

# Environmental fluctuations reshape an unexpected diversity-disturbance relationship in a microbial community

[Original paper](#) by: Christopher Mancuso et al (2021) [1]

Re-analysis by: Clay Swackhamer

## Abstract

Disturbances to an environment can influence the diversity of species it can support. However, the impact of a disturbance's intensity and the frequency that it occurs on the diversity of species in the environment is unclear. In this study, a soil microbial community was grown using an automated culturing system that applied controlled disturbances of predetermined intensity and frequency to the community using dilution as a model disturbance. 16s rRNA sequencing was used to determine the diversity of microbial species under each combination of disturbance intensity (mean dilution rate, % *volume hr*<sup>-1</sup>) and frequency (number of dilutions *day*<sup>-1</sup>). Unexpectedly, a U-shaped curve was observed between Shannon diversity and mean dilution rate at high dilution frequencies. However, at low dilution frequencies the U-shape curve was not present, indicating that diversity was preserved across a wide range of mean dilution intensity. The microbial community was modeled using a Monod consumer resource model, which was able to reproduce the effect observed by experiments. These findings could help understand and predict the effect of disturbances on environments in varying contexts, from forests affected by occasional wildfires to gut microbiomes affected by antibiotic usage.

## 1 Introduction

Biodiversity is a key feature of ecosystems. Higher diversity in an environment has been associated with greater stability and overall productivity [2]. Disturbances to an environment can influence the abundance of species in the environment, however, there has not yet been a systematic means to make testable predictions regarding the effect of a given disturbance on environmental diversity - an effect that has been referred to as a diversity-disturbance relationship. Here, two characteristics of an environmental disturbance were considered: 1) its intensity, defined as the average mortality rate imposed by the disturbance when it occurs, and 2) its frequency, defined as the number of instances of the disturbance in a given period of time.

Previous researchers have proposed the intermediate disturbance hypothesis [3], which states that diversity peaks at intermediate disturbance intensity. According to this hypothesis a small number of species that have superior fitness to their peers will dominate undisturbed environments, where conditions are relatively constant, leading to low diversity at low disturbance intensity. At high disturbance intensity only robust species can persist, excluding those with naturally slower growth rates, once again leading to low diversity. As low diversity is predicted at both high and low disturbance intensity, intermediate disturbance intensity is predicted to maximize diversity.

However, the prediction of an "upside down U" shaped diversity-disturbance relationship by the intermediate disturbance hypothesis has not always been consistent with experimental observations [4, 5]. This could point to a missing variable in the hypothesis, namely, the disturbance frequency. For example, a occasional disturbance could introduce temporal niches similar to seasonal effects [6, 7]. This could allow certain species to thrive when the disturbance is present or in its immediate aftermath, which may transition to dormancy as the disturbance recedes and the ecosystem returns to its undisturbed state, allowing a different set of species to thrive. In essence, an infrequent disturbance could create two pseudo-steady states in the ecosystem, an undisturbed state and a disturbed state, with the possibility for different species to thrive in each state and greater overall diversity as a result.

Previous studies have considered the impact on diversity due to disturbances such as fluctuating nutrient levels [8], sonication [9], ultraviolet light [10], osmotic pressure [11] or the presence of toxic compounds [12]. In this study, diversity disturbance relationships were evaluated for a soil-derived microbial community cultured using eVOLVER [13], an automated miniature bioreactor array which can be used to tightly regulate environmental conditions. Compared to larger-scale ecosystems in nature, this *in vitro* approach has the advantages of more rapid and inexpensive experiments, better control of environmental variables, and access to a 16s rRNA sequencing pipeline to determine diversity under the different experimental conditions that were imposed.

Dilution was chosen as a model disturbance as it causes species-independent mortality but also replenishes the environment with fresh resources, which may have parallels to environments such as the gut. Diversity was determined for the soil microbial community under different mean dilution rates and dilution frequencies. A canonical model of microbial population dynamics was applied to determine whether Monod growth kinetics would be sufficient to predict the diversity disturbance relationships experienced by the microbial community under disturbances of varying intensity and frequency.

## 2 Materials and methods

### 2.1 Experiments

A soil sample (2 g) were obtained from the Communications Lawn at Boston University on September 15, 2018 (coordinates: 42.349222 N, 71.101878 W). The sample was vortexed with 10 mL Phosphate Buffered Saline (PBS) solution + 200  $\mu\text{g}/\text{mL}$  cycloheximide and then incubated for 48 hours in the dark. Next, 350  $\mu\text{L}$  of the soil + PBS mixture were used to inoculate 25 mL of 0.1 X Nutrient Broth media (0.3 g/L yeast extract + 0.5 g/L peptone, Fisherbrand) with 200  $\mu\text{g}/\text{mL}$  cycloheximide, which was incubated for 18 hr at 25°C and used to prepare a cryostock in 15% glycerol which was frozen at -80 °C.

Cryostocks were used to inoculate mini bioreactors in the eVOLVER system [13], an automated, high-throughput system that was used to incubate samples at predetermined average dilution rates (0.1, 0.2, 0.3, or 0.4  $\text{hr}^{-1}$ ) and dilution frequencies (1, 4, 16, or 600 dilutions per day) for 6 days. For example, if a mean dilution rate of 0.2  $\text{hr}^{-1}$  and frequency of 16 were applied, then the culture was diluted 16 times per day, with 30% of the culture exchanged for fresh media on each dilution (0.3 = 0.2 \* 24/16).

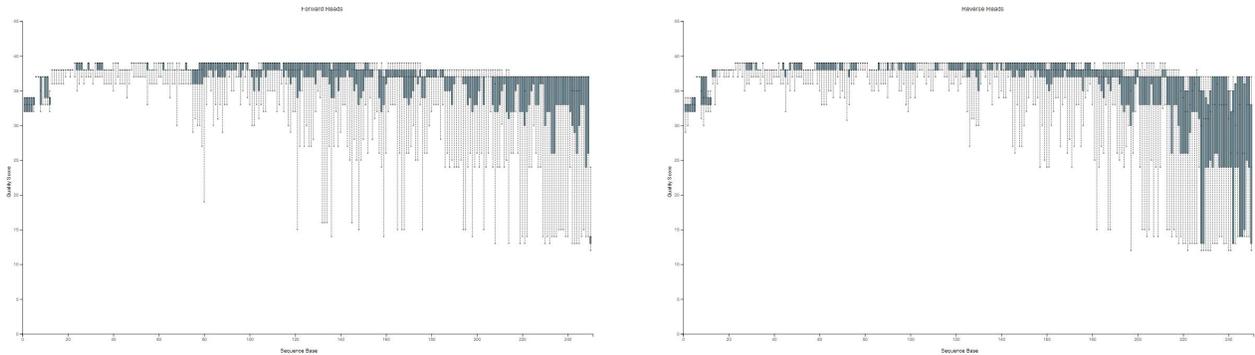
On day 6, 2 mL culture was extracted from each bioreactor, centrifuged, frozen at  $-20^{\circ}\text{C}$  for up to 72 hr, lysed at  $37^{\circ}\text{C}$  in 200  $\mu\text{L}$  lysozyme buffer, processed using the DNEasy Blood and Tissue Kit and then normalized to 5 mg DNA/ $\mu\text{L}$ .

The 16s v4 region was sequenced according to a previously described method [15] with primers prCM543 and prCM544. Sequencing was carried out at the Harvard University Biopolymers Facility. Samples were demultiplexed using the Illumina BaseSpace tool. All preceding steps were carried out by the authors of the original study [1]. Subsequent steps were carried out as part of the re-analysis.

## 2.2 Data analysis

Data from the original study were obtained from the NCBI Sequence Read Archive under BioProject accession code PRJNA719465. Samples were obtained for all combinations of mean dilution rate and frequency (4 dilution rates \* 4 frequencies \* 4 replicates =64 samples). One sample was excluded due to contamination, leaving 63 samples that were part of the re-analysis.

Samples were demultiplexed using Qiime2 (version 2021.11) using the command dada2 [16] to determine Amplicon Sequence Variants (ASVs). A left trim of 15 bases and truncation length of 200 bases was applied to all forward and reverse reads.

