figure S7: Shannon: Adjusted $R^2 = 0.52$, 6×10^{-5}; Inverse Simpson: Adjusted $R^2 = 0.48$, $p = 2 \times 10^{-4}$; Simpson's Evenness: $p = 0.48$).

Phylogenetic diversity correlated with per-capita heterotrophic production but not with community-wide production

Abundance-weighted phylogenetic diversity was not correlated with community or per-capita heterotrophic production (Figure S8C - S8D) and therefore no further analyses were performed with this diversity metric.

There was a moderate, negative, linear relationship when particle-associated and free-living samples were combined together into one linear model to test an overall relationship between unweighted phylogenetic diversity and observed richness (Figure 3B; Adjusted $R^2 = 0.35$, $p = 0.001$). To further validate this trend, randomized communities were generated with an equal richness as the samples but with OTUs randomly picked across the dataset. The unweighted phylogenetic diversity was then calculated and regressed against each the randomized richness and there was no relationship (Figure S9; Adjusted $R^2 = -0.02$, $p = 0.44$), verifying the negative relationship in the actual samples. When particle-associated and free-living samples were individually run in separate linear models to test for habitat-specific

282 relationships between unweighted phylogenetic diversity and observed richness, no trend was
283 found in either particle-associated or free-living models (Figure 3B; Particle: Adjusted $R^2 = 0.14$,
284 $p = 0.12$; Free = Adjusted $R^2 = -0.10$, $p = 0.97$). In other words, particle-associated and free-
285 living diversities did not have individual effects on community-wide or per-capita heterotrophic
286 production but rather, all samples were necessary for a correlation between per-capita
287 heterotrophic production and unweighted phylogenetic diversity.
288 There was no correlation between phylogenetic diversity and community-wide
289 heterotrophic production (Figure 3C). However, a negative correlation was found when particle-
290 associated and free-living samples were combined into one linear model to test an overall
291 relationship between unweighted phylogenetic diversity and per-capita heterotrophic production
292 (Figure 3D; $R^2 = 0.42$, $p = 5 \times 10^{-4}$). Therefore, these two results in combination indicated that
293 communities composed of more phylogenetically similar OTUs had a higher per-capita
294 heterotrophic production rate.
295 ***Diversity, and not environmental variation, is the best predictor of particle-associated***
296 ***heterotrophic production***
297 To identify variables that best predicted community-wide and per-capita heterotrophic
298 production (*i.e.* remove variables that were correlated with each other and/or uninformative
299 variables), we performed lasso regression with all samples and individually with particle-
300 associated and free-living samples. For prediction of community-wide heterotrophic production,
301 only the inverse Simpson's index was selected for particle-associated samples whereas pH and
302 PC5 were selected for free-living samples, and no variables were selected when all samples were
303 included in the lasso regression. In contrast, for per-capita heterotrophic production, temperature
304 and the inverse Simpson's index were selected for particle-associated samples whereas pH was

the only predictor for free-living samples, and observed richness was the only predictor for all samples (plotted in Figure S7A). Therefore, the best model for particle-associated microhabitats *always* included inverse Simpson's index whereas free-living samples only included environmental variables, such as pH.

To further verify that there were no confounding impacts of seasonal and environmental variables on community-wide and per-capita heterotrophic production, we performed ordinary least square (OLS) regressions and a dimension-reduction analysis of the environmental variables through a principal components analysis (Table S1 & S2; Figure S10). Specifically, the first 2 environmental axes explained ~70% of the environmental variation in the sampling sites (Figure S10). Next, we predicted community-wide and per-capita heterotrophic production with all environmental variables and the first six principal components as predictor variables with individual particle-associated and free-living samples, and combined (*i.e.* all samples) models (Table S1 & S2). The best single predictor of community-wide heterotrophic production was Inverse Simpson for particle-associated samples (AIC = 74.34; $R^2 = 0.69$), pH for the free-living samples (AIC =98.43; $R^2 = 0.49$, $p = 0.006$), and pH for all samples (AIC = 192.16; $R^2 = 0.35$) (Table S1). Whereas, the best single predictor of per-capita heterotrophic production was Inverse Simpson for particle-associated samples (AIC = 8.29; $R^2 = 0.69$), pH for the free-living samples (AIC = -2.39; $R^2 = 0.78$), and observed richness for all samples (AIC = 24.72; $R^2 = 0.63$) (Table S2). Thus, the OLS regressions are in agreement with the lasso regressions.

Discussion

We examined bacterial biodiversity-ecosystem function (BEF) relationships in relation to two microhabitats within freshwater lakes: particulate matter and the surrounding water. First, we found that community-wide and per-capita heterotrophic productivity of particle-associated

328 but not free-living bacterial communities showed a positive, linear BEF relationship with both
329 richness and evenness contributing. Second, particle-associated heterotrophic production was
330 better explained by diversity (*i.e.* inverse Simpson's index) than by environmental parameters.
331 Third, across both particle-associated and free-living communities, higher richness was
332 associated with lower phylogenetic diversity which, in turn, was associated with higher per-
333 capita heterotrophic bacterial production but not associated with community-wide heterotrophic
334 production.

