Plant Pathogens Consistently Reduce the Relative Abundance of Bacterial Order Sphingomonadales upon Infection

Fleur S.A. Wallis[1], **Nick C. Snelders**[1,2] **and Michael F. Seidl**[1]

[1] *Department of Fungal Evolutionary Genomics, Utrecht University, Utrecht, The Netherlands;*
[2] *Laboratory of Evolutionary Microbiology, University of Cologne, Cologne, Germany*

Key Words: Metagenomics - 16S data analysis - Plants - Pathogens - Abiotic stress - Immune mutants - Sphingomonadales

Plants are inhabited by a collection of microorganisms, together called the microbiome. Some of these microorganisms can positively influence plant functioning, while others negatively impact plant health. Therefore, the overall composition of the microbiome is important for the plants' well-being. Pathogens have been shown to alter the microbiome composition through microbe-pathogen interactions. Furthermore, plants themselves alter the microbiome upon pathogen detection. There is no knowledge on the generality of the microbial change; infections with different pathogen sources on different plants might affect the microbial community in similar ways. In this study, it was shown that there is a consistent decrease in the relative abundance of bacterial order Sphingomonadales upon pathogen infection. This shift happens irrespective of the present pathogen type and the infected host. Furthermore, the specific shift was localized to the pathogen residence, showing that systemic signalling is not responsible for the Sphingomonadales decrease. It appears that general plant stress invoked by successful pathogen invasion does not necessarily contribute towards the Sphingomonadales reduction. Immune system suppression on the other hand might partially explain the pathogenic effects on the Sphingomonadales, as *A. thaliana* mutants with altered immune system functioning show a similar shift in the relative abundances of the bacterial order. It is probable that direct microbe-pathogen interactions also play a role in the Sphingomonadales decrease, which should be confirmed in further research.

Introduction

All plant species live in symbiosis with a unique and diverse set of microorganisms, collectively referred to as the microbiome. Most abundant in this microbiome are the bacteria and fungi. The plant host with their microbes is together called the holobiont.[1] The composition of bacterial and fungal species within holobionts depends on many different factors: the plant species, the plant growth stage, the plant tissue type, the surrounding microbiota in the environment and the abiotic circumstances.[2]

The different microorganisms in a holobiont can have a large influence on the plant health and functioning.[1] Certain microbes can stimulate growth by producing phytohormones, or by facilitating nutrient uptake for production of new plant material.[2] Others can provide protection from harmful biotic influences, for example through production of antibiotics or through induced systemic resistance, priming the plant for invading pathogens.[3] Furthermore, microbes around and within plants can protect them from abiotic influences such as drought and salt stress.[4,5] Besides the positive influences, certain microorganisms can negatively impact a plant's health. Some microbes produce toxins, while others cause downright disease in plant tissue.[6]

Plants actively shape the microbial community to maximize their fitness.[7] This shaping is most visible around the roots, where the so-called rhizosphere effect takes place; root exudates that include organic acids, amino acids, fatty acids, nucleotides, sugars, and vitamins attract certain beneficial microbes.[3] Besides nutrient excretion, plants can secrete defensive secondary metabolites that inhibit the growth of certain detrimental microorganisms.[8]

Pathogens are microorganisms that cause disease, which need to evade the defense mechanisms of the holobiont for successful plant invasion.[8] Usually, pathogens express effector proteins to aid in their colonization of a specific plant species. These proteins can alter plant processes, usually related to plant immunity.[9] During successful invasion, the residing plant microbiome is concomitantly altered. A lot is unknown about this microbiome shift. For example, it has not been investigated whether different pathogens cause similar shifts, and how shifts are brought about. Potentially, pathogens need to remove certain general protective microbial species before being able to infect plants, due to the secretion of inhibitory antimicrobial compounds by the microbes. Certain other microbes might be stimulated by pathogen-presence if they produce nutrients and aid in hijacking the plant immune system. On the other hand, it could be that indirect effects

Experiment Name	Host plant	Microbiome Location	Pathogen	Infection type	Environment	Bacterial community
tomato root verticillium	*Solanum lycopersicum*	Root	*Verticillium dahliae*	Manual	Greenhouse	Synthetic
chili xylem fusarium	*Capsicum annuum L.*	Xylem	*Fusarium oxysporum*	Natural	Field	Natural
wheat leaf zymoseptoria	*Triticum aestivum*	Leaf	*Zymoseptoria tritici*	Manual	Phytochamber	Synthetic
wheat adjacentleaf zymoseptoria	*Triticum aestivum*	Adjacent leaf (control)	*Zymoseptoria tritici*	Manual	Phytochamber	Synthetic
arabidopsis leaf mildew	*Arabidopsis thaliana*	Leaf	*Golovinomyces orontii*	Manual	Greenhouse	Natural
arabidopsis root mildew	*Arabidopsis thaliana*	Root (control)	*Golovinomyces orontii*	Manual	Greenhouse	Natural
tomato rhizosphere phytophthora	*Solanum lycopersicum*	*Phytophthora parasitica*	Rhizosphere	Manual	Greenhouse	Natural
banana pseudostem ralstonia	*Musa paradisiaca*	Pseudostem	*Ralstonia*	Natural	Field	Natural
rice stem dickeya	*Oryza sativa*	Stem	*Dickeya zeae*	Natural	Field	Natural

Table 1: Characteristics of selected pathogen datasets. Displayed for each experiment are the infected host plant, the microbiome sampling location, the infecting pathogen, how the infection came about, in what environment the study was conducted and which bacterial community type was present on the plants.

like plant stress and immune system suppression are responsible for microbial alterations, as shifts in the microbiome have previously been reported for both these situations.[10,11]

In this research project we investigate the commonalities in microbiome alterations upon pathogen infection. We focus our analysis on the alterations in the relative abundances of bacterial orders. To research whether observed pathogen-induced shifts could be explained by indirect pathogen effects, we subsequently analyze the influence of abiotic stresses and immune system manipulation on the microbial community.

Materials and Methods

Pathogen invasion datasets

A literature search was performed to obtain publicly available datasets with 16S ribosomal DNA sequencing data of both pathogen-affected and healthy plant tissue. The influence of confounding factors on the microbiome composition was considered by including datasets with varying pathogen types, plant hosts and abiotic experimental circumstances. Furthermore, the plant tissue used for microbiome analysis was varied, thereby further varying the input microbiome. In the end, seven different datasets were included to visualize a potential general trend in microbiome alteration.[12–18] The BioProject accession numbers associated to these datasets are PRJEB34281[12], PRJNA667302[13], PRJNA549447.[14], PRJEB43139[15], PRJNA354847[16], PRJNA277904[17] and PRJNA602829.[18] For the study by Seybold et al., the data was unfortunately not demultiplexed, so that the reads could not be used to generate an OTU-table, a taxonomy table and a phylogenetic tree. Instead, these data structures were directly downloaded from https://github.com/hmamine/ZIHJE/ located under the directory community_analysis.

The most important characteristics of these datasets are displayed in Table 1. In total, the seven datasets include seven different pathogen types, infecting six different plant hosts. The microbiome samples were taken from pathogen-infected plant tissue, which was a different tissue for the different pathogen types. Two datasets additionally included microbiome samples taken from distal plant tissue, away from the pathogen-infection site. These samples were analyzed to look into the difference between direct and systemic pathogen effects.

The infection type varied in the datasets; In some experiments, pathogen infections were naturally occurring in the sampled plants, while other plants were manually infected with a pathogen. Another important discriminant for the datasets was the sampling location; certain experiments were performed in greenhouses, while others were sampled in the field. Lastly, the bacterial community type varied between the different experimental setups; certain datasets used a synthetic community of bacteria, while others contained natural bacterial communities.

For most datasets, subsets were created to limit the variables to only the pathogen presence. The used samples per BioProject are displayed in Table S1.

