Artificial ecosystem selection to deduce the mapping between microbial community structure and ecosystem function

Andrew H. Morris and Brendan J. M. Bohannan

# Abstract

# Introduction

One of the fundamental questions in Ecology is To what extent does biodiversity drive variation in ecosystem functioning? Variation in abiotic conditions regulates the rate of ecosystem processes, but it remains unclear to what extent variation in community structure regulates the rate of ecosystem processes. Many ecosystem functions are mediated by organisms. In particular, microorganisms such as bacteria mediate the major biogeochemical cycles such as the carbon and nitrogen cycles. Despite the central role of bacteria in performing these processes, we have not found a consistent relationship between microbial community structure and the rate of any ecosystem process. In order to understand the importance of biodiversity to ecosystem functioning and to better predict and manage specific ecosystem processes, we need to determine the relationship between microbial community structure and ecosystem function.

In macroorganismal communities, there appears to be a relationship between the diversity of a community and the resilience of ecosystem functioning to environmental perturbations (Tilman, ????… Hooper). Typically, this is the diversity (or more specifically, the richness) of a particular functional group, such as photosynthetic plants. Similar studies in microbial communities rarely find a relationship between diversity and function (Graham, ). However, these studies typically define diversity much more broadly than plant diversity studies by looking at the phylogenetic diversity of the entire bacterial and archaeal community, not just of a particular functional group. Studies that focus more narrowly on a particular ecosystem process and a functional group involved in that process still rarely find a significant correlation between the abundance of a functional group (or a functional protein-coding gene) and the rate of the corresponding process (Rocca, ????). For certain ecosystem functions in certain ecosystems, these relationships can be informative. For example, the abundance of a single methanogenic species (*G. species*) is correlated with methane emissions in a permafrost ecosystem (cite). In addition, bacterial and archaeal community composition is correlated with methane emissions in forests and pastures of the Brazilian Amazon (Meyer et al. 2020). Therefore, there appears to be some indication that microbial community structure drives variation in methane cycling in soils.

To further investigate the relationship between microbial community structure and the rate of methane cycling in soils, we want to determine which organisms are important for regulating methane cycling. This ecological question is analogous to the problem of genotype-phenotype mapping in quantitative genetics (Morris 2020). Variation in an organimsal trait is partially determined by environmental conditions and partially determined by genetic varition. One goal of quantitative genetics is to determine the proportion of the variation attributable to genetics and to identify which alleles regulate the phenotypic trait. These goals are commonly achieved through artificial selection experiments and association mapping studies, such as genome-wide association studies. To determine the mapping between microbial community structure and ecosystem function, we applied these approaches of genotype-phenotype mapping to whole microbial communities. In this case, the “alleles” are different microbial taxa and the “phenotype” is the rate of ecosystem function at the whole-ecosystem level.

To do this, we performed an artificial ecosystem selection experiment to identify communities of bacteria and archaea that collectively perform a high rate of methane oxidation (Swenson et al. 2000). (Goodnight et al. 1997, Williams and Lenton 2007)we generated twenty-four soil ecosystems in the lab using a sterilized potting mix as the substrate. We then inoculated these ecosystems with a small amount of living soil to generate variation in community structure across the twenty-four ecosystems. These ecosystems were then assigned to one of two selection treatments: neutral selection with ecosystems chosen for reproduction at random or positive selection with ecosystems chosen for reproduction based on their methane oxidation rate. Each ecosystem was maintained at 1000 ppm CH4 over several weeks to allow for colonization and growth of the microbial community and to enrich for methane oxidizers. We then determined the methane oxidation rate for each ecosystem and selected three jars in each treatment to inoculate the next set of twenty-four jars. This process continued for five passages. At the end of the experiment, we extracted DNA from soils in the second and fifth passage for both the neutral and positive treatments. By comparing the response to selection within the community between the two treatments, we can identify microbial species that contribute to variation in ecosystem methane emissions.

This approach is powerful because by using a common soil substrate and maintaining a constant headspace concentration of methane, we can eliminate much of the environmental variation that would be present in an observational study in the wild. In addition, by applying selection on methane oxidation rate in the positive treatment, we can enrich for taxa involved in methane cycling making it easier to perform association mapping between microbial taxa and ecosystem methane emissions. In this paper, we address whether there is a response to selection on methane oxidation rate at the whole ecosystem level. We then quantify the amount of variation in methane oxidation attributable to variation in community composition within our laboratory environment. Finally, we identify microbial markers of soil methane cycling using association mapping in order to deduce the mapping between microbial community structure and ecosystem function.

# Materials & Methods

## Selection experiment

We performed an artificial ecosystem selection experiment similar to Swenson et al. (2000) by passaging replicate soil microbial communities. The trait we selected on was CH4 oxidation rate. Our experiment had two selection lines with twelve jars each for a total of twenty-four jars per passage. One line was a positive selection line where the two or three jars with the highest CH4 oxidation rate were chosen to inoculate the next set of jars. The other line was a neutral selection line where an equal number of jars as the positive line were chosen at random among the twelve. The number of jars chosen was based on the distribution of fluxes among the positive jars, i.e., we chose the top three jars unless only two jars had considerably greater oxidation rates based on visual inspection of histograms.