335 Microbes have a large diversity of metabolisms and the choice of which to focus on may
336 inherently affect the BEF relationship. Indeed, "narrow" metabolic processes that are catalyzed
337 by a small subset of taxa within bacterial communities, such as nitrogen and sulfur cycling, have
338 been found to display BEF relationships (Levine et al., 2011; Delgado-Baquerizo et al., 2016). In
339 contrast, for "broad" processes that are performed by the majority of taxa within a bacterial
340 community, such as heterotrophic production (focus of the present study) and respiration,
341 functional redundancy appears to weaken or remove the presence of BEF relationships (Griffiths
342 et al., 2000; Langenheder et al., 2006; Wertz et al., 2006; Levine et al., 2011; Peter et al., 2011,
343 Galand et al, 2015). These findings are in line with the absence of a BEF relationship for free-
344 living bacterial communities in our study.

345 However, the above results and hypotheses surrounding narrow and broad processes are
346 in conflict with the strong BEF relationship we observed in particle-associated bacterial
347 communities. As such, our study signifies that microhabitats or habitat heterogeneity can
348 influence bacterial BEF relationships, in agreement with previous research in eukaryotic systems
349 across a variety of ecosystems (Tylianakis et al., 2008; Cardinale 2011; Zeppilli et al., 2016). A
350 study using controlled stream mesocosms by Cardinale (2011) found that niche complementarity

351 effects are particularly important in more heterogeneous environments. In more heterogeneous
352 streams, algal populations used different nutrients and avoided direct competition for resources,
353 resulting in unique species occupying distinct and local microhabitats.

354 Our observational study could not directly test the role of niche complementarity effects.
355 However, support for niche complementarity alone or in combination with species selection as
356 the mechanism underlying the BEF relationship in particle-associated habitats is provided by the
357 inverse Simpson's index being the strongest predictor of community-wide heterotrophic
358 production. As the inverse Simpson's index represents a measure of species dominance, it is
359 strongly affected by the evenness of abundant species. Communities that are more even have an
360 increased likelihood for complementary species to neighbor each other.

361 In our study, there are several reasons why heterogeneity of particulate matter may allow
362 for niche complementarity effects to occur and result in BEF relationships. First, particles have a
363 two-fold layer of heterogeneity as they (A) may be composed of different substrates such as
364 organic matter from terrestrial or aquatic environments and either heterotrophically or
365 photosynthetically derived (Grossart, 2010), and (B) each particle may comprise
366 physicochemical gradients as well (Simon et al., 2002). Second, microbial interactions are more
367 likely to occur between cells aggregated on particles as the interaction distances are usually
368 much shorter (Cordero & Datta, 2016) compared to free-living bacterial cells. In fact, genes
369 mediating social interactions, such as motility, adhesion, cell-to-cell transfer, antibiotic
370 resistance, mobile element activity, and transposases, have been found to be more abundant in
371 marine particles than compared to the surrounding water (Ganesh et al., 2014).

372 The importance of niche complementarity in microbial communities can also be deduced
373 from recent findings in the field of microbiology, which have shown widespread metabolic

374 interdependence among bacterial community members. First, a 2016 study that reconstructed
375 2,540 draft genomes of microbes found that most bacteria specialize in one particular step in
376 sulfur and nitrogen pathways and "hand-off" their metabolic byproducts to nearby organisms
377 (Anantharaman et al., 2016). It is likely that metabolic hand-offs, a specific form of bacterial
378 facilitation, will occur more in particle-associated compared to free-living communities. Indeed,
379 Datta and Cordero's (2016) work on model marine particles found that taxa that are incapable of
380 breaking down particles and instead rely on carbon produced by primary degraders thrive in later
381 phases of particle degradation. Second, Lilja and Johnson (2016) demonstrated that different
382 microbial cell types eliminate inter-enzyme competition by cross feeding, which increases
383 substrate consumption by allowing intracellular resources to go towards a single enzyme, rather
384 than having two enzymes that perform two separate reactions compete for nutrients within a cell.
385 Third, some bacteria are unable to grow in laboratory cultures unless they are in co-culture with
386 other organisms, which may be due to metabolic hand-offs or to growth factors such as
387 siderophores or catalases (Stewart, 2012).

388 Taking into account that (i) closely related taxa share more genes and metabolic
389 pathways than distantly related bacterial taxa (Konstantinidis & Tiedje, 2005; Kim et al., 2014)
390 and (ii) bacteria commonly have incomplete metabolic pathways, we propose that closely related
391 bacteria may be most likely to hand-off their metabolic byproducts. This may be why we found
392 that new taxa added to the community represented taxonomic clades similar to or already present
393 in the community, and that these communities with lower phylogenetic diversity (relative to
394 expected) had higher productivities. This result is in line with a recent study using freshwater
395 algae and vascular plants that reject predictions from the phylogenetic limiting similarity
396 hypothesis (Narwani et al., 2017). However, recent bacteria-focused studies from Russel et al.