Abiotic stress datasets

A literature search was performed to obtain publicly available datasets with 16S ribosomal DNA sequencing data of plant tissue in the presence or absence of different abiotic stresses. In the end, four different datasets were included to visualize a potential general trend in microbiome alteration.[19–22] The BioProject accession numbers associated to these datasets are PRJNA741547,[19] PRJNA551661,[20] PRJNA690819[21] and PRJEB47399.[22]

The most important characteristics of these datasets are displayed in Table 2. The datasets included plant hosts that were present in the pathogen invasion datasets, to be able to extrapolate the findings to plant stress upon pathogen infection. For all included datasets, plant growth hindering by the experimental circumstances was used as an indication of plant stress presence. In total, three different abiotic stresses were included as experimental conditions. The microbiome samples were again taken from different plant tissues affected by the abiotic stresses.

Subsets were again created for certain datasets to reduce

the number of variables. The used samples per BioProject are displayed in Table S2.

Experiment Name	Host plant	Microbiome Location	Abiotic Stress
tomato leaf drought	*Solanum lycopersicum*	Leaf	Drought
rice endosphere drought	*Oryza sativa*	Endosphere	Drought
rice endosphere salt	*Oryza sativa*	Endosphere	Salt
wheat root flood	*Triticum aestivum*	Root	Flooding

Table 2: Characteristics of selected abiotic stress datasets. Displayed for each experiment are the infected host plant, the microbiome sampling location and the imposed abiotic stress.

Immune mutants dataset

A dataset with 16S ribosomal DNA sequencing data of the endosphere of a collection of A. thaliana immune system mutants and a wildtype (WT) control was obtained. The immune system mutants were characterized by one or more gene knockouts of proteins present in important immune system signalling pathways.[11] In the end, 17 different immune mutants were included to visualize a potential general trend in microbiome alteration. The different mutation-combinations interfered with different parts of the immune system. A full description of the different mutant types can be read in the article by Pfeilmeier et al.[11]

Metagenomics analysis

16S ribosomal DNA sequencing data was downloaded from the NCBI database and subsequently converted to fastq-data using the SRA-toolkit (version 2.8.0). Most datasets contained Miseq data, while two datasets contained Hiseq data. For the different datasets, different parts of the 16S gene were analyzed. Also, different read lengths were present in the different datasets due to different sequencing setups. Each dataset was analyzed using the workflow from 'Workflow for Microbiome Data Analysis: from raw reads to community analyses'.[23]

Reads were first trimmed up to a suitable length with a quality score of ~ 30, and subsequently filtered based on the number of errors present in the sequence (maximum of two errors). The exact settings used for trimming the reads are displayed in Table S3 for each 'pathogen invasion' dataset, and in Table S4 for each 'abiotic stress' dataset. Chimeric sequences were removed, and for each unique read it was checked whether their uniqueness could be attributed to sequencing errors. If so, the reads were removed. Subsequently, the forward and reverse sequencing reads were joined for the paired-end sequencing data, and all total reads were demultiplexed. With the sequences, an OTU-table, taxonomy table and phylogenetic tree were created and combined in a phyloseq-object. For certain datasets, the phyloseq-objects were obtained from the respective study's authors. These datasets were 'tomato root verticillium', 'tomato leaf drought' and 'rice endosphere drought'. The sample counts were transformed into relative abundances, and subsequently agglomerated on the bacterial order level. Relative abundance plots and Principal Coordinate Analysis (PCoA) plots with Bray-Curtis dissimilarities (β-diversity) were generated for quality control before initiating the comparisons.

Software

The whole microbiome data analysis workflow was performed with R software[24] (version 4.1.2) using packages dada2[25](version 1.22.0), phyloseq[26] (version 1.38.0), DECIPHER[27] (version 2.22.0) and phangorn[28,29] (version 2.8.1). All packages were downloaded from Bioconductor using the BiocManager package[30] (version 3.14). Further packages used for data manipulation and visualization were ggplot2[31] (version 3.3.5), magrittr[32] (version 2.0.3), dplyr[33](version 1.0.8), tidyverse[34] (version 1.3.1), cowplot[35] (version 1.1.1), ggpubr[36] (version 0.4.0), gridExtra[37] (version 2.3), and wesanderson[38] (version 0.3.6).

Experimental Comparison and Statistical Analysis

The relative abundances for each bacterial order were compared between samples affected by a specific experimental condition and healthy controls. For the pathogen invasion and distal tissues experiments, the experimental condition was a certain pathogen infection. For the abiotic stress experiments, the experimental condition was the presence of a certain abiotic stress. For the immune mutants experiments, the experimental condition was the knockout of certain immune system proteins.

A Welch's t-test[39] was performed to compare the relative abundances for each bacterial order. No multiple testing correction was included, as the goal of the research was to visualize a general trend across conditions. $P < 0.05$ was considered statistically significant, while $P < 0.1$ was considered near-significant.

Results

Pathogens consistently decrease the local relative abundance of Sphingomonadales

Metagenomic datasets that included microbiomes from pathogen-affected plants and healthy plants were analyzed for their bacterial composition. After determining the relative abundances of bacterial orders, the results were compared between the pathogen-affected and healthy plants. For all these comparisons, plant tissue from which the microbiome was isolated corresponded to the infection-site in the pathogen-affected plants, showing the local impact of pathogens on the microbiome. To determine whether systemic pathogen effects were present as well, plant tissues' microbiomes from distal plant tissues without direct pathogen presence were analyzed. Again, the bacterial composition in tissues from pathogen-

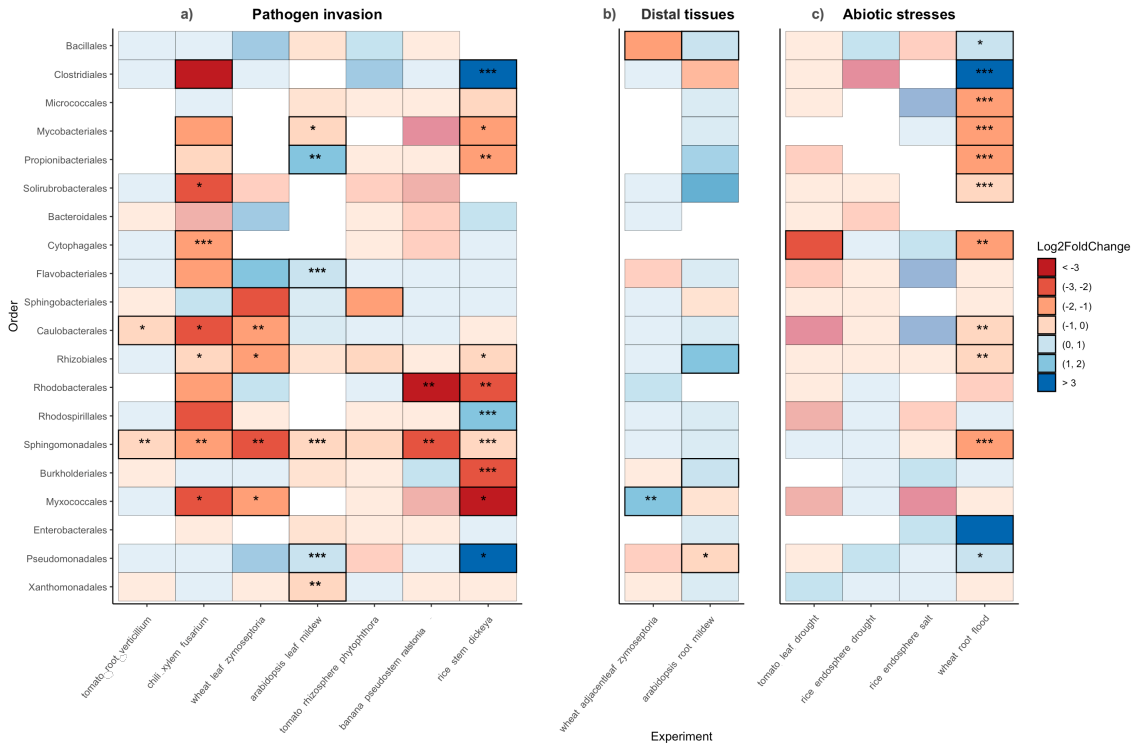

Figure 1: Bacterial composition change upon imposition of different experimental conditions. a-c) The relative abundances of different bacterial orders according to 16S ribosomal DNA profiling were compared in situations with and without an imposed experimental condition. The results are displayed in heatmaps showing the log2-fold changes of different bacterial orders and the significance of the change, as determined by a Welch's t-test. Significant differences are shown by * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$). Differences that are near-significant ($p < 0.1$) are displayed by a thickened outline of the associated heatmap box. Only the bacterial orders that were present in at least five out of seven pathogen-invasion studies were included in the plots. The experiment name specifies which plant host was investigated, which plant tissue the microbiome was isolated from, and which experimental condition the plant was subjected to. **a)** Local bacterial composition change upon pathogen infection. **b)** Distal bacterial composition change upon pathogen infection. The samples were obtained from two experiments of the pathogen invasion conditions from **a)**, but looked at tissues where pathogens were not localized. **c)** Bacterial composition change upon imposition of abiotic stress conditions.

affected and healthy plants were compared.

