Roles of biotic and abiotic factors in shaping plant microbiota diversity and plant health

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.Sc. Maryam Mahmoudi

aus Ijroud/Iran

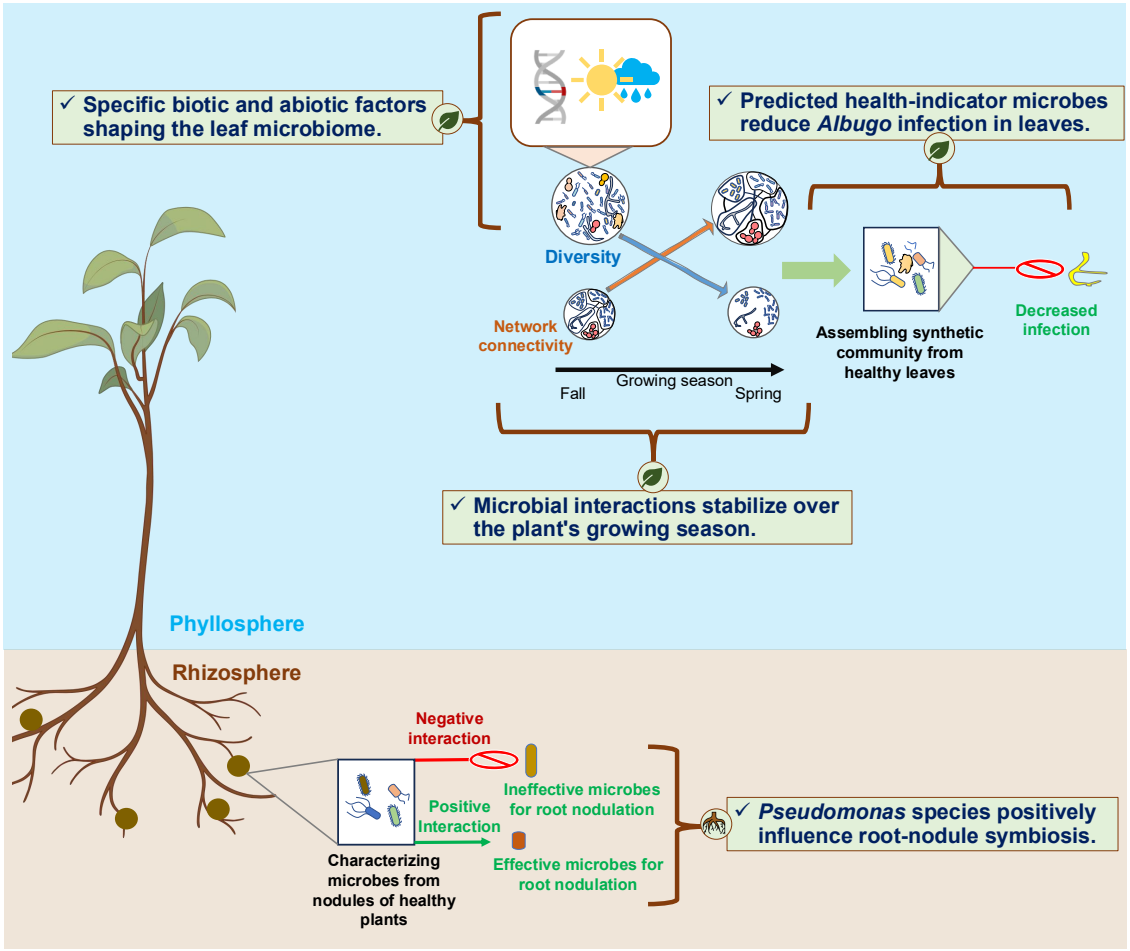
Tübingen

2024

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	08.05.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Eric Kemen
2. Berichterstatter:	Prof. Dr. Sven Nahnsen
3. Berichterstatter:	Prof. Dr. Bart Thomma

Graphical summary



Summary

Plants, like animals and humans, have a diverse microbiome crucial in maintaining host health and helping them cope with biotic and abiotic stresses. Interactions within these microbial communities, from mutualism to antagonism, add layers of complexity and variability. However, it is challenging to determine the specific roles of different interactions in the formation and stability of the microbiome. To address this issue, it is essential to study plant microbial communities in different states and under fluctuating environmental conditions. Achieving this goal requires collecting sufficient number of samples at different times and locations to picture diverse microbiomes.

An example of such a diverse and fluctuating environment is the phyllosphere. It is the aerial part of a plant and is directly exposed to various environmental factors such as temperature and rain. These abiotic factors can alter the leaf environment within hours or days. In **Chapter 2** of this study, I analyzed the impact of various environmental factors on the microbiome formation of different compartments of the model plant, *Arabidopsis thaliana*. Over a period of five years, we collected samples from natural ecosystems and studied their microbiome. We examined the different microbial groups, including bacteria, fungi and non-fungal eukaryotes. I identified microbes that specifically respond to environmental factors in epiphytic and endophytic compartments. Additionally, I found a correlation between environmental factors and variations in the connectivity of microbial co-abundance networks.

In addition to being influenced by environmental factors, the plant microbiome undergoes distinct changes as its host develops. As a result, it has become essential to determine the degree of stability and variability exhibited by the microbiome over time. Answering these questions requires a detailed study of its dynamics throughout the developmental stages of the host. This investigation extends to **Chapter 3**, where we have recorded monthly changes in the leaf microbiome of *Arabidopsis thaliana* throughout its growing season. This observational study was conducted in a garden experiment between November and March. Our results revealed the existence of conserved temporal patterns: at the beginning of the plant growing season, microbial communities and networks undergo a stabilization phase with reduced diversity and variability. Despite the high turnover rate, we identified some core taxa that remain on *Arabidopsis* leaves throughout the season, and the core microbiome is not necessarily the hub microbes in the network. The stability of the associations between the core and/or hub microbes and the host plant indicates a high degree of microbial adaptation to the leaf niche. This is also the case for traits associated with dysbiosis and pathogen infection.

Studying the core and hub microbiome functions is important to better understand the

causes of dysbiosis in plants. This knowledge allows us to assist plants in the management of dysbiosis and the effective control of pathogens. During my research, I identified *Albugo laibachii*, a plant pathogen, as a potential core microbe influenced by environmental factors with high impact. In **Chapter 4** of my thesis, I focused on identifying communities and microbiomes related to plant health and used them to protect plants from pathogens. To achieve this, I employed machine learning techniques to distinguish between infected and uninfected plants. I ranked the microbes based on their importance scores and identified important taxa that effectively discriminate between infected and uninfected *Arabidopsis* plants. To investigate this hypothesis further, I examined the pairwise interactions between these health-associated microbes and the pathogen *Albugo laibachii*. In conclusion, I could identify functional redundancy in complex microbial communities that stabilize protective equilibria to the causal agent *Albugo laibachii*.

Besides protecting plants from pathogens, microbes are critical for plant health as they help provide essential nutrients like nitrogen. Some plants have nodules in their roots where nitrogen fixation occurs. *Rhizobia* bacteria play a role in nitrogen fixation, but the interactions among other microbial components involved in the process are poorly understood. In **Chapter 5**, we aimed to investigate how the soil microbiome affects the growth of healthy *Lotus* plants as a model, specifically how it affects nodulation, where nitrogen fixation occurs. In our study, I used a machine learning model to discriminate between two different phenotypes of *Lotus* plants: healthy and starved. Our results showed that non-rhizobial microbes negatively interact with ineffective *Rhizobium* or what we call cheater microbes. These interactions were confirmed through correlation networks and in planta experiments. Our research shows that the microbiome in nodules has a significant impact on the outcome of root-nodule symbiosis.

In summary, my project has provided deep insights into plant microbe-microbe interactions. It has focused on identifying critical ecological patterns for microbiome stability and assembly. These findings help us to understand the complex relationships within the plant microbiome, particularly under varying environmental conditions. In addition, these findings have important implications for biotechnology. In particular, improving plant robustness to biotic and abiotic factors by manipulating plant microbiomes as probiotics through microbiome engineering.

Zusammenfassung

Pflanzen haben, wie Tiere und Menschen, ein vielfältiges Mikrobiom, welches entscheidend für die Aufrechterhaltung der Gesundheit des Wirts ist und ihm hilft mit biotischen und abiotischen Stressfaktoren umzugehen. Die Interaktionen innerhalb dieser mikrobiellen Gemeinschaften, von Mutualismus bis Antagonismus, verleihen dem Mikrobiom zusätzliche Komplexität und Variabilität. Es ist jedoch eine Herausforderung, die spezifische Rolle der verschiedenen Interaktionen bei der Bildung und Stabilität mikrobieller Gemeinschaften zu bestimmen. Um dies zu erreichen, müssen wir die verschiedenen Zustände des Pflanzenmikrobioms unter schwankenden Umweltbedingungen untersuchen, was durch die Sammlung einer ausreichenden Anzahl von Proben zu verschiedenen Zeitpunkten and unterschiedlichen Orten erreicht werden kann.

Ein Beispiel für eine solch vielfältige und schwankende Umgebung ist die Phyllosphäre. Sie ist der oberirdische Teil einer Pflanze und ein vielfältiger mikrobieller Lebensraum, der direkt verschiedenen Umweltfaktoren wie Temperatur und Regen ausgesetzt ist. Diese abiotischen Faktoren können das Blattmilieu innerhalb von Minuten, Stunden oder Tagen schnell verändern. In Kapitel 2 dieser Studie analysiere ich die Auswirkungen verschiedener Umweltfaktoren auf die Mikrobiombildung in verschiedenen Kompartimenten der Modellpflanze *Arabidopsis thaliana*. Im Laufe von fünf Jahren sammelten wir Proben aus natürlichen Ökosystemen und untersuchten das Mikrobiom auf verschiedene mikrobielle Gruppen, darunter Bakterien, Pilze und nicht-pilzlichen Eukaryoten. Ich identifizierte epiphytische und endophytische Mikroben, die spezifisch auf Umweltfaktoren reagieren, wobei sich eine Korrelation zwischen Umweltfaktoren und Variation mikrobieller Netzwerke herausstellte.

Das pflanzliche Mikrobiom wird nicht nur von der Umwelt beeinflusst, sondern verändert sich auch im Laufe der Entwicklung seines Wirtes. Daher ist es wichtig, den Grad der Stabilität und Variabilität des Mikrobioms im Laufe der Zeit zu bestimmen. Die Beantwortung dieser Frage erfordert eine detaillierte Untersuchung der Dynamik des Mikrobioms während der verschiedenen Entwicklungsstadien des Wirts. Im Rahmen meines Forschungsprojekts erstreckt sich diese Untersuchung auf Kapitel 3, in dem wir monatliche Veränderungen im Blattmikrobiom von *Arabidopsis thaliana* während der gesamten Wachstumsperiode aufgezeichnet haben. Diese Beobachtungsstudie wurde in einem Gartenexperiment zwischen November und März durchgeführt. Unsere Ergebnisse zeigten die Existenz konservierter zeitlicher Muster: Zu Beginn der Vegetationsperiode durchlaufen die mikrobiellen Gemeinschaften und Netzwerke eine Stabilisierungsphase mit reduzierter Diversität und Variabilität. Trotz der hohen Fluktuationsrate konnten wir einige Kerntaxa identifizieren, die während des gesamten Zeitraumes auf den Blättern von

Arabidopsis verblieben. Gleichzeitig müssen die Kernmikroben nicht notwendigerweise den mikrobiellen Knotenpunkt im Netzwerk entsprechen. Die Stabilität der Assoziationen zwischen den Kern- und/oder Hub-Mikroben und der Wirtspflanze deutet auf ein hohes Maß mikrobieller Anpassung an die Blattnische hin. Dies gilt auch für Merkmale, die mit Dysbiose und Pathogeninfektion verbunden sind.

Die Untersuchung der Kern- und Knotenpunktfunktionen des Mikrobioms ist wichtig, um die Ursachen der Dysbiose bei Pflanzen besser zu verstehen. Dieses Wissen kann helfen, Pflanzen vor Dysbiose und Krankheitserreger besser zu schützen. Während meiner Forschung habe ich *Albugo laibachii*, ein Pflanzenpathogen, als eine potenzielle Kernmikrobe identifiziert, die von Umweltfaktoren stark beeinflusst wird. In Kapital 4 meiner Dissertation konzentrierte ich mich auf die Identifizierung von Gemeinschaften und Mikrobiomen, die mit der Pflanzengesundheit in Zusammenhang stehen, und nutzte sie, um Pflanzen vor Krankheitserregern zu schützen. Zu diesem Zweck setzte ich maschinelle Lernverfahren ein, um zwischen infizierten und nicht infizierten Pflanzen zu unterscheiden. Ich erstellte eine Rangliste der Mikroben auf der Grundlage ihrer Wichtigkeit und identifizierte Taxa, die effektiv zwischen infizierten und nicht infizierten Arabidopsispflanzen unterscheiden. Um diese Hypothese zu testen, untersuchte ich die paarweisen Wechselwirkungen zwischen diesen gesundheitsassoziierten Mikroben und dem Pathogen *Albugo laibachii*. Abschließend konnten wir eine funktionelle Redundanz in komplexen mikrobiellen Gemeinschaften feststellen, die schützende Gleichgewichte gegenüber dem Erreger *Albugo laibachii* herstellen.

Mikroben schützen die Pflanzen nicht nur vor Krankheitserregern, sondern sind auch entscheidend für die Gesundheit, da sie dazu beitragen, wichtige Nährstoffe wie Stickstoff bereitzustellen. Einige Pflanzen haben Knöllchen in ihren Wurzeln, in denen die Stickstofffixierung stattfindet. *Rhizobien* spielen bei der Stickstofffixierung eine große Rolle, die Wechselwirkungen zwischen den anderen mikrobiellen Komponenten, die an diesem Prozess beteiligt sind, sind jedoch kaum bekannt. In Kapital 5 wollten wir untersuchen, wie sich das Bodenmikrobiom auf das Wachstum gesunder Lotuspflanzen auswirkt, insbesondere auf die Knöllchenbildung, bei der die Stickstofffixierung stattfindet. In unserer Studie verwendete ich ein maschinelles Lernmodell, um zwischen zwei verschiedenen Phänotypen von Lotus-Pflanzen zu unterscheiden: dem gesunden und dem ausgehungerten Phänotyp. Unsere Ergebnisse zeigten, dass nicht-rhizobiale Mikroben negativ mit ineffektiven *Rhizobien* oder, wie wir sie nennen, mit betrügerischen Mikroben interagieren. Diese Interaktionen wurden durch Korrelationsnetzwerke und in planta-Experimente bestätigt. Unsere Forschung zeigt, dass das Mikrobiom in den Knöllchen einen erheblichen Einfluss auf das Ergebnis der Wurzel-Knöllchen-Symbiose hat.

Zusammenfassend lässt sich sagen, dass mein Projekt tiefe Einblicke in die Interaktionen zwischen Pflanzenmikroben und Mikroben ermöglicht hat. Es konzentrierte sich auf die Identifizierung kritischer ökologischer Muster für die Stabilität und den Aufbau des Mikrobioms. Diese Erkenntnisse helfen uns, die komplexen Beziehungen innerhalb des pflanzlichen Mikrobioms zu verstehen, insbesondere unter verschiedenen Umweltbedingungen. Darüber hinaus haben diese Erkenntnisse wichtige Auswirkungen

auf die Biotechnologie. Die Manipulation des Pflanzenmikrobiomen durch Probiotika durch Mikrobiom-Engineering kann die Robustheit von Pflanzen gegenüber biotischen und abiotischen Faktoren deutlich verbessern.

Acknowledgments

At this point, I would like to take this opportunity to express my sincere gratitude to all those who have supported me throughout my Ph.D. journey. They have made this experience a most valuable and memorable one.

First, I am deeply grateful to Prof. Dr. Eric Kemen, my supervisor, for allowing me to do my Ph.D. in his group, being patient and giving me the freedom to explore my interests, and for his guidance and support all these years.

I want to thank my second supervisor, Prof. Dr. Kay Nieselt, for her invaluable advice and support in the computational aspects of my work and for the time she devoted to reviewing my manuscripts.

I sincerely thank Juliana for her insightful advice and the time she spent with me on my project.

I thank our collaborators, Dr. Macarena Marín and Dr. Duncan B. Crosbie, for their helpful collaboration leading to our joint work.

I deeply appreciate all my colleagues and friends in the Kemen lab, with special mention to Ariane for all the scientific insight and advice on improving my presentations and Vasvi for her guidance and in the wet lab part of my project. I also want to thank Elke for all the help and assistance in the wet lab experiments in my project. I would like to express my gratitude to Alfredo, Paul, Daniel, Vlad, Katrina, Monja, Yiheng, Bilal, Samuel, Leonie, Lucas, Hamed, and Paolo. I am deeply thankful for the scientific discussions and memorable times we shared during lab trips, parties, and other occasions.
I am grateful for the opportunity to supervise Maja, Bengt, and Lynn-Marie during their M.Sc. and B.Sc. studies. Thank you for giving me this wonderful experience.

I acknowledge the Interfaculty Graduate School of Infection Biology and Microbiology (IGIM) for providing an excellent study program and the SPP DECRyPT consortia for supporting scientific networks and programs. I would also like to thank the European Research Council (ERC) for funding this project under the DeCoCt research program (grant agreement: ERC-2018-COG 820124), the Cluster of Excellence “Controlling Microbes to Fight Infections” (CMFI; EXC 2124), and the DFG’s SPP 2125 DECRyPT program for their support. I further thank the de.NBI Cloud Storage and bwForCluster

BinAC Tübingen provide resources for storing and analyzing the data.

I want to express my heartfelt appreciation to my boyfriend for his unwavering emotional support, for sharing enjoyable moments together, and for dedicating time to discussing scientific matters related to my project.

To my parents, Ramezanali and Rana, your unwavering love and support have kept me going. In times when physical presence was challenging, and I found comfort in scientific pursuits, I am deeply grateful for your patience. I hope that my accomplishments bring you joy.

To my siblings, thank you for always being there, filling in when I'm absent, and your kindness. You make me proud, and I feel blessed to have you in my life.

With deepest appreciation,

Maryam

Contents

Abbreviations

A. thaliana	*Arabidopsis thaliana*
A. laibachii	*Albugo laibachii*
ASV	Amplicon sequence variant
AUC	Area under the curve
BSA	Bovine serum albumin
CCA	Canonical correspondence analysis
FAB	Nitrogen-limiting Fabaceae
LB	Lysogeny broth
LR	Logistic regression
L. japonicus	*Lotus japonicus*
L. corniculatus	*Lotus corniculatus*
L. burttii	*Lotus burttii*
MLP	Multi-layer perceptron
M. truncatula	*Medicago truncatula*
NMDS	Nonmetric multidimensional scaling ordination
OTU	Operational taxonomic unit
PMM	*Pseudomonas* minimal medium
PCoA	Principal coordinate analysis
PerMANOVA	Permutational multivariate ANOVA
ROS	Reactive oxygen species
RF	Random forest
SVM	Support vector machine
SynCom	Synthetic community
TS	Tryptone soy
YM	Yeast mannitol

Chapter 1

Introduction

1.1 The microbiome of natural plants: diversity and assembly processes

Higher organisms, including humans, animals, and plants, are closely associated with diverse microorganisms (Berg et al., 2016, Dastogeer et al., 2020). Together, these microorganisms form what is known as their microbiome. This microbiome contains a broad range of organisms, including bacteria, fungi, oomycetes, nematodes, algae, protozoa, viruses, and archaea (Dastogeer et al., 2020, Chaudhry et al., 2021). The plant harbors a microbiome above and below the ground. The microbiome plays a key role in plant growth and production, potentially increasing productivity, nutrient absorption, and resistance to pathogens (Kroll et al., 2017). Therefore, the overall fitness of the plant is affected not only by the plant itself but also by its microbiome. This symbiotic association between the plant and its microbiome is known as the holobiont (Zilber-Rosenberg and Rosenberg, 2008).

To achieve biotechnological goals such as developing plant-protective probiotics, increasing productivity, and modifying ecosystems, it is essential to identify the primary factors that influence the composition and diversity of the plant microbiome and regulate the complex microbe-microbe interactions (Chaudhry et al., 2021, Van Der Heijden et al., 1998). The mechanisms essential for community assembly, diversity, and function can be categorized into stochastic and deterministic processes (Yuan et al., 2019). Stochastic processes (e.g., birth, death and immigration shape the microbial community structure (Sloan et al., 2006, Chen et al., 2019). These processes are challenging due to difficulties defining and measuring stochasticity (Zhou and Ning, 2017). The second critical factor is the deterministic (Yuan et al., 2019). I will first present current knowledge on some key deterministic factors involved in microbiome assembly. They fall into three main categories: host, abiotic (environmental factors), and biotic (microbe-microbe interactions).

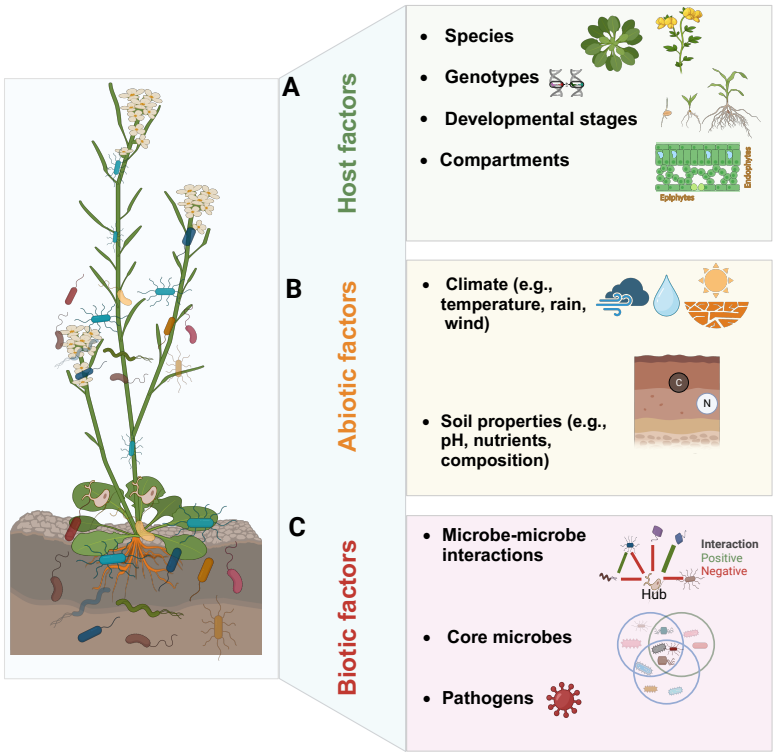


Figure 1.1: Factors that are involved in the assembly of the plant microbiome. (A) Host factors. (B) Abiotic factors. (C) Biotic microbial factors. Figure created with BioRender.com

1.1.1 Host factors that drive microbial community formation

The host has a great influence on its microbiota. These factors include genotype, species, organ, and developmental stage, which have been studied to demonstrate their impact on microbiome formation (Fig. 1A). It has been observed that different plant species can have unique microbiomes. For example, the microbiome composition of maize, sorghum, and wheat crops differs significantly (Dastogeer et al., 2020). The plant model organisms *Arabidopsis* and *Lotus* showed distinct microbial communities in their roots when grown in identical soils (Wippel et al., 2021). In addition, different microbial communities between different genotypes of the four *Tomato* species were observed in another study (Runge et al., 2023). There is also considerable variability in microbial diversity between different plant organs, such as the phyllosphere (above ground), rhizosphere (below ground), and endosphere (inside of the plant) (Gupta et al., 2021). Plants support a rich microbiota in a variety of different compartments and tissues. These include vegetative organs (e.g., roots and leaves). They also include reproductive and dispersal parts such as flowers and seeds. However, in the case of plant endophytes, it is important to consider that the host immune system controls the amount of endophytes and keeps the

most beneficial bacterial density for the plant in all of these organs. The immune system can prepare the plant for a faster and stronger defense response when pathogens attack (Liu et al., 2017). But even so, the pathogens can still outcompete the plant's defense mechanisms and increase in abundance compared to non-pathogenic strains (Compant et al., 2021, Brader et al., 2017). Furthermore, the microbiome undergoes significant changes throughout the host's life cycle, both above and below ground (Hamonts et al., 2018, Almario et al., 2022). It has been reported that seasonal patterns show that the roots of maize plants have a locally distinct microbiome. This microbiome changes as the plant grows and in response to environmental factors (Walters et al., 2018).

1.1.2 Variation in the microbiome of plants as a result of environmental factors

The sensitivity of the plant microbiome to changes in environmental factors can significantly affect many aspects of plants, particularly the phyllosphere (Kadivar and Stapleton, 2003, Vorholt, 2012). Various environmental factors significantly influence microbial communities in different ecosystems, including climatic elements such as precipitation, temperature, drought, geographical influences, and soil properties (Fig. 1B) (Griffiths et al., 2011, Sardans et al., 2008, Trivedi et al., 2022). For example, heat-triggered transposition of a low-copy element in *Arabidopsis thaliana* has been shown to enhance phenotypic diversity and can promote drought tolerance (Thieme et al., 2022). Analysis of climate data and microbiome within populations of *Arabidopsis thaliana* showed drought as a crucial factor affecting microbiome composition (Karasov et al., 2022). In addition, studies have shown that drought perturbs microbial community dynamics in the root microbiomes (Naylor et al., 2017). Furthermore, precipitation patterns are crucial in forming microbial communities in the soil (De Vries et al., 2012). Another important factor is rising temperature that can cause a decline in beneficial microbial populations or increase pathogen transmission (Aydogan et al., 2018, Zhu et al., 2022). Global climate change has the potential to affect the efficiency of agriculture worldwide, with reduced precipitation resulting in lower accessibility of water to crops (Xu et al., 2021, Romero et al., 2022). Additionally, the soil microbiome is critical in cycling carbon and other minerals, helping plants grow, and degrading pollutants (Jansson et al., 2023). The microbiome regulates root growth through the release of hormones like auxins and the absorption of essential nutrients such as iron and nitrogen. Furthermore, nodules are found in certain parts of the root system of leguminous plants, such as *Fabales*. Nodules play an important role in the process of nitrogen fixation. *Rhizobia* bacteria are primarily responsible for nitrogen fixation. This relationship is known as the root-nodule symbiosis (Venado et al., 2020).