(2017) and Venail and Vives (2013) found higher levels of antagonism (Russel et al., 2017) or more bacterial productivity (measured through colony forming units per mL;Venail and Vives, 2013) with more distantly related taxa. Both of these studies were performed in the lab with r-selected (*i.e.* copiotrophic) species grown in stable, warm, aerobic, agar plate conditions. Thus, Venail and Vives (2013) and Russel et al. (2017) inherently break up potential interdependent relationship between bacteria either by creating artificial communities or evaluating pairwise interactions and remove the natural effect of spatial heterogeneity, environmental fluctuations, and the rest of the bacterial community. As a result, future studies on bacterial interactions and the role of phylogenetic diversity will need to maintain natural structure and complexity in bacterial communities.

Previous studies on bacterial BEF relationships have used three approaches to manipulate bacterial diversity (Krause et al., 2014): (1) dilution to extinction in which complex communities are diluted to more simple communities (Wertz et al., 2006; Peter et al., 2011; Philippot et al., 2013; see Roger et al., 2016 for a review of this approach), (2) manually assembled communities in culture (Tan et al., 2012; Salles et al., 2009), or (3) natural or manipulated environmental communities (Griffiths et al., 2000; Levine et al., 2011; Galand et al., 2015). In this study, we took the latter approach. In contrast to the other two approaches, this had the benefit of (1) maintaining high diversity with both abundant and rare taxa, (2) including both r- and k-selected organisms, (3) allowing natural environmental and ecological forcings to shape the community, and (4) evaluating BEF relationships in diversity and productivity ranges that reflect natural communities. Admittedly, three inherent weaknesses to our approach were that (1) we cannot measure all the potential variables that influence heterotrophic productivity, (2) we only have 24 samples for a 12 versus 12 study, and (3) our analysis is correlational and we cannot manipulate

420 the system to unequivocally separate causes and consequences of bacterial production. For
421 example, strong correlations with heterotrophic production and pH in the free-living samples
422 (Table S1 & S2) may point to pH being a consequence of rather than a cause of varying
423 production levels. This is because bacterial production and bacterial respiration are positively
424 correlated (del Giorgio & Cole, 1998) and with increased respiration, pH may decrease due to
425 CO_2 dissolution into the water.

426 Finally, we acknowledge that the typical sampling of bacterial communities and analysis
427 using DNA sequencing reflects *all* bacteria present in the community and not necessarily only
428 the *active* members of the community contributing to a given ecosystem function. In freshwater
429 systems, up to 40% of cells from the total community have been shown to be inactive or dormant
430 (Jones and Lennon, 2010). If one were to sample plant communities in an analogous way to
431 bacterial systems, one would measure the diversity of all the above- and below-ground plant
432 biomass including seeds, pollen, and detrital biomass. In this context, it is interesting to reflect
433 on the richness in absence of function (i.e. x-intercept) of the observed BEF relationship which is
434 295. This could be interpreted as a baseline level of 295 inactive (either dead or dormant)
435 bacterial OTUs and in the case of particulate material, environmental DNA adhered the
436 substrate, in the community. This value represents 35-85% of the total particle-associated
437 communities and may obscure the *actual* diversity (and BEF relationship) of the bacterial
438 community (Carini et al., 2016).

439 In conclusion, we show that increased bacterial diversity, especially when measured by
440 the inverse Simpson's index, leads to increased community-wide and per-capita bacterial
441 heterotrophic production in particle-associated but not in free-living communities. As such, we
442 extend the validity of principles of the impact of microhabitat on BEF relationships from

443 Eukarya to Bacteria, contributing to current efforts to integrate ecological theories into the field
444 of microbiology (Barberán et al., 2014). Additionally, we show that communities with low
445 phylogenetic diversity have higher per-capita heterotrophic production rates, which we
446 hypothesize to be related to genome evolutionary patterns specific to bacteria that result in the
447 dependence on metabolic hand-offs. Differences between Bacteria and Eukarya in patterns of
448 genome evolution and its ecological consequences, as well as in how active and dormant
449 fractions of the community are measured need to be taken into account when trying to integrate
450 BEF studies across all domains of life.
451

452 **Acknowledgements**

453 This work was supported by the National Science Foundation Graduate Research Fellowship
454 Grant No. DGE 1256260 (MLS), the University of Michigan Office for Research MCubed
455 program (VJD), the American Society of Microbiology-Undergraduate Research Fellowship, the
456 University of Michigan Honors Summer Fellowship, and the Beckman Scholars Program (EC).
457 RP was supported by Ghent University (BOFDOC2015000601) and a Sofina Gustave-Boël grant
458 from the Belgian American Educational Foundation.We are grateful to the crew of the R/V W.G.
459 Jackson and the Grand Valley State University Robert B. Annis Water Resources Institute
460 science staff, and the generous help we received in the field from Amelia Waters and Daniel
461 S.W. Katz. Thank you to Kyle Buffin and Amadeus Twu for help with DNA extractions. Finally,
462 we thank Deborah Goldberg, George Kling, and members of the Denef, Dick, and Duhaime
463 laboratories for their comments on the manuscript.