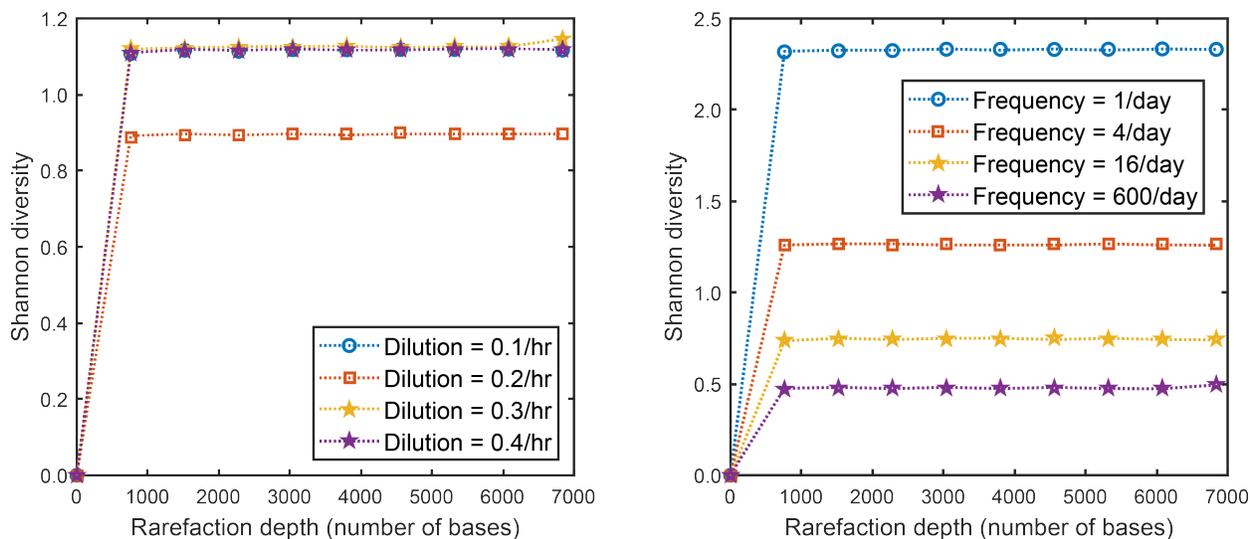


**Figure 1.** Sequence quality plots used to determine trim and truncation parameters.

**Table 1.** Sequence accounting for DADA2.

DADA2 Step	# Sequences	Percent retained
Input	2,652,996	100%
Filtered	2,528,182	95.3%
Denoised	2,521,952	95.1%
Merged and chimeras removed	2,323,800	87.6%

A phylogenetic tree was constructed using phylogeny align-to-tree-mafft-fasttree and samples were rarefied to 6840 features using the command diversity alpha-rarefaction, retaining 62 of the 63 samples.



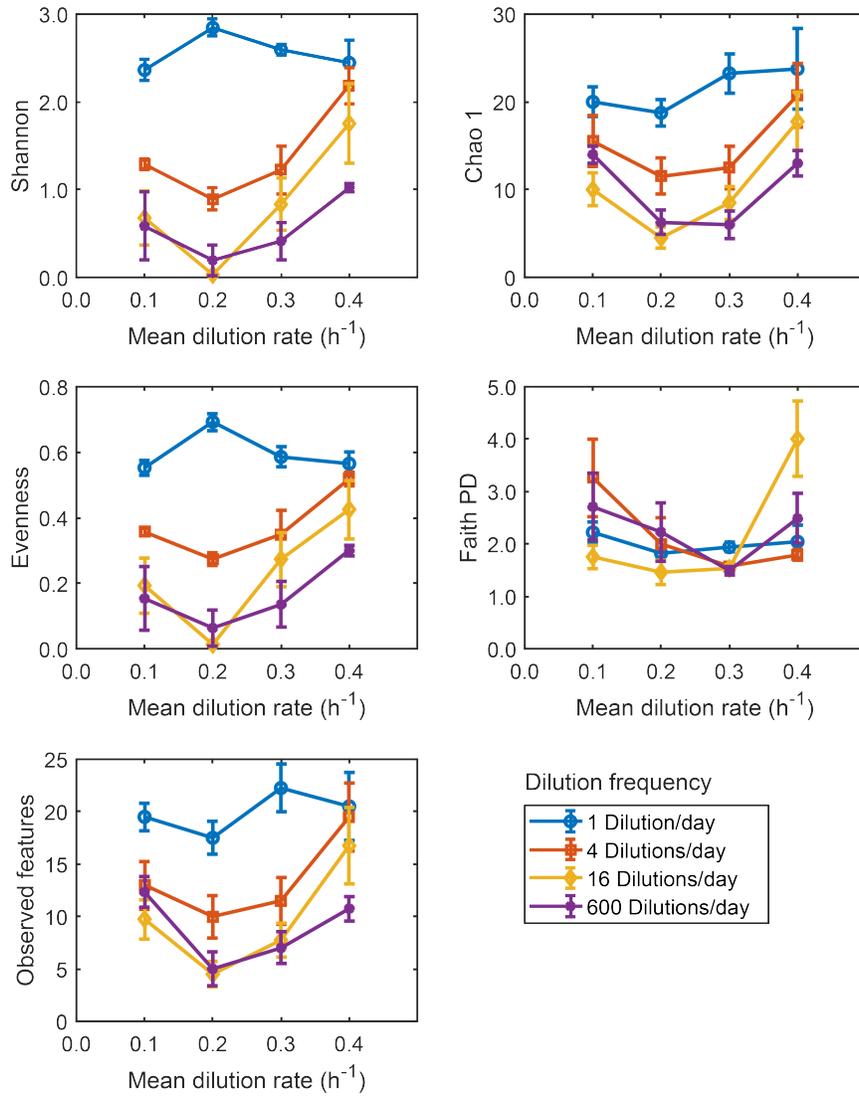
**Figure 2.** Shannon diversity vs rarefaction depth. Each value is the average from 10 iterations.

Taxonomy of each ASV was assigned using `feature-classifier classify-sklearn` with a pretrained Naive Bayes classifier trained on the Greengenes 13\_8 database: [gg-13-8-99-515-806-nb-classifier.qza](#).

## 3 Results

### 3.1 Alpha diversity

Core metrics of alpha diversity revealed an unexpected, U-shaped relationship between the diversity of the microbial community and the mean dilution rate (Figure 3), but only when dilution was applied somewhat frequently (4, 16, or 600 times per day), with the effect most evident at higher dilution frequencies. Alpha diversity metrics thus showed a loss of diversity at intermediate dilution rates, contrary to the intermediate disturbance hypothesis. However, diversity was preserved (loss of U-shape) when dilution was infrequent, most evident in the case where dilution was carried out only once per day. It should be noted that all samples were cultured independently under the prescribed experimental conditions, that is, the community exposed to a dilution rate of  $0.1 \text{ hr}^{-1}$  at a frequency of  $1 \text{ day}^{-1}$  was grown under those conditions in a separate bioreactor vessel than the community exposed to a dilution rate of  $0.2 \text{ hr}^{-1}$  at a frequency of  $1 \text{ day}^{-1}$ . Statistical analysis of alpha diversity metrics was carried out using the Kruskal-Wallis test (Table 2).