1.1.3 Biotic microbial interactions in plant

More than 3.5 billion years ago, microbes first emerged on Earth. This statement refers to a time long before plants inhabited land around 450 million years ago. This shows that, over this vast span of time, microbe-microbe interactions have continued to evolve and diversify. As a result, the formation of plant-associated microbial communities has likely been shaped by both intra- and interkingdom intermicrobial relationships (Hassani et al., 2018). These long-term associations give rise to different types of microbe-microbe and microbe-host interactions, including cooperative, symbiotic, and antagonistic relationships (Chaudhry et al., 2021). Cooperative interactions result from nutrient interdependencies. Metabolites are exchanged between nutrient-dependent microbes to overcome metabolic deficiencies (Morris et al., 2013). This exchange allows microbes to thrive in metabolically poor environments and obtain costly metabolites more efficiently (Wintermute and Silver, 2010). Another interesting form of cooperative interaction is the formation of biofilms by microbial groups that provide protection against environmental stresses and pathogens (Braga et al., 2016). In plant-microbe interactions, symbiotic relationships between plants and nitrogen-fixing bacteria and mycorrhizal fungi enhance plant adaptation to nutrient-poor soils and provide defense against pathogens (Goh et al., 2013). Conversely, competitive interactions result from resource competition and the secretion of antimicrobial compounds. Microbes develop strategies to use limited resources efficiently, while some produce antimicrobial compounds that directly suppress the growth of other microbes (Raaijmakers and Mazzola, 2012, Hassani et al., 2018).

1.2 Using time series data for exploring microbial temporal dynamics

Microbiomes change frequently. This makes it challenging to control ecosystem functioning. Therefore, predicting and identifying the most stable state (Fujita et al., 2023) can be aided by studying microbial dynamics. The phyllosphere is an example of such a dynamic system. A huge bacterial population of up to 10^{26} cells could potentially exist on the vast leaf surface, which is about twice the size of the land surface (Vorholt, 2012). Leaf microbial communities are constantly exposed to the arrival of new microbes transported by wind and water. They exhibit a high degree of stochasticity, i.e., high unpredictability and variability. Moreover, leaf communities have been shown to change over time, reaching different stable states (Copeland et al., 2015). The dynamics of microbiome assembly in leaves (Meyer et al., 2022, Copeland et al., 2015, Rastogi et al., 2012, Grady et al., 2019) and roots (Stopnisek and Shade, 2021) have been followed in recent studies. Temporal dynamics of leaf microbiomes using correlation networks over the growing cycle of a plant have been the source of observable patterns. In particular, temporal trends are visible, and the communities and networks go through a stabilization phase, with reduced diversity and variability at the beginning of the growing season

(Agler et al., 2016, Almario et al., 2022). By capturing temporal variation within the microbiome, longitudinal studies can address several key questions: How does the microbiome change and vary over time? How do microbial networks develop over time? Within these microbial networks, which microbes are acting as the driving force? How does the stability of the microbial networks change over time? To address these questions, I will first discuss the concept of stable components within a microbial time series. I will then discuss methods for analyzing microbial interaction networks, identifying key microbial drivers, and defining the stability of microbial networks.

1.2.1 Identifying the core microbiome and approaches for characterization

Due to the complex nature of microbial communities, determining the ecological role and importance of individual taxa is a significant challenge (Almario et al., 2022). However, to identify functional microorganisms of importance within leaf communities, concepts based on the persistence of specific microbes (known as core taxa) and their centrality within microbial networks (known as hub taxa) have been used (Fig. 1C) (Agler et al., 2016, Carlström et al., 2019). The search for a core microbiome has become increasingly important in microbial ecology. A core microbiome is a set of microbial species persistently found in a specific environment or in association with a certain host (Wang et al., 2023a). Although the term core microbiome is popular and gaining increased application, there is still limited consensus regarding its practical quantification. In practice, the range of spatial and temporal scales at which a core microbiome can be quantified varies, from single-site samples to global sampling of a given host category and from single to multiple sampling over days to years. One of three methods is commonly used to assess core microbiomes: Examining different samples from the same host or environment for the presence of microbial taxa. Within these samples, analyze the relative abundance of microbial taxa or use a combination of these metrics. These measurements can be obtained from standard operational taxonomic units (OTUs), which are sequences grouped at 97-99% sequence similarity, or amplicon sequence variants (ASVs), which are sequences grouped at 100% sequence similarity and include the number of occurrences of each OTU/ASV in individual samples (Neu et al., 2021). To study the critical role of core microbes in improving plant health, a synthetic microbial community (SynCom) could be created. These communities would consist of microbes with well-defined taxonomic status and functional traits, allowing for comprehensive testing and analysis (Wang et al., 2023b). SynCom serves as a model system to study microbial community performance and stability. The construction of SynCom enables the generation of controlled systems of reduced complexity, whereas natural microbial communities comprise a range of microbes, often characterized by unknown functions (Großkopf and Soyer, 2014).

1.2.2 Inferring and interpreting microbial interaction networks

Microbial networks are a popular tool for studying community interactions. They provide a broad picture of ecosystem behavior at the system level (Röttjers and Faust, 2018). Constructing correlation-based networks is a powerful strategy for transforming microbiome composition data into microbial interactions. In such networks, microbial taxa are represented by nodes connected by edges (Kishore et al., 2023). Statistical significance calculations for microbial correlations often use null models. One approach is to construct a null model by permuting the count matrix. The correlations are then recalculated and p-values are calculated against this null model using the permuted association matrices. This process helps determine the significance of the inferred correlations (Röttjers and Faust, 2018). Furthermore, to keep the pairwise correlations, more stringent thresholds for corrected p-values, such as 0.001, would be more appropriate for preserving boundaries. It is clear that to achieve high accuracy detection with most tools, the conventional corrected p-value of 0.05 is too low. In addition, factors such as sequencing technology and data normalization significantly impact the identification of correlations (Weiss et al., 2016). Although correlation-based approaches can help construct microbial interaction networks, challenges associated with microbiome data, such as compositionality and sparsity, should be considered when calculating correlations. Sequencing microbiome data yields compositional results. The counts represent the fraction of counts per characteristic, such as OTUs or genes, per sample, multiplied by the sequencing depth. Therefore, only relative abundances are accessible (Gloor et al., 2017). New methods have emerged to address the compositionality of microbiome data for the inference of robust networks. For example, the SparCC (Friedman and Alm, 2012) and SPieCeasi (Kurtz et al., 2015) approach make it possible to estimate correlations from compositional data; the sparse data matrix is assumed for both (Friedman and Alm, 2012).

In interpreting and analyzing a constructed network, one of the approaches to finding the important microbes in the network is hub identification. Hubs are species that have a high number of connections to other individuals, giving them a high degree of centrality. However, the ecological role of hub species remains unclear (Röttjers and Faust, 2018). Several studies have shown that highly connected species act as mediators between plants and microbiomes. These mediators are critical for maintaining plant health. In a study focused on *Arabidopsis* leaves, host genetics related to microbial hubs contributed significantly to the variation in seed production. One of these microbial hubs was successfully cultured, and its effect on plant growth under sterile conditions was confirmed. (Brachi et al., 2022). Another study on root-associated fungi in terrestrial plants suggests that network hubs in symbiont-symbiont networks can organize the overall dynamics of the root-microbiome (Toju et al., 2016). Microbial hubs can benefit the entire system by recruiting helpful organisms or suppressing pathogen invasion to increase overall fitness. For example, the obligate biotrophic oomycete pathogen *Albugo* has been studied as a hub microbe with significant effects on both epiphytic and endophytic bacterial colo-

nization. *Albugo* infection significantly reduces alpha diversity, highlighting its impact on the configuration of microbial communities (Van Der Heijden and Hartmann, 2016, Agler et al., 2016). The results demonstrate microbial networks' complex and significant role in shaping microbiomes. They can have direct application to ecosystem health and stability.

1.3 Digging into the characterization of microbes impacting plant health

The health of a plant is closely linked to the presence and diversity of co-existing microbes. This can be evidenced by comparing the microbial diversity in healthy and diseased plants (Sweet and Bulling, 2017). Developing new techniques for better analysis of the microbiome results from sequencing can help predict the microbiome that can help to cope with disease and improve plant health and productivity (Yuan et al., 2020). Among those techniques, machine learning has become an essential tool in microbiome research with many applications. These applications include predicting host and environmental phenotypes, classifying microbial properties, investigating interactions within the microbiome, and tracking changing microbiome composition. The successful implementation of these methods has been thoroughly reported (Hernández Medina et al., 2022). Here, I look at machine learning techniques and their significant impact on microbiome research.

1.3.1 Machine learning techniques and their applications in microbiome research

Advances in sequencing technology have made it possible for microbiome studies to be conducted with large sample sizes. These large datasets allow the use of machine-learning techniques to investigate associations between microbiomes and multiple traits (Namkung, 2020). The use of microbiome information to diagnose disease is becoming increasingly popular. However, due to the highly variable architecture of the microbiome among individuals, identifying the bacterial populations associated with disease using traditional statistical models is challenging (Topçuoğlu et al., 2020). Machine learning algorithms encompass a variety of computational and statistical techniques for analyzing data to construct and enhance predictive models (Hernández Medina et al., 2022).

Machine learning methods can be broadly divided into two categories based on their approach to modeling: unsupervised learning and supervised learning. Unsupervised learning serves two main purposes in microbiome research: grouping data points according to pairwise similarity measures, known as clustering, or deriving representative features from various variables, known as dimension reduction. Particularly in biology, and especially for genetic data, hierarchical clustering (Namkung et al., 2016) and K-means clustering (Bhalla et al., 2017) have been widely applied. For instance, hi-

erarchical clustering produces dendrogram-like structures that group samples based on similarities.

To predict disease risk, supervised learning uses microbiomes as biomarkers associated with an increased risk of certain diseases, such as cancer or diabetes. These supervised learning methods can be categorized based on the types of trait values they predict, specifically as classifying or regressing. Classification is the prediction of the category to which a sample will be assigned, while regression is the prediction of continuous numbers. Among the many supervised learning methods, some of the most commonly used ones are support vector machine and random forest (Namkung, 2020). For example, classifiers have revealed microbial community patterns in soils affected by *Fusarium wilt* under field conditions (Yuan et al., 2020). Similarly, by identifying specific taxa involved in nitrogen utilization, random forest machine learning accurately predicted productivity based on microbiome composition at the order level (Chang et al., 2017). In studying the nodule microbiome in healthy and starved plants, we investigated how it interacts with nodulation, a process influenced by the soil microbiome. Metabarcoding techniques were used to profile nodule communities. Bacteria of interest were identified using machine learning and network analysis (Crosbie et al., 2022). Another study revealed the composition of bacterial communities in disease-suppressed soils using a combination of metadata analysis and machine learning. *Firmicutes* and *Actinobacteria* increased significantly in disease-suppressed soils. At the same time, *Proteobacteria* and *Bacteroidetes* decreased. Using machine learning to identify potentially valuable biomarker genera, the models effectively discriminated between disease-suppressed and disease-supportive soils (Zhang et al., 2022).

1.4 Aim of the thesis and objectives

For my Ph.D., I aimed to improve the understanding of multi-kingdom microbial interactions associated with plants. I collected environmental plant microbiomes and used various computational methods to analyze them. I aimed to analyze the microbiome and identify patterns that may emerge under varying environmental conditions. Subsequently, I conducted controlled experiments to investigate the microbiome analysis results and aimed to improve plant health. It is known that different factors are involved in forming the phyllosphere microbiome (Meyer et al., 2022, Bowsher et al., 2021). However, studies on multi-kingdom surveys in long-term studies are limited (Almario et al., 2022).

The first objective of my project was to answer how different abiotic and biotic factors are involved in the formation of the phyllosphere microbiome in the long term. To study this objective, which is further discussed in Chapter 2, my colleagues and I collected the microbiome data from *Arabidopsis thalina* model plant in different time points from the natural ecosystem. First, I examined the ecological abiotic forces forming the leaf microbiome. In the face of global change, plant species are exposed to extreme and often unpredictable environmental stresses such as temperature and precipitation. Understanding the relationship between these factors and the plant microbiome can provide guidance for helping plants adapt to these stressors (Karasov et al., 2022, Desaint et al., 2021). Toward this direction, I examined the influence of different environmental factors on microbial interactions.

While most leaf microbes are highly variable across environments and time periods, certain microorganisms are consistent in plant populations (Karasov et al., 2018). Previous studies have used concepts based on a microbe's persistence core taxa and/or its importance in microbial networks as hub taxa to identify key microorganisms in leaf communities (Agler et al., 2016, Carlström et al., 2019). Therefore, I turned to Chapter 3 of my thesis to examine the temporal dynamics of leaf microbiomes over the growing season of the plant. Furthermore, it is important to determine which microbial groups persist during the growing season of plants in microbial interaction networks. To achieve this objective, we designed a common garden experiment and collected microbiome samples from November to March across various stages of *Arabidopsis* growth. I used microbial co-abundance networks to study the interactions and variability of microorganisms over time. Subsequently, I analyzed stable core microbes to see how microbial interactions affect stability.

Plant health is highly connected to its microbiome. Understanding the relationships between microbes is necessary to uncover potential mechanisms that protect plants from pathogens. *Albugo laibachii*, an oomycete pathogen of natural populations of *Arabidopsis thaliana*. Identified in my analysis in Chapter 3 as potential core and hub taxa (Agler et al., 2016). I investigated whether the microbiome's structure changes in the presence of this pathogen. To achieve this objective, I used various machine learning predictive models to differentiate between infected and uninfected plants. The study aimed to deter-

mine whether the microbiome suppresses pathogens that affect plant health. The analysis results were tested in a controlled laboratory condition by performing pairwise interactions between the microbiome associated with plant health and the pathogen; this part covers the Chapter 4 of this thesis.

The plant microbiome is also critical for making nutrients (such as nitrogen) available to the plant. Legumes have evolved a mutualistic interaction with nitrogen-fixing *Rhizobia*. The bacteria are harbored and fed in root organs called nodules in exchange for ammonia (Crosbie et al., 2022). In Chapter 5, we explored the research question of which microbes are involved in effective nodulation processes and thus contribute to better nitrogen fixation for plant health. I used a machine learning model to compare microbiome data from healthy and starved *Lotus* plants. Our goal was to identify the microbes that influence these predictions. In addition, we focused on studying the role of the discriminating microbes in an experimental investigation.

In summary, my study in this thesis used various microbiome data from different environmental conditions to explore the role of biotic and abiotic factors in shaping the plant microbiome. I investigated the use of microbes for plant health. These findings are essential in designing complex probiotics that protect microbial communities from dysbiosis.

Chapter 2

Environmental factors shape the leaf microbiome of *Arabidopsis*

- **Status** Published
- **URL** https://doi.org/10.1093/ismeco/ycae103
- **Own contribution**

  Designed research (with co-authors)

  Performed the experiments (with co-authors)

  Interpreted the data (with co-authors)

  Analyzed and visualized the data

  Wrote the first draft of the manuscript

  Final editing of the manuscript (with co-authors)

Seasonality adaptation patterns of the natural *Arabidopsis* leaf microbiome over several plant generations are shaped by environmental factors

Maryam Mahmoudi[1], Juliana Almario[2], Katrina Lutap[1], Kay Nieselt[3] and Eric Kemen[1]

[1]Microbial Interactions in Plant Ecosystems, IMIT/ZMBP, Eberhard Karls University of Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany

[2]Université Claude Bernard Lyon 1, CNRS, INRA, VetAgro Sup, UMR5557 Ecologie Microbienne, F-69622 Villeurbanne, France

[3]Institute for Bioinformatics and Medical Informatics, Eberhard Karls University of Tübingen, Sand 14, 72076, Tübingen, Germany

Corresponding author:
Eric Kemen
eric.kemen@uni-tuebingen.de
Current address: Microbial Interactions in Plant Ecosystems, IMIT/ZMBP, Eberhard Karls University of Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany

Abstract

Leaf-associated microbial communities can promote plant health and resistance to biotic and abiotic stresses. However, the importance of environmental cues in the assembly of the leaf endo- and epi-microbiota remains elusive. Here we aimed to investigate the impact of seasonal environmental variations, on the establishment of the leaf microbiome, focusing on long-term changes (five years) in bacterial, fungal, and non-fungal eukaryotic communities colonizing the surface and endosphere of six wild *Arabidopsis thaliana* populations. While leaf-microbial communities were found to be highly stochastic, the leaf niche had a predominant importance with endophytic microbial communities consistently exhibiting a lower diversity and variability. Furthermore, our analyses revealed that among environmental factors, radiation and humidity-related factors are the most important drivers of diversity patterns in the leaf, albeit with stronger effects on epiphytic communities. Using linear models, we further identified 30 important genera whose relative abundance in leaf compartments could be modeled from environmental variables, suggesting specific niche preferences for these taxa. With the hypothesis that these environmental factors could impact interactions within microbial communities, we analyzed the seasonal patterns of microbial interaction networks across leaf compartments. We showed that epiphytic networks are more complex than endophytic, and that the complexity and connectivity of these networks are partially correlated with the mentioned environmental cues. Our results indicate that humidity and solar radiation function as major environmental cues shaping the phyllosphere microbiome at the micro-scale (leaf compartment) and macro-

scale (site). These findings could have practical implications for selecting and developing field-adapted microbes in the face of, and for predicting microbial invasions in response to global change.

Keywords: Leaf microbiome, leaf compartments, microbial network, environmental factors.

Introduction

Leaves are colonized by various microbes including bacteria, fungi, oomycetes, and protists [1]. This leaf microbiota can play a beneficial role in protecting plants against biotic and abiotic stressors, thus ultimately promoting plant growth and fitness [2–6]. Within the microcosm of the leaf, distinct compartments emerge, primarily characterized as epiphytic (surface) and endophytic (internal tissues). Despite their proximity, these zones have different characteristics. The leaf surface is covered by a hydrophobic cuticle layer that prevents water loss from the leaf surface. This environment comprises components like wax and cutin [7], along with trichomes, which can protect against UV light and mediate leaf temperature [8]. Whereas, within the leaf interior, a vast area known as the apoplast that facilitates gas and water exchanges for photosynthesis. This environment, with its higher humidity, is also subjected to microbial colonization [9]. In addition, the apoplast can be subject to pH fluctuations in response to biotic stresses such as pathogen attack and abiotic factors such as salinity or drought [10]. Due to their different characteristics, these two niches may favor certain microbes and make it difficult for others to survive. In this context, it is not well known how environmental cues differentially shape the microbial communities occupying these two niches.

Environmental factors (e.g., light and humidity) significantly affect all microbial communities in different ecosystems and the phyllosphere [11, 12]. Light, as a fundamental element of living organisms, is essential not only for plant photosynthesis and growth but it also mediates plant-microbe interactions. The presence of photoreceptor proteins in microorganisms enables them to detect light to use for adhesion to host tissues for colonization, and for DNA repair [13]. In particular, UV light can enhance plant defense mechanisms by stimulating the production of defense-related compounds such as salicylic acid and jasmonic acid, which strengthen plants against pathogens [14]. In addition, the availability of water and nutrients is critical for plant health and ecosystem balance [15]. Precipitation has been shown to play a significant role as a primary driver in shaping fungal communities and facilitating the spread of fungal plant pathogens via rain droplets [16, 17]. Accordingly, precipitation has a significant effect on the composition of soil microbial communities [18]. Although some research has investigated the effects of environmental factors on leaf microbiomes, there remains a lack of studies focusing on the effects of such factors on different microbial communities within leaf compartments.

Microbes often interact with each other through various relationship types, such as mutualism or antagonism, and can develop complex plant-associated communities that can change throughout the growing season of the host [19, 20]. The use of microbial interaction network analysis has been useful in understanding the variability and stability of these communities under changing environments [21–23]. For example, the complexity of

microbial networks has been linked to community stability, as seen in a long-term study of grassland soil microbiome, showing that warming increases the complexity of microbial network (e.g., size and connectivity) [24]. Environmental stresses (decreasing water availability, nutrients, and vegetation) have been shown to have impact the stability of microbial communities by decreasing richness and altering the ratio of positive and negative interactions [25]. However, how plant-associated microbial communities respond to changing environmental conditions has rarely been studied using multi-kingdom microbial interaction networks.

Microbiota associated with plants and animals undergo seasonal fluctuations shaped by environmental cues and perturbations [26, 27]. In plants, research has shown that the microbiome tends to become more tissue-specific throughout the host developmental stages [28], but these communities remain highly stochastic. For example, tracking the leaf microbiome of *Arabidopsis thaliana* during its growing season from November to March in common garden experiments revealed overall high variability with some conserved patterns [29]. These conserved patterns were characterized by identifying more persistent microbes known as core microbes. Among these core microbes, some members of plant pathogens, such as *Peronosporales*, increased throughout the growing season, reaching maximum values in March which aligns with the disease dynamics of downy mildew in *Brassicaceae*, known to be favored by cold, wet weather [29]. However, despite the recognized importance of longitudinal microbiome data, few studies have used environmental data to explain observed temporal dynamics in the plant-associated microbiota.

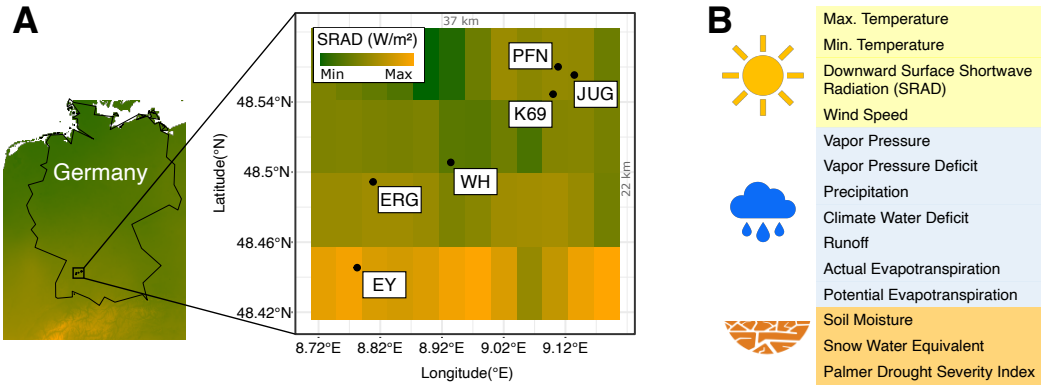
The objective of this research was to link temporal changes in the leaf microbiome of natural *Arabidopsis thaliana* populations, with naturally occurring environmental factors, in a long-term study. We hypothesized that leaf-associated microbiomes occupying the epiphytic or the endophytic compartment would respond differently to environmental cues. Using amplicon sequencing, we tracked leaf microbial communities (bacteria, fungi, and non-fungal eukaryotes) during fall and spring seasons, over five consecutive years. Our results revealed that while many environmental factors shaped these communities, the leaf niche emerged as the most important factor. Endo- and epiphytic microbial communities exhibited distinct responses to environmental cues, with radiation- and humidity-related factors appearing to have a greater influence on the diversity and structure of these communities. We further identified 30 microbial taxa showing distinct responses to certain environmental cues, suggesting some level of niche preference among leaf-colonizing taxa. By examining microbial interaction networks between epiphytic and endophytic communities, we further found that community cohesion, a measurement of connectivity, could be correlated with specific environmental factors suggesting that certain environmental cues can drive community stability.

Results

Leaf epiphytic and endophytic microbial communities differ in diversity and structure

With the aim to study the impact of environmental cues on the temporal dynamics of the leaf endo- and epi-phytic microbiota, we collected samples from six locations (sites) with stable *A. thaliana* populations in the proximity to Tu¨bingen (south Germany)(Fig. 1A) [30] over two seasons (fall and spring), and five years (Fig. 1C; see Table. S1). We then correlated

changes in the leaf microbiota with environmental variables measured locally (14 environmental factors with monthly resolution, Table. S2). Fall sampling covered the early growth phase of *A. thaliana* while Spring sampling covered the early reproductive stage. From each sample, we recovered epiphytic and endophytic microbial communities, extracted genomic DNA, and performed bacterial 16S rRNA, fungal ITS2, and eukaryotic 18S rRNA amplicon sequencing (Fig. 1D), as described in [30]. In the analysis of the 18S eukaryotic data, all microbes that belonged to the kingdom fungi were excluded. We refer to this data as non-fungal eukaryotes (NFEuk).