The quality of the final included metagenomic samples was checked by generating relative abundance plots, portraying the respective distributions of the bacterial orders (Figure S1). Furthermore, the influence of the pathogen-presence on the microbiome composition was characterized in PCoA based on β-diversity scores (Figure S2).

The relative abundances for all bacterial orders were compared between the pathogen-affected and healthy plants by both looking at the size of the log2-fold change, and the significance of the alteration. The results are displayed in Figure 1a. As the included plant hosts and plant tissues had different input microbiomes, the bacterial orders that were present in each experiment varied. To obtain a clear overview, only bacterial orders that were present in at least five out of the seven pathogen invasion datasets were included.

For most bacterial orders, the resulting heatmap displays a variable pattern over the different experiments. The only order that shows a consistent pattern are the Sphingomonadales; in all experiments, there was a significant or near-significant decrease in the relative abundance. This decrease varied from relatively small (< 2 fold) to large (4-8 fold). The relative abundance of the Sphingomonadales was more than 1% for all experiments but 'wheat leaf zymoseptoria' (Figure S3), showing the bacterial order encompasses a non-negligible portion of the bacteria in the analyzed microbiomes. Furthermore, the general abundance shows that the height of the fold-changes of the relative abundances of the Sphingomonadales in the experimental conditions was not caused by a minimal bacterial presence; the lower the relative abundance, the easier it is to obtain a large fold-change in either direction.

In Figure 2, the relative abundances of the Sphingomonadales per experiment, split up into pathogen-affected and healthy samples, are zoomed in on. In most experiments,

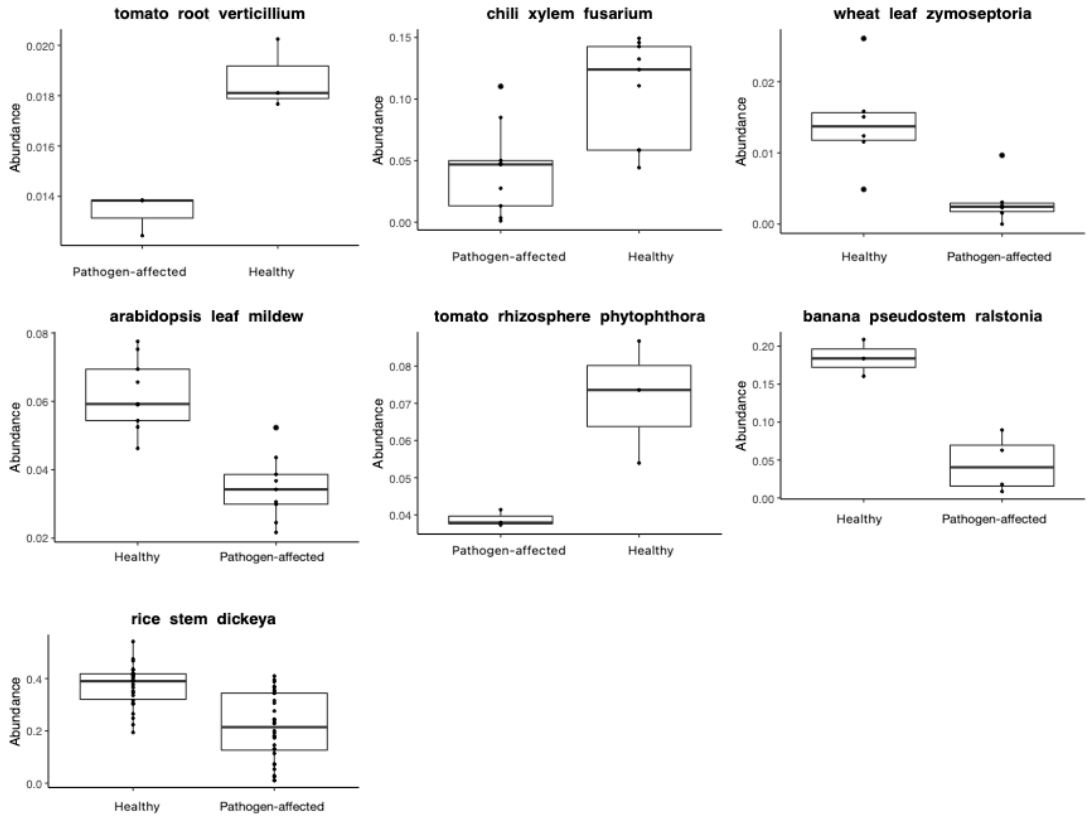

Figure 2: Boxplot with relative abundances Sphingomonadales per experimental condition in pathogen invasion experiments. The dots show the individual measurements of the relative abundances in each experimental condition. The separate boxplots per experiment show the relative abundances in pathogen-affected and healthy samples. Unaffected samples do not necessarily contain no pathogen, but are effective at suppressing pathogen invasion.

the Sphingomonadales relative abundance portrays a normal distribution, meeting the assumptions of the chosen statistical test. Furthermore, the boxplots confirm that the average value of the relative abundance of the Sphingomonadales is higher for the healthy samples than for the pathogen-affected samples, again suggesting the existence of a general trend of Sphingomonadales decrease upon pathogen-infection. Besides a higher average value, a clear separation of the abundances in the different experimental conditions could be visualized, increasing the confidence in the significance of the observed downshift.

In contrast to the local influence of pathogen invasion on the microbiome, there were only few significant changes visible for distal tissues without direct pathogen contact. As the samples were taken from the same plants as in their respective pathogen invasion experiments, a direct comparison between the two results is possible, suggesting that local pathogen effects on the microbiome are much larger than systemic effects. Furthermore, no significant shift in the Sphingomonadales was observed, making it likely that local pathogen effects are responsible for the consistent decrease of this specific bacterial order.