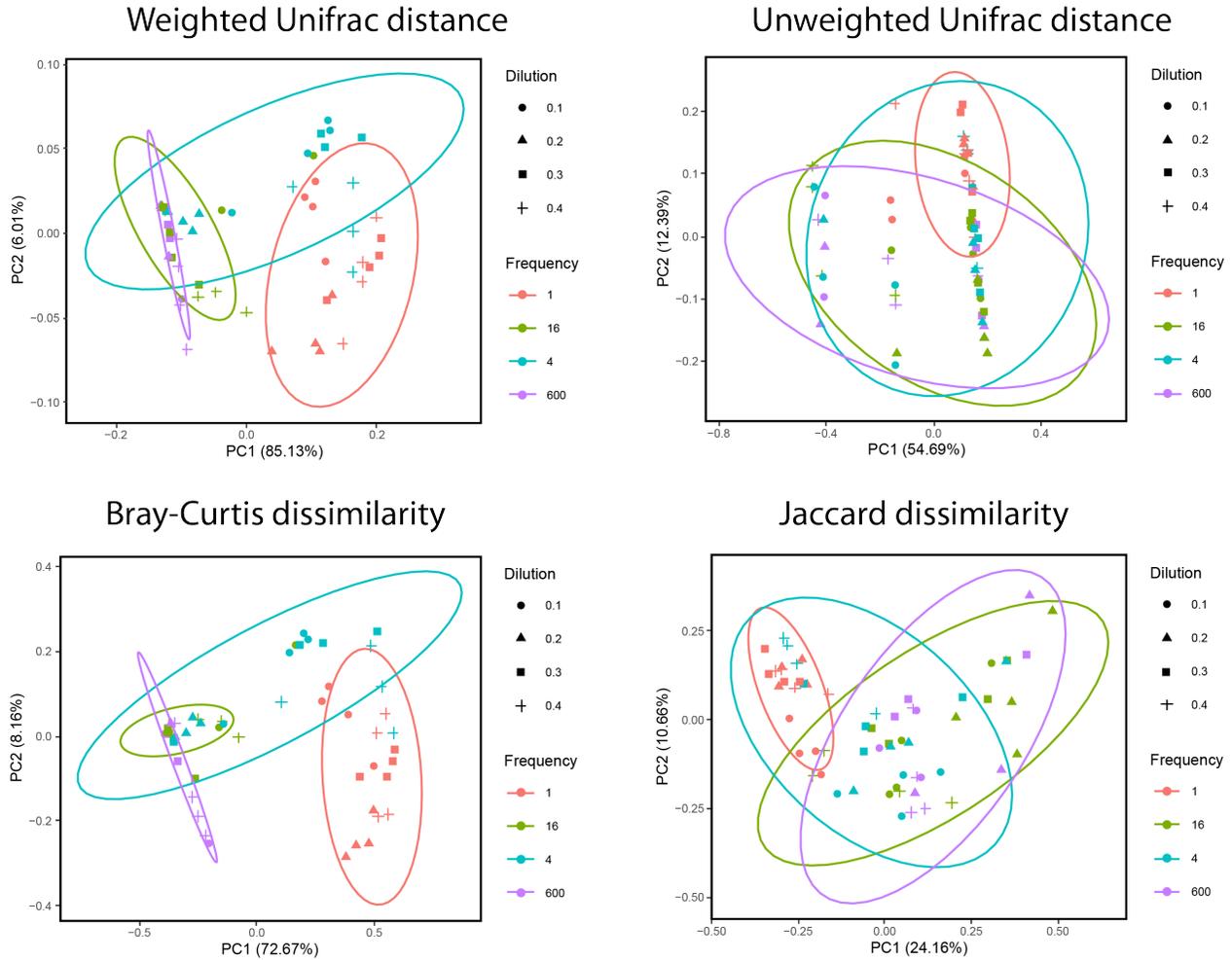
**Figure 3.** Core metrics of alpha diversity for the soil microbial community grown under varying mean dilution rates and frequencies.

**Table 2.** Statistical analysis of core metrics of alpha diversity. p-value classifications are shown for each main effect in the experiment (mean dilution rate and frequency) for each metric.

Diversity metric	$p_{\text{mean dilution rate}}$	$p_{\text{frequency}}$
Shannon	> 0.05	> 0.05
Chao1	< 0.01	< 0.01
Evenness	> 0.05	< 0.01
Faith phylogenetic diversity	< 0.01	> 0.05
Observed features	< 0.05	< 0.01

### 3.2 Beta diversity

Core metrics of beta diversity showed that the samples were clustered strongly by frequency and weakly by mean dilution rate (Figure 4). Beta diversity was analyzed using principal coordinate analysis (diversity beta-group-significance) in Qiime2 followed by PERMANOVA with the pairwise-adonis function in R (R version 4.2.2 using R studio version 2023.0.3). Significance levels of main effects are shown in Table 3. All metrics of beta diversity were significantly influenced by the mean dilution rate and frequency.



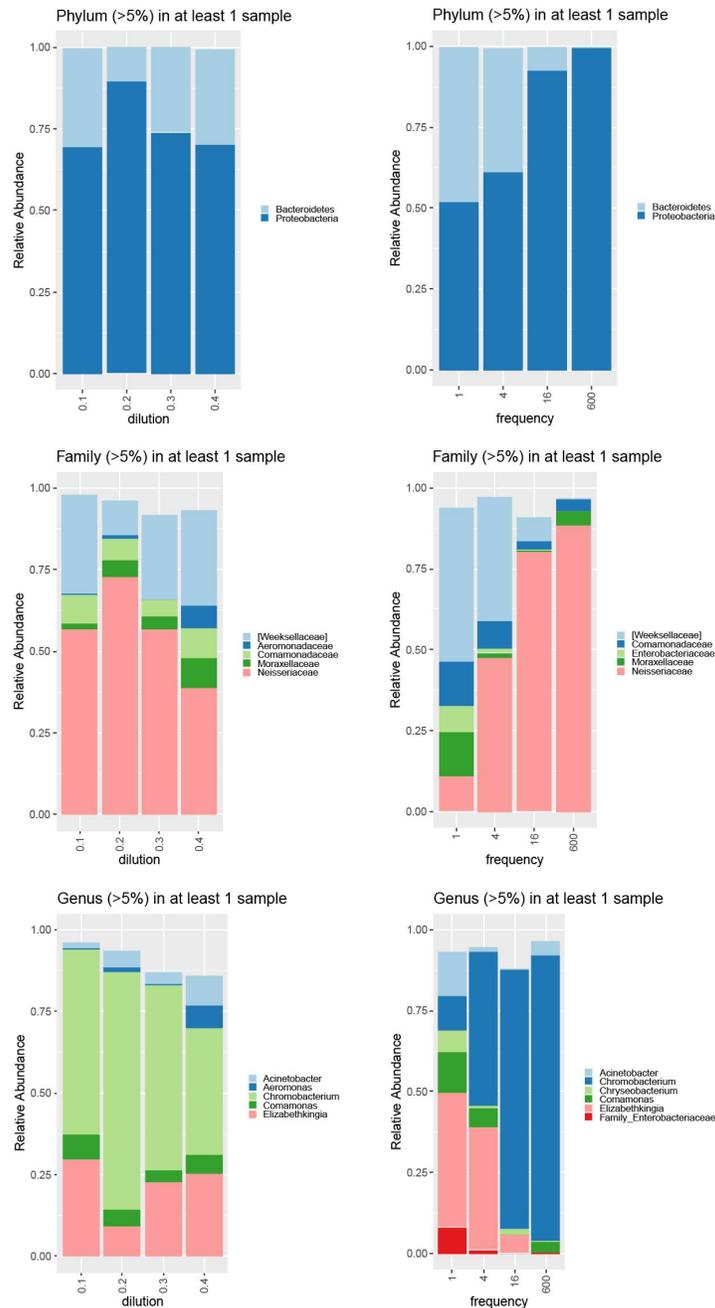
**Figure 4.** Core metrics of beta diversity for the soil microbial community grown under varying mean dilution rates and frequencies.

**Table 3.** Statistical analysis of core metrics of beta diversity. p-value classifications are shown for each main effect in the experiment (mean dilution rate and frequency) for each metric.