C

Year	Season	Sites					
		ERG	EY	JUG	K69	PFN	WH
2014	Spring	4	6	6	0	3	6
	Fall	8	7	7	0	3	6
2015	Spring	6	5	6	3	3	6
	Fall	6	6	5	4	6	6
2016	Spring	6	6	6	6	5	6
	Fall	6	5	6	6	1	6
2017	Spring	6	6	8	6	6	8
	Fall	6	9	6	5	0	8
2018	Spring	6	8	5	4	4	6
	Fall	5	6	6	0	5	6
2019	Spring	4	8	4	4	4	8

Total: 351 plants

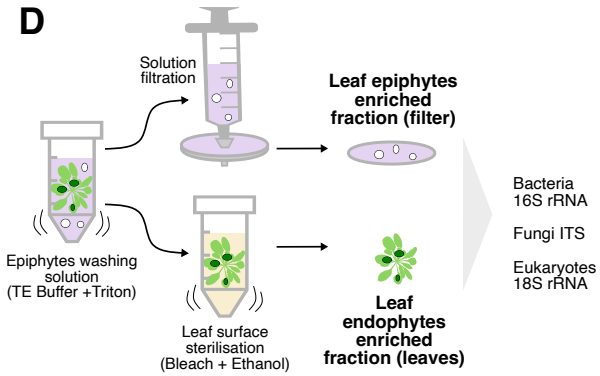


Figure 1. Microbial community collection in natural *A. thaliana* populations over time. (A) Map showing the six sampling locations of natural *A. thaliana* in southern Germany near Tübingen [30]. The heatmap on the map represents average variation in solar radiation of sampling locations (downward surface shortwave radiation (srad)). **(B)** Environmental variables (14; Table. S2) used in this study were obtained from the TerraClimate database [47]. **(C)** Plants (n = 351) were collected in the fall and spring of five consecutive years (starting spring 2014, ending spring 2019, 11 time points). **(D)** Leaf epiphytic and endophytic fractions collected from each sampled rosette (Table. S1). Microbiome analysis was conducted via Illumina-based amplicon sequencing (Miseq 2×300 bases). Taxonomic markers

included the bacterial 16S rRNA V5-V7 region, fungal ITS2, and 18S rRNA V9 region of eukaryotes.

To investigate the effect of the ”compartment” (endophytic vs epiphytic fractions; Fig. 1), the ”site” and the ”season” on leaf-associated microbial communities, we conducted multiple diversity analyses. Permutational multi-variate analysis of variance (PERMANOVA) results show that the leaf ”compartment” emerges as the primary driver of the structure of microbial communities (Fig. 2A). In particular, it exerts a major influence on the structure of bacterial (8.4% explained variance) and non-fungal eukaryotic communities (11.8%). Fungal communities appeared much less constrained by the leaf compartment (2.3% explained variance) and were more influenced by the sampling site (5.7%). These analyses further revealed a marginal effect of the ”season” (explaining 0.8% to 3% of the variance). Accordingly, non-metric multidimensional scaling (NMDS) plots showed a clear separation between epiphytic and endophytic samples for bacterial and non-fungal communities, while fungal communities exhibited the smallest separation (Fig. 2B).

Alpha-diversity measures (Shannon’s index, related to the number of taxa in the community) show that leaf endophytic communities were 1.6 to 2.6 times less diverse than their epiphytic counterparts (Fig. 2C). Additionally, the diversity of the endophytic communities appeared less variable between seasons, than that of epiphytic communities which showed significant changes between seasons (Fig. S1A). A notable exception were endophytic bacterial communities which were significantly more diverse in the fall than in the spring (Wilcoxon’s test, $P < 0.0001$) (Fig. S1A).

Differences in epiphytic and endophytic communities were associated with the enrichment of major microbial orders (overall most abundant orders; Fig. 2D). Among bacteria, *Rhizobiales* and *Flavobacteriales* were more abundant among epiphytes (1.1 times and 1.2 times, respectively), while *Burkholderiales* and *Pseudomonadales* were more abundant among endophytes (1.9 and 5 times more, respectively). Among fungi, *Tremellales* basidiomycetes and *Capnodiales* ascomycetes were more abundant in the epiphytic fraction (1.2 times and 1.4 times, respectively), while ascomycetes from the *Helotiales* were enriched in the endophytic fraction (1.1 times). In addition, non-fungal eukaryotic orders enriched in the epiphytic compartment included green algae *Watanabeales* and *Klebsormidiales*, as well as the cercozoan *Cryomonadida* (3.5 times, 7.8 times and 2.9 times, respectively). Surprisingly, green algae from the *Chlamydomonadales* were 1.3 times more abundant in endophytes. Finally, *Albuginales*, known to harbor the plant biotrophic pathogen *Albugo*, were 6.4 times more abundant among endophytes. These results illustrate the extent of preference that major leaf-associated microbes have for either the epiphytic niche or the endophytic niche.

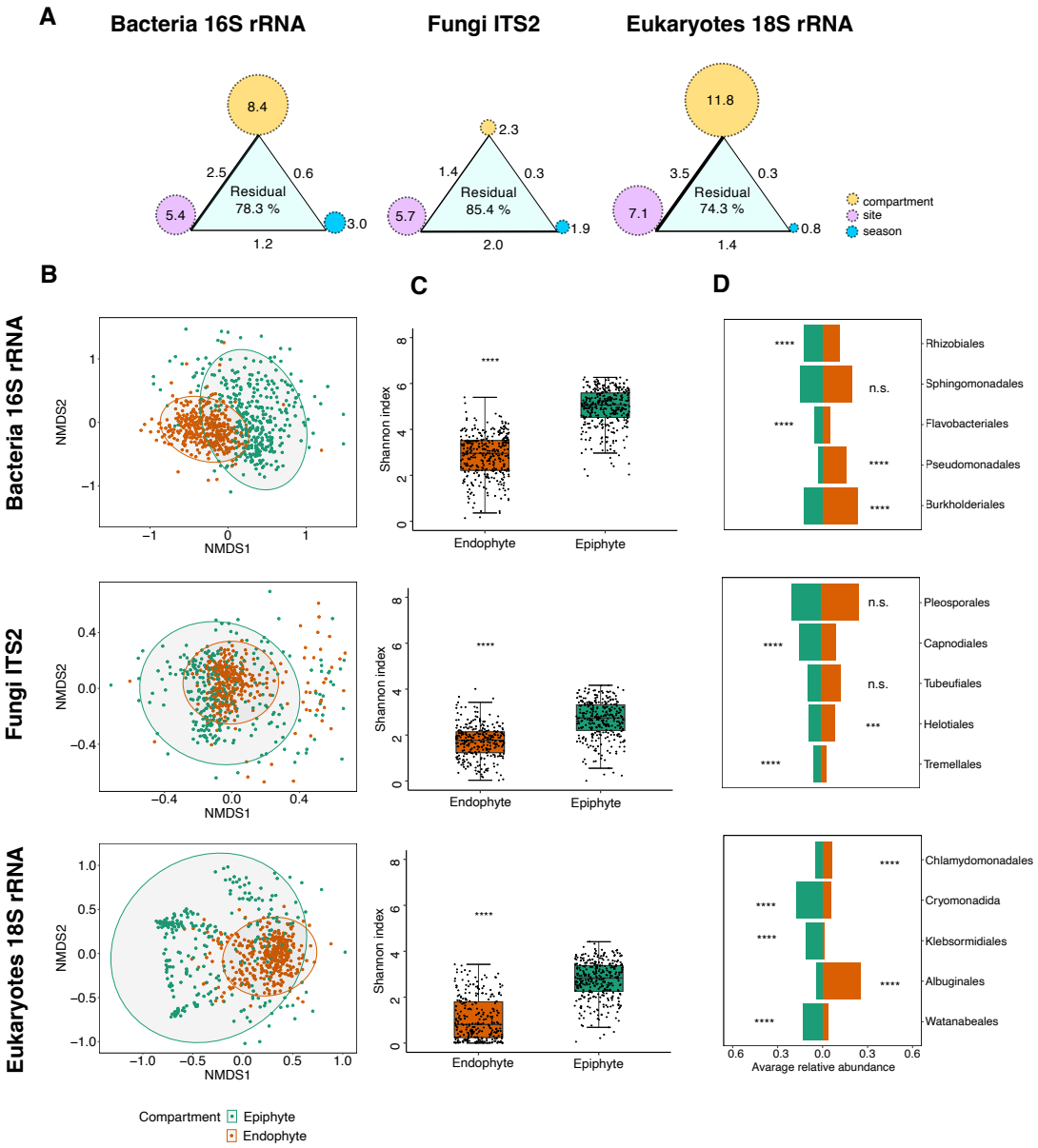


Figure 2. Multivariate analysis on factors structuring leaf communities. (A) A PERMANOVA analysis on Bray-Curtis dissimilarities. Circles depict the percentage of variance explained by factors 'compartment', 'site', and 'season'; connecting lines depict the percentage of variance explained by interactions between factors, and the thickness of lines shows the strength of explained variation. Only significant effects are shown (permutations 10,000, $P \leq 0.05$, explanatory categorical variables: compartment x site x season). **(B)** Non-metric multidimensional scaling ordination (NMDS) analysis of epiphytic and endophytic samples measured by Bray-Curtis dissimilarities in bacterial, fungal, and non-fungal eukaryotes. **(C)** Alpha-diversity measures (Shannon's H index) of epiphyte and endophyte samples. The box plots display individual samples as dots. **(D)** Average relative abundances of the top five most abundant microbial orders in epiphytic and endophytic samples. Asterisks indicate significant

differences based on Wilcoxon's test: n.s. ($P > 0.05$), * ($P \leq 0.05$), ** ($P \leq 0.01$), *** ($P \leq 0.001$), and **** ($P \leq 0.0001$).

Endophytic and epiphytic microbial communities respond differently to environmental cues

We hypothesized that the major differences observed between endo- and epiphytic communities are partially explained by the fact that these communities respond differently to major environmental cues. To test this hypothesis, we evaluated the effect of 14 selected environmental factors on community structure (PERMANOVA on Bray-Curtis dissimilarities) and alpha-diversity (correlation of environmental factors with communities alpha-diversity), in each of these niches. The fourteen environmental variables selected (Fig. 1B and Table. S2) showed variability across seasons, years and/or sampling sites (Fig. S2). While all environmental factors significantly impacted the structure of leafassociated microbial communities, marginal effects were observed for most of the factors with very low percentages of variance explained (0.5 to 1.9, $P < 0.05$) (Fig. 3A, Table. S3). Notably, bacterial communities were more affected by solar radiation and humidityassociated factors like vapor pressure, precipitation and evapotranspiration (actual and potential) than micro-eukaryotic communities (fungal and non-fungal) (Fig. 3A). More pronounced effects were found when considering the correlations between these environmental factors and per−sample microbial alpha-diversity. While factors associated with temperature (minimum and maximum) and high humidity (vapor pressure, precipitation, and soil moisture) had overall positive effects on microbial alpha-diversity, solar radiation had an overall negative effect. A notable exception was that solar radiation was positively correlated to higher fungal diversity, specifically on the leaf surface (epiphytic fraction) (Fig. 3B). This suggests that increased radiation levels may stimulate the growth or proliferation of certain fungal species adapted to thrive under such conditions.

Interestingly, certain factors had stronger effects on one niche or one microbial group. For example, in comparison to their epiphytic counterparts, the alpha-diversity of endophytic micro-eukaryotic communities (fungal and non-fungal), showed overall fewer significant correlations with the analyzed environmental factors (5 vs 17 significant correlations). This aligns with the previous observation that micro-eukaryotic alpha diversity is less variable inside the leaf (endophytic) than on the leaf surface (Fig. 2C and Fig. S1), suggesting that these communities are more resistant and/or resilient to environmental perturbations. Interestingly, these trends did not hold for bacterial alpha-diversity which correlated with several factors both for endophytic and epiphytic communities (10 vs 7 significant correlations). Some factors associated with water loss from the plant (wind speed and actual/potential evapotranspiration) had contrasting effects on microbial diversity depending on the niche considered, with negative effects on endophytic diversity (fungal and/or bacterial communities) and positive effects on epiphytic diversity (bacterial and/or micro-eukaryotic communities). Taken together these results suggest that environmental factors influence microbial communities differently depending on their habitat and their broad phylogenetic group.

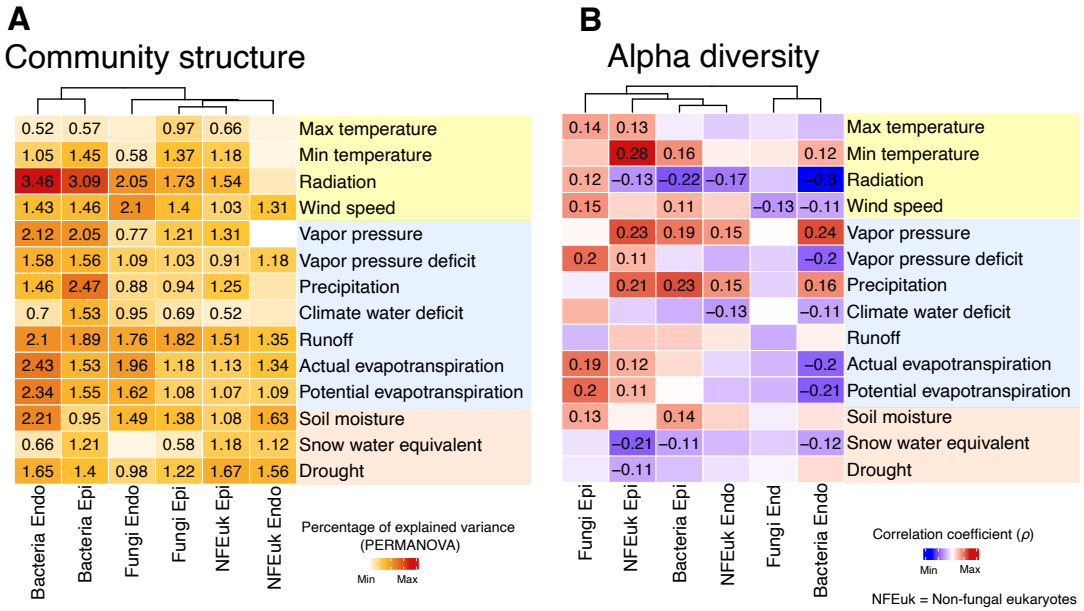


Figure 3. Effects of environmental factors on microbial community structure and alpha diversity. (A) Heatmap displaying explained variance in PERMANOVA models (R^2 value; BrayCurtis dissimilarities) indicating the influence of individual environmental factors on microbial community structure, within each leaf compartment. **(B)** Heatmap showing Spearman correlation coefficients (ρ) between alpha-diversity (Shannon's H index) and environmental factors, in each compartment. Only significant results are displayed ($P < 0.05$).

The abundance of specific taxa in different leaf compartments can be inferred from certain environmental data

Further analyses aimed to assess the impact of key environmental factors on the relative abundance of major microbial genera. To this end we used a GLM approach (generalized linear models) to assess the response of selected genera, in each leaf compartment. These analyses revealed significant effects for at least one environmental factoron the relative abundance of most taxa: 91% of the bacterial genera, 85% of the fungal genera and 86% of the non-fungal genera (FDR-corrected $P < 0.05$; Table. S4). When examining the 30 most responsive genera shared between epiphytic and endophytic compartments (those with the highest average coefficient values in the GLMs)), we found that they were mainly impacted by precipitation, soil moisture, maximum temperature, drought, radiation, and vapor pressure. Yet, no single factor was significant for all these taxa (Fig. 4). In this subset, the relative abundance of the considered bacterial genera on the leaf surface (epiphytes) was more often impacted by the considered factors than their endophytic counterparts, but the effects were marginal (coefficients below 0.01). Overall, radiation had a positive effect on the relative abundance of most of these taxa in the epiphytic compartment, while factors associated to high humidity (precipitation, vapor pressure and potential evapotranspiration) had mixed results with negative effects on *Sphingomonas* and positive effects on *Flavobacterium*. For endophytic bacteria, the strongest effects were observed for

Pseudomonas relative abundance which increased with higher radiation and lower humidity (lower soil moisture and lower potential evapotranspiration), and *Sphingomonas* relative abundance which increased with higher humidity (higher soil moisture and lower wind speed). These differential responses are probably associated with different niche preferences for these taxa.

Like bacteria, the relative abundance of major fungal taxa on the leaf surface (epiphytes) was more frequently influenced by the considered factors than their endophytic counterparts. However, the effects observed were mostly marginal (coefficients below 0.01). Notably, radiation and precipitation yielded mixed results. High radiation and humidity (precipitation) negatively impacted the abundance of fungal taxa *Cladosporium*, *Boeremia*, and *Vishniacozyma*, while increasing the abundance of *Titaea* (*Tetracladium*), a *Helotiales* fungus which is known to thrive in water environments [31].

Similarly, the relative abundance of major non-fungal eukaryotic taxa on the leaf surface (epiphytes) was more responsive to environmental factors than their endophytic counterparts, albeit with mostly marginal effects (coefficients below 0.01). Overall, temperature and precipitation had a positive impact on the relative abundance of most taxa in the epiphytic compartment. Specifically, precipitation increased the relative abundance of green algae *Chloridium* and *Klebsormidium* in the leaf surface (epiphytic), which is in line with the fact that these organisms proliferate in light-exposed high-humidity environments. The most striking results were observed with the pathogenic biotroph *Albugo* (oomycete) whose abundance inside the leaf was negatively impacted by high humidity indicators (high potential evapotranspiration, low wind speed) and promoted by high maximum temperatures, suggesting this pathogen invades the leaf under dry heat conditions.

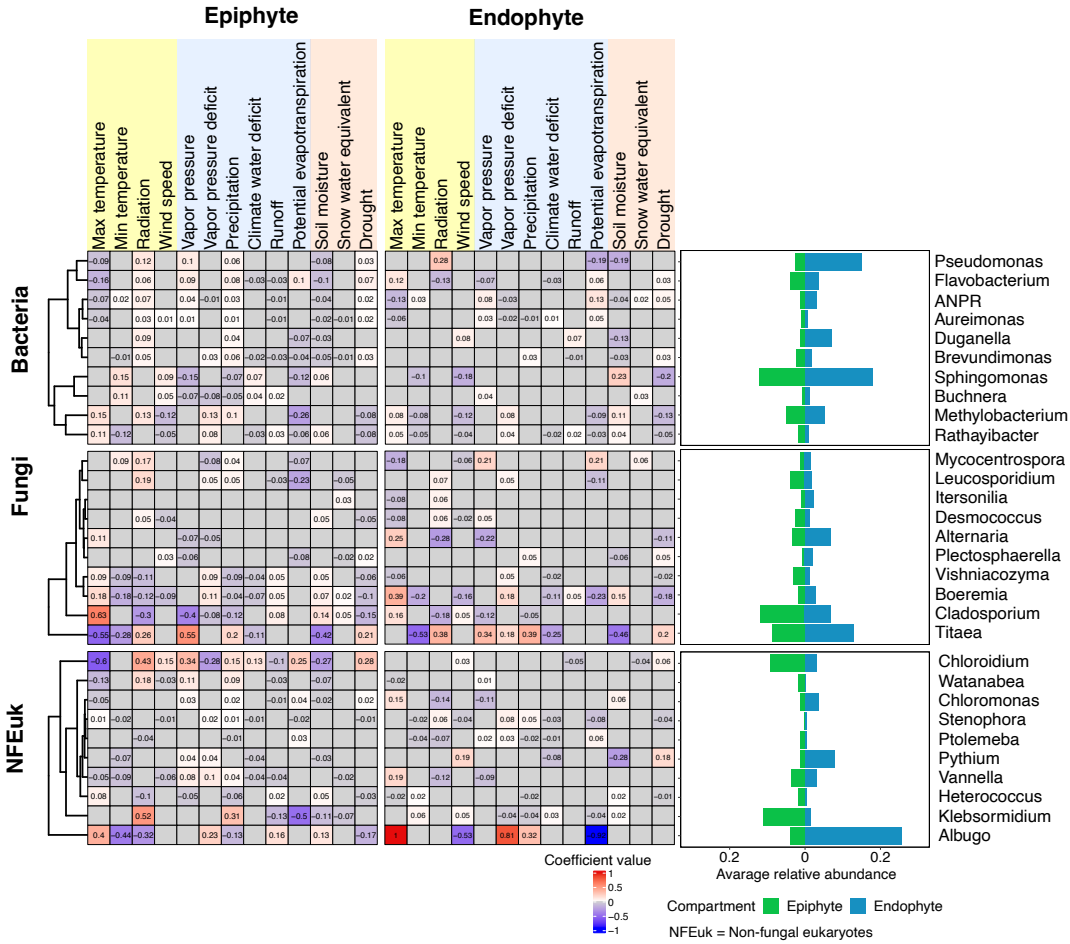


Figure 4. Association between environmental factors and relative abundance of microbial genera among epiphytic and endophytic compartments. (A) The values of the heatmaps show coefficient values of environmental factors in modeling the relative abundance of microbial genera using GLM. Negative values indicate genera that decrease with these environmental factors, while positive values indicate genera that increase. Only significant relations ($P < 0.05$, FDR-corrected) are displayed. (B) The histograms display the average relative abundances of selected microbial genera in each compartment. **(B)** Average relative abundance of selected genera present in epiphytic and endophytic compartments, and highly influenced by environmental factors (highest average coefficient values).

Microbial networks and community cohesion are driven by major environmental cues

We conducted a microbial network analysis to explore changes in the interactions among microbes in the epiphytic and endophytic compartment, aiming to assess the impact of environmental factors on the connectivity of these communities. Microbial networks were constructed for each of the eleven sampling times, and a comparative examination was carried out between the epiphytic and endophytic networks (Fig. 5A). It is worth noting that the epiphytic network had a greater complexity than the endophytic network with a larger

number of nodes (OTUs) and edges (correlations between taxa). On average, the epiphytic compartment contained 15.1 times more nodes and 79.7 times more edges (559 nodes and 3348 edges) than the endophytic compartment (37 nodes and 42 edges).

Further analyses were conducted to investigate the potential correlation between the complexity of microbial interaction networks (number of nodes and connectivity) and environmental factors. The findings revealed that from the 14 factors analyzed only two factors: precipitation and soil moisture, significantly correlated with the number of nodes in epiphytic networks (r = 0.7, P < 0.05) (Fig. 5B), while no significant correlations could be drawn for endophytic networks ($P \geq 0.05$) (Fig. 5C, Table. S5).

We further investigated whether the connectivity of microbial communities, which considers the strength of positive and negative interactions, could be explained by these environmental factors. To this end, we computed a community cohesion metric previously proposed [32]. Our analysis unveiled that epiphytic communities exhibited significantly higher cohesion levels compared to endophytic communities (P < 0.001) (Fig. 5D). Among environmental factors shaping community cohesion (Fig. S3), radiation exhibited the highest positive effect both for epiphytic and endophytic networks, along with factors associated with low humidity such as low vapor pressure and low precipitation (Fig. 5E). Taken together these results suggest that higher humidity is associated with bigger (more nodes) microbial networks on leaf surfaces, while increased solar radiation and low humidity are associated with overall more connected networks (higher cohesion).

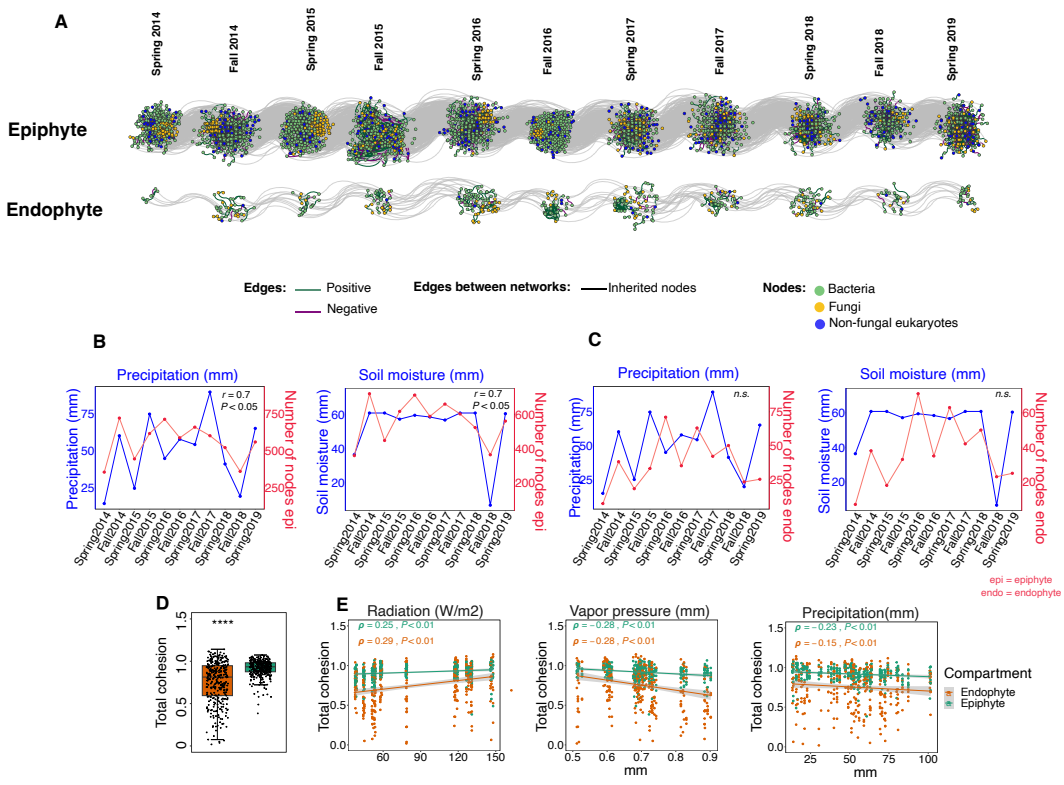


Figure 5. Correlating microbial network complexity and environmental factors. (A) Data from each time point was used to reconstruct co-abundance networks for epiphytic and endophytic compartments. The nodes (dots) represent OTUs, and the edges (colored lines) depict potential positive and negative interactions between OTUs (connections). Gray lines (connecting the networks) show nodes conserved in networks from one-time point to the next (inherited nodes). **(B-C)** Correlation between the number of nodes and monthly precipitation and soil moisture in the epiphytic and endophytic compartments across time points (Pearson correlation coefficient r, $P < 0.05$). **(D)** Total cohesion (sum of positive cohesion and the absolute value of negative cohesion) in epiphytic and endophytic samples. **(E)** Correlations between total cohesion and solar radiation, vapor pressure or precipitation, across leaf compartments (Spearman correlation coefficient ρ, $P < 0.05$). The grey lines indicate 95% confidence intervals. Individual samples are represented by dots and colored by compartments.