Abiotic stresses do not induce changes in the relative abundance of Sphingomonadales

To examine whether plant stress in general, caused by local pathogen presence, could partially be responsible for a decrease in the relative abundances of the Sphingomonadales, datasets were selected in which plants were subjected to different abiotic stresses. In all datasets, reduced plant growth was present,[19–22] confirming the presence of plant stress in general. Unfortunately, most abiotic stresses do not only generate plant stress, but also directly affect microbial composition by stressing the present microbes.[40] Therefore, direct abi-

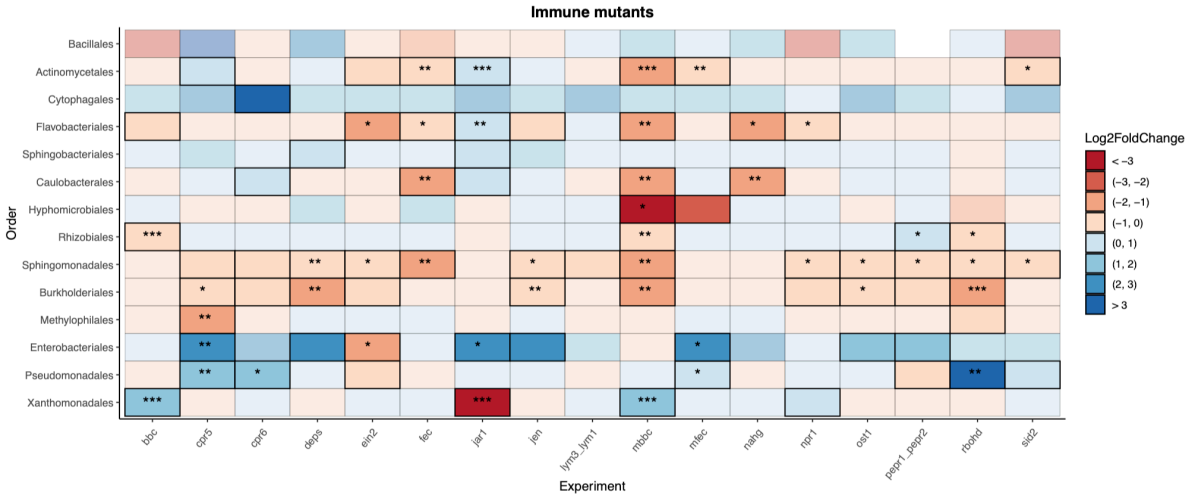

Figure 3: Bacterial composition change upon immune system protein knockout(s) in *A. thaliana*. The relative abundances of different bacterial orders were compared between *A. thaliana* wildtype plants and immune system mutants, where one or more proteins important for immune system functioning were knocked out. A heatmap was created showing the log2-fold changes of different bacterial orders and the significance of the change. Significant differences are shown by * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$). Differences that are near-significant ($p < 0.1$) are displayed by a thickened outline of the associated heatmap box.

otic stress effects on microbes can confound the results of the analyzed conditions. For the 'rice endosphere salt' dataset, the researchers tested multiple salt concentrations, of which the concentration was chosen where the salt-infused soil did not show large and significant alterations in the relative abundances of the Sphingomonadales (data not shown). For the other experiments, there was no opportunity to control for direct effects of the abiotic stress on the microbial community.

The quality of the final included metagenomic samples was again checked by generating relative abundance plots (Figure S4). Likewise, the influence of the abiotic stress on the microbiome composition was again characterized in PCoA based on β-diversities (Figure S5). A less clear separation of microbiomes was observed for most abiotic stresses than for the pathogen-invasion datasets. For certain datasets, this decreased separation was caused by the presence of other determinants that had an influence on the sample discrimination, like the phosphorous concentration for the 'tomato leaf drought' experiment. However, no determining confounders could be determined for both 'rice' experiments, suggesting here that the abiotic stress might have been a smaller determinant for the microbiome composition than the pathogen invasion.

In the different selected datasets, the relative abundances for all bacterial orders were compared between abiotic stress-affected and healthy plants in a similar fashion as for the pathogen invasion experiments. In the resulting heatmap (Figure 1c), the same bacterial orders were displayed as for the pathogen invasion. Visible is that the trend of decreased relative abundances of the Sphingomonadales does not hold for the different abiotic stresses. Neither drought stress or saline treatment had a significant impact on this specific bacterial order. The flooding treatment was the only condition that did significantly decrease the relative abundance of the Sphingomonadales. However, the direct effect of the flooding condition on the present microbes in the soil could not be determined, for which reason it is uncertain whether the significant decrease was caused by plant stress or direct microbe stress caused by the abiotic condition.

Immune system manipulations decrease the relative abundance of Sphingomonadales

Plants respond to the presence of a perceived pathogen by attracting certain beneficial microbes that can help defend holobionts. When the immune system functioning is impaired, for example through immune system manipulation by pathogen invasion, the normal microbe attraction might not function well anymore, likely lowering the resilience of the holobiont. Furthermore, the impaired immune system functioning may facilitate the growth of certain other microbes that are normally repressed, which in turn may impact the growth of others through resource competition or toxin production. Therefore, immune system alterations may have a profound effect on the bacterial composition, and might partially explain the influence of pathogen invasion on the microbiome.

To examine whether immune system manipulation by the investigated pathogens could explain the general decrease in the relative abundance of Sphingomonadales, a dataset with

different *A. thaliana* mutants that had compromised immune systems was obtained and analyzed. Within this dataset, there were 17 different mutant plants present with knockouts of different proteins important to immune system functioning. For all these mutant plants, the bacterial composition was compared to *A. thaliana* WT.

Again, the same methodology was used to create the heatmap of Figure 3, except now all measured bacterial orders were included. These bacterial orders do not completely overlap with the bacterial orders from Figure 1. The heatmap portrays that the log2-fold change of the relative abundance of the Sphingomonadales is consistent again across the different immune system mutants. Similar to the pattern associated with pathogen invasion, there was a general significant decrease in the relative abundance. The decrease was generally relatively small (< 2 fold), and not significant or near-significant in all conditions, in contrast to the pattern associated to the pathogen presence. Therefore, certain immune system function alterations appear to have a more pronounced effect on the Sphingomonadales than others.

Discussion & Conclusion

The data presented in this research project demonstrates for the first time that pathogens consistently and significantly lower the relative abundance of the Sphingomonadales. The decrease happens across different experimental conditions that vary in pathogen type and plant host. No other bacterial orders showed a consistent bacterial alteration upon infection across conditions, with various bacterial orders displaying significant changes in only one or two experiments. The inconsistency of the changes likely depends on confounding factors present in the experimental conditions, like the infected plant host or abiotic circumstances. Boxplots of the relative abundances of the Sphingomonadales in both pathogen-affected and healthy plants confirmed that there was a consistent downshift of the specific bacterial order upon infection. Therefore, it appeared that the change in the relative abundance of the Sphingomonadales was linked to general pathogen infection in plants.

Significant microbiome-alterations were mostly present at the pathogen site, showing that the largest microbiome disturbance is local. For the Sphingomonadales, the decrease in relative abundance was specifically linked to local presence of pathogens as well. This shows that systemic signalling by plants upon infection is not responsible for a decreased presence of the bacterial order.

Pathogens could locally impact the Sphingomonadales both in a direct and indirect way. With direct manipulation, pathogens could actively suppress Sphingomonadales. From literature it is known that Fungal pathogen *V. dahliae*, included in the datasets, can directly suppress Sphingomonadales growth with its effector protein VdAve1.[12] Therefore, it appears plausible that other pathogens might also secrete currently unknown effector proteins that specifically interact with the Sphingomonadales. Further research is needed to specify whether other included pathogen types also can directly suppress members of this bacterial order. For this purpose, proteins secreted by pathogens could be administered directly to isolates of Sphingomonadales, to look at the growth rates of the bacterial order as a whole upon encountering the pathogen secretome. Alternately, specific strains or families of the Sphingomonadales could be subjected to the so-called pathogen culture filtrate, to see more specifically which bacteria of the order are being targeted. An important assumption for this approach is that the pathogens have a similar secretome in culture and in plant tissue. Therefore, there is a probability to obtain a false absence of influence of the culture filtrate on the bacterial growth rate. However, if one of the pathogen culture filtrates has an effect on the growth rate of the bacteria, the presence of one or more growth-reducing pathogen proteins in the mixture is probable. This would provide opportunities to subsequently identify the protein responsible for the decrease.

Apart from local impact, pathogens could potentially affect local bacterial presence indirectly by altering plant functioning, through increasing plant stress, and locally manipulating the immune system. Different abiotic stresses that instigated plant stress did not alter the Sphingomonadales presence in a general way, with only the flooding of wheat plants causing a significant decrease in the bacterial order. For this significant decrease, it could furthermore not be determined whether it was caused by plant stress, or by the flooding process itself; oxygen deprivation stemming from flooding can have a significant impact on the bacterial composition as well, increasing anaerobic bacteria and concomitantly reducing other bacterial orders.[22] As a result, the conclusions from this abiotic stress condition should be approached with caution. Because two out of three abiotic stress conditions did not instigate significant downshifts in the relative abundances of the Sphingomonadales, it appears plausible that general plant stress is not one of the main pillars for Sphingomonadales suppression by pathogens upon invasion.