References

Anantharaman, K. et al. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nature Communications 7:13219.

Barberán, A., E. O. Casamayor, and N. Fierer. 2014. The microbial contribution to macroecology. Frontiers in Microbiology 5:1–8.

Blaser, M. J. 2014. Missing Microbes: How the Overuse of Antibiotics Is Fueling Our Modern Plagues. Henry Holt and Company LLC, New York 35:261.

Bushnell B. 2016. BBMap short read aligner. https://sourceforge.net/projects /bbmap/.

Cadotte, M. W., B. J. Cardinale, and T. H. Oakley. 2008. Evolutionary history and the effect of biodiversity on plant productivity. Proceedings of the National Academy of Sciences 105:17012–17017.

Caporaso, J. G. et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal 6:1621–1624.

Cardinale, B. J. 2011. Biodiversity improves water quality through niche partitioning. Nature 472:86–89.

Cardinale, B. J. et al. 2012. Biodiversity loss and its impact on humanity. Nature 489:326–326.

Carini, P., P. J. Marsden, J. W. Leff, E. E. Morgan, M. S. Strickland, and N. Fierer. 2016. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nature Microbiology 2:16242.

Cordero, O. X., and M. S. Datta. 2016. ScienceDirect Microbial interactions and community assembly at microscales. Current Opinion in Microbiology 31:227–234.

Cotner, J. B., and B. A. Biddanda. 2002. Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. Ecosystems 5:105–121.

Datta, M. S., E. Sliwerska, J. Gore, M. F. Polz, and O. X. Cordero. 2016. Microbial interactions lead to rapid micro-scale successions on model marine particles. Nature Communications 7:11965.

Delgado-Baquerizo, M., L. et al. 2016. Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. Journal of Ecology 104:936–946.

del Giorgio, P. A., and J. J. Cole. 1998. Bacterial Growth Efficiency in Natural Aquatic Systems. Annual Review of Ecology and Systematics 29:503–541.

494 Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation
495 61:1–10.
496 Friedman, J., T. Hastie, and R. Tibshirani. 2010. Regularization Paths for Generalized Linear
497 Models via Coordinate Descent. Journal of Statistical Software, 33:1-22.
498 Fritschie, K. J., B. J. Cardinale, M. A. Alexandrou, and T. H. Oakley. 2014. Evolutionary history
499 and the strength of species interactions: Testing the phylogenetic limiting similarity
500 hypothesis. Ecology 95:1407–1417.
501 Galand, P. E., I. Salter, and D. Kalenitchenko. 2015. Ecosystem productivity is associated with
502 bacterial phylogenetic distance in surface marine waters. Molecular Ecology 24:5785–5795.
503 Ganesh, S., D. J. Parris, E. F. DeLong, and F. J. Stewart. 2014. Metagenomic analysis of size-
504 fractionated picoplankton in a marine oxygen minimum zone. The ISME journal 8:187–211.
505 Griffiths, B. S. et al. 2000. Ecosystem response of pasture soil communities to fumigation-
506 induced microbial diversity reductions: an examination of the biodiversity-ecosystem
507 function relationship. Oikos 90:279–294.
508 Grossart, H. P. 2010. Ecological consequences of bacterioplankton lifestyles: Changes in
509 concepts are needed. Environmental Microbiology Reports 2:706–714.
510 Hobbie, J. E., R. J. Daley, and S. Jasper. 1977. Use of nuclepore filter counting bacteria by
511 fluoroscence microscopy. Applied and Environmental Microbiology 33:1225–1228.
512 Jiang, L., J. Tan, and Z. Pu. 2010. An Experimental Test of Darwin's Naturalization Hypothesis.
513 The American Naturalist 175:415–423.
514 Jones, S. E., and J. T. Lennon. 2010. Dormancy contributes to the maintenance of microbial
515 diversity. Proceedings of the National Academy of Sciences 107:5881–5886.
516 Kembel, S. W. 2009. Disentangling niche and neutral influences on community assembly:
517 Assessing the performance of community phylogenetic structure tests. Ecology Letters
518 12:949–960.
519 Kembel, S.W. et al.. 2010. Picante: R tools for integrating phylogenies and ecology.
520 Bioinformatics 26:1463-1464.
521 Kim, M., H. S. Oh, S. C. Park, and J. Chun. 2014. Towards a taxonomic coherence between
522 average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation
523 of prokaryotes. International Journal of Systematic and Evolutionary Microbiology 64:346–
524 351.

