

Supplementary File 1. Additional methods and experimental details

Details of data analysis

For bacteria, our analytical dataset consisted of 295 taxa across 240 samples, corresponding to the relative abundance of OTUs aggregated to the family level. For all analyses other than the heatmap visualizations, where we employed counts of OTUs with unknown classification as a reference for computing additive log ratios, we dropped all OTUs with missing annotation at the family and every more specific taxonomic level. Similarly, for fungi, we represented taxa at the genus level, the resulting dataset including 408 taxa across 240 samples, dropping any OTU with missing annotation at the family and every more specific taxonomic level. Among the most abundant bacterial phyla were *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Chloroflexi*, *Planctomycetes*, *Firmicutes*, and *Acidobacteria*; the most abundant fungal phyla were *Ascomycota*, *Basidiomycota*, *Rozellomycota*, *Chytridiomycota*, and *Monoblepharomycota*.

Eighty of our samples were taken from soil that was not fumigated. The rest of the data was subdivided into samples that were taken from grass immediately following fumigation in the summer, consisting of 32 samples, and samples that were taken from the grass at least a month after fumigation in the fall and spring, which consisted of 128 samples. We expected the composition of microbes in the soil to be affected by fumigation, but we also expected that composition to shift over time following fumigation and the planting of new seeds. To capture this, we defined a grouping variable that we call “fumigation status” with three levels:

- Never fumigated (denoted “never” in the figures): samples that were taken from grass that was not fumigated or taken from grass before fumigation.
- Recently fumigated (denoted “recent” in the figures): samples that were taken from fumigated grass immediately following the fumigation in summer.
- Fumigated more than a month ago (denoted “past” in the figures): Samples that were taken from fumigated grass at least one month after fumigation.

In an effort to discover any trends in the composition of the microbial community, we first visualized our data with clustered heatmaps and non-metric multidimensional scaling (NMDS). Following this initial exploration, we investigated the existence of community differences across sample groups via permutational multivariate analysis of variance (PERMANOVA). We additionally evaluated differences across groups in alpha diversity with boxplots and significance testing. Finally, we explored the impact of fumigation from a network perspective, analyzing our samples to infer differential microbial networks under different fumigation statuses.

Clustered Heatmaps

We first visualized the entire count dataset as a heatmap. Our visualization software, the R package pheatmap ^[1], clusters similar samples together using a hierarchical clustering algorithm.

This helped us visualize which samples have similar microbial communities. In order to produce an informative heatmap, the data must first be transformed via log ratio normalization. We used the log ratio normalization using the count of taxa with unknown classification at the family level as the reference (similarly, for fungi, we used the count of OTUs with unknown classification at the genus level as the reference). Following standard practice to handle zeroes, we added a pseudo-count of 0.5 to all entries before taking the log ratio transformation. Formally, M represented the matrix of read counts with M_{ij} representing the abundance of taxa i at sample j . Then, we applied the following transformation to each entry:

$$\hat{M}_{ij} = M_{ij} + 0.5$$

$$\tilde{M}_{ij} = \log \left(\frac{\hat{M}_{ij}}{\hat{M}_{uj}} \right)$$

where $\hat{M}_{uj}$ represented the count of OTUs with unknown families for sample j . We chose this normalization because it creates better visualization compared to standard library size normalization, with the majority of the count being close to zero and the heatmap would have the majority being the same color. In addition to visualizing all the family counts in all the samples, we aggregated the data in two ways and visualized these reduced datasets: 1. Aggregating the taxa into phyla, reducing the number of taxa from 295 families to 43 phyla. 2. Aggregating the samples into their days and fumigation status and taking the average count of the samples. For each fumigation status, we added all the samples on the same day, and then divided by the number of samples on that day. This reduced the samples from 240 into 6 days for never, 4 days for past, and 1 day for the recent.

