

Factors influencing the fungal and bacterial diversity of soil micro-life.

Microbial biodiversity can be calculated in several different ways. Each approach captures slightly different aspects of the diversity. Key differences relate to the extent that richness (number different organisms) and evenness (even abundance of detected organisms) is considered and how important the measure considers the more abundant versus the low abundant organisms. A subgroup of diversity measures considers phylogenetic diversity (how related are the detected organisms).

Most of these measures will correlate and thus analysing all will be redundant and thus uninformative.

We calculated a number of different such diversity measures (in the field of sequencing-based microbiome profiling called alpha diversity measures), and evaluated how much they correlated. From these we selected two measures: Faith's phylogenetic diversity that consider the phylogenetic relationship between organisms and finds a sample with organisms from different branches of the phylogenetic tree more diverse than samples with organisms from the same branches. And Shannon diversity that calculate diversity considering both evenness and richness.

For Shannon and Faith's D (FD) we did statistical analyses of their association with all measure's variables (leaving those out with too little data for such linear statistical modeling). We did this for both fungi and bacteria communities and reviewed if they detected different associations. We found that they largely agreed and thus we selected to proceed with Shannon diversity as the measure for biodiversity in this study.

For each dataset (fungi and bacteria) we have systemized a step-by-step process to evaluate how Shannon diversity associates with soil properties/farming practises and which of the different recorded variables are most important for dictating the level of diversity. (We use the same process as used for the functional GO processes.)

First, we look if key variables of interest that group samples into max 2 groups (JB groups, organic farming and livestock manure) have significant and consistent different levels

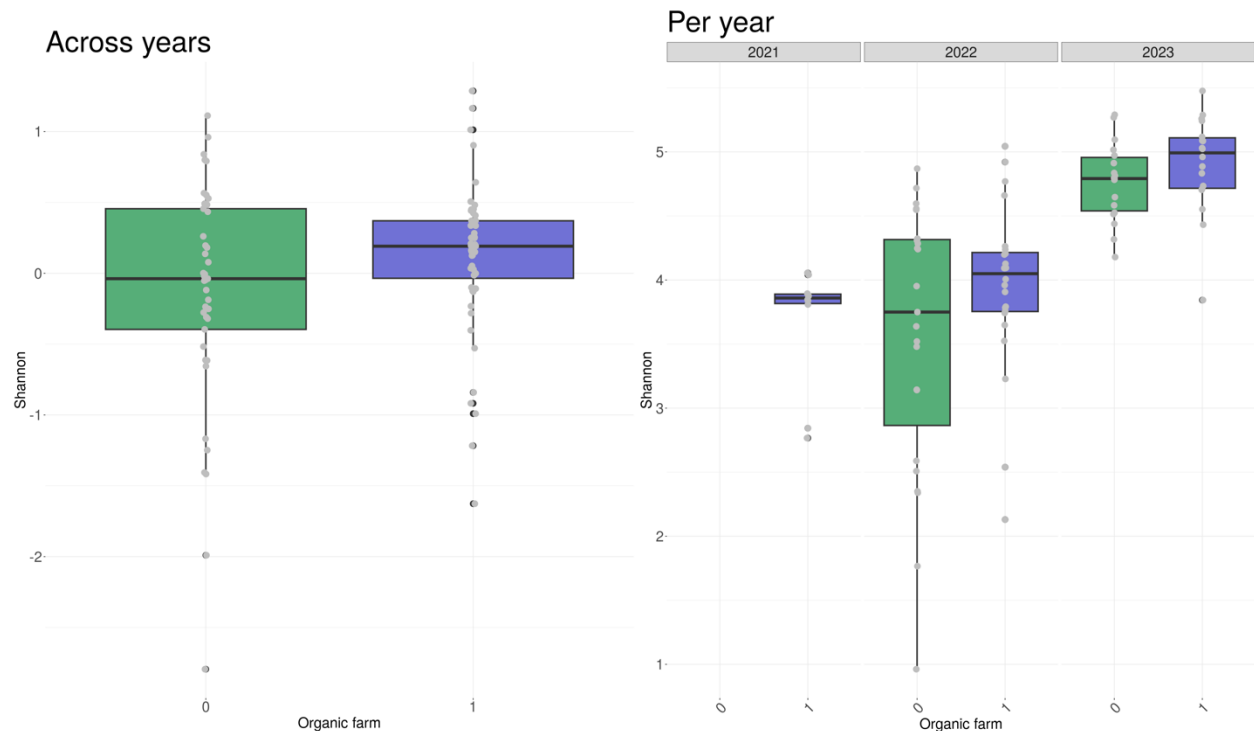
across the years of sampling. Then, we use a machine learning approach (random forest) to identify which variables are most important to predict diversity, first across all samples, and then within the key groups if any of these were significant.