525 Kirchman, D., E. K'nees, and R. Hodson. 1985. Leucine incorporation and its potential as a
526 measure of protein synthesis by bacteria in natural aquatic systems. Applied and
527 Environmental Microbiology 49:599–607.
528 Konstantinidis, K. T., and J. M. Tiedje. 2005. Genomic insights that advance the species
529 definition for prokaryotes. Proceedings of the National Academy of Sciences 102:2567–
530 2572.
531 Kozich, J. J., S. L. Westcott, N. T. Baxter, S. K. Highlander, and P. D. Schloss. 2013.
532 Development of a dual-index sequencing strategy and curation pipeline for analyzing
533 amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and
534 Environmental Microbiology 79:5112–5120.
535 Krause, S. et al. 2014. Trait-based approaches for understanding microbial biodiversity and
536 ecosystem functioning. Frontiers in Microbiology 5:1–10.
537 Langenheder, S., E. S. Lindström, and L. J. Tranvik. 2006. Structure and Function of Bacterial
538 Communities Emerging from Different Sources under Identical Conditions. Applied and
539 environmental microbiology 72:212–220.
540 Levine, U. Y., T. K. Teal, G. P. Robertson, and T. M. Schmidt. 2011. Agriculture's impact on
541 microbial diversity and associated fluxes of carbon dioxide and methane. The ISME Journal
542 5:1683–1691.
543 Lilja, E. E., and D. R. Johnson. 2016. Segregating metabolic processes into different microbial
544 cells accelerates the consumption of inhibitory substrates. The ISME Journal 10:1–11.
545 Magurran, A. E. 2004. Chapter four: An index of diversity... in Measuring Biological Diversity,
546 Wiley-Blackwell, Hoboken, NJ.
547 McMurdie, P. J., and S. Holmes. 2013. phyloseq: An R Package for Reproducible Interactive
548 Analysis and Graphics of Microbiome Census Data. PLoS ONE 8:e61217.
549 Narwani, A. et al. 2017. Ecological interactions and coexistence are predicted by gene
550 expression similarity in freshwater green algae. Journal of Ecology 105:580–591.
551 Oksanen, A. J. et al. 2015. vegan: Community Ecology Package. R package version 2.3-0.
552 Peter, H., S. Beier, S. Bertilsson, E. S. Lindström, S. Langenheder, and L. J. Tranvik. 2011.
553 Function-specific response to depletion of microbial diversity. The ISME Journal 5:351–361.
554 Philippot, L. et al. 2013. Loss in microbial diversity affects nitrogen cycling in soil. The ISME
555 Journal 7:1609–1619.

556 Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. FastTree 2 - Approximately maximum-
557 likelihood trees for large alignments. PLoS ONE 5.
558 Quast, C. et al. 2013. The SILVA ribosomal RNA gene database project: Improved data
559 processing and web-based tools. Nucleic Acids Research 41:590–596.
560 R Core Team (2017) R: A Language and Environment for Statistical computing. Vienna,
561 Austria: R Foundation for Statistical Computing. https:// www.R-project.org/.
562 Roger, F., S. Bertilsson, S. Langenheder, O. A. Osman, and L. Gamfeldt. 2016. Effects of
563 multiple dimensions of bacterial diversity on functioning, stability and multifunctionality.
564 Ecology 97:2716–2728.
565 Rohwer, R. R., J. J. Hamilton, R. J. Newton, and K. D. McMahon. 2017. TaxAss: Leveraging
566 Custom Databases Achieves Fine-Scale Taxonomic Resolution. bioRxiv:214288.
567 Russel, J., H. L. Røder, J. S. Madsen, M. Burmølle, and S. J. Sørensen. 2017. Antagonism
568 correlates with metabolic similarity in diverse bacteria. Proceedings of the National Academy
569 of Sciences:201706016.
570 Salles, J. F. et al. 2009. Community niche predicts the functioning of denitrifying bacterial
571 assemblages. Ecology, 90:3324–3332.
572 Schloss, P. D. et al. 2009. Introducing mothur: Open-source, platform-independent, community-
573 supported software for describing and comparing microbial communities. Applied and
574 Environmental Microbiology 75:7537–7541.
575 Shade, A., S. E. Jones, and K. D. McMahon. 2008. The influence of habitat heterogeneity on
576 freshwater bacterial community composition and dynamics. Environmental Microbiology
577 10:1057–1067.
578 Simon, M., and F. Azam. 1989. Protein content and protein synthesis rates of planktonic marine
579 bacteria. Marine Ecology Progress Series 51:201–213.
580 Simon, M., H. P. Grossart, B. Schweitzer, and H. Ploug. 2002. Microbial ecology of organic
581 aggregates in aquatic ecosystems. Aquatic Microbial Ecology 28:175–211.
582 Singh, B. K. et al. 2014. Loss of microbial diversity in soils is coincident with reductions in some
583 specialized functions. Environmental Microbiology 16:2408–2420.
584 Steinman, A. D., M. Ogdahl, R. Rediske, C. R. Ruetz III, B. a Biddanda, and L. Nemeth. 2008.
585 Current status and trends in Muskegon Lake, Michigan. Journal of Great Lakes Research
586 34:169–188.