Diversity metric	$p_{\text{mean dilution rate}}$	$p_{\text{frequency}}$
Weighted Unifrac	< 0.01	< 0.01
Unweighted Unifrac	< 0.01	< 0.01
Bray-Curtis	< 0.01	< 0.01
Jaccard	< 0.01	< 0.01

### 3.3 Taxonomic composition

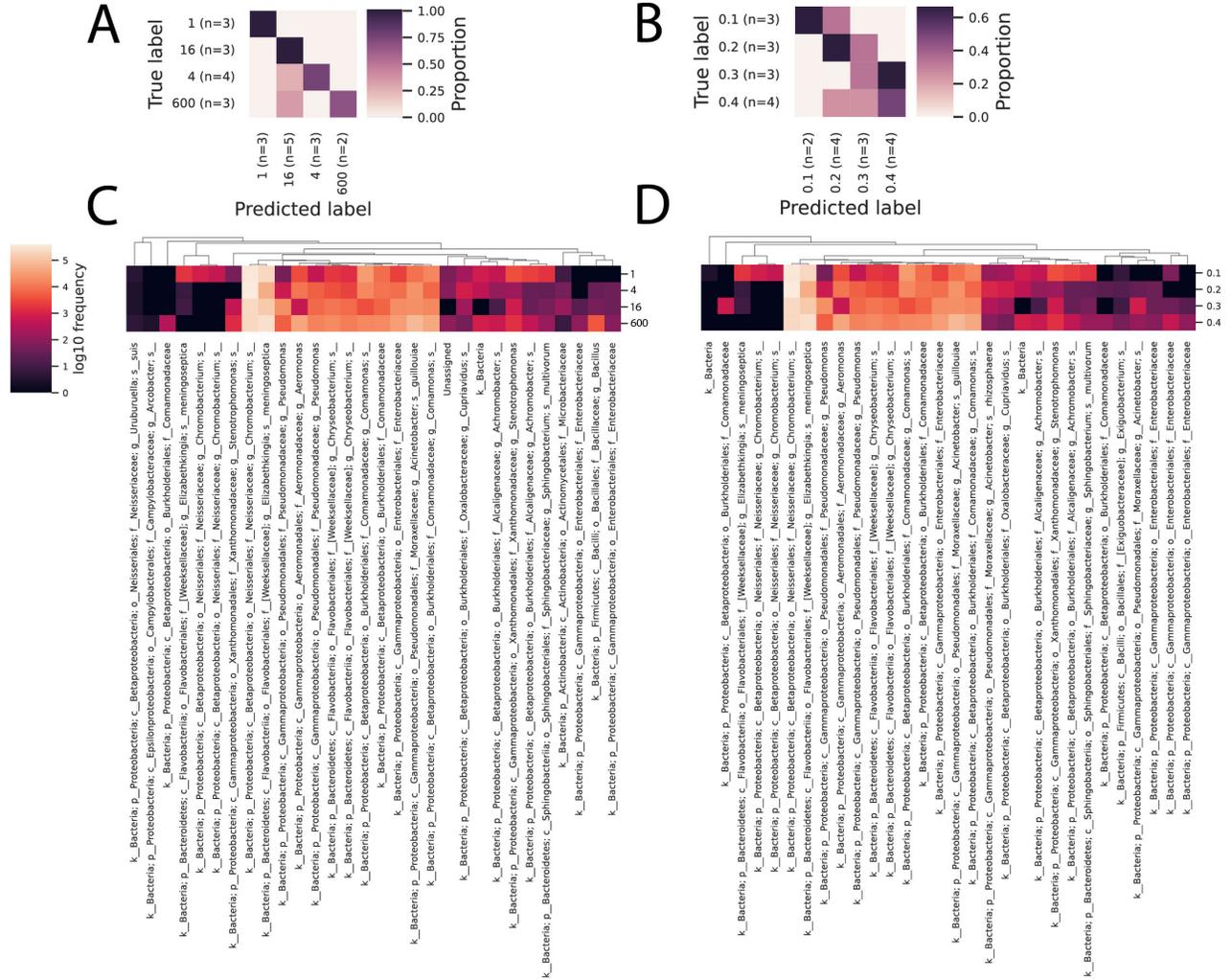
Taxonomy was analyzed by determining the relative abundance of different microorganisms at the phylum, family, and genus levels (Figure 5). The soil microbial community consisted of mostly Bacteroidetes and Proteobacteria. The relative abundance of Proteobacteria did not increase with the mean dilution rate but increased with frequency. Alpha diversity analysis revealed a U-shaped relationship between community diversity and dilution rate at high frequency, but no relationship at low frequency. Here, comparing the relative abundance bar graphs for a frequency of 600 dilutions per day to the graphs for frequency of 1 dilution per day supported that the community was more diverse at both the family and genus level when dilution was infrequent (1 dilution per day).



**Figure 5.** Relative abundance of microorganisms determined at different taxonomic levels.

### 3.4 Sample classification using machine learning

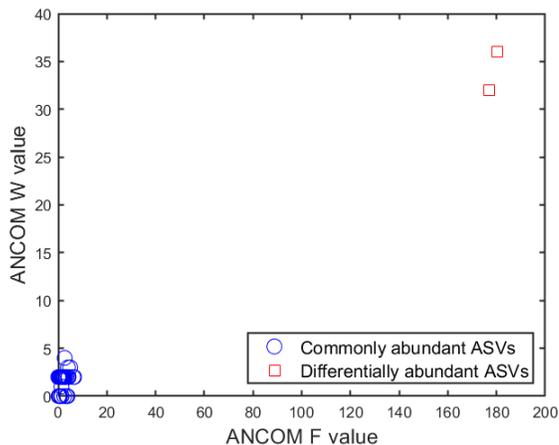
Samples were classified using the Qiime2 command `sample-classifier classify-samples`. A random forest classifier with 20 estimators and 123 random states was used. The model predicted frequency with 84.6 % accuracy (Figure 6 A), and mean dilution rate with 53.8 % accuracy (Figure 6 B). A feature importance heatmap was generated using the top 30 features (Figure 6 C and D). One microorganism that appeared to be influential to predictive models of both frequency and dilution rate was *Elizabethkingia Meningoseptica*, a soil-derived, gram negative bacteria that can cause nosocomial infections [17].



**Figure 6.** Classification of microbial community using Random Forest. Confusion matrix for classification by frequency of dilution (A). Confusion matrix for classification by mean dilution rate (B). Feature importance heatmap for classification by frequency of dilution (C). Feature importance heatmap for classification by mean dilution rate (D).

### 3.5 ANCOM

As a follow up analysis, samples that were grown with 600 dilutions per day were selected for Analysis of Composition of Microbiomes (ANCOM) using Qiime commands `composition ancom`. Since these samples displayed significant differences in alpha diversity across the range of mean dilution rates that were studied, it was desired to identify any differentially abundant taxa between these samples. Two ASVs had a sufficiently high W-statistic to reject the null hypothesis and be recognized by ANCOM as differentially abundant. These two ASVs are shown as red squares in Figure 7 and their taxonomy is shown in Table 4. Intriguingly, both ASVs belonged to the same taxonomic group to the genus level (Chromobacterium) and both were unclassified at the species level.



**Figure 7.** ANCOM plot showing that two ASVs had much higher leverage than the others.

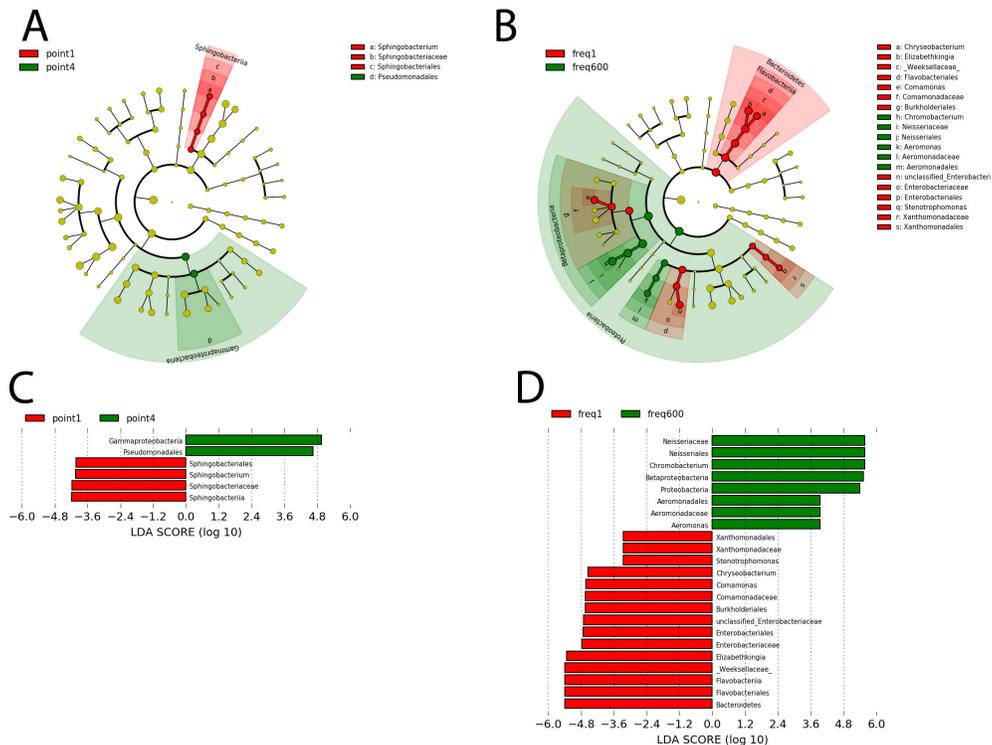
**Table 4.** Taxonomic identity of two ASVs that were differentially abundant among samples diluted at a frequency of  $600 \text{ day}^{-1}$ :