Fungi diversity

Significant differences between key group variables?

Var	Est	p
JB_groups	0.0633269	0.6548849
Organic_farm	0.2536310	0.0726330
Livestock_manure	-0.2924232	0.0720786

We see that fungi diversity does not differ between JB group (JB 5-6-7, versus JB1-2). There is a trend for higher diversity in organic farming and lower when using livestock manure. But none are significant ($p < 0.05$).



Which variables predict the GO process?

First we include all samples and most recorded variables.

We use %IncMSE to identify the most important variables. %IncMSE stands for the percentage increase in Mean Squared Error. It is a measure of how much the prediction error increases when a particular variable is permuted (i.e., its values are randomly shuffled) while all others are left unchanged.

Top predictors based on %IncMSE (most important in bottom):

	%IncMSE	IncNodePurity
Mulching_of_straw	1.167442	0.20629623
Nitrogen_perc	1.215669	2.43774661
No_plough	1.404138	0.58540658
Vinasse	1.852638	0.04158029
Commercial.fertilizer	1.918423	0.34259058
Organic_farm	1.919318	0.39214527
ConservationAgriculture	2.313712	0.90605501
Clay_perc	2.591975	2.43837783
Years_since_plowing	3.846720	2.49664918
Organic_material_perc	4.893072	3.16255542
Potassium	6.615807	7.62507339

We see that potassium is the most important predictor, followed by organic material and years since plowing. Organic farming is also important here.

Now we select the most important variables (%IncMSE>1) and look how much of the variation in the diversity these variables explain:

% Variance explained: 2%

Are the important variables different between the organic and conventional fields?

Below is a table of the variables selected as important in organic and conventional fields, with overlapping variables in bold text. And below the table is the % variance explained in each group.

We see that two variables are the same, but that more of the diversity can be explained by evaluated variables in conventional farming than in organic. In conventional fields, potassium is very important.

Possible reasoning: In ecosystems where potassium is a limiting factor, the availability of this nutrient can directly influence microbial communities, including fungi. Fungi that are better adapted to low-potassium environments may dominate when potassium is scarce, while higher potassium levels can support a broader range of fungal species. Thus, potassium levels in the soil can shape the competitive dynamics of fungal communities, influencing diversity.

Nitrogen is also important in both, but perhaps even more in organic farming. Low nitrogen environments tend to support more diverse and specialized fungal communities, while high nitrogen levels, especially from fertilization or deposition, often reduce fungal diversity by favoring fast-growing, opportunistic species.

Interestingly, we see no correlation across all samples between nitrogen percentage and fungi Shannon diversity. This again highlights that the one-to-one analyses likely oversimplify the situation, and that nitrogen is important in context of other variables.

Conv_predictors	Organic_predictors
NA	Haveparkaffald
OrganicMaterial_category	Cast
ConservationAgriculture	Phosphorus
Nitrogen_perc	Potassium
Cold_soil	Organic_material_perc
Magnesium	Clay_perc
Potassium	Nitrogen_perc