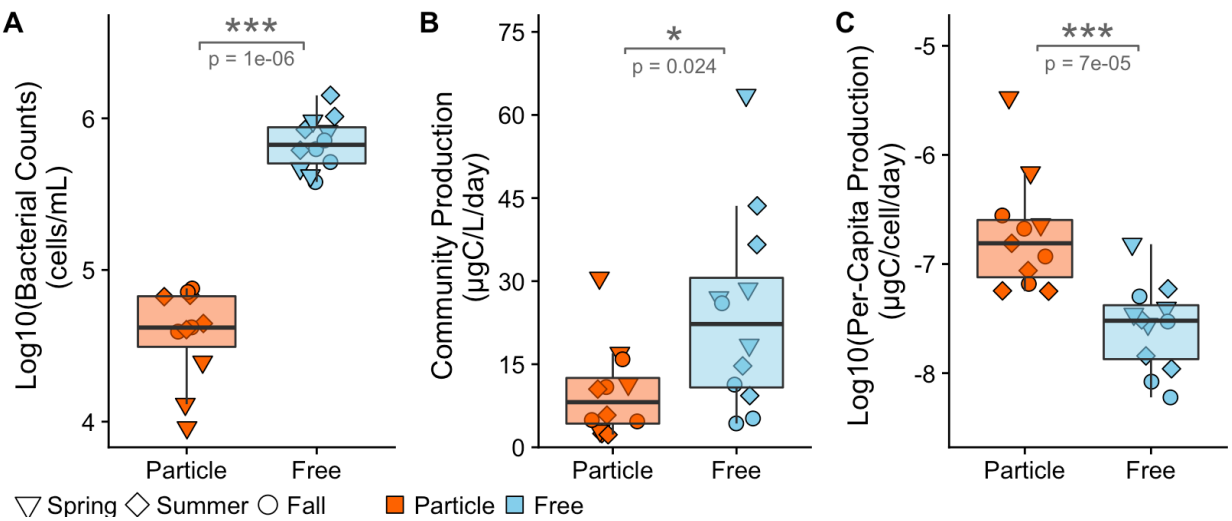
587 Stewart, E. J. 2012. Growing unculturable bacteria. Journal of Bacteriology 194:4151–4160.
588 Tan, J., Z. Pu, W. A. Ryberg, and L. Jiang. 2012. Species phylogenetic relatedness, priority
589 effects, and ecosystem functioning. Ecology 93:1164–1172.
590 Thomas, C. D. et al. 2004. Extinction risk from climate change. Nature 427:145–8.
591 Tilman, D., F. Isbell, and J. M. Cowles. 2014. Biodiversity and Ecosystem Functioning. Annual
592 Review of Ecology, Evolution, and Systematics 45:471–493.
593 Tylianakis, J. M. et al. 2008. Resource heterogeneity moderates the biodiversity-function
594 relationship in real world ecosystems. PLoS Biology 6:0947–0956.
595 Venail, P. A., A. Narwani, K. Fritschie, M. A. Alexandrou, T. H. Oakley, and B. J. Cardinale.
596 2014. The influence of phylogenetic relatedness on species interactions among freshwater
597 green algae in a mesocosm experiment. Journal of Ecology 102:1288–1299.
598 Venail, P. A., and M. J. Vives. 2013. Phylogenetic distance and species richness interactively
599 affect the productivity of bacterial communities. Ecology 94:2529–2536.
600 Violle, C., D. R. Nemergut, Z. Pu, and L. Jiang. 2011. Phylogenetic limiting similarity and
601 competitive exclusion. Ecology Letters 14:782–787.
602 Wake, D. B., and V. T. Vredenburg. 2008. Are we in the midst of the sixth mass extinction? A
603 view from the world of amphibians. Proceedings of the National Academy of Sciences
604 105:11466–11473.
605 Wertz, S., V. et al. 2006. Maintenance of soil functioning following erosion of microbial
606 diversity. Environmental Microbiology 8:2162–2169.
607 Zeppilli, D., A. Pusceddu, F. Trincardi, and R. Danovaro. 2016. Seafloor heterogeneity
608 influences the biodiversity–ecosystem functioning relationships in the deep sea. Scientific
609 Reports 6:26352.
610 Zhou, J., S. et al. 2008. Spatial scaling of functional gene diversity across various microbial taxa.
611 Proceedings of the National Academy of Sciences 105:7768–7773.
612

Figure 1: Bacterial counts, community-wide and per-capita heterotrophic production differ between microhabitats. Particle-associated and free-living samples were taken from four stations within Muskegon Lake during 2015 in May, July, and September. **(A)** Free-living bacteria were an order of magnitude (10^6 cells/mL) more abundant compared to particle-associated bacteria. **(B)** Free-living bacteria were more heterotrophically productive compared to particle-associated bacteria. **(C)** Particle-associated bacteria were disproportionately heterotrophically productive per cell compared to free-living bacteria.