Discussion

The phyllosphere is a system directly exposed to various environmental factors such as light and humidity. All these factors can cause significant perturbations to the microbiome of the leaf [29, 33]. This leads to a fundamental question: to what extent do seasonal environmental factors determine the different dynamics of epiphytic and endophytic microbial communities? To address this fundamental ecological question, we have conducted a comprehensive, five-year investigation of the leaf microbiome of *Arabidopsis thaliana* from natural populations in six different geographical locations [30] (Fig. 1). Our results highlight the significant influence of plant compartment, site location, and sampling season on shaping microbial communities, elucidating 14.6-25.7% of their variability, including bacteria, fungi, and non-fungal eukaryotes (Fig. 2), consistent with results from previous studies [30, 34–39].

In our analysis, we identified the leaf compartment as the primary factor driving the variation in the bacteria and non-fungal eukaryotes, resulting in a lower level of alpha diversity among the endophytes (Fig. 2). This finding supports the hypothesis that the diversity gap might result from different conditions within these niches. Endophytic microbiomes face obstacles such as apoplastic acidity and oxidative stress [40], as well as nutrient deficiencies [41]. These are likely to affect their diversity patterns. These challenges may therefore be responsible for the observed lower endophyte diversity. Conversely, the observed higher diversity among epiphytes suggests that the leaf surface, which is more exposed to environmental elements, provides more favorable conditions for microbial proliferation than the protected environment within the plant. In addition, it's important to recognize the potential influence of environmental factors on microbial communities distributed within these compartments.

We found that solar radiation correlated negatively with microbial alpha diversity (Fig. 3). This effect could occur directly by damaging microbial DNA, especially on the leaf surface. Alternatively, it could affect diversity indirectly by promoting the production of reactive oxygen species (ROS) that inhibit the growth and diversity of sensitive species. Long-term low-dose ionizing radiation has been shown to affect soil microbial communities by inhibiting predatory or parasitic fungi [42]. For example, we found a reduction in the abundance of the fungus *Cladosporium* in response to solar radiation. While *Cladosporium* species are melanized filamentous fungi, and melanin can typically protect them from UV radiation, our

results contradicted this expectation [43]. Interestingly, we observed an increase in endophytic *Pseudomonas* species favored by solar radiation, possibly due to their pigment-producing abilities. This suggests that bacteria in the plant microbiome may use pigments as a protective mechanism against ROS, which is particularly important under high UV radiation or intense light [44, 45]. While analyzing microbial interaction networks, solar radiation emerged as an important factor positively correlated with their cohesion. This suggests that radiation may increase the strength of interactions. This may represent a survival strategy whereby microbes form stronger bonds or dependencies to cope with the environmental stress imposed by radiation exposure, thereby increasing resilience to external perturbations.

Humidity-related factors, such as precipitation and vapor pressure, emerged as significant contributors to higher microbial alpha-diversity (Fig. 3). This was expected as humidity influences microbial diversity by modifying substrate diffusion [46] and facilitating microbial dispersal via rain, as demonstrated for fungal diversity [16, 17]. However, we found negative correlations between high humidity parameters (precipitation and vapor pressure), with microbial network cohesion. This suggests that under conditions such as precipitation, microbes may choose strategies such as adhesion over motility, potentially reducing the connectivity of microbial interactions. In addition, intense precipitation may physically disrupt microbial habitats and structures, such as biofilms or microbial aggregates, leading to temporary disintegration of microbial networks and reduced cohesion. Alternatively, during periods of high precipitation or humidity, microbial communities may allocate more resources to survival mechanisms such as biofilm formation or stress response pathways rather than investing in microbial interactions.

Conclusions

Our study conducted a comprehensive analysis of leaf microbiomes over an extended period, revealing the leaf compartment as the primary determinant shaping microbial communities. In addition, we highlighted the critical role of environmental cues in shaping the diversity, composition, and interactions among microbes within leaf compartments. In particular, we identified specific microbial communities that respond to these environmental cues. By using cohesion as a metric to quantify microbial community connectivity [32], we illuminated how external environmental factors can alter internal microbial interactions. Our study introduces a novel approach for investigating temporal community dynamics in natural, host-associated microbiomes. In addition, our findings hold promise for advancing the modeling and prediction of microbial community dynamics over time using insights into environmental influences. Understanding these processes could potentially guide efforts to direct microbial communities toward desired states, with valuable implications for ecosystem management and sustainability.

Method

Collection of *Arabidopsis thaliana* samples and environmental data

Wild *Arabidopsis thaliana* samples were collected from six sites near Tübingen. In the fall and spring over five years (2014-2019, 11 time points, Table. S1). Epiphytic and endophytic

microorganisms were collected from each sample, as described in Agler *et al.* [30]. In brief, rosettes were washed gently with water for 30 sec, then in 5 ml of epiphyte wash solution (0.1% Triton X-100 in 1x TE buffer) for 1 min. Epiphytic microorganisms were collected by filtering the solution through a 0.2 um nitrocellulose membrane filter (Whatman, Piscataway, NJ, USA). The filter was placed in a screw-cap tube and frozen in dry ice. For collecting endophytic fractions, the rosette was then surface-sterilized by washing with 80% ethanol for 15 sec, followed by 2% bleach (sodium hypochlorite) for 30 sec. Rosettes were rinsed thrice with sterile autoclaved water for 10 sec, before placing them in a screw-cap tube and freezing them on dry ice. Phenol-chloroform-based DNA extraction was performed according to a custom protocol as described in Agler *et al.* [30]. The extracted DNA was used for two-step PCR amplification of the V5-V7 region of bacterial 16S rRNA (primers 799F/1192R), the fungal ITS2 region (primers fITS7/ITS4), and V9 region of eukaryotic 18S rRNA (primers F1422/R1797) (Table. S6). Blocking oligos were used to reduce amplification of plant DNA (Table. S6). Purified PCR products were pooled in equimolar amounts before sequencing in Illumina MiSeq runs (Miseq 2×300 bases) spiked with PhiX genomic DAN to ensure high enough sequence diversity. Fourteen environmental factors (Fig. 2B) were collected from TerraClimate [47] for each sampling month (Table. S3). The TerraClimate database has a monthly temporal resolution and approximately 4 km (1/24th degree) spatial resolution [47].

Amplicon sequencing data analysis

Amplicon sequencing data was processed in Mothur (version 1.42.3) [48, 49] as described in Almario *et al*. [29]. Single-end reads were combined to make paired-end reads (make.contigs command), and paired reads with less than five bases overlap between the forward and reverse reads were removed. Only 100-600 bases long reads were kept (screen.seqs). Chimeric sequences were detected and removed using Vsearch [50] in Mothur (chimera.vsearch, remove.seqs). Cutadapt 2.10 [51] was used to trim primer sequences from 16S rRNA and 18S reads. For fungal reads, we used ITSx 1.1b [52] to trim reads to only the ITS2 region. Sequences were clustered into Operational Taxonomic Units (OTUs) at 97% similarity threshold (cluster, dgc method), and abundance filtering was applied to retain OTUs with more than 50 reads (split.abund) and OTU tables were generated (make.shared). OTUs were taxonomically classified (classify.otu) based on the Silva database [53] (version 138.1) for bacterial 16S rRNA data, the UNITE_public database [54] (version 02_02_2019) for fungal ITS2, and the Pr2 [55] (version 4.12.0) for eukaryotic 18S rRNA. The PhiX genome was included in each of the databases to improve the detection of remaining PhiX reads. OTUs classified as chloroplast, mitochondria, Arabidopsis, Embryophyceae, unknown, and PhiX were removed (remove.lineage).

Diversity and multivariate analysis

OTU tables (bacteria, fungi and non-fungal eukaryotes) were modified by removing samples with less than 50 reads. OTU abundance tables were used to calculate Shannon's H−diversity index (estimate richness function in Phyloseq [56] R package) to estimate alpha−diversity. To calculate between-sample diversity, relative abundance OTU tables were computed and transformed (log10 (x + 1)), to calculate Bray-Curtis dissimilarities used for nonmetric multidimensional scaling ordination (NMDS, 'ordinate' function, Phyloseq [56] R package). A

PERMANOVA analysis on Bray-Curtis dissimilarities was performed to identify the main factors influencing the structure of the leaf microbiome ('adonis2' function, Vegan package [57], 10 000 permutations, $P < 0.05$, explanatory categorical variables: Compartment x Site x Season). To facilitate comparability, all quantitative environmental variables (e.g., Temperature and Precipitation) were z-transformed to have a mean of zero and a standard deviation of one. These data were correlated to the measured alpha-diversity (as mentioned above) of each compartment ('cor.test' function, spearman method, Stats [58] package, $P < 0.05$). Environmental data was then used in PERMANOVA analyses to assess the effect of each factor on Bray-Curtis dissimilarities. In detail, for each microbial group and compartment, 14 models were performed ('adonis2', Vegan package [57], 10 000 permutations, $P < 0.05$, explanatory categorical variables: one environmental factor). Unless otherwise stated, data normality was checked ('shapiro.test', Stats [58] package), and means were compared using the nonparametric multivariate test for multiple groups ('dunnTest' function, FSA [59] package, Benjamini-Hochberg $Padj < 0.05$) and two groups ('wilcox.test' function, stats [58] package, $P < 0.05$). All analyses were performed in R (version 4.1.2) [60].

Linear Model Analysis

The association between independent variables (environmental factors) and the dependent variable (relative abundance of genera) was investigated using linear models. Original abundance OTU tables (samples with > 1 read) were aggregated at the taxonomic genus level (aggregate function in R). Rare genera (those with < 50 reads) were excluded, and the table was converted to relative abundance. Additionally, highly correlated environmental factors were identified, and one factor, actual evapotranspiration, was removed from the analysis. A z-transformed environmental table was utilized for consistency. Linear models ('lm' function, stats [58] package) were executed per compartment per genus using the formula:

$$\text{lm}(\text{Genus} \sim e_1 + e_2 + e_3 + ... + e_{13})$$

where 'e' denotes environmental factors. The resulting models were employed to identify the most influential environmental factors (ols−step−best−subset function, olsrr package [61]). . Models with the lowest estimated prediction error (msep parameter) were selected. To estimate the coefficient values of selected factors per genus, a generalized linear model was performed using the formula:

$$glm(Genus \sim subsetofenvironmentalfactors)$$

The significance of each factor individually (by dropping it from the model) was assessed ('drop1' function, 'Chisq' test, lme4 [62] package). The obtained P-values from the chi-squared tests were adjusted for false discovery rate ('p.adjust' function method='fdr', stats [58] package, $P < 0.05$). Coefficient values (of selected environmental factors, demonstrating their strength in predicting the relative abundance of microbes, were utilized to subset some of the genera that exhibited differential effects between both compartments for further visualization. To do this, absolute values of coefficient values of environmental factors per genus were averaged, and the top overlapping genera between epiphytic and endophytic compartments were selected. All analyses were performed in R (version 4.1.2) [60].

Microbial network calculations and properties

Bacteria, fungi, and non-fungal eukaryotes OTU abundance tables were merged and used for correlation calculation using the SparCC algorithm [63], which relies on Aitchison's logratio analysis and is designed to deal with compositional data with high sparsity. OTU tables were filtered to OTUs in at least 5 samples with ≥ 10 reads per OTU per time point per compartment. The filtered OTU tables (OTU raw abundances) were used to calculate SparCC correlation scores (with default parameters) in FastSpar platform [64]. Pseudo p-values were inferred from 1000 bootstraps. Only correlations with $P \leq 0.001$ and absolute correlation > 0 were kept for further analyses. Cytoscape (version 3.7.1) [65] was used for network visualization. A "Cohesion" metric [32] was calculated to quantify the connectivity of each network. For each sample (j), a positive and a negative cohesion metric (equation1) were calculated by multiplying OTUs relative abundances to the average of the OTU's positive or negative correlations.

equation1:

$$Cohesion_j = \sum_{i=1}^{n} RA_i * \overline{cor}_{cor,i}$$

Where RA_i is relative abundance of OTU_i in sample j and $\overline{cor}_{cor,i}$ is average of significant positive (range from 0 to +1) or negative (range from -1 to 0) correlations for OTU_i. Total cohesion per sample is then measured by the sum of the positive and negative cohesions. Total cohesion is correlated with environmental factors using Spearman correlation (cor function in R).

Availability of data and material:

Sequencing data are available under NCBI Bioproject PRJNA961058. OTU tables and scripts are available here https://gitlab.nfdi4plants.de/maryam mahmoudi/AbioticAraMicrobe

Competing interests:

Authors declare no competing financial interests in relation to the work.

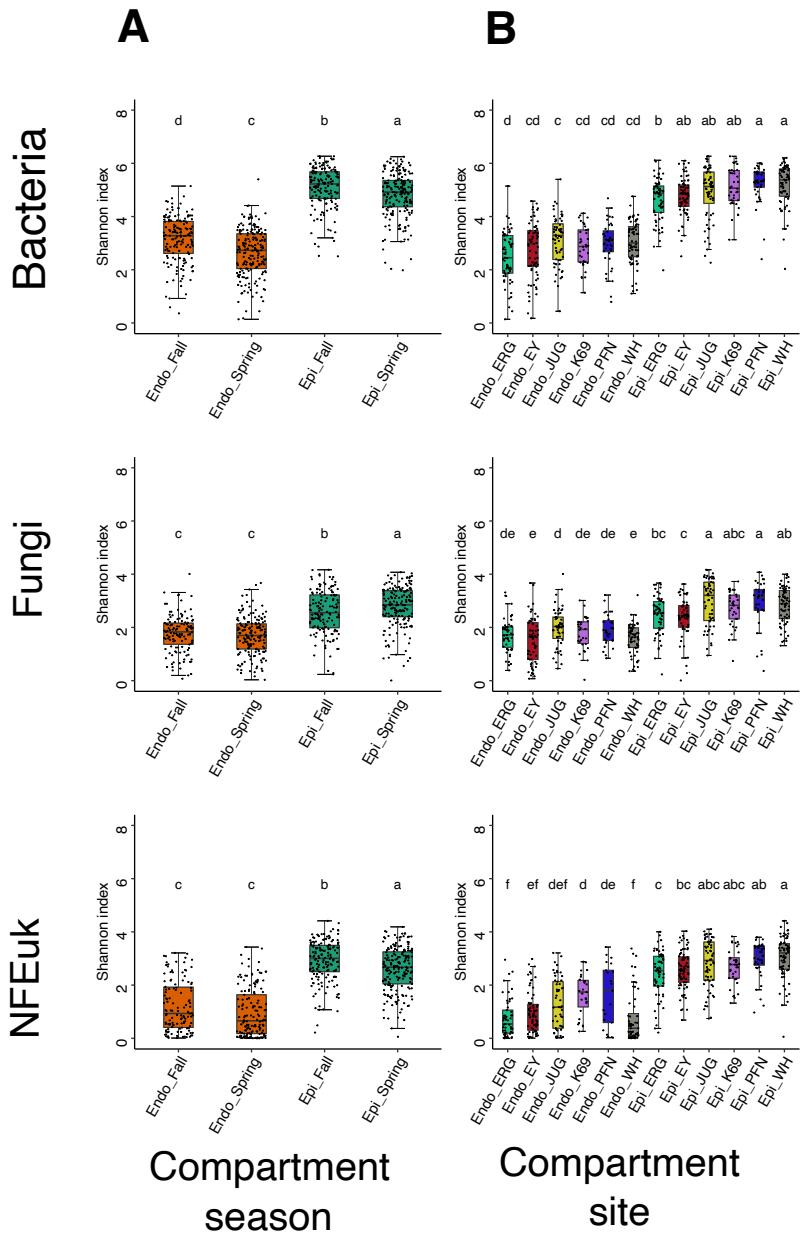
Author Contributions:

MM, KN and EK conceived and devised the study. MM, JA and KL performed the experiments. MM and JA analyzed the data. MM, JA, KN and EK contributed to writing and preparation of the manuscript. All authors read and approved the final manuscript.

Acknowledgements:

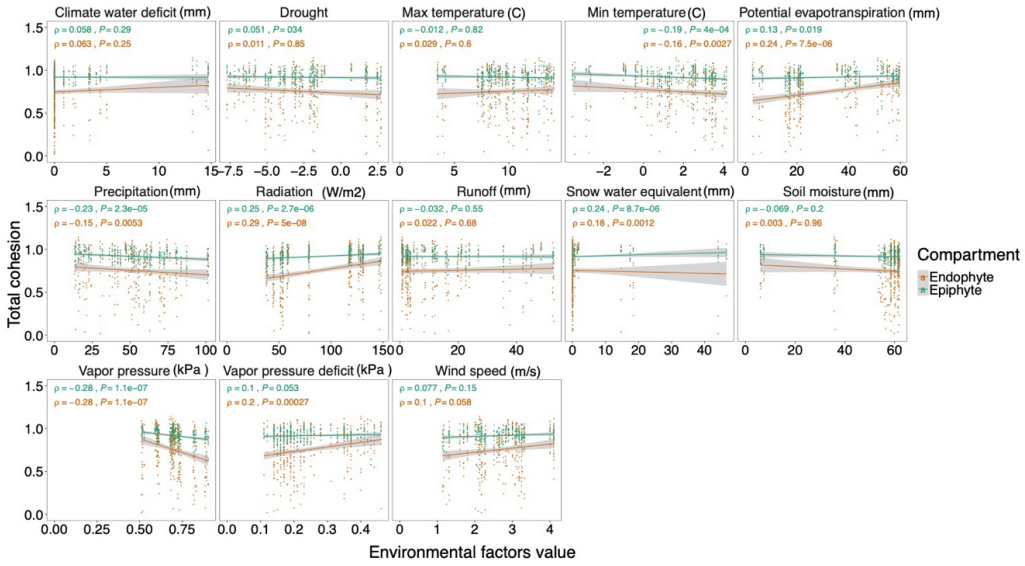
We thank KemenLabSamplingTeam for organizing and participating in several sampling trips and Elke Klenk for helping in MiSeq sequencing preparation. This project has been funded by the European Research Council (ERC) under the DeCoCt research program (grant agreement: ERC-2018-COG 820124), the Cluster of Excellence "Controlling Microbes to Fight Infections" (CMFI; Exc 2124) and the SPP 2125 DECRyPT program from the DFG.

Supplementary figures



Supplementary figure 1. Within-sample diversity of leaf compartments across sampling sites and sampling seasons. Alpha diversity of epiphytic and endophytic samples across spring and fall samples **(A)** and sampling sites **(B)** in bacteria, fungi, and non-fungal eukaryotic communities. The box plots display individual samples as dots. Different letters indicate significant group differences (Dunn test, $P < 0.05$).

Supplementary figure 2. Changes in environmental variables over sampling time points and sampling sites. (A) Lines show the average values for each environmental factor for each sampling month and site. **(B)** Spring vs fall differences for each environmental factor **(C)** Differences between sites for each environmental factor. Asterisks indicate significant differences based on Wilcoxon's test: n.s. ($P > 0.05$), * ($P \leq 0.05$), ** ($P \leq 0.01$), *** ($P \leq 0.001$), and **** ($P \leq 0.0001$).



Supplementary figure 3. Relationship between total cohesion and environmental factors. Each plot shows a linear regression model fit to the data to show the association of total cohesion and environmental factors, across compartments. Samples are represented by dots and colored by compartments. The grey lines indicate 95% confidence intervals, and the Spearman correlation coefficient ρ and actual *P*-values are shown.

Supplementary tables are in https://gitlab.nfdi4plants.de/maryam mahmoudi/AbioticAraMicrobe

Table S1. Experimental set-up with sampling locations and sampling time points.

Numbers indicate the number of samples collected from epiphytic and endophytic samples at different time points (season and year). In addition, the exact date of each sampling event and the coordinates of the sampling locations are given.

Table S2. Fourteen environmental factors (Fig. 1B) used in this study.

The environmental variables were obtained from TerraClimate [47], a database with monthly temporal resolution and approximately 4 km spatial resolution.

Table S3. PERMANOVA results indicating the impact of individual environmental factors on microbial communities within each leaf compartment.

Table S4. Results of linear modeling association of environmental factors and microbial genera.

Table S5. Pearson correlation analysis results depicting the relationship between the number of nodes in microbial networks and the environmental factors.

Table S6. Primers and blocking oligos used in this study.

References

[1] Melanie Sapp, Sebastian Ploch, Anna M Fiore-Donno, Michael Bonkowski, and Laura E Rose. Protists are an integral part of the arabidopsis thaliana microbiome. *Environmental microbiology*, 20(1):30–43, 2018.

[2] Eric JN Helfrich, Christine M Vogel, Reiko Ueoka, Martin Sch¨afer, Florian Ryffel, Daniel B Mu¨ller, Silke Probst, Markus Kreuzer, J¨orn Piel, and Julia A Vorholt. Bipartite interactions, antibiotic production and biosynthetic potential of the arabidopsis leaf microbiome. *Nature Microbiology*, 3(8):909–919, 2018.

[3] Julia A Vorholt. Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10 (12):828–840, 2012.

[4] Gerd Innerebner, Claudia Knief, and Julia A Vorholt. Protection of arabidopsis thaliana against leaf-pathogenic pseudomonas syringae by sphingomonas strains in a controlled model system. *Applied and environmental microbiology*, 77(10):3202–3210, 2011.

[5] Tsunashi Kamo, Syuntaro Hiradate, Ken Suzuki, Ichiro Fujita, Shinji Yamaki, Tadashi Yoneda, Motoo Koitabashi, and Shigenobu Yoshida. Methylobamine, a uvaabsorbing compound from the plant-associated bacteria methylobacterium sp. *Natural Product Communications*, 13(2):1934578X1801300208, 2018.

[6] Unyarat Ritpitakphong, Laurent Falquet, Artit Vimoltust, Antoine Berger, JeanPierre M´etraux, and Floriane L'Haridon. The microbiome of the leaf surface of arabidopsis protects against a fungal pathogen. *New Phytologist*, 210(3):1033–1043, 2016.

[7] Vasvi Chaudhry, Paul Runge, Priyamedha Sengupta, Gunther Doehlemann, Jane E Parker, and Eric Kemen. Shaping the leaf microbiota: plant–microbe–microbe interactions. *Journal of Experimental Botany*, 72(1):36–56, 2021.

[8] Corinne Vacher, Arndt Hampe, Annabel J Port´e, Ursula Sauer, St´ephane Compant, and Cindy E Morris. The phyllosphere: microbial jungle at the plant–climate interface. *Annual review of ecology, evolution, and systematics*, 47:1–24, 2016.

[9] Vidhi Chaudhary, Radha Prasanna, Lata Nain, SC Dubey, Vishal Gupta, Rajendra Singh, Seema Jaggi, and Ashok Kumar Bhatnagar. Bioefficacy of novel cyanobacteriaamended formulations in suppressing damping off disease in tomato seedlings. *World Journal of Microbiology and Biotechnology*, 28:3301–3310, 2012.

[10] Christoph-Martin Geilfus. The ph of the apoplast: dynamic factor with functional impact under stress. *Molecular plant*, 10(11):1371–1386, 2017.

[11] Robert I Griffiths, Bruce C Thomson, Phillip James, Thomas Bell, Mark Bailey, and Andrew S Whiteley. The bacterial biogeography of british soils. *Environmental microbiology*, 13(6):1642–1654, 2011.

[12] Jordi Sardans, Josep PeNUuelas, Marc Estiarte, and Patricia Prieto. Warming and drought alter c and n concentration, allocation and accumulation in a mediterranean shrubland. *Global Change Biology*, 14(10):2304–2316, 2008.

[13] Josefina Tano, Mar´ıa Bel´en Ripa, Mar´ıa Laura Tondo, Anal´ıa Carrau, Silvana Petrocelli, Mar´ıa Victoria Rodriguez, Virginia Ferreira, Mar´ıa In´es Siri, Laura Piskulic, and Elena Graciela Orellano. Light modulates important physiological features of ralstonia pseudosolanacearum during the colonization of tomato plants. *Scientific Reports*, 11(1):14531, 2021.

[14] Prisca Meyer, Bram Van de Poel, and Barbara De Coninck. Uv-b light and its application potential to reduce disease and pest incidence in crops. *Horticulture Research*, 8, 2021.

[15] Reena Debray, Yvonne Socolar, Griffin Kaulbach, Aidee Guzman, Catherine A Hernandez, Rose Curley, Alexander Dhond, Timothy Bowles, and Britt Koskella. Water stress and disruption of mycorrhizas induce parallel shifts in phyllosphere microbiome composition. *New Phytologist*, 234(6):2018–2031, 2022.