A pitfall of the included datasets could be that the instigated abiotic stresses might not have been stressful enough for plants to subsequently alter the microbiome. The PCoA analysis would support this claim for certain datasets, as the abiotic stress had less influence on the bacterial composition than pathogen invasion. However, as all included datasets had significantly decreased plant growth as a characteristic, plants from the studies were per definition stressed. Furthermore, both drought-stress datasets showed a non-signifant increase instead of a decrease in the relative abundance of the Sphingomonadales. Therefore, it seems unlikely that more stressful abiotic circumstances would cause a pattern of Sphingomonadales reduction as in the pathogen-affected samples.

Immune system suppression is another indirect way for pathogens to manipulate the associated microbiome. Surprisingly, almost all *A. thaliana* immune system mutants showed a significant decrease in the relative abundance of the Sphingomonadales. This suggests that immune system manipulation by pathogens could contribute to this microbiome shift. With the knowledge that immune system manipulation by pathogens could alter the relative abundance of Sphingomonadales upon infection, it remains a mystery what is the exact nature of the connection between the bacterial order and the immune system. Potentially, in a normal situation the immune system allows for active recruitment of Sphingomonadales, with immune manipulation leading to decreased attraction. For active recruitment, plants secrete certain molecules that allow growth of specific microbes that have enzymes for utilization of the molecules in their metabolism. For the Sphingomonadales, it is already known they have enzymes for the conversion of certain plant-generated molecules.[41] However, more research is needed to figure out whether Sphingomonadales are actively attracted by the immune system, for example by bacterial attraction studies upon plant infection.[42] In these studies, it is researched which soil bacteria are drawn into the rhizosphere from a long distance due to root exudation. If immune system activation leads to active attraction of Sphingomonadales, the root exudates excreted upon infection should attract more Sphingomonadales than the healthy plant exudate. Furthermore, attraction upon pathogen infection could be researched in a similar fashion, where root-associated pathogens could potentially cause a decrease in the attraction of the Sphingomonadales due to local immune system manipulation.

Another explanation for the decrease of the relative abundance of the Sphingomonadales upon immune system manipulation by pathogens might be that the immune system normally suppresses other microorganisms that are in resource competition with the Sphingomonadales. With reduced immune suppression, these microorganisms can grow more rapidly and decrease the available food sources. Resource competition could be identified using metabolic profiling studies, where different microorganisms can portray similar metabolic pathways to metabolize particular metabolites. Further research should provide a more definitive answer on the reason why altered immune system functioning decreases the Sphingomonadales presence.

With the current results, it was displayed that there is a significant reduction in the relative abundances of the Sphingomonadales upon pathogen invasion. However, the significance of this reduction on plant functioning is currently still unknown. Potentially, the reduced presence of Sphingomonadales reduces the resilience of the holobiont and leaves the plant host more vulnerable to infection by several pathogen types. Different synthetic bacterial communities could be created with increased and decreased amounts of Sphingomonadales, to research the effect of the Sphingomonadales presence on the resilience of the holobiont. In this setup, it could be determined how successful the different synthetic communities are at fighting infection with different pathogen types. If the Sphingomonadales have a protective function in plant holobionts, an increased abundance should protect from infections, while a decreased abundance should give rise to more infected plants. It was already shown before that a Sphingomonas strain, isolated from *A. thaliana*, activates a subset of plant defense genes upon invasion of bacterial pathogen *P. syringae* DC3000, promoting immunity.[43] Therefore, it is not unlikely that specific other strains of the Sphingomonadales order may also assist immunity in different plant hosts against different pathogen types.

The search for bacterial strains that may assist in plant immunity may eventually lead to a better way to control pathogen infections in plants. Namely, if specific beneficial bacterial strains that reduce the infection potential of pathogens are boosted in plants, crops might become infected less often, causing less plant wasting. As certain pathogens are currently detrimental to crop harvesting yields,[17] with no known control options, other methods to reduce pathogen invasion potential are much needed. Therefore, the results of this studies may potentially provide a new direction to limiting crop wasting, where artificially increasing Sphingomonadales might boost immune system functioning.

References

1 Zilber-Rosenberg, I. & Rosenberg, E. Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. *FEMS Microbiology Reviews* **32**, 723–735 (2008).

2 Dastogeer, K. M. G., Tumpa, F. H., Sultana, A., Akter, A. & Chakraborty, A. Plant microbiome-an account of the factors that shape community composition and diversity (2020).

3 Compant, S., Samad, A., Faist, H. & Sessitsch, A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research* **19**, 29–37 (2019).

4 Fitzpatrick, C. R. *et al.* The plant microbiome: From ecology to reductionism and beyond. *Annual Review of Microbiology* **74**, 81–100 (2020).

5 Siddiqui, Z. S. *et al.* Scrutinizing the application of saline endophyte to enhance salt tolerance in rice and maize plants. *Frontiers in Plant Science* **12**, 3334 (2022).

6 Li, P., Lu, Y. J., Chen, H. & Day, B. The lifecycle of the plant immune system. *Critical reviews in plant sciences* **39**, 72 (2020).

7 Bakker, P. A., Pieterse, C. M., de Jonge, R. & Berendsen, R. L. The soil-borne legacy. *Cell* **172**, 1178–1180 (2018).

8 Nejat, N. & Mantri, N. Plant immune system: Crosstalk between responses to biotic and abiotic stresses the missing link in understanding plant defence. *Current issues in molecular biology* **23**, 1–16 (2017).

9 Toruño, T. Y., Stergiopoulos, I. & Coaker, G. Plant-pathogen effectors: Cellular probes interfering with plant defenses in spatial and temporal manners. *Annual review of phytopathology* **54**, 419 (2016).

10 Rovenich, H., Boshoven, J. C. & Thomma, B. P. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology* **20**, 96–103 (2014).

11 Pfeilmeier, S. *et al.* The plant nadph oxidase rbohd is required for microbiota homeostasis in leaves. *Nature Microbiology 2021 6:7* **6**, 852–864 (2021).

12 Snelders, N. C. *et al.* Microbiome manipulation by a soil-borne fungal plant pathogen using effector proteins. *Nature Plants* (2020).

13 Gao, M. *et al.* Disease-induced changes in plant microbiome assembly and functional adaptation (2021).

14 Seybold, H. *et al.* A fungal pathogen induces systemic susceptibility and systemic shifts in wheat metabolome and microbiome composition (2020).

15 Durán, P. *et al.* A fungal powdery mildew pathogen induces extensive local and marginal systemic changes in the arabidopsis thaliana microbiota. *Environmental Microbiology* **23**, 6292–6308 (2021).

16 Larousse, M. *et al.* Tomato root microbiota and phytophthora parasitica-associated disease. *Microbiome* **5**, 1–11 (2017).

17 Suhaimi, N. S. M. *et al.* Diversity of microbiota associated with symptomatic and non-symptomatic bacterial wilt-diseased banana plants determined using 16s rrna metagenome sequencing. *World Journal of Microbiology and Biotechnology* **33** (2017).

18 Bez, C. *et al.* The rice foot rot pathogen dickeya zeae alters the in-field plant microbiome. *Environmental Microbiology* **23**, 7671–7687 (2021).

19 Debray, R. *et al.* Water stress and disruption of mycorrhizas induce parallel shifts in phyllosphere microbiome composition. *New Phytologist* **234**, 2018–2031 (2022).

20 Santos-Medellín, C. *et al.* Prolonged drought imparts lasting compositional changes to the rice root microbiome. *Nature Plants* **7**, 1065–1077 (2021).

21 Santos, S. S. *et al.* Specialized microbiomes facilitate natural rhizosphere microbiome interactions counteracting high salinity stress in plants (2021).