Taxa	ANCOM $F$	ANCOM $W$
Bacteria-Proteobacteria-Betaproteobacteria Neisseriales-Neisseriaceae-Chromobacterium	180.2	36
Bacteria-Proteobacteria-Betaproteobacteria Neisseriales-Neisseriaceae-Chromobacterium	177.0	32

### 3.6 LEfSe

Another analysis of differentially abundant taxa was carried out using LEfSe (Linear discriminant analysis Effect Size) [18]. Considering all samples simultaneously (across all ranges of mean dilution rate and frequency) no significant differentially abundant taxa were identified. However, the effect of dilution rate and frequency on alpha diversity was found to be nonlinear and likely contained a significant interaction between the main effects. Therefore, a follow up analysis was carried out by selecting only samples at extreme values of dilution rate or frequency, and then considering each analysis separately. For the first case, dilution was set as the main class with frequency as the subclass, and samples with either dilution rates of  $0.1 \text{ hr}^{-1}$  or  $0.4 \text{ hr}^{-1}$  but with all four frequency values still represented were analyzed (Figure 7 A and C). For the second case, frequency was set

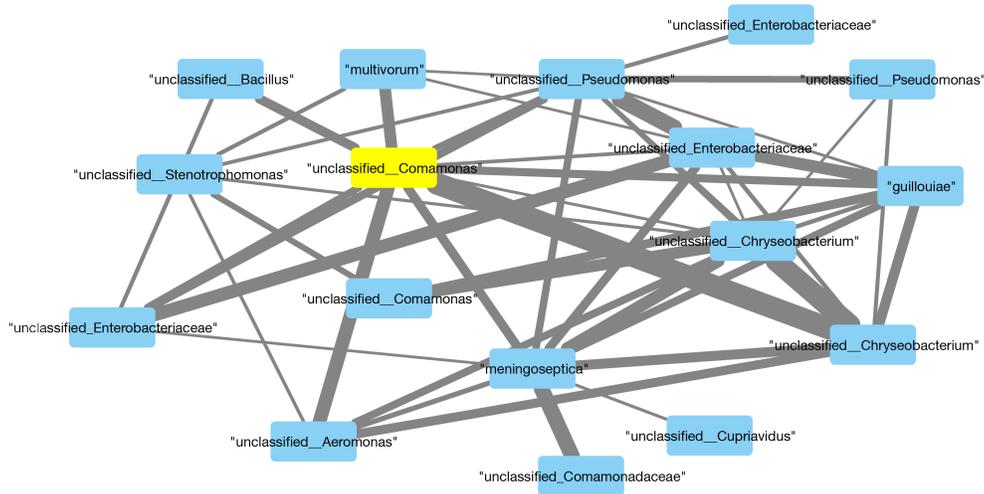
as the main class with dilution as the subclass, and samples with either frequencies of  $1 \text{ day}^{-1}$  or  $600 \text{ day}^{-1}$  but with all four dilution values still represented were analyzed (Figure 7 B and D). In both cases, the alpha value for factorial Kruskal-Wallis tests among classes and the alpha value for pairwise Wilcoxon test between subclasses were set to 0.05, with a threshold for the logarithmic LDA score for discriminative features set to 2.0 and the comparison strategy set to "all against all."



**Figure 7.** Results from LefSe. Cladogram showing differentially abundant taxa between samples at the extremes of dilution rate ( $0.1 \text{ hr}^{-1}$  vs  $0.4 \text{ hr}^{-1}$ ) (A). Cladogram showing differentially abundant taxa between samples at the extremes of dilution frequency ( $1 \text{ day}^{-1}$  vs  $600 \text{ day}^{-1}$ ) (B). Bar plot showing taxa that significantly differed in the samples at extremes of dilution rate (C) and frequency (D).

### 3.7 Cooccurrence

Cooccurrence analysis was carried out by filtering out sparse taxa, which were defined as those having total abundance below 500 and average abundance  $< 2$  across all samples, in agreement with previous researchers [19]. Co-correlated taxa with  $\rho > 0.75$  were included in the network, which was visualized using Cytoscape (Figure 8).



**Figure 8.** Cooccurrence network generated using all taxa across all samples with correlation  $\rho > 0.75$ .

## 4 Mathematical model of the microbial community

### 4.1 Modeling overview

Mathematical modeling was used to determine whether the U-shaped diversity-disturbance relationship observed in experiments would result from simulation of microbial growth under varying mean dilution rates and dilution frequencies chosen to match the values used in the experiments. A Monod consumer resource model [14] was chosen as it is capable of predicting the growth of multiple species of microorganisms in a system with several resources, where each species can have a unique growth rate on each resource. The Monod model assumes that the growth rates of a microorganism can be represented by the sum of its specific growth rate on each resource in the system, and allows for growth rates to saturate at high resource concentrations. The Monod model couples microbial populations to resource concentrations, allowing the concentration of resources to change as they are depleted by microorganisms.

The inputs to the Monod model consist of an initial value for the relative abundance of different species of microorganisms in the system, the initial concentration of each resource, the specific growth rate of each microorganism on each resource, and the Monod constant for each microorganism with respect to each resource. The outputs of the Monod model were the relative abundance of each species of microorganism in the system, from which Shannon diversity could be calculated, and the concentration of each resource in the system at each timepoint.

The hypothesis was that the U-shaped diversity-disturbance relationship would result from simulation of multiple microorganisms growing in a system with multiple resources, demonstrating that Monod-style growth kinetics are sufficient to predict such an effect, even in a complex microbial community.

Since the specific growth rates and Monod constants for all the microorganisms in the community that was studied experimentally were not known, they were randomly generated according to distributions that would cast them into ranges similar to those determined in previous experiments [1].

This introduced a limitation of the model, which is that it cannot predict the absolute abundance of any particular microorganism or resource.

In essence, the model did not incorporate biological data from any microorganism, but instead described how the interactions between multiple (theoretical) microorganisms growing in a system with multiple resources could be influenced by different mean dilution rates and dilution frequencies. If the model were to recapitulate the U-shaped diversity-disturbance relationship observed in experiments, it could enable researchers to make testable predictions about such relationships in other systems, for example between forest fire intensity and frequency on the diversity of forests, or the effect of antibiotic severity and dose regime on the diversity of gut microbiomes.