[16] Qing-Lin Chen, Hang-Wei Hu, Zhen-Zhen Yan, Chao-Yu Li, Bao-Anh Thi Nguyen, Yong-Guan Zhu, and Ji-Zheng He. Precipitation increases the abundance of fungal plant pathogens in eucalyptus phyllosphere. *Environmental Microbiology*, 23(12): 7688–7700, 2021.

[17] Christophe Gigot, Claude de Vallavieille-Pope, Laurent Huber, and S´ebastien SaintJean. Using virtual 3-d plant architecture to assess fungal pathogen splash dispersal in heterogeneous canopies: a case study with cultivar mixtures and a non-specialized disease causal agent. *Annals of Botany*, 114(4):863–876, 2014.

[18] Franciska T De Vries, Pete Manning, Jerry RB Tallowin, Simon R Mortimer, Emma S Pilgrim, Kathryn A Harrison, Phil J Hobbs, Helen Quirk, Bill Shipley, Johannes HC Cornelissen, et al. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology letters*, 15(11):1230–1239, 2012.

[19] Rudolf O Schlechter, Moritz Miebach, and Mitja NP Remus-Emsermann. Driving factors of epiphytic bacterial communities: a review. *Journal of advanced research*, 19:57–65, 2019.

[20] Julia K Copeland, Lijie Yuan, Mehdi Layeghifard, Pauline W Wang, and David S Guttman. Seasonal community succession of the phyllosphere microbiome. *Molecular Plant-Microbe Interactions*, 28(3):274–285, 2015.

[21] Marcel GA Van Der Heijden and Martin Hartmann. Networking in the plant microbiome. *PLoS biology*, 14(2):e1002378, 2016.

[22] Karoline Faust and Jeroen Raes. Microbial interactions: from networks to models. *Nature Reviews Microbiology*, 10(8):538–550, 2012.

[23] Lisa R¨ottjers and Karoline Faust. From hairballs to hypotheses–biological insights from microbial networks. *FEMS microbiology reviews*, 42(6):761–780, 2018.

[24] Mengting Maggie Yuan, Xue Guo, Linwei Wu, YA Zhang, Naijia Xiao, Daliang Ning, Zhou Shi, Xishu Zhou, Liyou Wu, Yunfeng Yang, et al. Climate warming enhances microbial network complexity and stability. *Nature Climate Change*, 11(4):343–348, 2021.

[25] Damian J Hernandez, Aaron S David, Eric S Menges, Christopher A Searcy, and Michelle E Afkhami. Environmental stress destabilizes microbial networks. *The ISME Journal*, 15(6):1722–1734, 2021.

[26] Keara L Grady, Jackson W Sorensen, Nejc Stopnisek, John Guittar, and Ashley Shade. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nature communications*, 10(1):4135, 2019.

[27] Alice Baniel, Katherine R Amato, Jacinta C Beehner, Thore J Bergman, Arianne Mercer, Rachel F Perlman, Lauren Petrullo, Laurie Reitsema, Sierra Sams, Amy Lu, et al. Seasonal shifts in the gut microbiome indicate plastic responses to diet in wild geladas. *Microbiome*, 9:1–20, 2021.

[28] Kathleen Beilsmith, Matthew Perisin, and Joy Bergelson. Natural bacterial assemblages in arabidopsis thaliana tissues become more distinguishable and diverse during host development. *Mbio*, 12(1):e02723–20, 2021.

[29] Juliana Almario, Maryam Mahmoudi, Samuel Kroll, Mathew Agler, Aleksandra Placzek, Alfredo Mari, and Eric Kemen. The leaf microbiome of arabidopsis displays reproducible dynamics and patterns throughout the growing season. *Mbio*, 13 (3):e02825–21, 2022.

[30] Matthew T Agler, Jonas Ruhe, Samuel Kroll, Constanze Morhenn, Sang-Tae Kim, Detlef Weigel, and Eric M Kemen. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology*, 14(1):e1002352, 2016.

[31] Pauline Bruyant, Yvan Mo¨enne-Loccoz, and Juliana Almario. Root-associated helotiales fungi: Overlooked players in plant nutrition. *Soil Biology and Biochemistry*, page 109363, 2024.

[32] Cristina M Herren and Katherine D McMahon. Cohesion: a method for quantifying the connectivity of microbial communities. *The ISME Journal*, 11(11):2426–2438, 2017.

[33] Yong-Guan Zhu, Chao Xiong, Zhong Wei, Qing-Lin Chen, Bin Ma, Shu-Yi-Dan Zhou, Jiaqi Tan, Li-Mei Zhang, Hui-Ling Cui, and Gui-Lan Duan. Impacts of global change on the phyllosphere microbiome. *New Phytologist*, 2022.

[34] Paloma Dur´an, Thorsten Thiergart, Ruben Garrido-Oter, Matthew Agler, Eric Kemen, Paul Schulze-Lefert, and St´ephane Hacquard. Microbial interkingdom interactions in roots promote arabidopsis survival. *Cell*, 175(4):973–983, 2018.

[35] Ting Ou, Wei-fang Xu, Fei Wang, Gary Strobel, Ze-yang Zhou, Zhong-huai Xiang, Jia Liu, and Jie Xie. A microbiome study reveals seasonal variation in endophytic bacteria among different mulberry cultivars. *Computational and Structural Biotechnology Journal*, 17:1091–1100, 2019.

[36] Ian M Ware, Michael E Van Nuland, Zamin K Yang, Christopher W Schadt, Jennifer A Schweitzer, and Joseph K Bailey. Climate-driven divergence in plantmicrobiome interactions generates range-wide variation in bud break phenology. *Communications Biology*, 4(1):748, 2021.

[37] Fabrice Roux, L´ea Frachon, and Claudia Bartoli. The genetic architecture of adaptation to leaf and root bacterial microbiota in arabidopsis thaliana. *Molecular Biology and Evolution*, 40(5):msad093, 2023.

[38] Joseph Edwards, Cameron Johnson, Christian Santos-Medell´ın, Eugene Lurie, Natraj Kumar Podishetty, Srijak Bhatnagar, Jonathan A Eisen, and Venkatesan Sundaresan. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences*, 112(8):E911–E920, 2015.

[39] Paul Runge, Freddy Ventura, Eric Kemen, and Remco Stam. Distinct phyllosphere microbiome of wild tomato species in central peru upon dysbiosis. *Microbial ecology*, 85(1):168–183, 2023.

[40] Xilan Yu, Steven P Lund, Russell A Scott, Jessica W Greenwald, Angela H Records, Dan Nettleton, Steven E Lindow, Dennis C Gross, and Gwyn A Beattie. Transcriptional responses of pseudomonas syringae to growth in epiphytic versus apoplastic leaf sites. *Proceedings of the National Academy of Sciences*, 110(5):E425–E434, 2013.

[41] Mariana Murillo-Roos, Hafiz Syed M Abdullah, Mossaab Debbar, Nico Ueberschaar, and Matthew T Agler. Cross-feeding niches among commensal leaf bacteria are shaped by the interaction of strain-level diversity and resource availability. *The ISME Journal*, 16(9):2280–2289, 2022.

[42] Feng Cheng, Xiaofei Huang, Qingao Qin, Zijian Chen, Fei Li, and Wenchen Song. The effect of aboveground long-term low-dose ionizing radiation on soil microbial diversity and structure. *Frontiers in Ecology and Evolution*, 11:1184582, 2023.

[43] Jesse Bland, Lisa Astuto Gribble, Michael C Hamel, Jeremy B Wright, Garrett Moormann, Marlene Bachand, Ginger Wright, and George D Bachand. Evaluating changes in growth and pigmentation of cladosporium cladosporioides and paecilomyces variotii in response to gamma and ultraviolet irradiation. *Scientific Reports*, 12(1):12142, 2022.

[44] Carlos A Ruiz-P´erez, Silvia Restrepo, and Mar´ıa Mercedes Zambrano. Microbial and functional diversity within the phyllosphere of espeletia species in an andean high-mountain ecosystem. *Applied and Environmental Microbiology*, 82(6):1807–1817, 2016.

[45] Wouter Buytaert, Rolando C´elleri, Bert De Bi`evre, Felipe Cisneros, Guido Wyseure, Jozef Deckers, and Robert Hofstede. Human impact on the hydrology of the andean p´aramos. *Earth-Science Reviews*, 79(1-2):53–72, 2006.

[46] Xuechen Yang, Kai Zhu, Michael E Loik, and Wei Sun. Differential responses of soil bacteria and fungi to altered precipitation in a meadow steppe. *Geoderma*, 384: 114812, 2021.

[47] John T Abatzoglou, Solomon Z Dobrowski, Sean A Parks, and Katherine C Hegewisch. Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data*, 5(1):1–12, 2018.

[48] Patrick D Schloss, Sarah L Westcott, Thomas Ryabin, Justine R Hall, Martin Hartmann, Emily B Hollister, Ryan A Lesniewski, Brian B Oakley, Donovan H Parks, Courtney J

Robinson, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.*, 75(23):7537–7541, 2009.

[49] Patrick D Schloss. Reintroducing mothur: 10 years later. *Applied and environmental microbiology*, 86(2):e02343–19, 2020.

[50] Torbjørn Rognes, Tom´aˇs Flouri, Ben Nichols, Christopher Quince, and Fr´ed´eric Mah´e. Vsearch: a versatile open source tool for metagenomics. *PeerJ*, 4:e2584, 2016.

[51] Marcel Martin. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, 17(1):10–12, 2011.

[52] Johan Bengtsson-Palme, Martin Ryberg, Martin Hartmann, Sara Branco, Zheng Wang, Anna Godhe, Pierre De Wit, Marisol S´anchez-Garc´ıa, Ingo Ebersberger, Filipe de Sousa, et al. Improved software detection and extraction of its1 and its 2 from ribosomal its sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in ecology and evolution*, 4(10):914–919, 2013.

[53] Christian Quast, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, J¨org Peplies, and Frank Oliver Gl¨ockner. The silva ribosomal rna gene database project: improved data processing and web-based tools. *Nucleic acids research*, 41 (D1):D590–D596, 2012.

[54] Kessy Abarenkov, R Henrik Nilsson, Karl-Henrik Larsson, Andy FS Taylor, Tom W May, Tobias Guldberg Frøslev, Julia Pawlowska, Bj¨orn Lindahl, Kadri P˜oldmaa, Camille Truong, et al. The unite database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. *Nucleic Acids Research*, 52(D1):D791–D797, 2024.

[55] Laure Guillou, Dipankar Bachar, St´ephane Audic, David Bass, C´edric Berney, Lucie Bittner, Christophe Boutte, Ga´etan Burgaud, Colomban de Vargas, Johan Decelle, et al. The protist ribosomal reference database (pr2): a catalog of unicellular eukaryote small sub-unit rrna sequences with curated taxonomy. *Nucleic acids research*, 41(D1): D597–D604, 2012.

[56] Paul J McMurdie and Susan Holmes. phyloseq: an r package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, 8(4):e61217, 2013.

[57] Jari Oksanen, F Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R Minchin, RB O'hara, Gavin L Simpson, Peter Solymos, M Henry H Stevens, Helene Wagner, et al. Package 'vegan'. *Community ecology package, version*, 2(9):1–295, 2013.

[58] R Core Team, Maintainer R Core Team, MASS Suggests, and S Matrix. Package stats. *The R Stats Package*, 2018.

[59] Derek Ogle and Maintainer Derek Ogle. Package 'fsa'. *Cran Repos*, 1:206, 2017.

[60] R R Core Team et al. R: A language and environment for statistical computing. 2013.

[61] Aravind Hebbali and Maintainer Aravind Hebbali. Package 'olsrr'. *Version 0.5*, 3, 2017.

[62] Douglas Bates, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Gabor Grothendieck, Peter Green, and M Ben Bolker. Package 'lme4'. *convergence*, 12(1):2, 2015.

[63] Jonathan Friedman and Eric J Alm. Inferring correlation networks from genomic survey data. *PLoS computational biology*, 8(9):e1002687, 2012.

[64] Stephen C Watts, Scott C Ritchie, Michael Inouye, and Kathryn E Holt. Fastspar: rapid and scalable correlation estimation for compositional data. *Bioinformatics*, 35 (6):1064–1066, 2019.

[65] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11):2498–2504, 2003.

Chapter 3

Dynamics and patterns of the leaf microbiome over the growing season

---

- **Status** Published
- **URL** https://doi.org/10.1128/mbio.02825-21
- **Own Contribution**
  - Designed the research (with co-authors)
  - Performed the experiments (with co-authors)
  - Interpreted the data (with co-authors)
  - Analyzed and visualized the data (with co-authors)
  - Wrote the manuscript (with co-authors)

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

The Leaf Microbiome of *Arabidopsis* Displays Reproducible Dynamics and Patterns throughout the Growing Season

Juliana Almario,[b] Maryam Mahmoudi,[a] Samuel Kroll,[a,c] Mathew Agler,[d] Aleksandra Placzek,[c] Alfredo Mari,[a,c] Eric Kemen[a,c]

[a]Microbial Interactions in Plant Ecosystems, IMIT/ZMBP, Eberhard Karls University of Tübingen, Tübingen, Germany
[b]Université Claude Bernard Lyon 1, CNRS, INRA, Villeurbanne, France
[c]Max Planck Research Group Fungal Biodiversity, Max Planck Institute for Plant Breeding Research, Cologne, Germany
[d]Plant Microbiosis Group, Institute for Microbiology, Friedrich Schiller University Jena, Jena, Germany

Juliana Almario, Maryam Mahmoudi, and Samuel Kroll contributed equally to this article. Author designation and order were based on contributions to writing the manuscript and data analysis and interpretation.

ABSTRACT Leaves are primarily responsible for the plant's photosynthetic activity. Thus, changes in the leaf microbiota, which includes deleterious and beneficial microbes, can have far-reaching effects on plant fitness and productivity. Identifying the processes and microorganisms that drive these changes over a plant's lifetime is, therefore, crucial. In this study, we analyzed the temporal dynamics in the leaf microbiome of *Arabidopsis thaliana*, integrating changes in both composition and microbe-microbe interactions via the study of microbial networks. Field-grown *Arabidopsis* were used to monitor leaf bacterial, fungal and oomycete communities throughout the plant's natural growing season (extending from November to March) over three consecutive years. Our results revealed the existence of conserved temporal patterns, with microbial communities and networks going through a stabilization phase of decreased diversity and variability at the beginning of the plant's growing season. Despite a high turnover in these communities, we identified 19 "core" taxa persisting on *Arabidopsis* leaves across time and plant generations. With the hypothesis these microbes could be playing key roles in the structuring of leaf microbial communities, we conducted a time-informed microbial network analysis which showed core taxa are not necessarily highly connected network "hubs," and "hubs" alternate with time. Our study shows that leaf microbial communities exhibit reproducible dynamics and patterns, suggesting the potential of using our understanding of temporal trajectories in microbial community composition to design experiments aimed at driving these communities toward desired states.

IMPORTANCE Utilizing plant microbiota to promote plant growth and plant health is key to more environmentally friendly agriculture. A major bottleneck in the engineering of plant-beneficial microbial communities is the low persistence of applied microbes under filed conditions, especially considering plant leaves. Indeed, although many leaf-associated microorganisms have the potential to promote plant growth and protect plants from pathogens, few of them are able to survive and thrive over time. In our study, we could show that leaf microbial communities are very variable at the beginning of the plant growing season but become more and more similar and less variable as the season progresses. We further identify a cohort of 19 "core" microbes, systematically present on plant leaves that would make these microbes exceptional candidates for future agricultural applications.

KEYWORDS leaf microbiome, time dynamics, microbial networks, microbial hubs, community dynamics, core microbial community, hub microbes, microbial communities, persistence, plant-microbe interactions

Editor David S. Guttman, University of Toronto

Address correspondence to Eric Kemen, eric.kemen@uni-tuebingen.de.

The authors declare no conflict of interest.

Received 13 October 2021
Accepted 28 February 2022
Published 14 April 2022

Leaves are primarily responsible for the plant's photosynthetic activity and gaseous exchange. Consequently, leaf health and performance have a direct effect on plant growth and fitness (1). Leaves are colonized by a wide range of microbes, including bacteria, archaea, and microeukaryotes such as fungi and oomycetes. While natural openings on leaves such as stomata, hydathodes, or wounds represent entry points for major plant pathogens, they also house commensal and even beneficial microbes (2, 3), leading to plant-protecting effects (4–6). There is increasing interest particularly by plant breeders in microbiota-engineering approaches to promote the growth and health of crops through beneficial microbes (7). In this context, it is essential to understand the processes that shape the composition of leaf microbiota.

There is a level of specificity between plants and their leaf microbiota as studies have repeatedly shown that different plant lineages tend to harbor quantitatively different microbial consortia in their leaves (8), with differences even observed between ecotypes of the same plant species (9). Although it is unclear how plants can selectively recruit certain microbial groups, the soil in which plants grow appears to be an important driver (9, 10). The study of plant microbiota over different developmental stages suggests that as the plant grows, the microbiota becomes more tissue-specific with major differences observed between root and shoot microbiota (11, 12). There is increasing awareness of the fact that plant-associated microbiota are not static but dynamic communities whose members engage in multiple layered interactions, such as mutualism, antagonism, or predation, which change through time under the convergent influence of environmental and host cues and neighboring plants (13). Indeed, leaf microbial communities are constantly exposed to the arrival of new microbes carried by soil, water, and wind and can thus show a high level of stochasticity, i.e., high unpredictability and high variability. Furthermore, leaf communities have been shown to change throughout time and reach different stable states, depending on early (random) events (priority effects) (14). Recent studies have followed the dynamics of microbiome formation in leaves (13–18) and roots (19), but few of them have conducted a cross-kingdom survey, integrating both bacterial and microeukaryotic communities, which means we only have partial views of microbial dynamics in leaves.

Correlation network analyses on the relative abundance of microbial taxa can inform us about potential interactions between community members, albeit with high rates of false-positive and false-negative interactions among the predictions (20). Still, the combination of in-depth analysis of microbial coabundance networks with hypothesis-testing experiments has led to the description of new biological interactions in host-associated microbiomes (21), including plant microbiomes (22). Moreover, the study of microbial networks over time can inform us about the dynamics of these potential interactions and how they relate to changes in the diversity and structure of microbial communities (23). However, such approaches have rarely been applied to investigate how plant-associated microbiome change through the plant's life.

Given the complexity of leaf microbial communities, assigning ecological roles and ecological importance to individual taxa is extremely challenging. Concepts based on the persistence of a microbe (core taxa) and/or its importance on microbial networks (hubs taxa) have been applied to identify microorganisms playing key roles in leaf communities (22, 24). Although the large majority of leaf microbes show scattered distributions with highly fluctuating occurrences in plant leaves across environments and time, some microorganisms achieve a stable presence in plant populations (25). It is unclear how these "core" microbes are able to systematically colonize the host plant, but their "persistence" could involve recolonization processes (26) or vertical inheritance via seeds (27). The stability of the associations between "core" microbes and the host-plant suggests a high level of adaptation to the leaf niche by microbes. This can involve traits associated with plant colonization and infection, as suggested for leaf-pathogenic *Pseudomonas viridiflava* (25), but it can also involve the capacity of the microorganism to reshape the leaf microbiota, as part of a "niche construction" strategy. Notably, Agler et al. (22) showed that the inoculation of the leaf-pathogenic oomycete *Albugo* on *Arabidopsis thaliana* plants translates into decreased microbial diversity on leaves and altered microbiome profiles. The analysis of microbial interaction networks within the

leaf microbiome showed *Albugo* acts as a network "hub," showing the highest level of connections (interactions) with other microbes, which would allow it to influence the structure of the leaf microbiota. Because of its hub characteristics and experimentally proven impact on leaf microbial communities, *Albugo* has been proposed as a "keystone" taxon of the leaf microbiota in *Arabidopsis*. However, it is still unclear whether reshaping the leaf microbiome contributes to persistence of core taxa.

The aim of this study was to analyze the temporal dynamics in the leaf microbiome of *Arabidopsis thaliana*, integrating both compositional changes and changes in microbe-microbe interactions via the study of microbial networks. Amplicon sequencing was used to follow leaf bacterial, fungal, and oomycete communities in a field experiment throughout the natural growing season of *Arabidopsis*, which, in the Cologne area, extends from November (seedling emergence) to March (beginning of flowering). The experiment was carried out in a common garden over three consecutive years in order to capture long-term dynamics, and four *Arabidopsis* ecotypes were included. Our results reveal seasonal/monthly patterns associated with reproducible changes in particular groups across kingdoms like *Sphingomonadales* and *Actinomycetales* bacteria, *Microbotryales* and *Sporidiobolales* fungi, and *Peronosporales* oomycetes. Despite a high level of stochasticity in microbial colonization of the leaf, we identified 19 taxa that were consistently present (core taxa), including putative pathogenic and beneficial taxa. Between November and February, the diversity and variability of leaf microbial communities decreased, as microbial networks stabilized (changed less) and exhibited decreasing complexity (number of nodes and connections). With the hypothesis that certain microbes play a predominant role in the structuring and stability of these communities, we focused on the identification of microbes having both a persistent presence on *Arabidopsis* leaves (core microbes) and a high connectivity in leaf microbiome networks (hub microbes).

RESULTS

The leaf microbiome is highly dynamic. To study the temporal dynamics of the leaf microbiome, we grew four *A. thaliana* ecotypes in a common garden and surveyed the changes in their leaf microbiome via amplicon-sequencing (bacteria, fungi, and oomycetes). Leaf samples were taken monthly between November and March (5 months), thus covering most of the plant's growing season over autumn and winter (Fig. 1; see also Table S1 in the supplemental material). To identify the main factors shaping leaf microbial communities, we used multivariate approaches, including nonmetric multidimensional scaling (NMDS; Fig. S1A) and permutational multivariate analysis of variance (ANOVA; Bray-Curtis dissimilarities, $P < 0.05$; see Fig. S1B) on the relative abundance of bacterial, fungal and oomycete taxa (operational taxonomic units [OTUs] defined at 97% similarity). These analyses showed a marginal effect of the plant ecotype (2 to 4% explained variance) but an important effect of the time of sampling (32 to 40% explained variance; factors "month," "experiment," and their interactions; see Fig. S1B), confirming that leaf microbial communities are highly variable in time (i.e., dynamic). Although variability between experiments was significant (4 to 13% explained variance), the "month" of sampling was an important factor (11 to 15% explained variance; Fig. S1A and B), suggesting the existence of seasonal/monthly patterns in these microbial communities. Such patterns were easily observable when considering changes in the relative abundance of highly abundant microbial orders (Fig. 1). For example, the relative abundances of *Sphingomonadales* and *Actinomycetales* increased throughout the plant's growing season, while the relative abundance of *Rhizobiales* tended to decrease. As for fungi, the relative abundance of *Microbotryales* increased, while that of *Sporidiobolales* decreased. Interestingly, the relative abundance of Peronosporales oomycetes, which include *A. thaliana*'s pathogen *Hyaloperonospora* spp., increased with time, reaching maximum values at the end of the plant's growing season (Dunn test, $P < 0.05$) (Fig. 1; see also Fig. S2).

Persistent (core) taxa in the leaf microbiome. We aimed to identify microbial groups showing a persistent presence throughout the plant's life, hypothesizing that they might play important roles in plant-microbe and microbe-microbe interactions within the microbiome. Highly persistent microbes (≥95% sample occurrence for fungi and oomycete, ≥98% for bacteria) varied considerably between experiments, with only 19 of 67 OTUs (28%) showing robust patterns across experiments and ecotypes (see Fig. S3 and Table S2).