The initial soil microbial community was sampled from the top 10 cm of an upland mineral soil under a deciduous forest ecosystem near the University of Oregon campus. Incubations were performed in 500 mL mason jars (Ball Corporation, Broomfield, Colorado, USA) with rubber septa installed in the lids. Jars were incubated at ambient temperature on the benchtop. To create the first set of jars, twenty-four jars were created in an identical manner and then randomly assigned to either the positive or negative selection line. Each jar was sterilized with 70% ethanol to which was added 45 g of autoclaved potting mix, 5 g of live soil, and 3.5 mL of sterile deionized water to bring them to 60%??? of field capacity . The jars were then homogenized, capped, and injected with 4.3 mL of 99% CH4 to bring the headspace concentration to approximately 1000 ppm CH4. Jars were flushed and respiked twice per week to maintain elevated CH4 concentrations until travel, holidays, and COVID-19 got in the way and then we just did it as often as we could. Jars were incubated for approximately three weeks and then flux measurements were taken.

Once CH4 oxidation rates were determined, the top two or three jars were chosen for selection in the positive line and an equal number of jars were randomly chosen from the neutral line. The selected jars for each line were homogenized and used as the 5 g live soil for the subsequent passage. Twelve new jars were created for each line in the same manner as above using the inoculum for that line. This procedure continued until a divergence was observed in the mean flux of each line for two consecutive generations. Jars in passage number 2 and 5 were homogenized and a sample was collected and stored at -80C for later DNA extraction.

## Methane flux measurements Methane oxidation rates were determined after

flushing and spiking jars to approximately 1000 ppm CH4. Head-space samples of 1 mL were collected from each jar immediately after spiking and then at time points 3, 6, 24, and 48 hours for a 5-point curve. Samples were immediately injected into a ??? gas chromatograph fitted with an electron capture detector (company, city, state, USA) to determine the headspace CH4 concentration. Fluxes were calculated from a first-order exponential decay function as decay constant k. Oxidation rates are presented as the additive inverse of k (i.e., -k) so that a more positive value represents a faster oxidation rate.

## Soil DNA extraction and sequencing

Soil DNA was extracted from a 0.25 g subsample of soil from the starting inoculum and from each jar in passage 2 and 5. Negative controls were extracted from autoclaved potting mix and DNase-free water. Extractions were performed using the DNeasy PowerSoil kit (QIAGEN, Hilden, Germany) and quantified using Qubit (company, city, state). In order to estimate the diversity and relative abundance of the bacterial and archaeal taxa in our soil ecosystems, we sequenced the V3-V4 region of the 16S rRNA gene using the Earth Microbiome Project protocol and primers???. In addition, to estimate the diversity and relative abundance of CH4 oxidizing taxa, we sequenced the particulate methane monooxygenase subunit A gene (*pmoA*) using primers???. Sequencing was performed on the Illumina MiSeq in v3 mode with 300 bp paired-end reads.

## Bioinformatics Bioinformatic analyses were conducted using the QIIME2

bioinformatics platform (Bolyen et al. 2019). Raw sequencing reads were demultiplexed using ??? (citation), denoised using DADA2 (citation). Taxonomic assignment for the 16S reads was performed using ??? based on the Silva database. All subsequent tatistical analyses were performed in the R statistical programming environment (R Core Team 2020). Richness estimates were made using the breakaway package and diversity indices using the DivNet package (Willis et al. ???).

## Statistical Analysis

### Response to selection

To test whether there was a significant change in CH4 oxidation rate (k) as a response to selection, we tested the deviation between the positive and neutral selection lines over multiple passages. We calculated deviation as the difference in the mean oxidation rate between the positive and neutral treatment within each passage. The standard errors of the difference were calculated as the square root of the sum of the squared errors of the positive and neutral lines. We then fit an ordinary least squares regression model taking into account the standard errors of the deviations. The model was fit using the rma function from the metafor package in r (Viechtbauer 2010).

### Ecosystem heritability

To estimate the proportion of variance in ecosystem function due to additive genetic variance of the microbial communtiy, we applied the Breeder’s equation to the results of our selection experiment. the heritability of CH4 oxidation rate as an ecosystem property, we calculated narrow-sense heritability () as the regression of offspring on mid-parent (Goodnight 2000, Falconer 1997). The mid-parent was the mean for all three selected jars and the offspring was the mean for all jars produced by those parents.

In addition, we calculated the inbreeding coefficient, , as:

where is the population size and is the number of generations (Goodnight 2000 or Falconer 1997).

### Taxa differentially abundant in the positive selection treatment

To identify taxa that responded to selection on CH4 oxidation rate, we tested differential abundance using a beta-binomial model (Martin et al. 2020). We compared taxa in passage 5 between the positive and neutral treatment to identify taxa that

Possible tests:

* Random forest or gradient boosting to predict Positive vs. Neutral.