Non-metric Multi-Dimensional Scaling

The heatmaps allowed us to visualize the original, non-reduced dataset simultaneously. In addition, we also constructed visualizations that relied on first reducing the dimensionality of our data. Non-metric Multidimensional Scaling (NMDS) is a dimensionality reduction method popular for microbial datasets as well as other non-normal, highly skewed datasets in that it allows arbitrary monotonic transformations to be applied on the original distances before finding the mapping to lower dimensional space that best preserves transformed distances. We applied NMDS to the original untransformed matrix of counts, as the method automatically works on the ranks of the counts, thereby avoiding any bias due to inherent compositionality. We used the implementation of NMDS found in the vegan R package^[2].

PERMANOVA

PERMANOVA is a non-parametric alternative to the MANOVA test for differences in a multivariate response between groups^[1]. Because of its robustness to skew and zero-inflation, PERMANOVA is commonly used in microbiome community analyses. By viewing the vector of taxa abundances in a sample as a multivariate response, PERMANOVA tests the null hypothesis

that the centroids of the groups, which are defined by a distance matrix, are equal. It partitions the distance matrix based on the grouping variable and then assesses the significance via permutations.

Before performing PERMANOVA, several pre-processing steps are needed to obtain a vector of relative abundances for each sample. After filtering out families with missing values, we perform PERMANOVA on three filtering levels: all families, families with over one thousand total reads across samples, and families with over ten thousand total reads across samples. We performed Total Sum Scaling (TSS) normalization by dividing each OTU's count by the total count of its sample:

$$\tilde{M}_{ij} = \frac{M_{ij}}{\sum_{k=1}^n M_{ik}}$$

Where M_{ij} is the raw count of taxa j in sample i , and n is the total number of taxa. We performed PERMANOVA to test for community differences across three different grouping variables: fumigation status, soil amendment treatment, and sample collection date. For each hypothesis test, we computed the distance matrix using Bray-Curtis Dissimilarity^[3] using the R package `vegan`^[3]. When we tested for differences across soil amendment treatments, we filtered out samples with a fumigation status of “never” as they did not receive any treatment. When we tested for differences across the sample collection date, we stratified by fumigation status and performed one PERMANOVA analysis for “never” samples and one for “past” samples.

Diversity Index

To investigate the relationship between alpha diversity and several experimental conditions, we used one-way ANOVA on ranks. The experimental conditions we explore are fumigation status and soil amendment treatment. We studied two measures of alpha diversity: Shannon's Diversity index and the Inverse Simpson index. Shannon's Diversity Index is a widely used measurement of species evenness and is defined by $H' = -\sum_{i=1}^R p_i \log(p_i)$. On the other hand, inverse Simpson index provides more information on species richness and is calculated by $\frac{1}{D}$, where D is the Simpson index, $D = \sum_{i=1}^R p_i^2$. A higher value of inverse Simpson index represents higher biodiversity.

We first graphically examined the differences in diversity with box plots for each diversity index. We then applied the Kruskal-Wallis test (ANOVA on ranks) to test whether the median of the alpha-diversity index differed across fumigation statuses and across soil amendment treatment. Following the Kruskal-Wallis test, Dunn's test for multiple comparisons with a Bonferroni correction was conducted to identify pairs of fumigation status/soil amendment groups that exhibited significant differences in alpha-diversity^[4]. Dunn's test compares the ranks of the data between each pair of groups, adjusting for multiple comparisons to control the family-wise error rate. Data analysis was performed using R statistical software (`kruskal.test` function (R Core Team, 2023) and `dunn.test`^[5,6]).

Microbiome Network Inference

Microbiome networks allow us to study the interactions among microbial taxa while accounting for fumigation status or other variables. Each node in the constructed network represents one OTU and an edge between two nodes represents an association between these microbial taxa.

There are various approaches to build microbiome networks. In this paper, we mainly employed the graphical LASSO (GLASSO)^[7,8], a method designed to infer conditional dependencies among normally distributed variables by estimating a sparse inverse covariance matrix. To evaluate and analyze the effect of fumigation, we constructed three different networks using the GLASSO method for each of the three fumigation statuses.