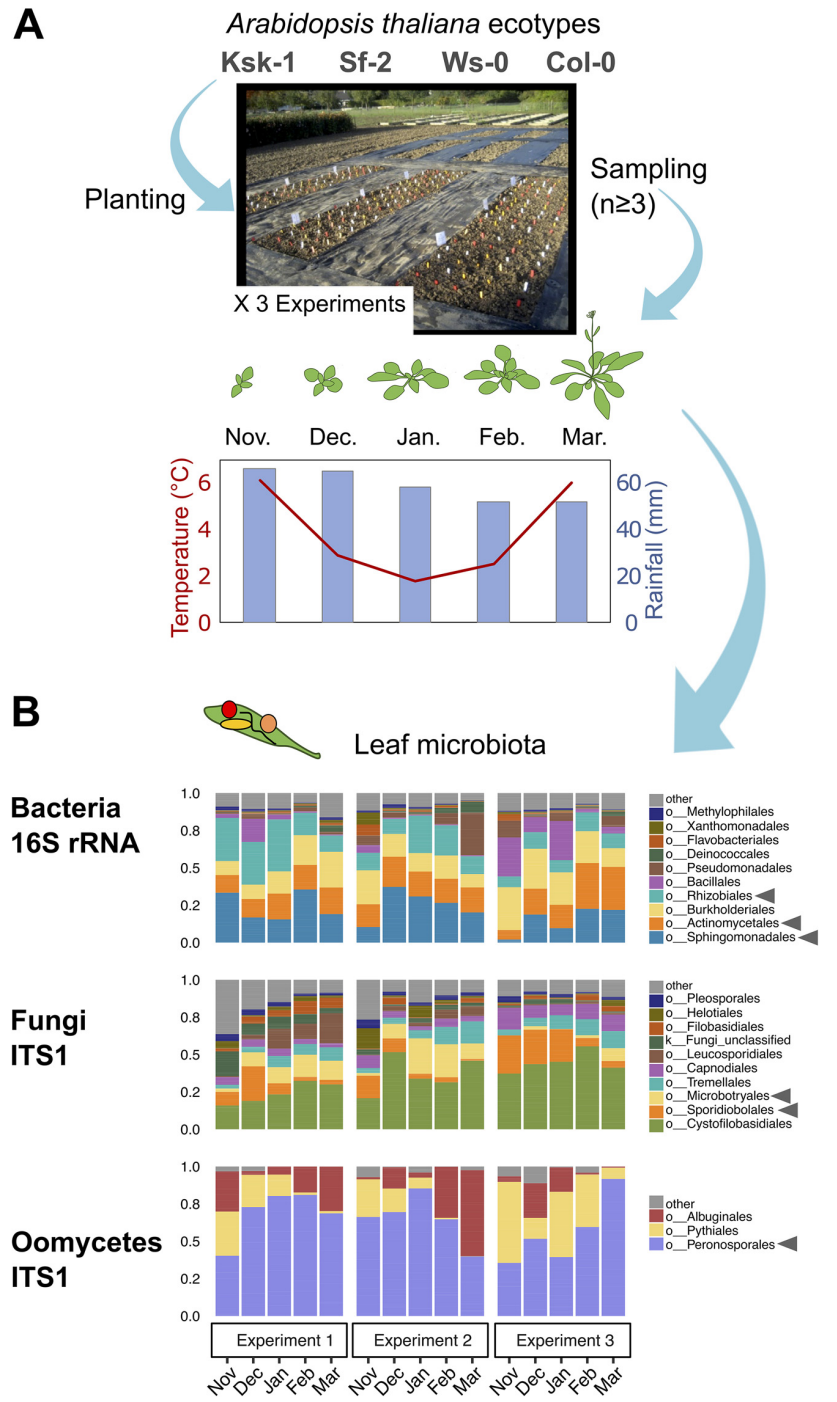


FIG 1 Monitoring leaf microbiome dynamics throughout the natural growing season of *A. thaliana*. (A) Experimental setup. The four global *Arabidopsis* accessions Ws-0, Col-0, Ksk-1, and Sf-2 were planted in a common garden (Max Planck Institute, Cologne, Germany). Every month from November to March, three individual plants per ecotype were collected, and leaf samples were taken for microbiome analysis (destructive sampling). The experiment was repeated three times over the years 2014 to 2015 (experiment 1), 2015 to 2016 (experiment 2), and 2016 to 2017 (experiment 3), with a total number of 206 plant leaf samples analyzed (see Table S1). Average temperature and rainfall during the sampling season are shown. (B) Composition of the leaf microbiome. Microbiome analysis was conducted via Illumina-based amplicon sequencing (Miseq 2 × 300 bases). Taxonomic markers included the bacterial 16S rRNA v5-v7 region, fungal ITS1, and the oomycete ITS1 region. Bar charts show the average relative abundance of the main microbial groups (order level) by months, across three experiments. Arrowheads indicate taxa exhibiting marked seasonal patterns (see Fig. S2).

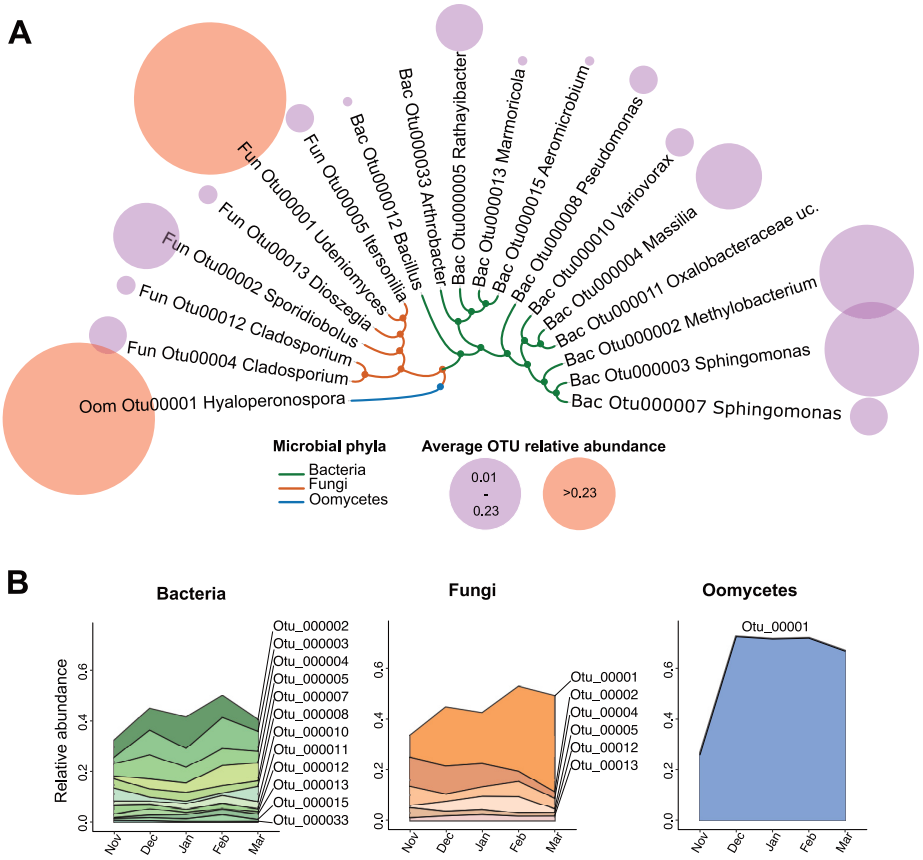


FIG 2 Persistent core members of the *Arabidopsis* leaf microbiome. (A) Core taxa were identified as OTUs showing high occurrence results (≥95% for fungi and oomycete, ≥98% for bacteria) in each of the three experiments. Bubbles depict the average relative abundance of each core OTU, per sample. The dendrogram depicts taxonomical distances between OTUs (hierarchical clustering on Gower distances from OTU taxonomy). (B) Changes in the relative abundance of core taxa over time (month averages; $n > 38$ samples per month).

Notably, these persistent core taxa (1 oomycete, 6 fungus, and 12 bacterial OTUs) included known *Arabidopsis* pathogens like the obligate biotrophic oomycete *Hyaloperonospora* sp. (Otu00001), as well as bacterial taxa known to colonize *Arabidopsis* leaves, including *Sphingomonas* spp. (OTUs), *Methylobacterium* sp. (Otu000002), and *Variovorax* (Otu000010). Persistent fungal taxa included two ascomycetes (*Cladosporium* spp. Otu00004 and Otu00012) and four basidiomycete yeast (*Dioszegia* sp. Otu00013, *Itersonilia* sp. Otu00005, *Sporidiobolus* sp. Otu00002, and *Udeniomyces* sp. Otu00001) (Fig. 2A; see also Table S2). The relative abundances of these core taxa changed throughout the plant's growing season, reaching a maximum in February, where it represented as much as 49, 52, and 71% of the bacterial, fungal, and oomycete communities, respectively (Fig. 2B). These results indicate that despite the highly dynamic and stochastic nature of the leaf microbiome, a limited number of microbes—only 19 of 3,058 OTUs (0.62%)—consistently cocolonize plant leaves. This suggests a high degree of adaptation to this niche but also frequent interactions with one another.

Diversity and variability of the leaf microbiome decrease throughout the plant's growing season as communities stabilize. With the hypothesis that leaf-associated microbial communities become increasingly stable throughout the plant's growing season, we analyzed their dynamics in terms of alpha diversity (number of taxa in the community), within-month variability (plant-to-plant differences in community composition), and variability between consecutive months (month-to-month differences in community composition). While the bacterial alpha diversity (Shannon's H index) remained unchanged, the fungal and oomycete alpha diversity decreased with significant differences observed between November and the last 2 months, February and March (Dunn test, $P < 0.05$) (Fig. 3). A similar trend was

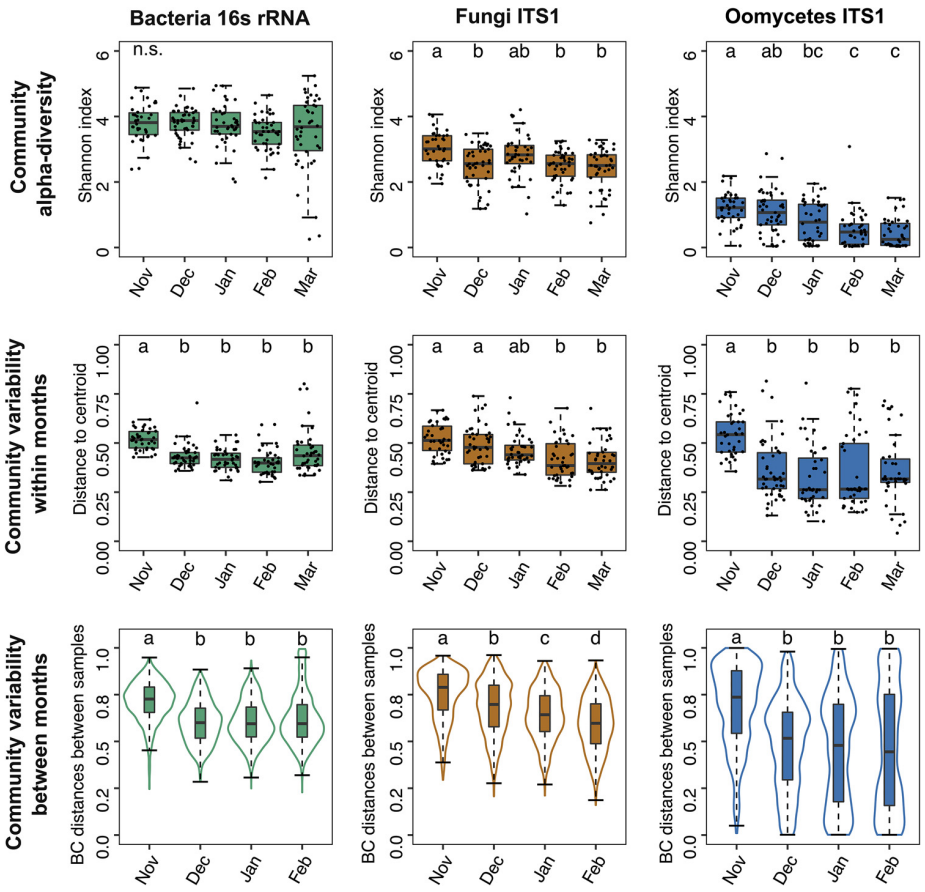


FIG 3 Changes in alpha diversity and variability in leaf microbial communities over time. The alpha diversity (Shannon's H index), within-month variability (distance to the group centroid; beta-dispersion), and between-month variability (Bray-Curtis distances between samples from consecutive months) in bacterial, fungal, and oomycete communities are shown. Each plot shows combined data from the three experiments, with $n > 38$ samples per month. Dots represent individual samples, whiskers depict the dispersion of the data (1.5 × interquartile range), and different letters indicate significant differences between groups (Dunn test, $P < 0.05$). Single BC distances between samples are not shown because of the high number of comparisons (>700).

observed for within-month variability (sample distance to the group centroid), as variability of bacterial and fungal communities decreased progressively from November to February (Dunn test, $P < 0.05$) (Fig. 3). Similarly, a progressive decrease in between-month variability (sample-to-sample distances between consecutive months) was observed for bacterial and fungal communities (Dunn test, $P < 0.05$; Fig. 3). Oomycete communities exhibited similar trends, but the dynamics were less pronounced due to higher data variability. Together, these results suggest that throughout the plant's growing season, leaf microbial communities become progressively less diverse, more similar between plant individuals, and less variable in time. This suggests leaf communities go through a consolidation and stabilization phase from November to February.

Interaction networks within the leaf microbiome stabilize over time. Microbial networks computed from correlation of species abundances, are used to infer potential interactions between microbes within a community. To determine if/how leaf microbial networks changed over time, we used taxa abundance data from each time point (month) to generate five "month" networks (Fig. 4A). Because the data were highly sparse (53% sparsity), the SparCC algorithm (optimized for sparse data) was used for network calculation (28). The five networks differed in terms of general characteristics such as the number of nodes (number of taxa) and edges (correlations between taxa; syn. connections) with no clear pattern, except for the month of "February," which had both the lowest number of nodes and the

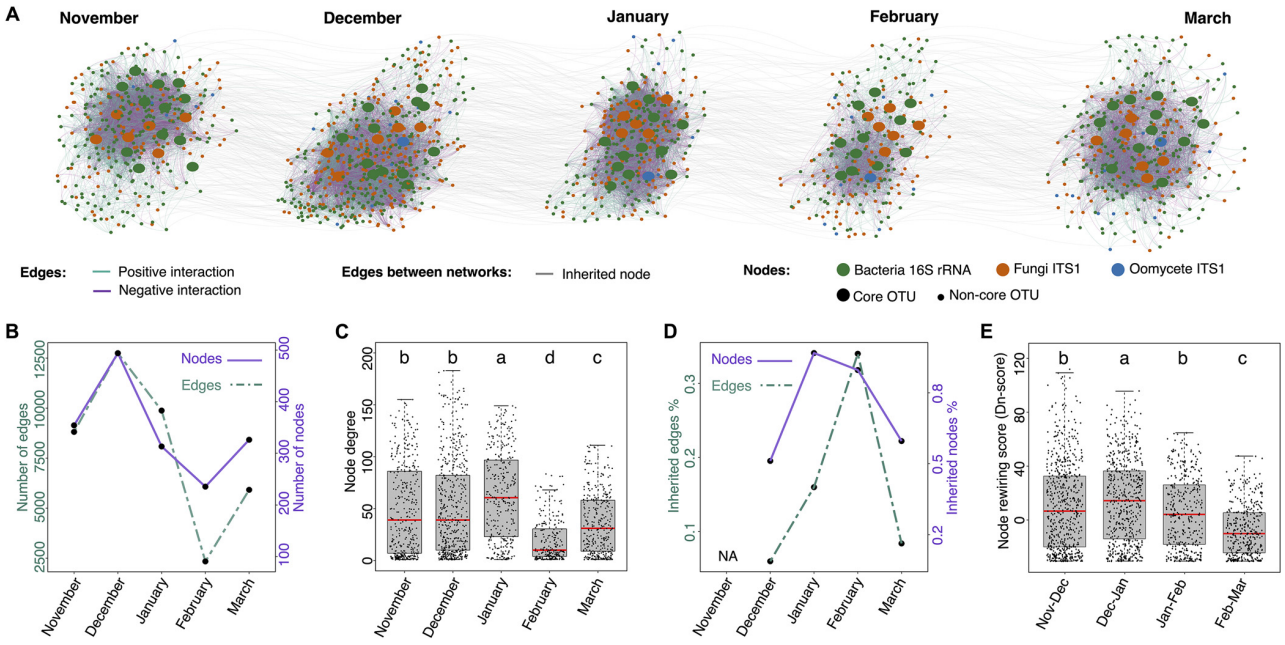


FIG 4 Changes in phyllosphere microbial interaction networks throughout *A. thaliana's* growing season. (A) Data from the three experiments were aggregated to reconstruct co-abundance networks for each time point (month) using the SparCC algorithm. Nodes (dots) represent OTUs; edges (colored lines) depict potential positive and negative interactions between OTUs (connections). Nodes from core microbes are indicated. Gray lines connecting networks show nodes conserved in networks from 1 month to the next (inherited nodes). (B) Number of nodes and edges in each month network. (C) Percentage of nodes and edges in a given month network which are inherited from (shared with) the previous month network. (D) Percentage of edges inherited for a given inherited node. (E) Node degree, i.e., number of edges per node in each month network. (F) Node-rewiring score (Dn-score) calculated in DyNet. For each node, its connected neighbors are compared between two networks (consecutive months) and the changes (rewiring) are quantified. Points represent rewiring scores from single nodes, high values indicate important changes in the node's connections between the compared networks. Different letters indicate significant differences between conditions (Dunn test, $P < 0.05$).

lowest number of edges (Fig. 4B). Similarly, the nodes of this network had the lowest number of interactions (node degree), going from 70 on average in January to only 10 on average in February (Fig. 4B and C; Dunn test, $P < 0.05$). This confirmed that microbial networks indeed changed throughout the plant's growing season and suggested major restructuring events around the month of February, when the network exhibited minimal complexity.

With the hypothesis that these changes were associated with an increased stability of the network's structure, we compared networks from consecutive months, recording similarities (inherited nodes/edges) and differences (node rewiring events) between them. Inherited nodes/edges were defined as those shared between consecutive months. The percentage of inherited nodes per network increased from 51% in December to 89% in January and 82% February, meaning the large majority (82%) of the nodes in the February network were already present in the January network (Fig. 4D). A similar trend was observed for the number of inherited edges, doubling from December (6%) to January (16%) and February (34%). To quantify changes between networks, taking into account the nodes and their connections, we calculated a node-rewiring score for each node in the network. This score reflects the changes in a node's connections between the compared networks (Dn-score in DyNet) (29). This analysis revealed that differences between networks tended to decrease through time, with minimum rewiring events between the months of February and March (Dunn test, $P < 0.05$) (Fig. 4E). These results suggest that throughout the beginning of the season (November to February) leaf microbial networks go through a stabilization phase, during which month-to-month changes tend to diminish (increasing numbers of shared nodes and edges, and decreasing node rewiring) as networks exhibit lowering complexity (lower numbers of nodes, edges, and connections), reaching minimum levels in February.

Identifying hubs among core microbes in *Arabidopsis* leaf microbiome. Time-based microbial networks were analyzed to determine whether potential "keystone"' microbes (i.e., hubs—taxa with high betweenness and high closeness centrality) in the leaf microbiome

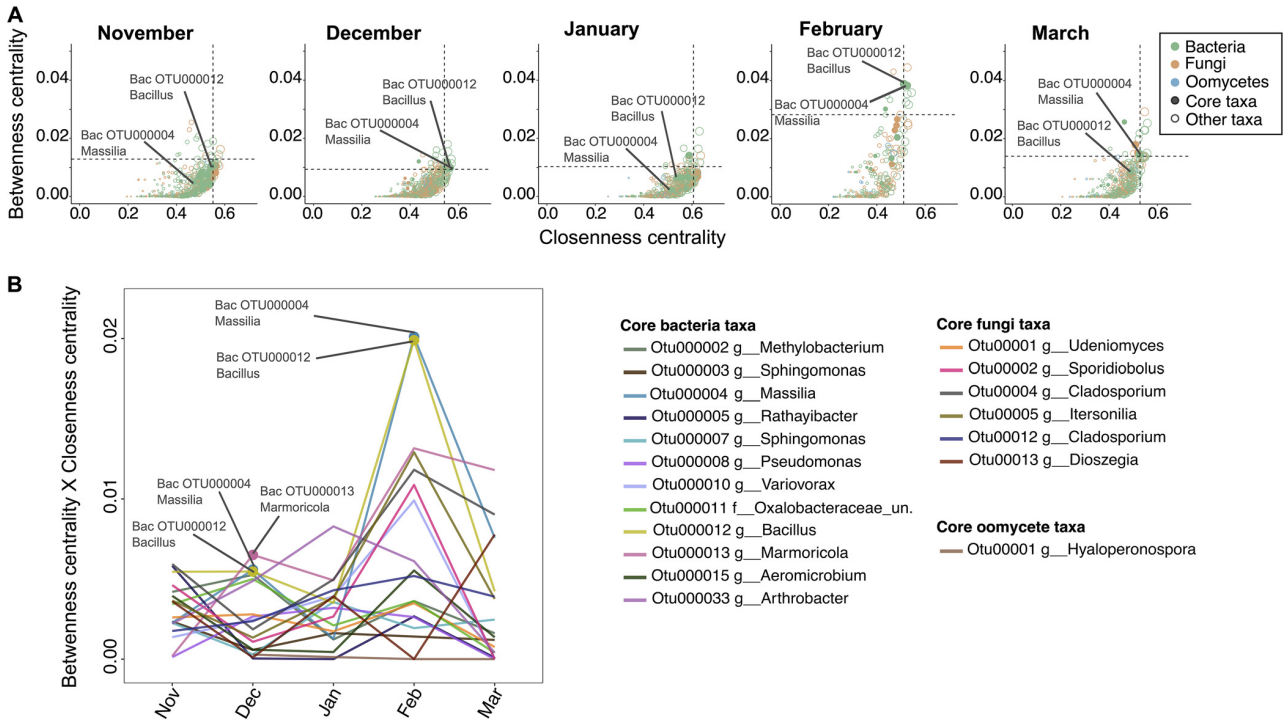


FIG 5 Identification of microbial hubs within *A. thaliana*'s core leaf microbiome. The correlation networks calculated with SparCC (Fig. 3) were used to identify microbial hubs as nodes with high betweenness centrality (i.e., the fraction of shortest paths passing through the given node) and high closeness-centrality (i.e., the average shortest distance from the given node to other nodes). (A) Values for single taxa, with dotted lines indicating the top 5% values. Circles are colored based on microbial phyla. Circle sizes depict the node's degree. Closed circles indicate taxa identified as part of the core leaf microbiome. Two core OTUs (12 and 4) are annotated. (B) Changes in the connectivity of core taxa. The product of "betweenness centrality × closeness centrality" was used to depict monthly changes in the connectivity of core OTUs. Hub taxa are indicated.

were also highly persistent core microbes. Connectivity analysis on individual month networks revealed few taxa exhibiting hub characteristics (4 to 10 OTUs, 1 to 3% of network OTU nodes) and a high turn-over between months, with no taxon systematically identified as hub in every month network (Fig. 5A; see also Table S1). Among the 19 "core" taxa identified previously (Fig. 2), only three bacterial OTUs, i.e., *Bacillus* OTU00012, *Massilia* OTU00004, and *Marmoricola* OTU000013, could be identified as hubs exhibiting high network connectivity in the months of December and February (Fig. 5B; see also Table S3).

As hub identification is highly dependent on network calculation approaches, we repeated these analyses on Spearman-based correlation networks calculated in Co-Net (see Fig. S4) with partially similar results. Approximately a third of the OTUs identified as hubs in the SparCC networks were also identified as hubs in the Spearman-based networks (see Table S1). Notably, this also included *Bacillus* OTU00012 and *Massilia* OTU00004. Taken together, these results indicate that, with the exception of one *Massilia* and one *Bacillus* lineage, "core" taxa in the *Arabidopsis* leaf microbiome are not major network hubs and that network hub microbes change over time.

DISCUSSION

The phyllosphere is a complex microbial habitat due to its direct exposure to a range of abiotic factors—light, humidity, and temperature—that can alter the leaf environment within minutes, hours, or days. Furthermore, leaf microbial communities are directly exposed to the arrival of new microbes disseminated by soil particles, water, and wind (30). In this context, key ecological questions are still unanswered: what is the relative importance of environmental filtering versus biotic interactions in shaping community structures and what is the impact of stochasticity (31)? Our limited understanding of the processes behind colonization of leaf surfaces by microbes and their assembly and persistence thereon under field conditions

constitutes a major drawback for the agricultural usage of plant-beneficial microbes (32). To address these fundamental questions, we have conducted a long-term experiment to follow month-to-month changes in the composition of the *Arabidopsis* leaf microbiome during its natural growing season, which, in the Cologne area, extends from November (seedling emergence) to March (beginning of flowering).

As expected for dynamic ecological systems (33), bacterial, fungal, and oomycete leaf-associated communities were highly stochastic, with factors such as the sampling time and the plant ecotype explaining only half of the variability observed (see Fig. S1). Despite high between-experiment variability, robust differences between months were observable for some microbial groups known to be relevant for plant-growth like *Peronosporales* oomycetes (see Fig. S2). *Hyaloperonospora*, the causal agent of downy mildew, was by far the most abundant *Peronosporales* in *Arabidopsis* leaves, as it has been described for various geographic locations elsewhere (26). Although our sampled plants exhibited no downy mildew disease symptoms at any time throughout the field experiments, the relative abundance of *Peronosporales* increased throughout the growing season reaching maximum values in March. This is in agreement with disease dynamics of downy mildew in *Brassicaceae* known to be favored by cold wet weather and could indicate that the pathogenic pressure on the plant increases over the early growing, which takes place in winter for *Arabidopsis* populations of the Cologne area.

The analysis of community dynamics indicates that from November to February leaf microbial communities go through a stabilization phase becoming less diverse and less variable, which results in microbial networks of decreased complexity (Fig. 3 and 4). This is likely associated with the fact that core microbes become increasingly dominant throughout the season (Fig. 3B) but contrasts with previous studies showing higher diversity in *Arabidopsis*' bacterial microbiome in spring (34). Seasonal dynamics have been described in microbiome associated with plants (17, 18, 35) and animals (36–39) and are thought to be driven partly by environmental cues and perturbations. By monitoring the *Arabidopsis* bacterial leaf microbiome under controlled conditions, Maignien et al. similarly showed that leaf communities become increasingly similar as the microbiome "matures" over time (14). However, the fact that in our study microbiome dynamics mirror decreases in temperature and rainfall associated with winter (Fig. 1) leads us to hypothesize that climatic conditions might be also driving the observed leaf microbiome dynamics, maybe via the selection of cold-resistant microorganisms. Indeed, a strong "winter effect" on microbial communities has been observed in a diversity of environments, including the bee's gut (39), lake water (40), and air (41). We hypothesize that winter conditions might apply a strong selective filter causing leaf microbial communities to reduce in complexity. Longer experiments are needed to determine whether different dynamics would be observed at later stages, e.g., during spring.