22 Francioli, D. *et al.* Flooding causes dramatic compositional shifts and depletion of putative beneficial bacteria on the spring wheat microbiota. *Frontiers in Microbiology* **12** (2021).

23 Callahan, B. J., Holmes, S. P., Sankaran, K., Fukuyama, J. A. & McMurdie, P. J. Workflow for microbiome data analysis: from raw reads to community analyses. (2017). Materials methods.

24 R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2021).

25 Callahan, B. J. *et al.* Dada2: High-resolution sample inference from illumina amplicon data. *Nature Methods* **13**, 581–583 (2016).

26 McMurdie, P. J. & Holmes, S. Phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8** (2013).

27 Wright, E. S. Using decipher v2.0 to analyze big biological sequence data in r. *The R Journal* **8**, 352–359 (2016).

28 Schliep, K. phangorn: phylogenetic analysis in r. *Bioinformatics* **27**, 592–593 (2011).

29 Schliep *et al.* Intertwining phylogenetic trees and networks. *Methods in Ecology and Evolution* **8**, 1212–1220 (2017).

30 Morgan, M. *BiocManager: Access the Bioconductor Project Package Repository* (2021). R package version 1.30.16.

31 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016).

32 Bache, S. M. & Wickham, H. *magrittr: A Forward-Pipe Operator for R* (2022). R package version 2.0.3.

33 Wickham, H., François, R., Henry, L. & Müller, K. *dplyr: A Grammar of Data Manipulation* (2022). R package version 1.0.8.

34 Wickham, H. *et al.* Welcome to the tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).

35 Wilke, C. O. *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'* (2020). R package version 1.1.1.

36 Kassambara, A. *ggpubr: 'ggplot2' Based Publication Ready Plots* (2020). R package version 0.4.0.

37 Auguie, B. *gridExtra: Miscellaneous Functions for "Grid" Graphics* (2017). R package version 2.3.

38 Ram, K. & Wickham, H. *wesanderson: A Wes Anderson Palette Generator* (2018). R package version 0.3.6.

39 Moser, B. K. & Stevens, G. R. Homogeneity of variance in the two-sample means test. *American Statistician* **46**, 19–21 (1992).

40 Yuan, Y., Brunel, C., van Kleunen, M., Li, J. & Jin, Z. Salinity-induced changes in the rhizosphere microbiome improve salt tolerance of hibiscus hamabo. *Plant and Soil* **443**, 525–537 (2019).

41 Glaeser, S. P. & Kämpfer, P. The family sphingomonadaceae. *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria* **9783642301971**, 641–707 (2014).

42 Schulz-Bohm, K. *et al.* Calling from distance: attraction of soil bacteria by plant root volatiles. *ISME Journal* **12**, 1252–1262 (2018).

43 Teixeira, P. J. P., Colaianni, N. R., Fitzpatrick, C. R. & Dangl, J. L. Beyond pathogens: microbiota interactions with the plant immune system. *Current Opinion in Microbiology* **49**, 7–17 (2019).

Supplementary material

Experiment Name	BioProject accession number	Diseased samples	Healthy samples
tomato root verticillium	PRJEB34281	ERR3509498 - ERR3509500	ERR3509495 - ERR3509497
chili xylem fusarium	PRJNA667302	SRR12773360 - SRR12773365, SRR12773367 - SRR12773369	SRR12773370 - SRR12773376, SRR12773378, SRR12773379
wheat leaf zymoseptoria	PRJNA549447	N/A	N/A
wheat adjacentleaf zymoseptoria	PRJNA549447	N/A	N/A
arabidopsis leaf mildew	PRJEB43139	ERR6617855 - ERR6617863	ERR6617846 - ERR6617854
arabidopsis root mildew	PRJEB43139	ERR6617837 - ERR6617845	ERR6617828 - ERR6617836
tomato rhizosphere phytophthora	PRJNA354847	SRR5079786-SRR5079788	SRR5079784, SRR5079785, SRR5079789
banana pseudostem ralstonia	PRJNA277904	SRR1920149 - SRR1920152	SRR1920144, SRR1920146, SRR1920147
rice stem dickeya	PRJNA602829	SRR10959410 - SRR10959426, SRR10959431, SRR10959442, SRR10959453 - SRR10959461, SRR10959464, SRR10959475, SRR10959573, SRR10959574	SRR10959388 - SRR10959391, SRR10959436 - SRR10959441, SRR10959443, SRR10959444, SRR10959471 - SRR10959474, SRR10959476, SRR10959484, SRR10959495, SRR10959506, SRR10959564 - SRR10959566, SRR10959569 - SRR10959571

Table S1: Selected samples for pathogen invasion experiments. For each dataset, displayed are the associated BioProject accession number, which samples were analyzed as pathogen-affected samples, and which samples were analyzed as healthy samples.

Experiment Name	BioProject accession number	Stressed samples	Healthy samples
tomato leaf drought	PRJNA741547	N/A	N/A
rice endosphere drought	PRJNA551661	SRR9612737, SRR9612745, SRR9612865, SRR9612870	SRR9612739, SRR9612647, SRR9612450, SRR9612456, SRR9612864, SRR9612867, SRR9612869, SRR9612571
rice endosphere salt	PRJNA690819	SRR13428751, SRR13428755, SRR13428987, SRR13428992, SRR13429003, SRR13429007, SRR13429011	SRR13428782, SRR13428803, SRR13428807, SRR13428875, SRR13428923, SRR13428958, SRR13428998
wheat root flood	PRJEB47399	ERR6717428 - ERR6717433, ERR6717440 - ERR6717445, ERR6717450 - ERR6717455	ERR6717434 - ERR6717439, ERR6717446 - ERR6717449, ERR6717456 - ERR6717463

Table S2: Selected samples for abiotic stress experiments. For each dataset, displayed are the associated BioProject accession number, which samples were analyzed as stressed samples, and which samples were analyzed as healthy samples.

Experiment Name	Sequencing type	Trim length forward read	Trim length reverse read	Trim front (both)
tomato root verticillium	Paired-end	N/A	N/A	N/A
chili xylem fusarium	Paired-end	235	235	5
wheat leaf zymoseptoria	N/A	N/A	N/A	N/A
wheat adjacentleaf zymoseptoria	N/A	N/A	N/A	N/A
arabidopsis leaf mildew	Paired-end	260	240	5
arabidopsis root mildew	Paired-end	260	240	5
tomato rhizosphere phytophthora	Single-end	220	N/A	0
banana pseudostem ralstonia	Paired-end	145	100	10
rice stem dickeya	Paired-end	250	250	20

Table S3: Characteristics of the read trimming procedures for each pathogen invasion dataset. Displayed is whether paired-end of single-end sequencing was performed, what the total trim length was of the forward and reverse reads, and how many bases were trimmed of the front end of the read for both the forward and reverse read. In case of single-end sequencing, the associated read length is displayed under the forward read trim length.

Experiment Name	Sequencing type	Trim length forward read	Trim length reverse read	Trim front (both)
tomato leaf drought	N/A	N/A	N/A	N/A
rice endosphere drought	N/A	N/A	N/A	N/A
rice endosphere salt	Paired-end	280	250	5
wheat root flood	Paired-end	250	200	5

Table S4: Characteristics of the read trimming procedures for each abiotic stress dataset. Displayed is whether paired-end of single-end sequencing was performed, what the total trim length was of the forward and reverse reads, and how many bases were trimmed of the front end of the read for both the forward and reverse read. In case of single-end sequencing, the associated read length is displayed under the forward read trim length.