## 4.2 Model definition

The basic form of the Monod model [14] defines the specific growth rate of a microorganism,  $\mu$  as a function of its maximum growth rate on a particular resource  $r$ , the concentration of the resource in the microorganism's environment,  $c$ , and the Monod constant,  $K$ , which gives the concentration of the resource at which the microorganisms growth rate is one-half of its maximum value:

$$\mu = \frac{rc}{K + c}$$

The specific growth rate,  $\mu$ , is defined as the rate of change in population of the microorganism  $dN/dt$ , relative to the total number of cells,  $N$ :

$$\mu = \frac{1}{N} \frac{dN}{dt}$$

The model can be extended to consider the growth of the microorganism when multiple resources are available by defining the overall growth rate of the microorganism as the sum of its specific growth rates on each resource. In this case, the resource concentration,  $c$ , becomes a vector with the number of entries equal to the number of resources:  $c_j = [c_{resource\ 1}, c_{resource\ 2}, \dots, c_{resource\ J}]$ . Substituting the definitions of the resource concentration vector,  $c_j$  and the specific growth rate,  $\mu$ , the model becomes:

$$\frac{1}{N} \frac{dN}{dt} = \sum_{j=1}^J \frac{r_j c_j}{K_j + c_j}$$

Multiple species of microorganisms can be considered by replacing the species abundance,  $N$ , with a vector giving the abundance of each species:  $N_i = [N_{species\ 1}, N_{species\ 2}, \dots, N_{species\ I}]$ . Making this substitution, the model for microbial growth is obtained:

$$\frac{1}{N_i} \frac{dN_i}{dt} = \sum_{j=1}^J \frac{r_{ij} c_j}{K_{ij} + c_j}$$

Where:

$N_i$  = Abundance of species  $i$  divided by its carrying capacity

$\frac{dN_i}{dt}$  = Rate of change of species  $i$  with respect to time

$r_{ij}$  = Growth rate of species  $i$  on resource  $j$

$c_i$  = Concentration of resource  $j$  in the system  
 $K_{ij}$  = Monod constant for species  $i$  on resource  $j$   
 $J$  = Total number of resources in the system

Next, the model was modified to include a dilution term ( $\delta$ ):

$$\frac{1}{N_i} \frac{dN_i}{dt} = \sum_{j=1}^7 \frac{r_{ij}c_j}{K_{ij} + c_j} - \delta$$

The dilution term in the model was calculated by dividing the mean dilution rate,  $D$ , (in units of % total volume per hour) over equally spaced 15 minute intervals at the relevant frequency ( $f$ ):

$$\delta = \frac{1}{1/4hr} \frac{D(1/hr)}{f(1/24hr)} = 96 \frac{D}{f}$$

Where:

$\delta$  = Dilution rate per 15 minutes (% volume per 1/4 hour)  
 $D$  = Dilution rate per hour (% volume per hour)  
 $f$  = Frequency of dilution (dilutions per day)

The rate of change of the concentrations of each resource in the system was calculated by subtracting the rate at which it was depleted by microorganisms (consumption rate) as well as the rate at which it was replenished by fresh media during dilutions (replenish rate):

$$\begin{aligned} \frac{dc_j}{dt} &= \text{replenish} - \text{consumption} \\ \frac{dc_j}{dt} &= \delta(c_{j0} - c_j) - \sum_{i=1}^{10} \frac{N_i r_{ij} c_j}{K_{ij} + c_j} \end{aligned}$$

Where:

$\frac{dc_j}{dt}$  = Rate of change of resource  $j$  with respect to time  
 $\delta$  = Dilution rate per 15 minutes (% volume per 1/4 hour)  
 $c_{j0}$  = concentration of resource  $j$  in the fresh media used for diluting  
 $c_i$  = Concentration of resource  $j$  in the system  
 $N_i$  = Abundance of species  $i$  divided by its carrying capacity  
 $r_{ij}$  = Growth rate of species  $i$  on resource  $j$   
 $K_{ij}$  = Monod constant for species  $i$  on resource  $j$

The final model consisted of two coupled, linear, ordinary differential equations which were solved simultaneously. The first governed the population abundances of different species of microorganisms in the system and the second governed the concentrations of resources in the system:

$$\begin{aligned} \frac{1}{N_i} \frac{dN_i}{dt} &= \sum_{j=1}^7 \frac{r_{ij}c_j}{K_{ij} + c_j} - 96 \frac{D}{f} \\ \frac{dc_j}{dt} &= 96 \frac{D}{f} (c_{j0} - c_j) - \sum_{i=1}^{10} \frac{N_i r_{ij} c_j}{K_{ij} + c_j} \end{aligned}$$

The output of the model consists of a vector giving the predicted relative abundances of the different species of bacteria ( $N_i$ ) as well as a vector of the concentrations of resources in the system ( $c_j$ ) at 15 minute intervals from zero to six days after inoculation. At day 6, the simulation was stopped and Shannon diversity was calculated using:

$$S = \sum_{i=1}^{10} -\rho_i \ln(\rho_i)$$

Where:

$S$  = Shannon diversity

$\rho_i$  = relative abundance of each species =  $N_i / \sum_{i=1}^{10} N_i$

Only species representing at least 0.01% of the total community ( $\rho_i > 0.0001$ ) were included in the calculation.

### 4.3 Parameter values

The simulation was limited to 10 species. The species abundance vector was thus:  $N_i = [N_1, N_2, \dots, N_{10}]$ , where  $N_1$  is the abundance of species 1 divided by its carrying capacity. All carrying capacities were set to one, therefore  $N_i \in [0, 1]$ . The initial values (at time zero)  $N_i$  all species were set to 1.

The simulation was limited to 7 resources, giving a resource concentration vector of:  $c_j = [c_1, c_2, \dots, c_{10}]$ , where  $c_1$  is the concentration of resource 1. The initial values (at time zero) of each resource were set to 1, except for  $c_1$ , which was arbitrarily set to 1.2.

The mean dilution rates,  $D(1/hr)$ , were chosen to match the range of values used in the experiment:  $D = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2]$ . Dilution frequencies,  $f(1/day)$ , were also chosen to cover the range studied experimentally:  $f = [1, 2, 4, 8, 16, 32, 64]$ . For example,  $D = 0.1$  and  $f = 16$  indicates that the average dilution rate was 10% of the total system volume per hour, and the dilution was carried out 16 times each day, meaning that each dilution event replaced 15% of the system with fresh media.

The maximum growth rates matrix,  $r_{ij}$  was generated in two steps. First, each entry was determined by randomly sampling from a normal distribution with mean of 1 and standard deviation of 0.1, and then normalized by dividing each entry  $r_{ij}$ , representing the maximal growth rate of species  $i$  on resource  $j$ , by the sum of the same species' growth rate on all resources ( $\sum_{j=1}^7 r_{ij}$ ). In other words, each value was first sampled from the normal distribution and then divided by the sum of all values in its row. An example growth rate matrix is:

$$r_{ij} = \begin{bmatrix} 0.1534 & 0.1259 & 0.1554 & 0.1585 & 0.1441 & 0.1330 & 0.1297 \\ 0.1635 & 0.1801 & 0.1215 & 0.1223 & 0.1348 & 0.1392 & 0.1386 \\ 0.1142 & 0.1583 & 0.1581 & 0.1318 & 0.1523 & 0.1296 & 0.1557 \\ 0.1510 & 0.1382 & 0.1617 & 0.1278 & 0.1434 & 0.1236 & 0.1543 \\ 0.1490 & 0.1547 & 0.1515 & 0.1019 & 0.1319 & 0.1443 & 0.1667 \\ 0.1198 & 0.1350 & 0.1521 & 0.1577 & 0.1374 & 0.1590 & 0.1390 \\ 0.1405 & 0.1451 & 0.1576 & 0.1517 & 0.1445 & 0.1356 & 0.1250 \\ 0.1456 & 0.1618 & 0.1365 & 0.1302 & 0.1496 & 0.1460 & 0.1303 \\ 0.1776 & 0.1492 & 0.1346 & 0.1487 & 0.1451 & 0.1278 & 0.1169 \\ 0.1674 & 0.1497 & 0.1208 & 0.1087 & 0.1457 & 0.1458 & 0.1619 \end{bmatrix}$$