Microbes with a stable presence in *Arabidopsis* leaves (core taxa; Fig. 2) accounted for only 0.62% of all detected leaf-taxa, indicating a high turnover in leaf microbial communities. Interestingly, most microbes identified as "core" in one experiment were not identified as "core" in subsequent experiments, suggesting that most dominant lineages change from year to year. This is in accordance with observations that leaf microbiomes are strongly structured by priority effects during early colonization events, meaning that communities can be alternatively dominated by different microorganisms or core taxa (14). In our study, core taxa included putative plant pathogens like *Hyaloperonospora* and *Cladosporium* (42, 43) but also plant beneficial microorganisms such as *Sphingomonas* and *Variovorax*, which could explain the asymptomatic state of the sampled plants. Leaf-inhabiting *Sphingomonas* bacteria have been shown to protect *Arabidopsis* from bacterial pathogens (4) and are hypothesized to participate in plant disease resistance against root fungal pathogens. *Variovorax* strains have been shown to modulate plant hormonal balance by degrading auxins, thus promoting plant growth under stress conditions (44). However, not only bacteria have been reported to interfere with plant hormone levels; there have been reports of yeasts on *A. thaliana* capable of producing auxin-like indolic compounds (45). We have identified four basidiomycete yeast taxa (*Udeniomyces*, *Sporidiobolus*, *Itersonilia*, and *Dioszegia*) as

systematic colonizers of *Arabidopsis* leaves. Although little is known about the associations between these yeasts and *Arabidopsis*, a recent study on a leaf basidiomycete yeast (*Moesziomyces bullatus*) suggests they can play important roles in plant protection by antagonizing pathogenic oomycetes through secretion of protein effectors (46). While previous studies have reported on the prevalence of some of the identified core taxa on *Arabidopsis*' leaves (12, 22, 47), we show here that these associations persist throughout the plant's life and between plant generations, suggesting some level of microbial adaptation to the leaf niche or even possible coevolution between core microbes, as well as with the host plant. Future isolation/reinoculation experiments will aim at understanding the ecological role of these microbes in *Arabidopsis* leaves.

Microbe-microbe interactions participate in the structuring of microbial communities, with certain microbes—hub and keystone microbes—playing central roles (48). We hypothesized that high connectivity within leaf microbial networks might explain the persistence of the identified core taxa. However, in contrast to our hypothesis, the connectivity level (hubness) of individual core taxa was highly variable from month to month, with no taxon maintaining high connectivity levels throughout the entirety of the growing season (Fig. 5). This indicates that high connectivity is not a prerequisite for high prevalence in the leaf microbiome as core taxa are not necessarily network hubs (19). Nevertheless, two microbes among the leaf core taxa within the *Bacillus* and *Massilia* lineages deviated from this rule and were identified as hubs. Interestingly, in the month of February when leaf microbial communities displayed the lowest levels of complexity, both *Bacillus* and *Massilia* reached maximum connectivity levels within leaf microbial networks (Fig. 5), while their relative abundances on leaves remained stable (Fig. 2). It is tempting to speculate that there might be a functional link between these hubs and community stability. Indeed, it has been shown that highly connected microbes can be good predictors of the stability of microbial communities (49). In the future, experimental evidence will be needed to improve predictions and to determine whether (and how) hub removal affects the stability of microbial communities over time.

Conclusions. Taken together, our results show that, despite a high level of stochasticity, leaf microbial communities exhibit detectable time patterns with stable and unstable components. This study opens a new field of research on time-informed community dynamics in natural host-associated microbiomes. In the long term, these types of studies could make it possible to model and predict microbial community dynamics. Understanding these processes could allow us to design experiments aimed at driving microbial communities toward desired states.

MATERIALS AND METHODS

Common garden experiment. To study the temporal dynamics of *A. thaliana*'s leaf microbiome, we conducted a common garden experiment wherein *A. thaliana* plants were sampled every month from November to March, covering the plant's natural growing season, including the vegetative and early reproductive growth phases (Fig. 1). The experiment was conducted as described in Agler et al. (22). Briefly, surface-sterilized seeds were germinated on Jiffy pellets for 10 days under greenhouse conditions before transfer to the field. To take into account host genetic variability, four *Arabidopsis thaliana* ecotypes covering different geographic origins were used (Ws-0 [Wassilewskija, Russia], Col-0 [Columbia, USA], Ksk-1 [Keswick, UK], and Sf-2 [San Feliu, Spain]), using the same seed batch for the three experiments. The field was divided into nine experimental plots which were planted with ten plants per ecotype, in a randomized setup. At each sampling point, whole leaf samples were taken from two to four randomly selected plants per ecotype. The whole experiment was repeated three times in 2014 to 2015, 2015 to 2016, and 2016 to 2017. The field is located at the Max Planck Institute for Plant Breeding Research (Cologne, Germany) (see Table S1).

DNA extraction and amplicon sequencing. Samples were processed exactly as described in Agler et al. (22). Briefly, whole-leaf samples were crushed and used for phenol-chloroform-based DNA extraction. The obtained DNA was used for two-step PCR amplification of the V5-V7 region of the bacterial 16S rRNA (primers B799F/B1194R), the fungal ITS1 region (primers ITS1F/ITS2), and the oomycete ITS1 region (primers ITS1O/5.8s-O-R). Blocking oligonucleotides were used to reduce plant DNA amplification (50). Purified PCR products were pooled in equimolar amounts before sequencing on three Illumina MiSeq runs (2 × 300-bp reads) with 10% PhiX control. Primers targeting the oomycete ITS1 region also produced "non-oomycete" reads but at a very marginal level (3%).

Amplicon sequencing data analysis. Amplicon sequencing data were processed in Mothur (51) as described in Karasov et al. (25). Single-end reads were paired (*make.contigs* command), and paired reads with more than 5 bases overlap between the forward and reverse reads were kept. Only 100 to 600 bases long reads were retained (*screen.seqs*). Chimeras were checked using Uchime in Mothur with more abundant sequences as reference (*chimera.uchime*, abskew = 1.9). Sequences were clustered into OTUs

at the 97% similarity threshold using the VSEARCH program in Mothur (*cluster*, dgc method). Individual sequences were taxonomically classified using the rdp classifier method (*classify.seqs*, consensus confidence threshold set to 80) and the greengenes database (13_8 release) for 16S rRNA data, the UNITE_public database (v12_2017) for fungal ITS1, and Pr2 (v4.10.0) for oomycete ITS1. The PhiX genome was included in each of the databases to improve the detection of remaining PhiX reads. Each OTU was then taxonomically classified (*classify.otu*, consensus confidence threshold set to 66); OTUs with unknown taxonomy at the kingdom level were removed, as were low-abundance OTUs (<50 reads, *split.abund*). This last step removed extreme low abundance (<0.0001%)/low occurrence (<0.48%) OTUs. The taxonomy of bacterial OTUs of interest was further verified using the silva database (v1.38; SINA Aligner). This allowed us to classify bacterial OTU00004 as *Massilia* sp.

Sample alpha-diversity analysis was conducted on OTU abundance tables using Shannon's H diversity index (estimate_richness function in phyloseq package). Data normality was checked (Shapiro-Wilk's test), and means were compared using a nonparametric multivariate test (Dunn's test, Bonferroni-corrected adjusted *P* value [P_{adj}] <0.05). Beta-diversity analyses were conducted on transformed [$\log_{10}(x + 1)$] OTU relative abundance tables. Bray-Curtis dissimilarities between samples were computed and used for nonmetric multidimensional scaling ordination (NMDS, function "ordinate"; Phyloseq package). A PerMANOVA analysis on Bray-Curtis dissimilarities was conducted to identify the main factors influencing the structure of the leaf microbiome (Adonis, Vegan package, 10,000 permutations, $P < 0.05$, explanatory categorical variables: experiment × month × ecotype). A beta-dispersion analysis on Bray-Curtis dissimilarities was conducted to compare sample-to-sample variation within each month of sampling (multivariate homogeneity of group dispersion analysis, "betadisper"; Vegan package). Differences between conditions were tested using a nonparametric multivariate test (Dunn's test, Bonferroni corrected, $P_{adj} < 0.05$). All analyses were conducted in R 3.6.1.

Identification of a core leaf microbiome in *A. thaliana*. Core taxa were identified as OTUs showing high occurrence over time (≥95% for fungi and oomycete, ≥98% for bacteria) in each of the three experiments analyzed. A higher cutoff was used for bacteria (98% occurrence) since they exhibited a higher average occurrence compared to fungi and oomycetes. The taxonomical classification of core OTUs was used to compute pairwise dissimilarities (distances) between OTUs ("daisy" function, Cluster package in R, Gower's distance) which were used for hierarchical clustering ("hclust" function, Cluster package in R). The obtained dendrogram was modified in the browser version of iTOL (v5.5.1) (52).

Network analysis. Bacteria, fungi and oomycete OTU tables were merged and used for correlations calculation using either the Spearman correlation coefficient in Co-Net (53) or the SparCC algorithm (20), which relies on Aitchison's log-ratio analysis and is designed to deal with compositional data with high sparsity like this data set (sparsity = 74%) (28). OTU tables were filtered to OTUs present in at least five samples with >10 reads per OTU (sparsity = 53%). For the Co-Net based analysis, the OTU relative abundances were calculated, and the obtained OTU tables were transformed [$\log_{10}(x + 1)$] before calculating Spearman correlation scores using Co-Net in Cytospace (54). The parameters included the selection of top 5% correlations (edge selection, quantile = 0.05, top and bottom) and the computing of *P* values by Fisher Z-score with multiple-test correction (Bonferroni, $P = 0.001$). For the SparCC-based analysis, the filtered OTU tables (OTU raw abundances) were used to calculate SparCC correlation scores (with default parameters). Pseudo *P* values were inferred from 1,000 bootstraps. Only correlations with $P < 0.001$ were kept for further analyses. Cytoscape (v3.7.1) was used for network visualization and determination of betweenness centrality (i.e., the fraction of shortest paths passing through a given node) and closeness centrality values (i.e., the average shortest distance from given node to each other node). Node-rewiring score (Dn-score) was calculated via the DyNet package in Cytoscape (29). For each node, its connected neighbors are compared between two networks and the changes (rewiring) are quantified. Differences between conditions were tested using a nonparametric multivariate test (Dunn's test, Bonferroni corrected, $P_{adj} < 0.05$). Microbial hubs were identified as top 5% OTUs showing maximum betweenness centrality and closeness centrality scores.

Data availability. Sequencing data are available under NCBI BioProject PRJNA438596. OTU tables and scripts are available (https://github.com/IshtarMM/Dynamic_LeafMicrobiome). All *A. thaliana* accession numbers used in this study have been published previously, and seeds are available from stock centers.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.8 MB.

FIG S2, TIF file, 1 MB.

FIG S3, TIF file, 1.2 MB.

FIG S4, TIF file, 2.6 MB.

TABLE S1, XLSX file, 0.01 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

This study was supported by the Max-Planck Gesellschaft, the University of Tübingen, the European Research Council (ERC) under the DeCoCt research program (grant agreement ERC-2018-COG 820124), the Cluster of Excellence on Plant Sciences (CEPLAS; Exc 1028), the

Cluster of Excellence "Controlling Microbes to Fight Infections" (CMFI; Exc 2124), and the SPP 2125 DECRyPT program from the DFG. We acknowledge support by Open Access Publishing Fund of University of Tübingen.

We thank Libera Lo Presti for critical reading of the manuscript.

We declare there are no competing financial interests in relation to this work.

All authors give their consent for publication of the manuscript, and all authors received the manuscript and had the opportunity to read it and comment. This project was initiated by S.K., M.A., and E.K. S.K., M.A., and A.M. conducted the experiments with help from A.P. J.A. and M.M. analyzed the data, and J.A., M.M., and E.K. wrote the manuscript, with contributions from all authors.

REFERENCES

1. Berger S, Sinha AK, Roitsch T. 2007. Plant physiology meets phytopathology: plant primary metabolism and plant-pathogen interactions. J Exp Bot 58:4019–4026. https://doi.org/10.1093/jxb/erm298.
2. Vorholt JA. 2012. Microbial life in the phyllosphere. Nat Rev Microbiol 10:828–840. https://doi.org/10.1038/nrmicro2910.
3. Müller DB, Vogel C, Bai Y, Vorholt JA. 2016. The plant microbiota: systems-level insights and perspectives. Annu Rev Genet 50:211–234. https://doi.org/10.1146/annurev-genet-120215-034952.
4. Ritpitakphong U, Falquet L, Vimoltust A, Berger A, Métraux J-P, L'Haridon F. 2016. The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. New Phytol 210:1033–1043. https://doi.org/10.1111/nph.13808.
5. Beattie GA, Lindow SE. 1999. Bacterial colonization of leaves: a spectrum of strategies. Phytopathology 89:353–359. https://doi.org/10.1094/PHYTO.1999.89.5.353.
6. Vogel CM, Potthoff DB, Schäfer M, Barandun N, Vorholt JA. 2021. Protective role of the *Arabidopsis* leaf microbiota against a bacterial pathogen. Nat Microbiol 6:1537–1548. https://doi.org/10.1038/s41564-021-00997-7.
7. Kroll S, Agler MT, Kemen E. 2017. Genomic dissection of host-microbe and microbe-microbe interactions for advanced plant breeding. Curr Opin Plant Biol 36:71–78. https://doi.org/10.1016/j.pbi.2017.01.004.
8. Massoni J, Bortfeld-Miller M, Jardillier L, Salazar G, Sunagawa S, Vorholt JA. 2020. Consistent host and organ occupancy of phyllosphere bacteria in a community of wild herbaceous plant species. ISME J 14:245–258. https://doi.org/10.1038/s41396-019-0531-8.
9. Wagner MR, Lundberg DS, del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun 7:12151. 1. https://doi.org/10.1038/ncomms12151.
10. Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, Bokulich NA, Mills DA, Martin G, Taghavi S, van der Lelie D, Gilbert JA. 2015. The soil microbiome influences grapevine-associated microbiota. mBio 6:e02527-14. https://doi.org/10.1128/mBio.02527-14.
11. Beilsmith K, Perisin M, Bergelson J. 2021. Natural bacterial assemblages in *Arabidopsis thaliana* tissues become more distinguishable and diverse during host development. mBio 12:e02723-14. https://doi.org/10.1128/mBio.02723-20.
12. Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, Dombrowski N, Münch PC, Spaepen S, Remus-Emsermann M, Hüttel B, McHardy AC, Vorholt JA, Schulze-Lefert P. 2015. Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature 528:364–369. https://doi.org/10.1038/nature16192.
13. Meyer KM, Porch R, Muscettola IE, Vasconcelos ALS, Sherman JK, Metcalf CJE, Lindow SE, Koskella B. 2022. Plant neighborhood shapes diversity and reduces interspecific variation of the phyllosphere microbiome. ISME J https://doi.org/10.1038/s41396-021-01184-6.
14. Maignien L, DeForce EA, Chafee ME, Eren AM, Simmons SL. 2014. Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. mBio 5:e00682-13. https://doi.org/10.1128/mBio.00682-13.
15. Rastogi G, Sbodio A, Tech JJ, Suslow TV, Coaker GL, Leveau JHJ. 2012. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. 10. ISME J 6:1812–1822. https://doi.org/10.1038/ismej.2012.32.
16. Copeland JK, Yuan L, Layeghifard M, Wang PW, Guttman DS. 2015. Seasonal community succession of the phyllosphere microbiome. Mol Plant Microbe Interact 28:274–285. https://doi.org/10.1094/MPMI-10-14-0331-FI.
17. Grady KL, Sorensen JW, Stopnisek N, Guittar J, Shade A. 2019. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. Nat Commun 10:4135. https://doi.org/10.1038/s41467-019-11974-4.
18. Bowsher AW, Benucci GMN, Bonito G, Shade A. 2020. Seasonal dynamics of core fungi in the switchgrass phyllosphere, and co-occurrence with leaf bacteria. Phytobiomes J PBIOMES-07–20-0051-R.
19. Stopnisek N, Shade A. 2021. Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. ISME J 15:2708–2715. https://doi.org/10.1038/s41396-021-00955-5.
20. Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. PLoS Comput Biol 8:e1002687. https://doi.org/10.1371/journal.pcbi.1002687.
21. Röttjers L, Faust K. 2018. From hairballs to hypotheses: biological insights from microbial networks. FEMS Microbiol Rev 42:761–780. https://doi.org/10.1093/femsre/fuy030.
22. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM. 2016. Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLoS Biol 14:e1002352. https://doi.org/10.1371/journal.pbio.1002352.
23. Faust K, Lahti L, Gonze D, de Vos WM, Raes J. 2015. Metagenomics meets time series analysis: unraveling microbial community dynamics. Curr Opin Microbiol 25:56–66. https://doi.org/10.1016/j.mib.2015.04.004.
24. Carlström CI, Field CM, Bortfeld-Miller M, Müller B, Sunagawa S, Vorholt JA. 2019. Synthetic microbiota reveal priority effects and keystone strains in the *Arabidopsis* phyllosphere. Nat Ecol Evol 3:1445–1454. https://doi.org/10.1038/s41559-019-0994-z.
25. Karasov TL, Almario J, Friedemann C, Ding W, Giolai M, Heavens D, Kersten S, Lundberg DS, Neumann M, Regalado J, Neher RA, Kemen E, Weigel D. 2018. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. Cell Host Microbe 24:168–179.e4. https://doi.org/10.1016/j.chom.2018.06.011.
26. Lundberg DS, Jové R de P, Ayutthaya PPN, Karasov TL, Shalev O, Poersch K, Ding W, Bollmann-Giolai A, Bezrukov I, Weigel D. 2021. Contrasting patterns of microbial dominance in the *Arabidopsis thaliana* phyllosphere. bioRxiv 2021.04.06.438366.
27. Rodríguez CE, Antonielli L, Mitter B, Trognitz F, Sessitsch A. 2020. Heritability and functional importance of the *Setaria viridis* bacterial seed microbiome. Phytobiomes J 4:40–52. https://doi.org/10.1094/PBIOMES-04-19-0023-R.
28. Weiss S, Van Treuren W, Lozupone C, Faust K, Friedman J, Deng Y, Xia LC, Xu ZZ, Ursell L, Alm EJ, Birmingham A, Cram JA, Fuhrman JA, Raes J, Sun F, Zhou J, Knight R. 2016. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME J 10:1669–1681. https://doi.org/10.1038/ismej.2015.235.
29. Goenawan IH, Bryan K, Lynn DJ. 2016. DyNet: visualization and analysis of dynamic molecular interaction networks. Bioinformatics 32:2713–2715. https://doi.org/10.1093/bioinformatics/btw187.
30. Chaudhry V, Runge P, Sengupta P, Doehlemann G, Parker JE, Kemen E. 2021. Shaping the leaf microbiota: plant-microbe-microbe interactions. J Exp Bot 72:36–56. https://doi.org/10.1093/jxb/eraa417.
31. Zhou J, Ning D. 2017. Stochastic community assembly: does it matter in microbial ecology? Microbiol Mol Biol Rev 81:e00002-17. https://doi.org/10.1128/MMBR.00002-17.
32. Finkel OM, Castrillo G, Herrera Paredes S, Salas González I, Dangl JL. 2017. Understanding and exploiting plant beneficial microbes. Curr Opin Plant Biol 38:155–163. https://doi.org/10.1016/j.pbi.2017.04.018.
33. Shoemaker LG, Sullivan LL, Donohue I, Cabral JS, Williams RJ, Mayfield MM, Chase JM, Chu C, Harpole WS, Huth A, HilleRisLambers J, James ARM, Kraft NJB, May F, Muthukrishnan R, Satterlee S, Taubert F, Wang X, Wiegand T, Yang Q, Abbott KC. 2020. Integrating the underlying structure

of stochasticity into community ecology. Ecology 101:e02922. https://doi.org/10.1002/ecy.2922.
34. Bartoli C, Frachon L, Barret M, Rigal M, Huard-Chauveau C, Mayjonade B, Zanchetta C, Bouchez O, Roby D, Carrère S, Roux F. 2018. *In situ* relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. ISME J 12:2024–2038. https://doi.org/10.1038/s41396-018-0152-7.
35. Liu D, Howell K. 2021. Community succession of the grapevine fungal microbiome in the annual growth cycle. Environ Microbiol 23:1842–1857. https://doi.org/10.1111/1462-2920.15172.
36. Marino S, Baxter NT, Huffnagle GB, Petrosino JF, Schloss PD. 2014. Mathematical modeling of primary succession of murine intestinal microbiota. Proc Natl Acad Sci U S A 111:439–444. https://doi.org/10.1073/pnas.1311322111.
37. Ren T, Boutin S, Humphries MM, Dantzer B, Gorrell JC, Coltman DW, McAdam AG, Wu M. 2017. Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels. Microbiome 5:163. https://doi.org/10.1186/s40168-017-0382-3.
38. Schlomann BH, Parthasarathy R. 2019. Timescales of gut microbiome dynamics. Curr Opin Microbiol 50:56–63. https://doi.org/10.1016/j.mib.2019.09.011.
39. Kešnerová L, Emery O, Troilo M, Liberti J, Erkosar B, Engel P. 2020. Gut microbiota structure differs between honeybees in winter and summer. ISME J 14:801–814. https://doi.org/10.1038/s41396-019-0568-8.
40. Vigneron A, Lovejoy C, Cruaud P, Kalenitchenko D, Culley A, Vincent WF. 2019. Contrasting winter versus summer microbial communities and metabolic functions in a permafrost thaw lake. Front Microbiol 10:1656. https://doi.org/10.3389/fmicb.2019.01656.
41. Cáliz J, Triadó-Margarit X, Camarero L, Casamayor EO. 2018. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. Proc Natl Acad Sci U S A 115:12229–12234. https://doi.org/10.1073/pnas.1812826115.
42. Coates ME, Beynon JL. 2010. Hyaloperonospora *Arabidopsidis* as a pathogen model. Annu Rev Phytopathol 48:329–345. https://doi.org/10.1146/annurev-phyto-080508-094422.
43. Thomma BPHJ, Esse HPV, Crous PW, Wit P. 2005. *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic *Mycosphaerellaceae*. Mol Plant Pathol 6:379–393. https://doi.org/10.1111/j.1364-3703.2005.00292.x.
44. Finkel OM, Salas-González I, Castrillo G, Conway JM, Law TF, Teixeira PJPL, Wilson ED, Fitzpatrick CR, Jones CD, Dangl JL. 2020. A single bacterial genus maintains root growth in a complex microbiome. Nature 587:103–108. https://doi.org/10.1038/s41586-020-2778-7.
45. Wang K, Sipilä TP, Overmyer K. 2016. The isolation and characterization of resident yeasts from the phylloplane of *Arabidopsis thaliana*. Sci Rep 6:39403. https://doi.org/10.1038/srep39403.
46. Eitzen K, Sengupta P, Kroll S, Kemen E, Doehlemann G. 2021. A fungal member of the *Arabidopsis thaliana* phyllosphere antagonizes *Albugo laibachii* via a GH25 lysozyme. Elife 10:e65306. https://doi.org/10.7554/eLife.65306.
47. Bodenhausen N, Horton MW, Bergelson J. 2013. Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. PLoS One 8:e56329. https://doi.org/10.1371/journal.pone.0056329.
48. Banerjee S, Schlaeppi K, van der Heijden MGA. 2018. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol 16:567–576. https://doi.org/10.1038/s41579-018-0024-1.
49. Wieczorek AS, Schmidt O, Chatzinotas A, von Bergen M, Gorissen A, Kolb S. 2019. Ecological functions of agricultural soil bacteria and microeukaryotes in chitin degradation: a case study. Front Microbiol 10:1293. https://doi.org/10.3389/fmicb.2019.01293.
50. Mayer T, Mari A, Almario J, Murillo-Roos M, Abdullah HSM, Dombrowski N, Hacquard S, Kemen EM, Agler MT. 2021. Obtaining deeper insights into microbiome diversity using a simple method to block host and nontargets in amplicon sequencing. Mol Ecol Resour 21:1952–1965. https://doi.org/10.1111/1755-0998.13408.
51. Schloss PD. 2020. Reintroducing mothur: 10 years later. Appl Environ Microbiol 86:e02343-19. https://doi.org/10.1128/AEM.02343-19.
52. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259. https://doi.org/10.1093/nar/gkz239.
53. Faust K, Raes J. 2016. CoNet app: inference of biological association networks using Cytoscape. F1000Res 5:1519. https://doi.org/10.12688/f1000research.9050.2.
54. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303.