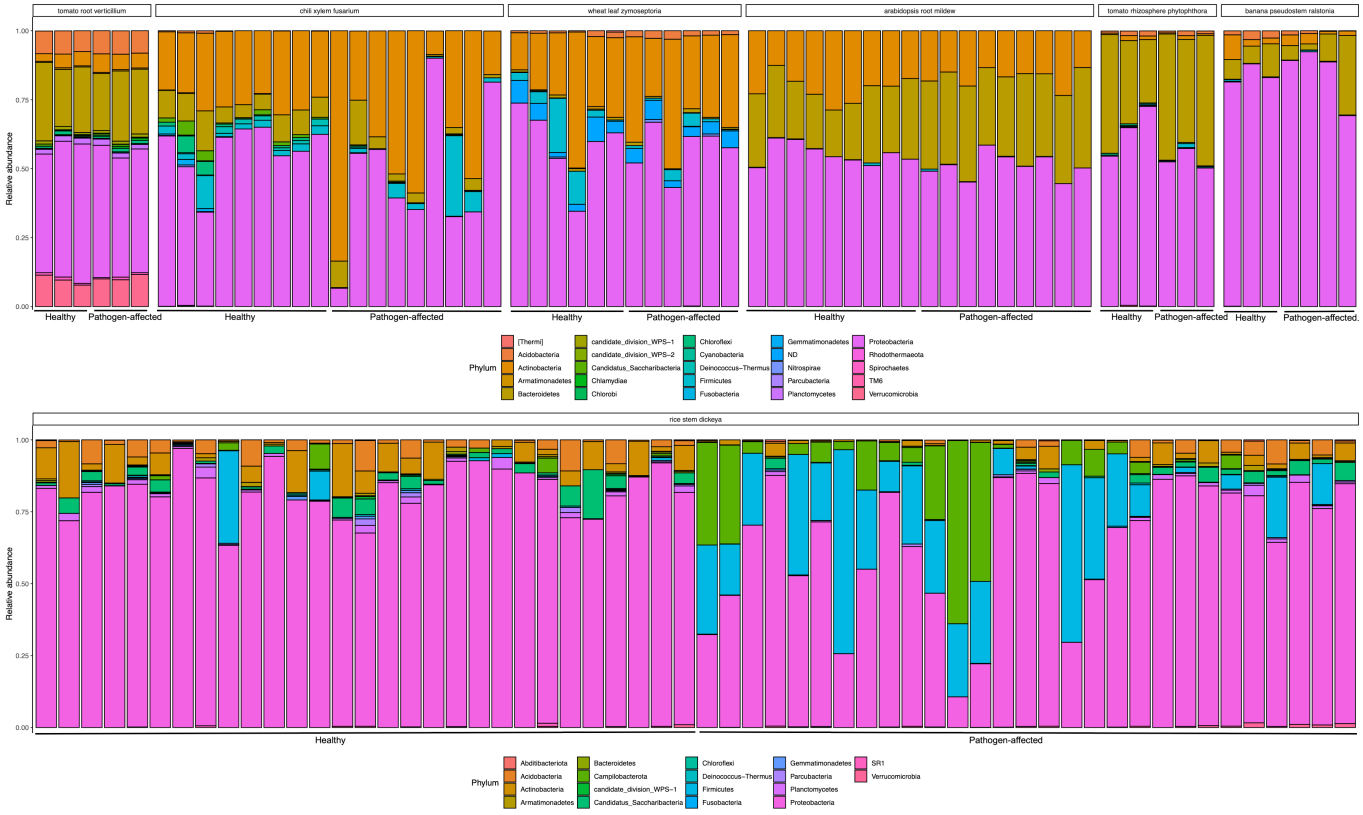

Figure S1: Relative abundances for all bacterial phyla in the pathogen invasion datasets. Pathogen-affected and healthy samples are shown together in one graph for each experiment.

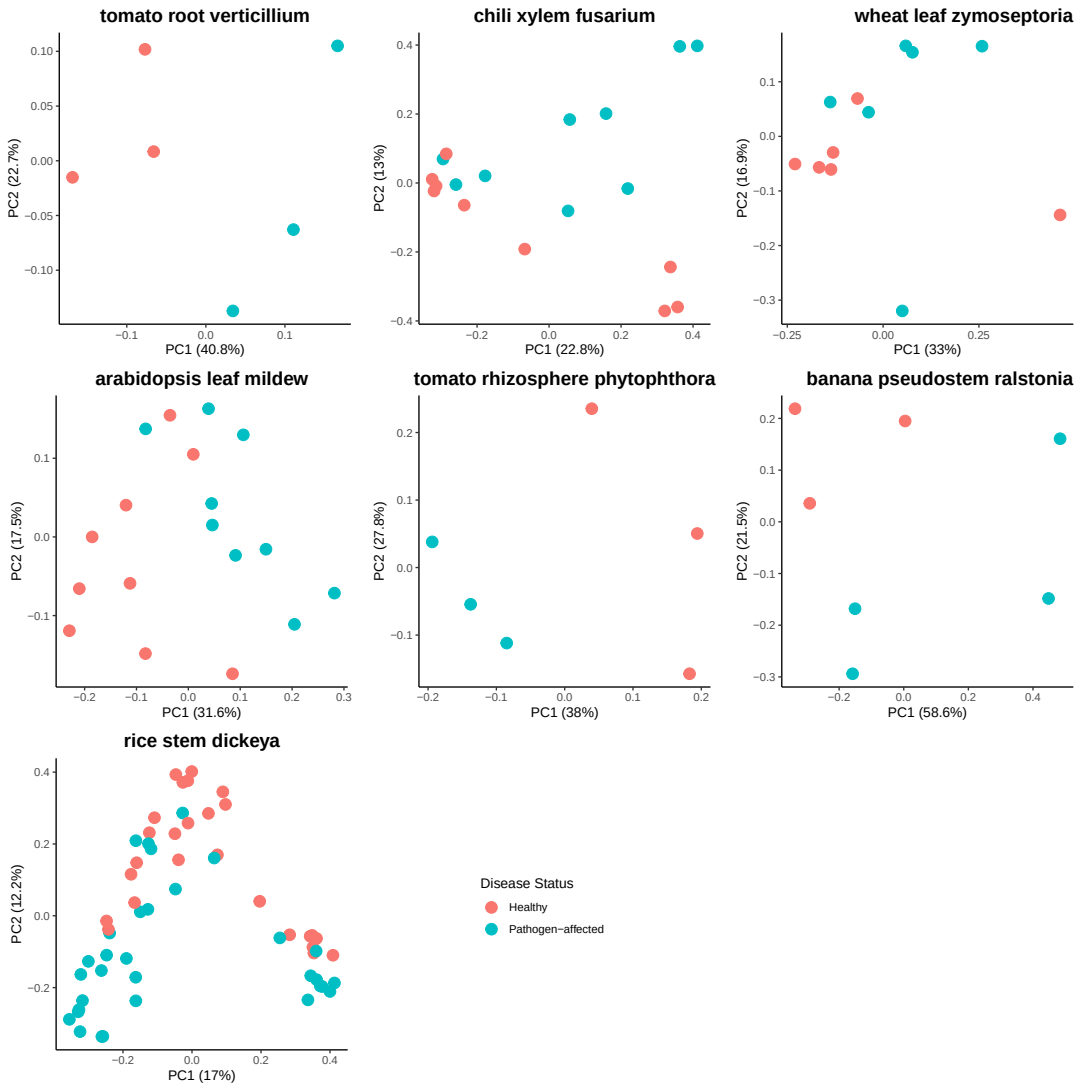

Figure S2: Principal coordinate analysis (PCoA) plot with Bray-Curtis dissimilarity for pathogen invasion datasets. Pathogen-affected and healthy samples are illustrated by the different colors and show separation in space.