We used the implementation of GLASSO found in the `glasso` package from R to estimate the inverse covariance matrix of the abundance data, normalized with the center log ratio (clr) transformation. Because of the relationship between partial covariances and entries of the inverse covariance matrix under an assumed normal distribution, adding edges between pairs of nodes whose corresponding entries in the estimated precision matrix are non-zero produces a network in which edges can be interpreted as conditional dependencies. A tuning parameter ρ is set to control the sparsity level. We used a threshold of 0.75 in this analysis. The partial correlation matrix P from the inverse covariance matrix C^{-1} output by the GLASSO software is calculated using the formula

$$p_{ij} = \frac{-C_{ij}^{-1}}{\sqrt{C_{ii}^{-1}C_{jj}^{-1}}}$$

Note that the partial correlation p_{ij} is 0 precisely when C_{ij}^{-1} is 0.

We can thus construct an undirected weighted network with weight in which an edge exists if $p_{ij} \neq 0$ and with edge weight $|p_{ij}|$. To reduce visual clutters, all analysis only displays strongest edges (top X% by $|p_{ij}|$); all analysis are based on the full set of nonzero $|p_{ij}|$.

To better analyze the networks, we compared the edge sets of the GLASSO networks. A Venn diagram was used to visualize the overlap and exclusiveness of the microbial associations between the never, recent, and past fumigation groups. In order to identify taxon-taxon relationships that survived fumigation, we visualized the subnetwork consisting of only the edges shared by all three networks. In order to reveal microbial interactions specifically induced or disrupted by fumigation and potentially identify the taxa driving these interactions, we visualized the subnetworks of edges unique to each fumigation condition by excluding any association found in the other two networks.

To further compare the microbiome networks, we used metrics such as the degree centrality and betweenness centrality. Degree centrality measures the number of nodes that are connected to a node in the graph. Betweenness centrality measures how frequently a node lies on the shortest path between two nodes. We used those two metrics, along with the network, to determine a subset of OTUs that differ the most among the networks for the different fumigation statuses. We denoted this subset of OTUs as the “differential OTUs”. Focusing on only these differential OTUs, we applied the CARlasso, a Bayesian chain graph model, to infer a network that includes both the interactions among OTUs and interactions between OTUs and variables such as fumigation status or treatment. We chose the 20 and 10 OTUs that differed the most among the networks and constructed two CARlasso graph models for each of them, one with both fumigation status and treatments as additional nodes in the network and one only with fumigation status. We used the R package CARlasso for our implementation^[9].

References

- [1].Kolde R. 2019. *pheatmap: Pretty heatmaps* (R package version 1.0.12). <https://CRAN.R-project.org/package=pheatmap>
- [2].Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, et al. 2022. *Vegan: Community ecology package* (Version 2.6-4). <https://CRAN.R-project.org/package=vegan>
- [3] Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* 27(4):325–349 <https://doi.org/10.2307/1942268>
- [4] Kruskal WH, Wallis WA. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47(260):583–621 <https://doi.org/10.1080/01621459.1952.10483441>
- [5].Dinno A. 2024. *dunn.test: Dunn’s test of multiple comparisons using rank sums* (R package version 1.3.6). <https://CRAN.R-project.org/package=dunn.test>
- [6] Dunn OJ. 1964. Multiple comparisons using rank sums. *Technometrics* 6(3):241–252 <https://doi.org/10.1080/00401706.1964.10490181>
- [7] Friedman J, Hastie T, Tibshirani R. 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9(3):432–441 <https://doi.org/10.1093/biostatistics/kxm045>
- [8].Friedman J, Hastie T, Tibshirani R. 2019. *glasso: Graphical lasso-estimation of Gaussian graphical models* (R package version 1.11). <https://CRAN.R-project.org/package=glasso>
- [9].Shen Y, Solis-Lemus C. 2022. Bayesian chain graph lasso models to learn sparse microbial networks with predictors. *arXiv Preprint: 2012.08397* <https://doi.org/10.48550/arXiv.2012.08397>