% Variance explained for organic farming: 0.3%

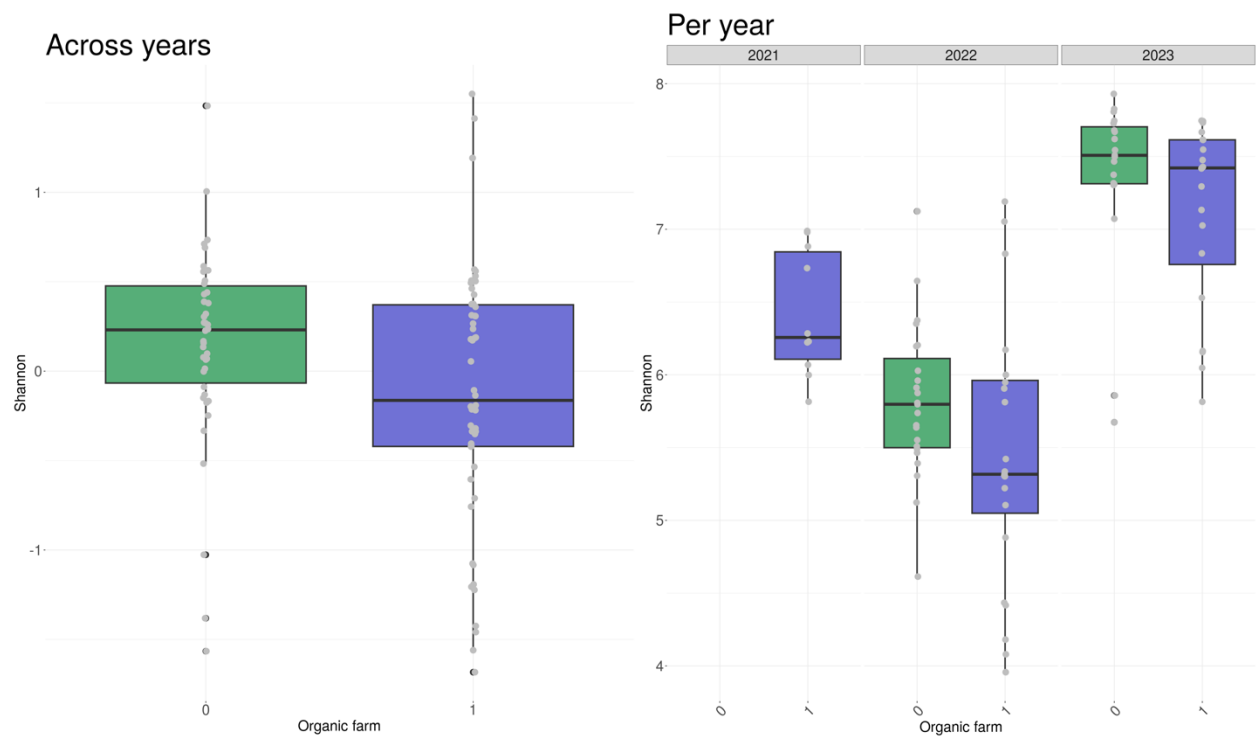
% Variance explained for conventional farming: 11.7%

Bacterial diversity

Significant differences between key group variables?

Var	Est	p
JB_groups	0.2410680	0.0815118
Organic_farm	-0.2715244	0.0502751
Livestock_manure	-0.3367074	0.0364643

We see that bacterial diversity does not differ significantly between the groups, but there are trends for all. Between JB group diversity is higher in JB 5-6-7, versus JB1-2. Surprisingly, there is a trend for lower diversity in organic farming and when using livestock manure.



Which variables predict the GO process?

First we include all samples and most recorded variables.

We use %IncMSE to identify the most important variables. %IncMSE stands for the percentage increase in Mean Squared Error. It is a measure of how much the prediction error increases when a particular variable is permuted (i.e., its values are randomly shuffled) while all others are left unchanged.

Top predictors based on %IncMSE (most important in bottom):

	%IncMSE	IncNodePurity
Earthworm_status	1.109227	0.3644813
Phosphorus	1.319326	3.1096453
Haveparkaffald	1.369234	0.1476811
Rt	1.607571	2.1956041
Nitrogen_perc	1.871677	2.3630166
Livestock	2.682811	0.4645716
Clovergrass_within_3_years	2.718631	0.4521693
Livestock_manure	2.770235	0.5622418
Organic_material_perc	3.025431	3.1206521
Cast	3.051868	0.2577092
JB_groups	3.280597	0.4162317
Clay_perc	3.370236	4.8325140
OrganicMaterial_category	3.639512	0.6422284
No_plough	3.685375	0.6085196
Potassium	3.803848	4.5906414
Years_since_plowing	3.804545	1.7752556
Organic_farm	3.969777	0.5786121
Commercial.fertilizer	4.402237	0.9126380
Cobber	4.608313	4.1491559
Magnesium	6.377552	3.9306634

We see that magnesium is the most important predictor. Organic farming is also important here.

Now we select the most important variables (%IncMSE>1) and look how much of the variation in the diversity these variables explain:

% Variance explained: 2.3%

Are the important variables different between the organic and conventional fields?

Below is a table of the variables selected as important in organic and conventional fields, with overlapping variables in bold text. And below the table is the % variance explained in each group.

We see that Clay% is important in both, and again can more of the diversity be explained by evaluated variables in conventional farming than in organic but now 7% can be explained in organic. In conventional fields, potassium is again very important as it was for fungi. And also for bacterial diversity is nitrogen important.

Conv_predictors	Organic_predictors
NA	Degassed.fertilizer
OrganicMaterial_category	Clovergrass_within_3_years
ConservationAgriculture	Mulching_of_straw
Degassed.fertilizer	Years_since_plowing
Magnesium	Haveparkaffald
Organic_material_perc	Cobber
JB_groups	Cast
Cobber	Organic_material_perc
Rt	OrganicMaterial_category
Nitrogen_perc	Nitrogen_perc
Potassium	Clay_perc
Clay_perc	Magnesium

% Variance explained for organic farming: 6.9%

% Variance explained for conventional farming: 11.2%