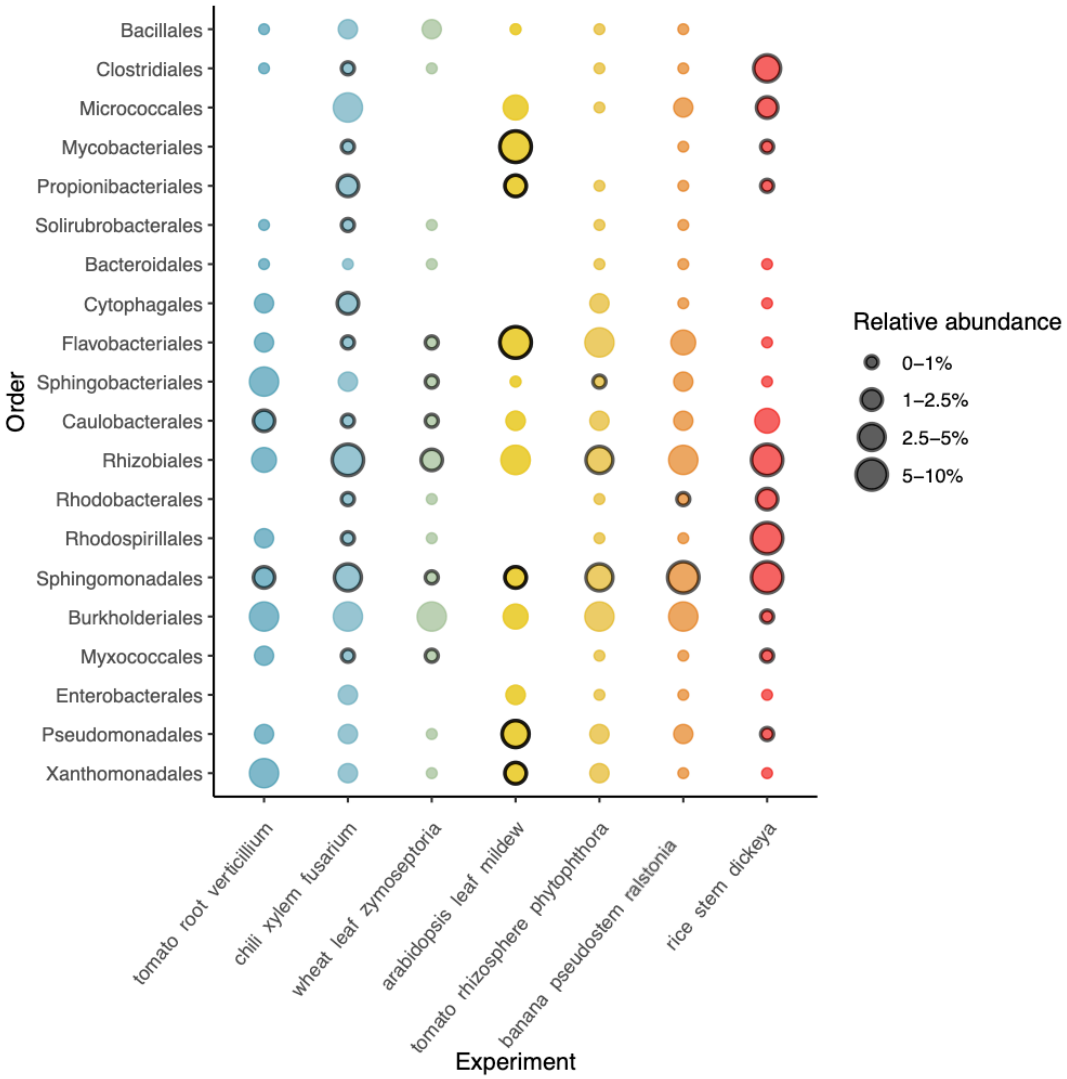

Figure S3: Mean relative abundances bacterial orders in pathogen invasion experiments. Significant and near-significant experimental conditions of the relative abundance comparison are highlighted with the thickened outline of the bubbles. For the Sphingomnonadales, the relative abundances varied mostly between 1-10%.

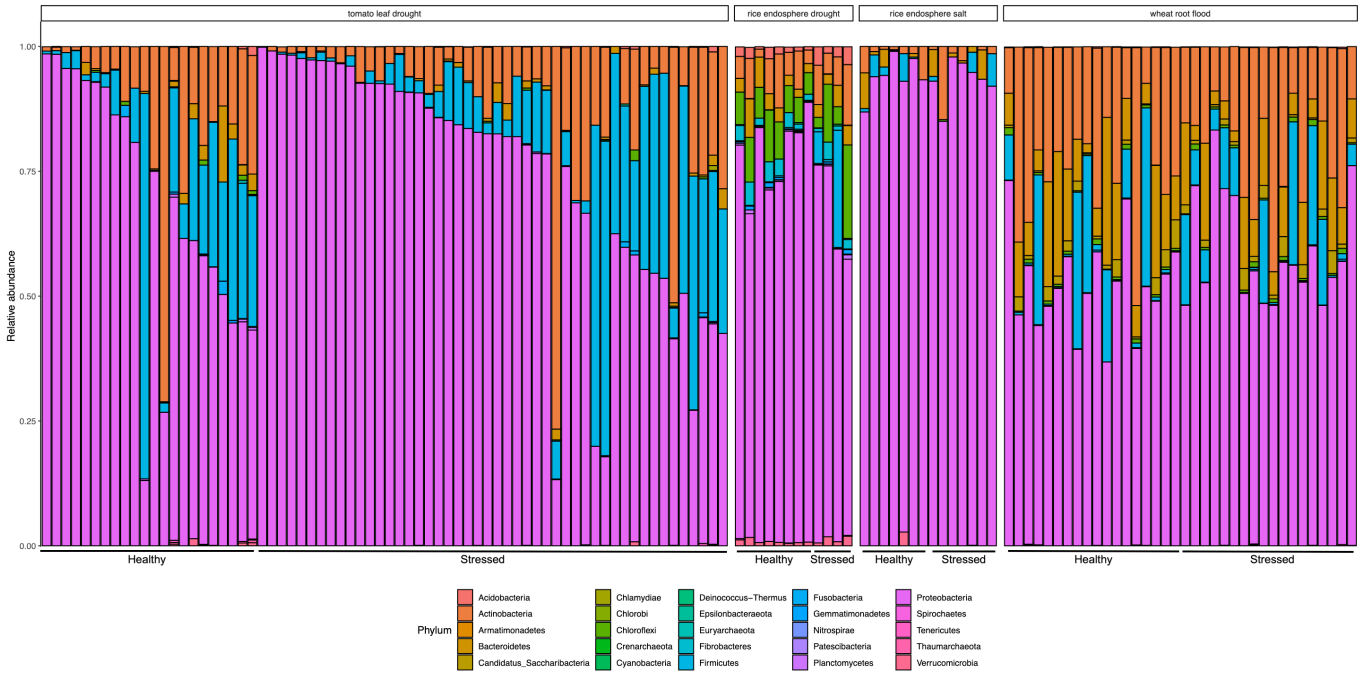

Figure S4: Relative abundances for all bacterial phyla in the abiotic stress datasets. Stressed and healthy samples are shown together in one graph for each experiment.

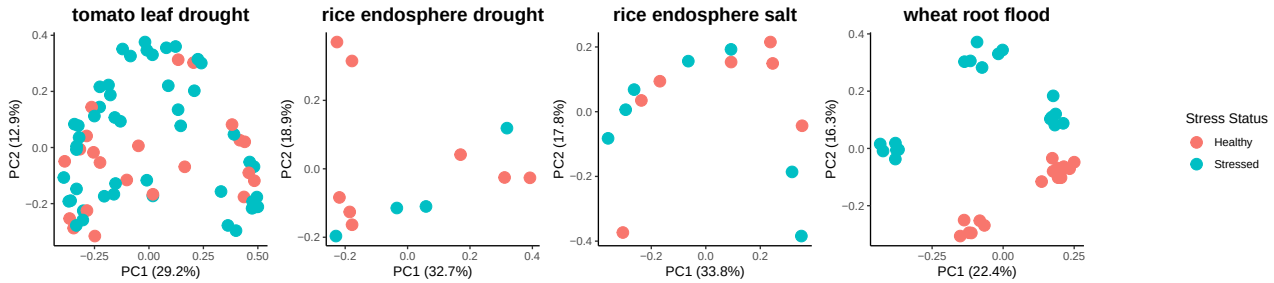

Figure S5: Principal coordinate analysis (PCoA) plot with Bray-Curtis dissimilarity for abiotic stress datasets. Stressed and healthy are illustrated by the different colors and show some separation in space.