The Monod constant matrix,  $K_{ij}$  gives the concentration of resource  $j$  at which species  $i$  achieves half of its maximal growth rate on that resource. This matrix was generated by randomly sampling from a uniform distribution with lower bound 0.001 and upper bound 0.01. An example Monod constant matrix is:

$$K_{ij} = \begin{bmatrix} 0.0033 & 0.0085 & 0.0015 & 0.0081 & 0.0018 & 0.0097 & 0.0092 \\ 0.0083 & 0.0063 & 0.0058 & 0.0038 & 0.0031 & 0.0010 & 0.0026 \\ 0.0032 & 0.0059 & 0.0080 & 0.0058 & 0.0092 & 0.0080 & 0.0034 \\ 0.0094 & 0.0093 & 0.0094 & 0.0025 & 0.0024 & 0.0084 & 0.0023 \\ 0.0041 & 0.0036 & 0.0022 & 0.0064 & 0.0084 & 0.0088 & 0.0022 \\ 0.0028 & 0.0078 & 0.0061 & 0.0034 & 0.0058 & 0.0018 & 0.0088 \\ 0.0033 & 0.0078 & 0.0052 & 0.0069 & 0.0100 & 0.0046 & 0.0062 \\ 0.0065 & 0.0044 & 0.0011 & 0.0072 & 0.0017 & 0.0033 & 0.0059 \\ 0.0053 & 0.0061 & 0.0040 & 0.0077 & 0.0050 & 0.0082 & 0.0023 \\ 0.0042 & 0.0017 & 0.0025 & 0.0051 & 0.0020 & 0.0049 & 0.0087 \end{bmatrix}$$

## 4.4 Solution method

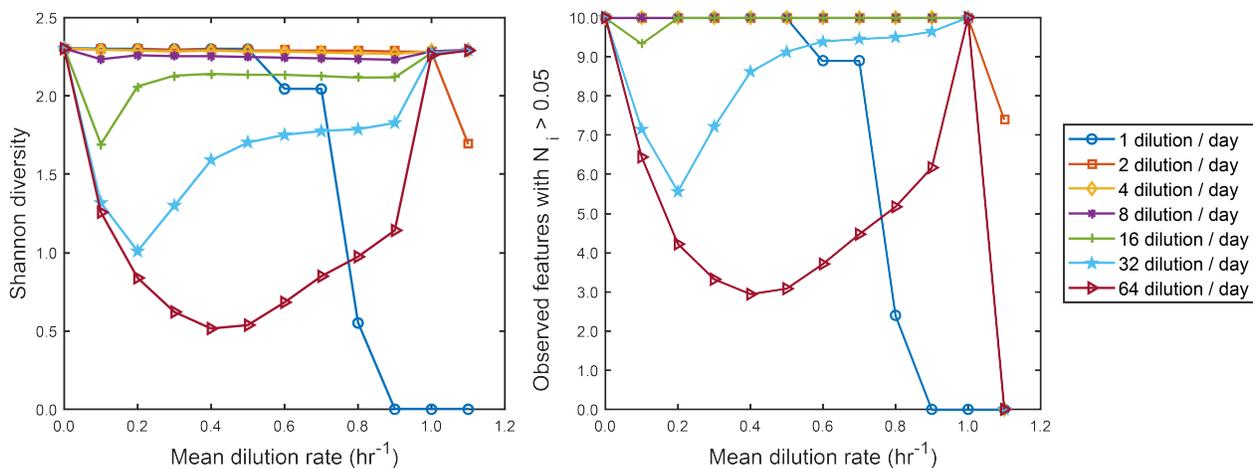
For each simulation, one growth rate matrix  $r_{ij}$  and one Monod constant matrix  $K_{ij}$ , both constant with respect to time, were generated as previously described. Mean dilution rate,  $D$ , and frequency of dilution,  $f$ , were selected and held constant throughout the simulation. The species abundance vector  $N_i$  and resource concentration vector  $c_j$  were determined at 15 minute intervals for a six day period by solving the equations in the box. At the end of the simulation, the species abundance vector was used to determine Shannon diversity as previously described. For each combination of factor levels (mean dilution rate and frequency) 100 simulations were run, each with a new, randomly generated growth rate matrix and Monod constant matrix. The Shannon diversity on day 6 for all 100 simulations were averaged to determine the overall estimate of Shannon diversity of the community under the mean dilution rate and frequency values that were used.

The model was solved numerically using Python (version 3.11.0) in Jupyter Lab (version 3.5.1) with `scipy` (version 1.9.3). Integration was carried out using the function `solve_ivp` with the explicit Runge-Kutta method of order 2 (`method=RK45`). Total solution time for 100 simulations of 6 days of microbial growth at all combinations of mean dilution rate and frequency was approximately 2 hr on a desktop computer (64 GB RAM, Intel i7-8700k CPU @ 3.7 GHz).

## 4.5 Model results

Results from the model were the Shannon diversity and number of observed features (species) at each combination of mean dilution rate and frequency (Figure 9). The unexpected U-shaped diversity disturbance relationship observed from experiments was predicted, suggesting that the consumer resource model with Monod growth kinetics can describe such an effect even when the model was parameterized without specific biological data for any particular microorganism.

At high mean dilution rates and low frequencies (ex. dilution =  $0.8 \text{ hr}^{-1}$  and frequency =  $1 \text{ day}^{-1}$ ) the Shannon diversity and number of species dropped rapidly to zero, indicating that extinction was predicted under these conditions. Such extreme values of dilution were not studied experimentally, however, under these conditions the system would be diluted once per day with a total volume of 192% of the bioreactor's volume in a single 15-minute time step, so it is reasonable that under these conditions no microorganism could withstand the dilution pressure. From the model results, it can be observed that a system diluted once per day would be predicted to become extinct if the mean dilution rate exceeds approximately  $0.8 \text{ hr}^{-1}$ , whereas a system diluted twice per day would only be driven to extinction if the mean dilution rate exceeded approximately  $1.2 \text{ hr}^{-1}$ , highlighting the strong impact of disturbance frequency on the diversity of the system.



## 5 Discussion

Results from this study showed that a U-shaped diversity disturbance relationship resulted from analysis of a soil-derived bacterial community grown under varying conditions of mean dilution rate and dilution frequency.

This relationship was present in all metrics of alpha diversity (Figure 1) but only for frequent dilutions. When dilution was infrequent (ex. 1 dilution per day) diversity was maintained across a wide range of mean dilution rates. Although the U-shaped relationship identified in this study is contrary to the prediction of the intermediate disturbance hypothesis [3], the preservation of diversity under low-frequency disturbance is consistent with previously proposed idea that such disturbances could lead to two quasi-steady states in the system, similar to temporal niches or seasonal effects [6, 7]. Overall, results from alpha diversity analysis suggests that the frequency of an environmental disturbance should be considered in addition to its intensity (here defined as the mean mortality

introduced by the disturbance).

Beta diversity analysis (Figure 4) showed that samples were well clustered by their frequency of dilution, and somewhat well clustered by their mean dilution rate, although both factors significantly influenced all four metrics of beta diversity that were calculated. This supports findings from alpha diversity analysis and suggests that disturbance frequency could be an even more influential feature of an ecosystem on its diversity than the disturbance intensity.

Taxonomic composition of samples (Figure 5) further demonstrated loss of diversity in samples that were diluted at high frequency. This analysis showed that the loss of diversity could be attributed to loss of bacteroidetes and increased dominance of proteobacteria. This may have implications for design of microbiomes, as proteobacteria have been found to be abundant in dysbiotic gut microbiomes and are associated with inflammation [20].

Samples were successfully classified using a machine learning model (random forest), although the accuracy was greater for prediction of dilution frequency (84.6 % accuracy) than mean dilution rate (53.8 % accuracy), again supporting that the frequency of an environmental disturbance may be a more salient feature than the disturbance intensity.

ANCOM and LEfSe were used to identify differentially abundant taxa in the samples, with the finding that bacteroidetes were lost during experiments with frequent dilution, in agreement with the taxonomic compositional analysis. Springobacteria were found to exhibit similar abundance patterns.

The mathematical model implemented in this work was capable of predicting the effect observed during experiments, despite relatively poor parameterization to the system that was studied. This analysis demonstrated that a linear consumer resource model can successfully predict diversity even in a complex microbial ecosystem subjected to disturbances. This holds promise for forward engineering of microbiomes, for example, by predicting the relative abundance of different species of microorganisms based on their preference for different substrates in the system, an idea that has been proposed to apply to the gut microbiome [21].

## 6 Limitations and future work

Only one, soil-derived microbial community was studied, pointing to the need for future research to determine whether the effects observed in this study are common to other microbial or larger scale ecosystems.