Figure 2: Richness and Inverse Simpson correlate with heterotrophic productivity. Top panel: Differences in **(A)** the observed richness and **(B)** the inverse Simpson diversity metrics between particle-associated (orange) and free-living (blue) habitats. **Middle panel:** Biodiversity and community-wide heterotrophic production (ugC/L/day) relationships. The y-axis between **(C)** and **(D)** is the same, however, the x-axis represents **(C)** richness and **(D)** Inverse Simpson. **Bottom panel:** Biodiversity and $\log_{10}$(per-capita heterotrophic production) (ugC/cell/day) relationships. The y-axis between **(E)** and **(F)** is the same, however, the x-axis represents **(E)** richness and **(F)** Inverse Simpson's index. Solid lines represent ordinary least squares models for the free-living (blue) and particle associated (orange) communities. All R^2 values represent the adjusted R^2 from an ordinary least squares model.

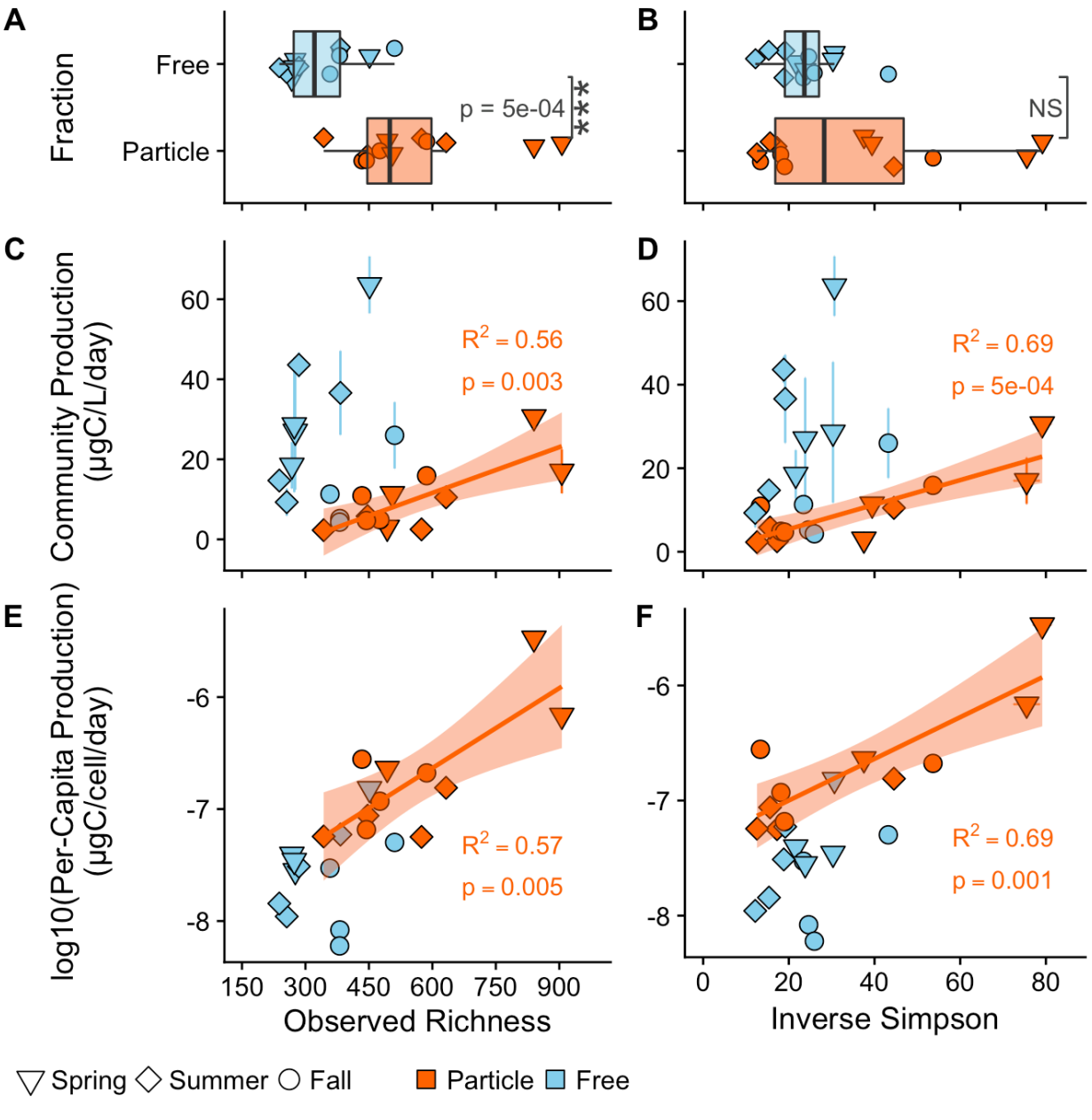
Figure 3: The relationship between heterotrophic productivity and unweighted phylogenetic diversity (SES_{MPD}; ses.mpd function in *picante* with null.model = "independentswap"). Positive phylogenetic diversity values represent communities that are phylogenetically diverse (*i.e.* overdispersed) while negative phylogenetic diversity values represent communities that are phylogenetically less diverse (*i.e.* clustered) compared to a null community with equal species richness. **(A)** Phylogenetic diversity was higher in free-living communities compared to particle-associated communities. **(B)** Negative relationship between observed richness and phylogenetic diversity. **(C)** Absence of phylogenetic diversity and community bulk heterotrophic production (µgC/L/day) relationships. **(D)** Negative phylogenetic diversity and per-capita heterotrophic production (µgC/cell/day) relationship. Linear models in figure **B** and **D** represent trends over all samples.