# Results

## Response to selection

The intercept for the Neutral treatment was significantly different from zero at -1.42 (SE = r myround(tidy(response\_model)[[1, 3]], 2), t = -12.56, p = 0.00). The Positive treatment started at a significantly lower methane oxidation rate than the Neutral treatment with a y-intercept of -0.34 relative to Neutral (SE = 0.16, t = r myround(tidy(response\_model)[[3, 4]], 2), p = 0.03). There was no change in methane oxidation rate of the Neutral treatment over the five passages (slope = -0.01, SE = 0.05, t = r myround(tidy(response\_model)[[2, 4]], 2), p = 0.80). By contrast, there was a significant increase in the Positive treatment of 0.18 per passage which is approximately a 51% increase in methane oxidation rate per passage (SE = 0.06, t = 2.76, p = 0.01).

Methane oxidation rate increased as a response to selection in the Positive selection treatment (Figure 3). This can be seen in the increase in deviation between the Positive and Neutral selection lines over the course of the experiment (z = , p = ). Mean CH4 oxidation rate for the Positive treatment increased by per passage relative to the Neutral treatment. The mean flux across all jars in all treatments in the first passage was so this represents a doubling of CH4 consumption in every passage due to selection.

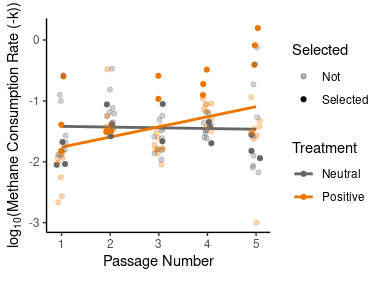


Figure 1: Deviance of CH4 oxidation rate for the Positive selection treatment relative to the Neutral selection treatment. Deviance is calculated as the mean of P within each passage minus the mean of N within each passage. Errorbars are the standard error of the difference between P and N for each passage. The best fit line is an ordinary least squares fit taking into account the standard errors of the deviations. The gray ribbon is the 95% confidence interval for the regression.

## Ecosystem heritability

Here, I fit two models. The first is simply offspring flux on selected parent flux, ignoring treatment. There is no correlation. Next, I added treatment and the interaction between treatment and parental flux. Here, there is a significant correlation between parent and offspring for both the Neutral treatment (slope = -0.58, SE = 0.20, t = -2.89, p = 0.04) and the Positive treatment (slope = 1.08, SE = 0.24, t = 4.43, p = 0.01). Notably, the sign of the effect is reversed: negative for the Neutral treatment and positive for the Positive treatment.

The negative heritability of methane oxidation rate in the Neutral selection treatment may result from selection on alternative characters in those populations. The strong, positive heritability in the Positive treatment … . It seems strange that the Methane oxidation rate was heritable between consecutive passages (Figure 2). Specifically, the narrow-sense heritability, calculated as the slope of the mid-parent and mid-offspring, was = (t = , p = ). With an inbreeding coefficient of ,

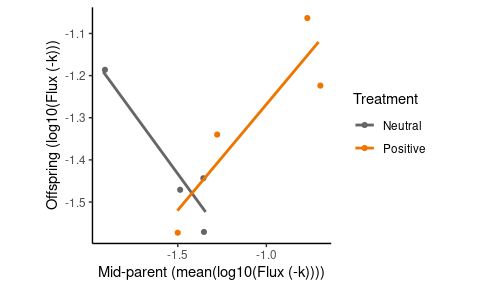


Figure 2: Regression of offspring CH4 flux on mid-parent CH4 flux. Mid-parent is the mean of jars selected to inoculate the next passage. The mid-offsprings are the mean of all twelve jars produced in one passage. The best fit line is an ordinary least squares regression line and the gray ribbon is the 95% confidence interval for the regression coefficient.

# Discussion

## Future Directions

This study focused on the practical implications of using AES as a tool to deduce how microbes map onto functions at the whole-ecosystem level. If one were interested in the evolutionary dynamics underlying this process, one could design a study to distinguish between individual selection on microbial strains and group selection or multilevel selection. That is, selection that cannot be recapitulated through selection at the individual level (Williams and Lenton 2007). This could be accomplished by culturing differentially abundant taxa in the positive selection treatment and then inoculating them in isolation (or in a non-selected background community) to see if the same level of ecosystem function could be recapitulated as in the final generation of the positive selection treatment.

Another avenue for research could be to look at correlated responses of ecosystem-level traits to selection. This would address questions such as, “Is methane oxidation rate correlated with other ecosystem processes due to pleiotropy or linkage disequilibrium?” At the community level, pleiotropy might be defined as “the production by a single taxon of two or more unrelated effects.” Further, linkage disequilbrium would represent correlated traits due to taxa who tend to co-associate due to shared environmental preferences or coupled metabolic pathways.

These experiments would further help lus understand the nature of the “map” between community structure and ecosystem function. This would not only reveal how best to predict ecosystem function in earth system models, but also demonstrate the mechanism underlying that relationship. This is akin to deducing the genetic control of trait variation in molecular biology.

Community epistasis is analogous to functional redundancy.

# Acknowledgments

This project was supported by the National Science Foundation Graduate Research Fellowship Program (grant no. DGE 1255832) and the ARCS Foundation Florence and Mike Nudelman Scholarship.

# Competing Interests

We declare we have no competing interests.

# References