The model implemented in this study was set up to consider the presence of different growth rates and Monod constants for the microorganisms that were simulated, but did not simulate the growth of any particular microorganism. This points to the need for future research to determine whether the specific microorganisms predicted to thrive by the model are actually the ones that thrive in the community that is simulated. For example, if microorganisms that have high growth rate on a substrate that is abundant in the environment are predicted to dominate, experimental results should test the same hypothesis and be compared to model predictions.

## 7 Conclusions

In this study, an unexpected U-shaped relationship between diversity of a soil-derived microbial community and mean dilution rate was observed, but the U-shape was removed when the dilution was applied infrequently. Results highlighted the influence of the frequency of an environmental disturbance (not only its intensity) on the diversity of a system which is subjected to it. This relationship was demonstrated by a Monod consumer resource model, suggesting that canonical bacterial growth kinetics are sufficient to predict such an effect. Furthermore, the relative agreement between the mathematical model and experimental results suggests that consumer resource models could be used to make testable predictions in other contexts, potentially allowing future researchers to predict the effect of disturbances or variable resource concentrations on the relative abundance of species in different environments such as the gut. Results may also have implications for the loss of biodiversity due to global climate change, which could affect ecosystems by modulating the intensity and frequency of disturbances such as storms or droughts.

## 8 Acknowledgments

Dr. Christopher Mancuso is acknowledged for providing data and code used in the original study [1]. This re-analysis was carried out for a class project in ANSC 595 at Purdue, taught by Dr. Tim Johnson.

## 9 Funding

Clay Swackhamer is funded by a USDA-NIFA postdoctoral fellowship (Award number: 2022-67012-37189)

## 10 Data and code availability

Sequencing data from the original study [1] are available at the Sequence Read Archive (SRA) with accession code PRJNA719465. Codes used to carry out the re-analysis and prepare figures are available at: <https://github.com/clayswack/Mancuso-re-analysis>. The pairwise Adonis code used to carry out PERMANOVA is available at <https://github.com/pmartinezarbizu/pairwiseAdonis/blob/master/pairwiseAdonis/R/pairwise.adonis.R/> Python code used to implement the mathematical model was based on the original version available at <https://github.com/khalillab/DDR-eVOLVER-model>. Qiime2 and R codes were modified from original versions developed by Dr. Tim Johnson at: <https://github.com/john2929/ANSC516/wiki>.

## References

- [1] Christopher P Mancuso, Hyunseok Lee, Clare I Abreu, Jeff Gore, Ahmad S Khalil. *Environmental fluctuations reshape an unexpected diversity-disturbance relationship in a microbial community*. eLife, 10:e67175. 2021. <https://doi.org/10.7554/eLife.67175>
- [2] David Tilman, Forest Isbell, Jane M. Cowles. *Biodiversity and Ecosystem Functioning*. Annual Review of Ecology, Evolution, and Systematics, 45:471-493. <https://doi.org/10.1146/annurev-ecolsys-120213-091917>

- [3] Joseph H. Connell. *Diversity in Tropical Rain Forests and Coral Reefs: High diversity of trees and corals is maintained only in a nonequilibrium state.* Science, 199:1302-1310. 1978. <https://doi.org/10.1126/science.199.4335.1302>
- [4] Robin L. Mackey, David J. Currie. *The diversity-disturbance relationship: Is it generally strong and peaked?* Ecology, 82:3479-3492. 2001. [https://doi.org/10.1890/0012-9658\(2001\)082\[3479:TDDRII\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2001)082[3479:TDDRII]2.0.CO;2)
- [5] A. Randall Hughes, Jarrett E. Byrnes, David L. Kimbro, John J. Stachowicz. *Reciprocal relationships and potential feedbacks between biodiversity and disturbance* Ecology Letters, 10:849-864. <https://doi.org/10.1111/j.1461-0248.2007.01075.x>
- [6] Ricahrd Levins. *Coexistence in a variable environment.* The American Naturalist, 114:765-783. <https://doi.org/10.1086/283527>
- [7] Peter Chesson. *Multispecies Competition in Variable Environments.* Theoretical Population Biology, 45:227-276. <https://doi.org/10.1006/tpbi.1994.1013>
- [8] James Grover. *Dynamics of Competition in a Variable Environment: Experiments with Two Diatom Species* Ecology, 69:408-417. <https://doi.org/10.2307/1940439>
- [9] Cyrille Violle, Zhichao Pu, Lin Jiang. *Experimental demonstration of the importance of competition under disturbance.* Proceedings of the National Academy of Science, 107:12925-12929. <https://doi.org/10.1073/pnas.1000699107>
- [10] Sean M. Gibbons, Monika Scholz, Alan L. Hutchison, Aaron R. Dinner, Jack A. Gilbert, Maureen L. Coleman. *Disturbance Regimes Predictably Alter Diversity in an Ecologically Complex Bacterial System.* mBio, 7:e01372-16. <https://doi.org/10.1128/mBio.01372-16>
- [11] Andrew D. Letten, Manpreet K. Dhami, Po-Ju Ke, Tadashi Fukami. *Species coexistence through simultaneous fluctuation-dependent mechanisms.* Proceedings of the National Academy of Science, 115:6745-6750. <https://doi.org/10.1073/pnas.1801846115>
- [12] Ezequiel Santillan, Hari Seshan, Florentin Constancias, Daniela I. Drautz-Moses, Stefan Wuertz. *Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities.* npj Biofilms and Microbiomes, 5:8. 2019. <https://doi.org/10.1038/s41522-019-0079-4>
- [13] Brandon G Wong, Christopher P Mancuso, Szilvia Kiriakov, Caleb J Bashor, Ahmad S Khalil. *Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER.* Nature Biotechnology, 36:614-623. 2018. <https://doi.org/10.1038/nbt.4151>
- [14] Jacques Monod. *The growth of bacterial cultures.* Annual Review of Microbiology, 3:1, 371-394. 1949. <https://doi.org/10.1146/annurev.mi.03.100149.002103>
- [15] Daryl M Gohl, Pajau Vangay, John Garbe, Allison MacLean, Adam Hauge, Aaron Becker, Trevor J Gould, Jonathan B Clayton, Timothy J Johnson, Ryan Hunter, Dan Knights, Kenneth B Beckman. *Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies.* Nature Biotechnology, 34:942-949. 2016. <https://doi.org/10.1038/nbt.3601>

- [16] Benjamin J Callahan, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, Susan P Holmes. *DADA2: High-resolution sample inference from Illumina amplicon data*. Nature methods, 13:581:583. 2016. <https://doi.org/10.1038/nmeth.3869>
- [17] Manoj Kumar Sahu, Uma Balasubramaniam, Bipin C, Sarvesh Pal Singh, Sachin Talwar. *Elizabethkingia Meningoseptica: An Emerging Nosocomial Pathogen Causing Sepsis in Critically Ill Patients*. Indian Journal of Critical Care Medicine, 23:104-105. <https://www.doi.org/10.5005/jp-journals-10071-23127>
- [18] Nicola Segata, Jacques Izard, Levi Waldron, Dirk Gevers, Larisa Miropolsky, Wendy S Garrett, Curtis Huttenhower. *Metagenomic biomarker discovery and explanation*. Genome Biology, 12: R60. <https://doi.org/10.1186/gb-2011-12-6-r60>
- [19] Jonathan Friedman, Eric J. Alm. *Inferring Correlation Networks from Genomic Survey Data*. PLOS Computational Biology, 8:e1002687. <https://doi.org/10.1371/journal.pcbi.1002687>
- [20] Na-Ri Shin, Tae Woong Whon, Jin-Woo Bae. *Proteobacteria: microbial signature of dysbiosis in gut microbiota*. Trends in Biotechnology, 33:496-503. <https://doi.org/10.1016/j.tibtech.2015.06.011>
- [21] Bruce Hamaker, Yunus Tuncil. *A perspective on the complexity of dietary fiber structures and their potential effect on the gut microbiota*. Journal of Molecular Biology, 426: 3838-3850. <https://doi.org/10.1016/j.jmb.2014.07.028>