644 **Figure 1**



645

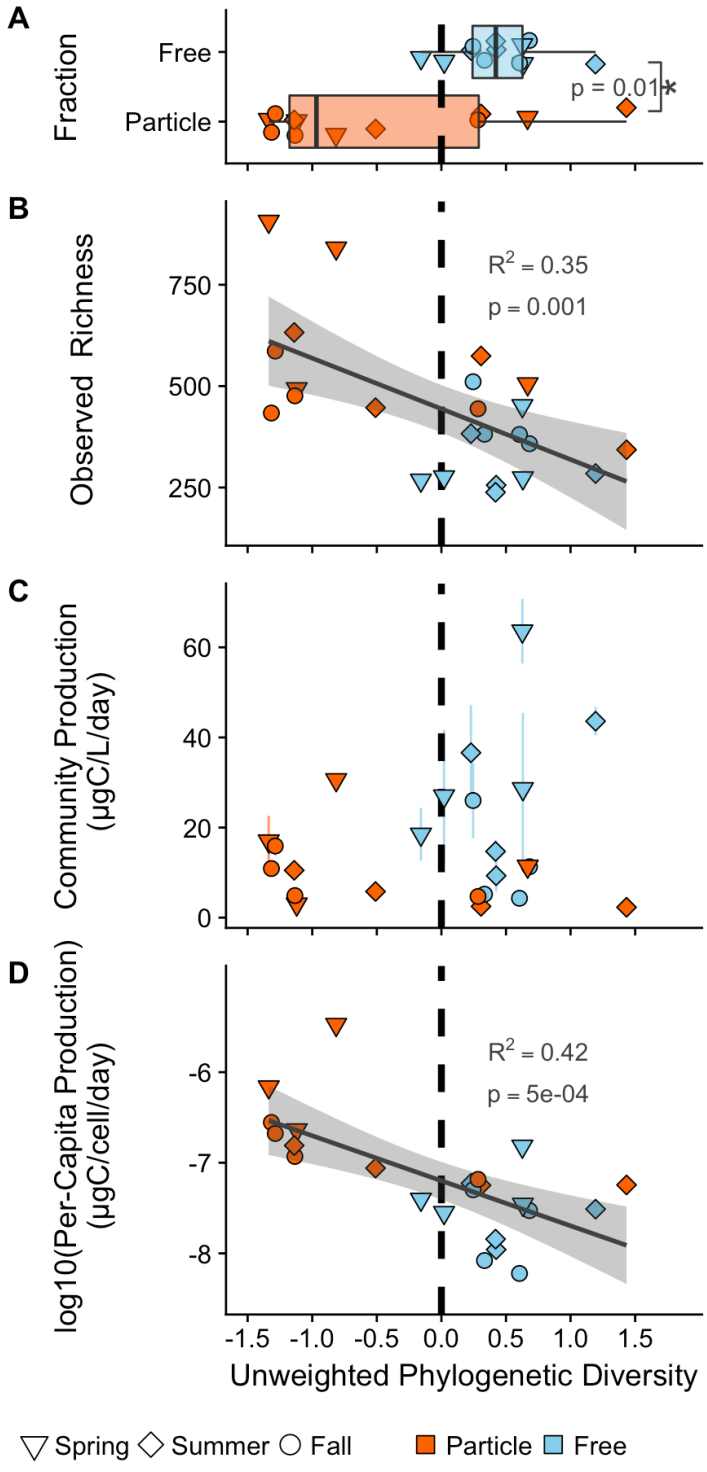
646 **Figure 2**



647

648

649 **Figure 3**



650