Chapter 4

Pairwise interactions with the native host microbiome inhibit infection by *Albugo*

- **Status** Under revision at *Nature Communications*
- **URL** https://doi.org/10.1101/2024.10.25.620230
- **Own contribution**

  Designed research (with co-authors)

  Performed the experiments (with co-authors)

  Interpreted the data

  Analyzed and visualized the data

  Wrote the first draft of the manuscript

  Final editing of the manuscript (with co-authors)

Pairwise interactions with the native host microbiome in natural *Arabidopsis thaliana* inhibit the infection of *Albugo laibachii*

Maryam Mahmoudi[1], Juliana Almario[2], Kay Nieselt[3] and Eric Kemen[1]

[1]Microbial Interactions in Plant Ecosystems, IMIT/ZMBP, Eberhard Karls University of Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany

[2]Université Claude Bernard Lyon 1, CNRS, INRA, VetAgro Sup, UMR5557 Ecologie Microbienne, F-69622 Villeurbanne, France

[3]Institute for Bioinformatics and Medical Informatics, Eberhard Karls University of Tübingen, Sand 14, 72076, Tübingen, Germany

Corresponding author:
Eric Kemen
eric.kemen@uni-tuebingen.de
Current address: Microbial Interactions in Plant Ecosystems, IMIT/ZMBP, Eberhard Karls University of Tübingen, Auf der Morgenstelle 32, 72076 Tübingen, Germany

Abstract

Plant health in terms of disease symptoms and fitness loss can be determined by the composition of the leaf microbiome. Uncovering potential mechanism that protect plants from pathogens requires not only an understanding in plant-microbe interactions but further an understanding of the relationships between microbes. *Albugo laibachii*, an oomycete pathogen and causal agent of white blister rust is commonly found in natural population of *Arabidopsis thaliana* and has been shown to significantly impact the host microbiome by reducing microbial diversity. However, little is known about specific interactions that happen within such complex microbial communities and the individual microbes that positively or negatively impact the pathogen. Here we propose that robustness of microbiome disease suppression under natural conditions relies on quantitative effects with multiple individual species being capable of disease suppression belonging to distinct phylogenetic groups. Consequently, the process of microbiome-mediated disease suppression could be likened to the cumulative result of pairwise interactions between microorganisms. This phenomenon underscore the functional redundancy present among microbes of different phylogenetic origins. To test this hypothesis, we analyzed the microbiome of *Arabidopsis thaliana* in both infected and uninfected plants in a six years study using machine learning algorithms. We

trained multiple supervised machine learning classification models to classify infected and uninfected plants based on the relative abundance of the microbiome. Furthermore, we ranked the microbes based on their importance scores and identified 80 taxa that effectively discriminate between infected and uninfected plants. For experimentally testing our hypothesis of pairwise interactions, we selected candidates from phylogenetic distinct groups of bacteria, fungi and protists to compete against *Albugo laibachii*. We found candidates that successfully antagonize the pathogen albeit with quantitative differences. In conclusion, our results based on machine learning algorithms enable us to identify functional redundancy in complex microbial communities that stabilizes protective equilibria to the causal agent of white rust, *A. laibachii*. Our findings are therefore an important stepstone to design complex probiotica that protect microbial communities from dysbiosis.

Keywords:
Leaf microbiome, infected and uninfected leaves, machine learning, plant's pathogen, microbe-microbe interaction

Introduction

Plant-associated microbiome play a crucial role in the physiology, fitness and defence mechanisms of plants against various biotic and abiotic stresses [1, 2]. In natural ecosystems, the plant microbiome is not static but rather a dynamic environment influenced by numerous factors shaping its structural composition. Therefore, exploring the dynamics of plant microbiomes is critical [3]. These contributing factors comprise geographical location, sampling time, neighbouring microorganisms, climate, and pathogens [4–8]. Therefore, to comprehend the mechanisms and microorganisms responsible for the compositional changes that take place throughout the life of a plant under varying conditions, it is critical. The complex nature of microbial communities on plant leaves makes it difficult to objectively assess the ecological role and importance of individual taxa [6]. Scientists have implemented a range of approaches to identify crucial players within leaf-associated communities, with attention on microbes that persist over time (known as core taxa) and their importance within microbial networks (known as hub taxa).

It is commonly assumed that communal assemblages of microbial species within a specific habitat can aid in maintaining important ecosystem functions. The notion of a core microbiome has been introduced [9], referring to the consistent prevalence of distinct phyla across different samples in a particular ecosystem [6, 10]. Simultaneously, the analysis of microbial networks has become a prominent technique in microbiome data analysis, providing insight into the identification and decoding of important microbial factors [11, 12]. Crucially, both

core microbiome and network analyses have demonstrated significance in identifying both beneficial and pathogenic microbes. For instance, the phyllosphere microbiome of *Arabidopsis* leaves contained *Sphingomonas* and *Methylobacterium*, which are recognised as plants' beneficial microbes, and *Hyaloperonospora* and *Albugo*, common pathogens of the *Brassicaceae* family, as core and/or hub microbes [4, 6].

Albugo laibachii, a significant hub microbe, had a significant impact on both epiphytic and endophytic bacterial colonisation [4, 6]. The impact of *Albugo* infection on alpha diversity decrease and beta diversity stabilization was recorded [4]. Additionally, recent research emphasized the direct involvement of *Albugo* in shaping the plant-associated microbial communities. This influence results from the secretion of proteins and peptides with antimicrobial properties into the apoplast [13]. If both pathogens and beneficial microbiota establish themselves as stable drivers of host microbes, it is crucial to comprehend how the host perceives its microbiome. Essential to both pathogenic and beneficial lifestyles is the ability to recognize and communicate with the host through the use of nutrients, and the ability to temporarily suppress the host defense response [14]. Disruption of the microbiota balance can lead to a dysbiotic phenotype in plants, characterized by an altered microbiota that is associated with host diseases [15].

Phyllosphere dysbiosis is characterized by features such as reduced alpha diversity, an overall increase in bacterial abundance, a transition from a Firmicutes-dominated community to one enriched in Proteobacteria, and the appearance of chlorosis/necrosis in leaf tissue [16]. In the context of dysbiosis, the concept of the "pathobiome" has been introduced, which integrates the pathogenic agents within biotic environments and applied to several pathosystems. Pathobiome analyses frequently expose the potential contribution of multiple pathogens and their followers (opportunistic microbes causing secondary infections) to severe dysbiosis. These discoveries emphasise the importance of microbial diversity in disease prevention in plants [17]. Plant pathogens, including fungi, bacteria, viruses and nematodes, result in substantial losses to agricultural production annually. To combat these pathogens, plants employ a range of strategies, notably utilizing beneficial microbes for control purposes [18].

Biological control is an effective and environmentally friendly replacement or supplement to conventional pesticides for the control of fungal and bacterial plant diseases [19]. These beneficial bacteria offer a wide range of biological control capabilities and hold great promise for agricultural applications. However, identifying and practically validating such bacteria can be challenging [18]. It is still uncertain what constitutes a healthy or beneficial microbiome. Despite being a fundamental inquiry, providing a comprehensive response remains challenging. Healthy plant communities are characterised by diverse and balanced microorganisms coexisting with the plant, which is evident when comparing the

microorganisms in healthy plants to those in diseased plants [20]. One promising way to address this issue is through the application of machine learning. Machine learning techniques have been employed in microbiome research to accomplish a wide range of tasks, such as predicting host or environmental phenotypes, categorising microbial characteristics, including measuring microbiota abundance, diversity or distribution, investigating complex chemical and physical interactions within the microbiome, and monitoring changes in microbiome composition [21]. For instance, Yuan et al. [22] employed classifiers to identify patterns in microbial communities in soil that were influenced by Fusarium wilt disease in field conditions. Similarly, the random forest method employing machine learning accurately anticipated productivity based on microbiome composition at the order level.

Significant differences in crop yield were found to be linked to the microbiome composition of bulk soil, with many taxa demonstrating importance in nitrogen utilisation [23]. In a study on the nodule microbiome of healthy and starved plants, a connection between the nodule microbiome and nodulation was discovered, which was influenced by the soil microbiome. Nodule communities were characterised using a metabarcoding technique, and significant nodular bacteria were identified through machine learning and network analyses. These bacteria were then experimentally validated by being re-inoculated onto plants under controlled conditions [24]. Another study used metadata analysis and machine learning to identify common bacterial community patterns in disease-suppressed soils. *Firmicutes* and *Actinobacteria* were increased and *Proteobacteria* and *Bacteroidetes* were decreased in these soils. The machine learning models successfully distinguished between disease-suppressive and disease-supportive soils, highlighting potentially useful genera [18].

In this study, we collected and studied the microbiome of Arabidopsis thaliana over a period of six years, including symptomatic and asymptomatic plants in the presence of *Albugo laibachii*. We used high-throughput sequencing to characterise bacteria, fungi and other eukaryotic microorganisms in the samples. Our main hypothesis suggested the presence of distinct microbial signatures that could discriminate between symptomatic and healthy states. To test this hypothesis, we used machine learning classification algorithms to identify and classify differentiating microbes in healthy and infected plants. During our investigation, we identified several potential candidate signatures within the microbiome that indicate health and infection status. These signatures came from a variety of microbial groups, including bacteria, fungi and eukaryotes. We identified microbial communities that structure microbial interaction networks in health and disease samples by analysing microbe-microbe interactions. In addition, our research investigated microbialplant interactions and the potential for health-related microbiota to mitigate the presence of pathogens. To evaluate our computational hypothesis and our analysis at a mechanistic level, we used co-inoculation assays in plant hosts to study pairwise interactions between

selected microbes and the pathogen *Albugo*. Our results confirm the significant potential of health-associated microorganisms to reduce *Albugo* infection in plants. In contrast, disease-associated microbiomes exhibited a range of functions from pathogen suppression to minimal effects on the pathogen. These findings highlight the functional redundancy of microbial communities from different phylogenetic groups in manipulating plant health and disease outcomes.

Results

Comparison of two microbiome conditions in the phyllospehre of natural Arabidopsis thaliana populations: Condition 1 infected by the causal agent of white rust A. laibachii; Condition 2 uninfected

To investigate the diversity and composition within the phyllosphere microbiome in the presence of the pathogen *Albugo laibachii*, we sampled six sites with stable *Arabidopsis thaliana* populations in Tu¨bingen [4, 25]. White rust, which is caused by *Albugo* sp., was a very common phenotype at most of the sites. Two seasons, autumn and spring, were used to collect leaf samples. The autumn samples represented the early growth phase of *A. thaliana* under short day conditions before entry into winter dormancy. The spring samples were collected just before the reproductive stage, during the increasingly long days of the [25]. Based on whether or not white rust was observed on the plants collected, we categorised the samples. During spring, infected plants were typically observed. However, there is no significant difference in the relative abundance of *Albugo* infected plants between spring and autumn. This suggests a dormant phase of *Albugo* prevalence during the colder period in autumn. This sampling was repeated for six consecutive years (Fig. 1A; see also Supplementary Material, Table S1). Genomic DNA was extracted from both epiphytic and endophytic microbes for each sample [4]. Amplicon sequencing was performed for the 16S bacterial, ITS2 fungal and 18S eukaryotic data [25]. For subsequent analysis, only the endophytic samples were used and all microbes belonging to the kingdom Fungi were removed from the 18S eukaryotic data as they were analysed separately. We refer to the 18S data as 'other eukaryotes' in the remainder of the paper to highlight that fungi are excluded.

We performed diversity analyses to assess the microbial communities of bacteria, fungi and other eukaryotes between infected and uninfected leaves. The within-sample diversity of bacterial and other eukaryotic communities was significantly lower in *Albugo* infected plants compared to uninfected plants ($P < 0.05$ for bacteria and ($P < 0.001$ for other eukaryotes), but no significant differences were observed in fungal communities (Fig. 1B). In both infected and uninfected groups, the average bacterial diversity was the highest (2.9 in uninfected, 2.6 in infected), followed by fungi (1.7 in uninfected, 1.6 in infected), and other eukaryotes (1.4 in uninfected, 0.4 in infected) (Fig. 1B). To assess the diversity between samples, we employed multivariate approaches including principal coordinate analysis (PCoA) and permutational multivariate analysis of variance (PERMANOVA). The principal

component analysis revealed a separation between infected and uninfected samples in bacterial and other eukaryotic communities, whereas the separation was less distinct in fungal communities (Fig. 1C). The PERMANOVA results demonstrated that infection stages explained 1.8% of the variation in bacteria, 0.5% in fungi, and 10.9% in other eukaryotes (Table. 2). Furthermore, we analysed the the relative abundances of highly abundant microbial taxa at the order level, as taxonomical orders often represent functionally related groups of microorganisms. The analysis revealed distinct compositions in infected and uninfected leaves over sampling years (Fig. 1D). The aggregate relative abundances indicated significant differences in 36% of the taxa. Uninfected bacterial samples exhibited an increase in *Sphingomonadales, Bacillales*, and *Enterobacterales*, along with a decrease in *Pseudomonadales*. Infected fungal samples showed an increase in *Cystofilobasidiales*. In uninfected other eukaryotes, there were increases in *Chlamydomonadales*, *Cryomonadida*, *Glissomonadida*, *Watanabea-Clade*, and *Xanthophyceae* along with a decrease in *Oomycota* (Fig. S1).

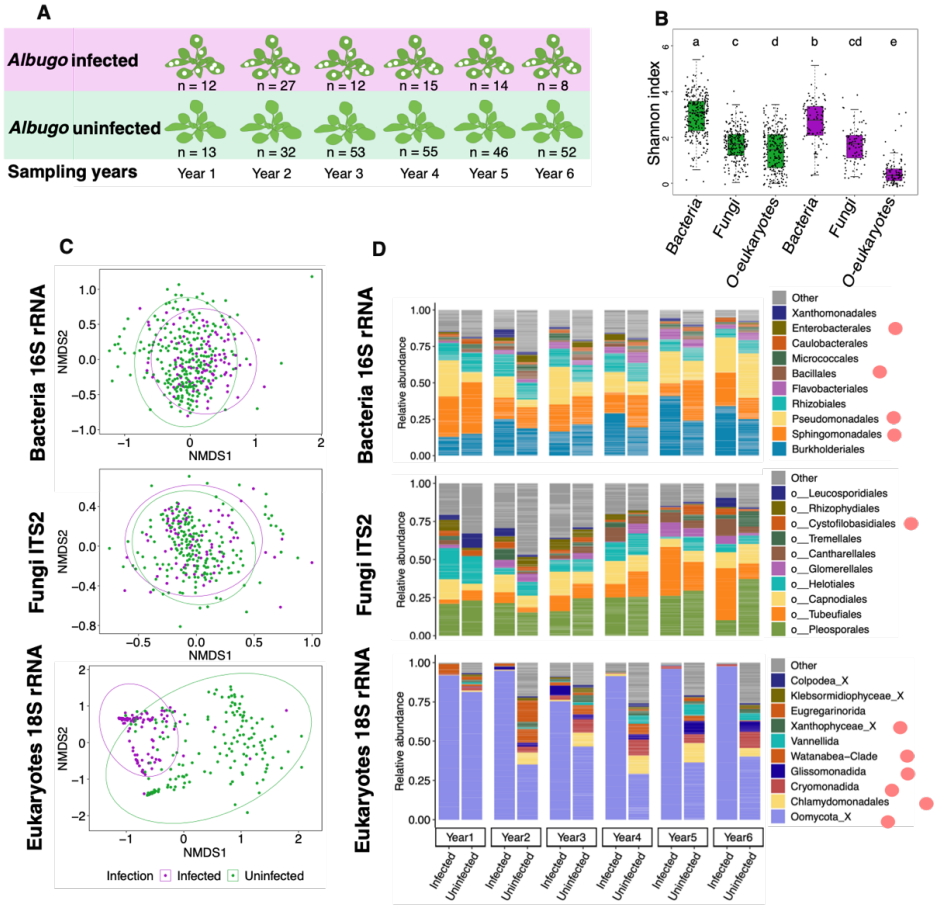


Figure 1. Diversity and composition of leaf microbial communities in infected and uninfected leaves. (A) The annual number of samples from infected and uninfected

plants. (B) Alpha diversity, measured by Shannon's H index, representing the variation in microbial diversity between infected and uninfected samples in bacteria, fungi and other eukaryotes. (C) Separation of infected (purple) and uninfected (green) samples in bacteria, fungi and eukaryotes using principal component analysis (PCoA) based on Bray-Curtis dissimilarities. Each of the dots represents a single sample. (D) Histograms show the relative abundance of bacteria, fungi and other eukaryotic communities at the order level, categorised by plant infection stage over six years for each microbial group. The most abundant microbial orders of each taxon are indicated. Circle indicate taxa showing significant differences between infection stages (see Fig. S1). Different letters indicate statistically significant differences between groups (Tukey HSD's test).

Microbial signature for predicting infected and uninfected leaves using machine learning models

Our hypothesis was that the plant microbiome contains distinct microbial signatures that can serve as robust indicators of infection, enabling accurate discrimination between infected and uninfected samples. To investigate this hypothesis, we utilized machine learning classification models, such as Random Forest (RF), Support Vector Machine (SVM), Multi-layer Perceptron (MLP), and Logistic Regression (LR), which are well-known for their interpretability (Fig. 2A). The training phase used 70% of the sample set, consisting of 169 uninfected and 66 infected samples, while the remaining 30% (73 uninfected and 29 infected samples) served as the test set to assess predictive performance. Four different evaluation metrics were employed, resulting in accuracies ranging from 75% to 86% (Fig. 2C). It is worth noting that LR obtained the highest accuracy at 86%, with an area under the curve (AUC) of 94% (Fig. 2B). SVM, MLP, and RF were evaluated and exhibited differing performance. RF achieved the lowest accuracy with a range of 57-75% across all measured metrics (Fig. 2B and C).

The potential of the microbiome as a biomarker for the discrimination of infected from uninfected samples and for the characterization of microbes for synthetic communities was then investigated. The role of each microbe was analyzed by calculating and comparing the feature importance of all 2543 OTUs in the trained classification models (SVM, RF, and LR) (Fig. 2A). After comparing the three models, we found that they shared 2253 OTUs, indicating consistent microbial signatures that contributed significantly to the classification process (Fig. S2A). To determine the most important microbes for classification, we used recursive feature elimination with cross-validation (Fig. 3A). The results showed that RF had the highest accuracy, achieving 91% accuracy with 40 selected features (Fig. 3B). When the experiments were completed, LR reached 87% accuracy with 4 selected characteristics (Fig. 3B), while SVM reached 86% accuracy with 53 selected characteristics (Fig. 3B). It is interesting to note that four OTUs were shared by all three models that originated from LR, and 13 OTUs were shared by SVM and RF (Fig. 3B and Fig. S2B).

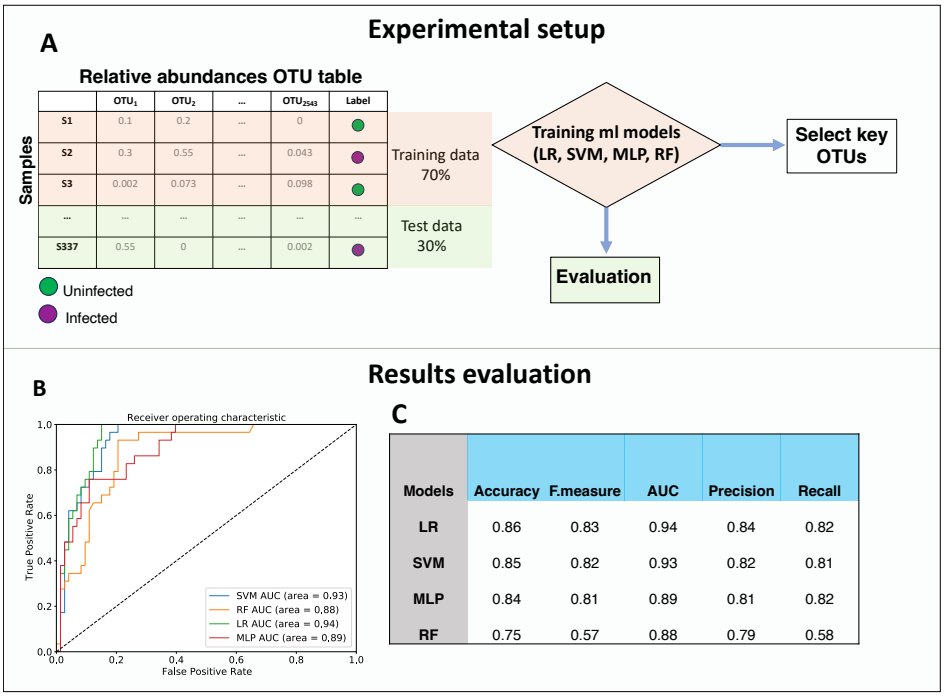


Figure 2. Classifying infected samples from uninfected using machine learning classifiers. (A) Workflow demonstrating the methodology employed to classify infected samples from uninfected using machine learning models. Four machine learning classifiers, namely Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), and Multilayer Perceptron (MLP), were trained on 70% of the sample. The goal was to classify uninfected leaves from infected leaves based on observed symptoms of *Albugo*, utilizing the relative abundance OTU table of bacteria, fungi, and other eukaryotes. The trained models with 70% of samples were evaluated using the remaining 30% of the dataset. Feature importance was extracted from the trained models. (B) Area Under Curve (AUC) plots representing the performance of the classifiers on the test set. (C) The table presents the accuracy, F-measure, precision, and recall metrics for each classifier on the test set.

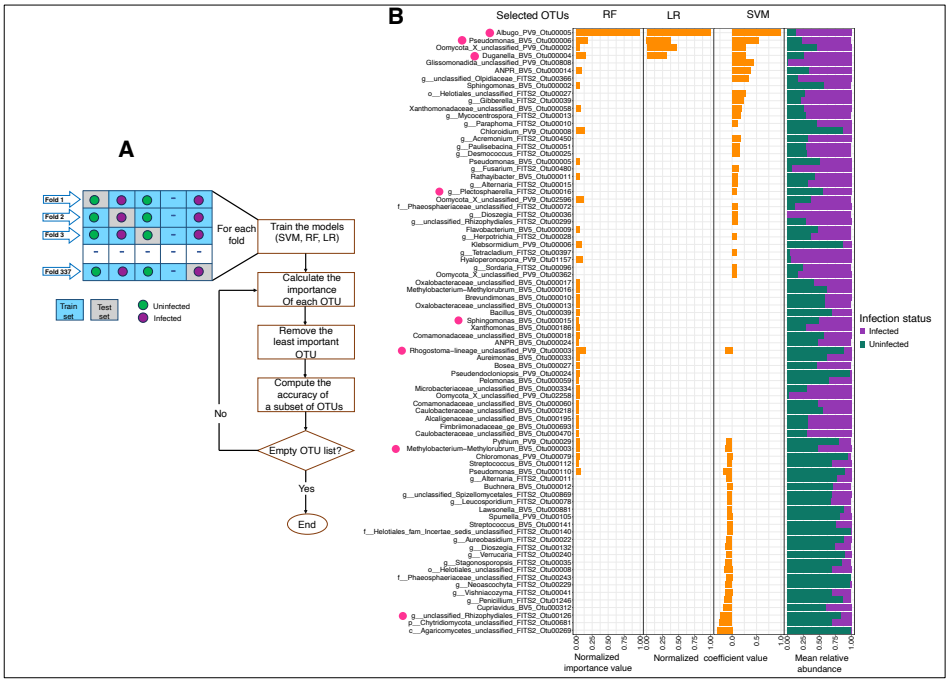


Figure 3. Selected microbes as indicators of uninfected and infected samples.

(A) Recursive feature elimination with k-fold cross-validation (k = total number of samples) was performed using random forest (RF), support vector machine (SVM), and logistic regression (LR) to identify indicators of infected samples from uninfected samples. (B) Histogram showing the normalized coefficient values of the 53 microbes selected for recursive feature elimination with k-fold cross-validation using SVM. (C) The 40 OTUs selected by recursive feature elimination with k-fold cross-validation using RF. (D) The 4 selected OTUs after recursive feature elimination with k-fold cross-validation using LR. Negative values (green bars) indicate OTUs with high scores in discriminating uninfected leaves, while positive values (purple bars) indicate high scores in discriminating infected samples. Bar plots illustrate the percentage of aggregated relative abundances of OTUs in infected and uninfected samples. Significance values indicate the differences in OTUs between infected and uninfected samples as determined by the Wilcoxon test (ns: p value > 0.05, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001). Highlighted microbes identified by recursive feature elimination with k-fold cross-validation were selected for further experimental analysis.

Microbe-microbe interactions in infected and uninfected plants

Microbial networks are a useful resource for identifying potential interactions among microorganisms in a given community, derived by correlating the abundances of different species. Our goal was to investigate the microbial networks present in both infected and uninfected samples. Surprisingly, the network resulting from the uninfected samples showed a larger number of nodes (OTUs) and edges (connections between OTUs) compared to the one generated from the infected samples (2024 nodes and 73511 edges vs. 1058 nodes and 22089 edges, respectively) (Fig. 4B vs. Fig. 4A). Notably, the topological characteristics of these networks showed higher degree and closeness centrality values ($P < 0.001$) in the uninfected network (Fig. 4E and 4F). We compared the modularity of the constructed networks to assess the effect of infection on the microbial communities in the networks. Both networks had the same number of

modules (seven). However, the modularity value of the network derived from infected samples (0.31) was higher than that of the network derived from uninfected samples (0.23) (Fig. 4A and 4B). Each module contained OTUs from different taxonomic categories such as bacteria, fungi, and other eukaryotes (Fig. 4C and 4D).

In addition, we examined the interactions between the OTUs that were specifically selected as discriminators between infected and uninfected samples (Fig. 3). In contrast to the network generated from infected samples, which contained 64 nodes and 193 edges (Fig. S3A), these OTUs exhibited a greater number of connections within the uninfected network, resulting in 72 OTUs with 493 edges (Fig. S3C). Furthermore, these OTUs were spread across modules and taxa, including bacteria, fungi and other eukaryotes (Fig. S3B and S3D). Notably, in contrast to the average correlation of 0.21 in the uninfected network, an average correlation of 0.30 was observed between these OTUs in the network of infected samples.

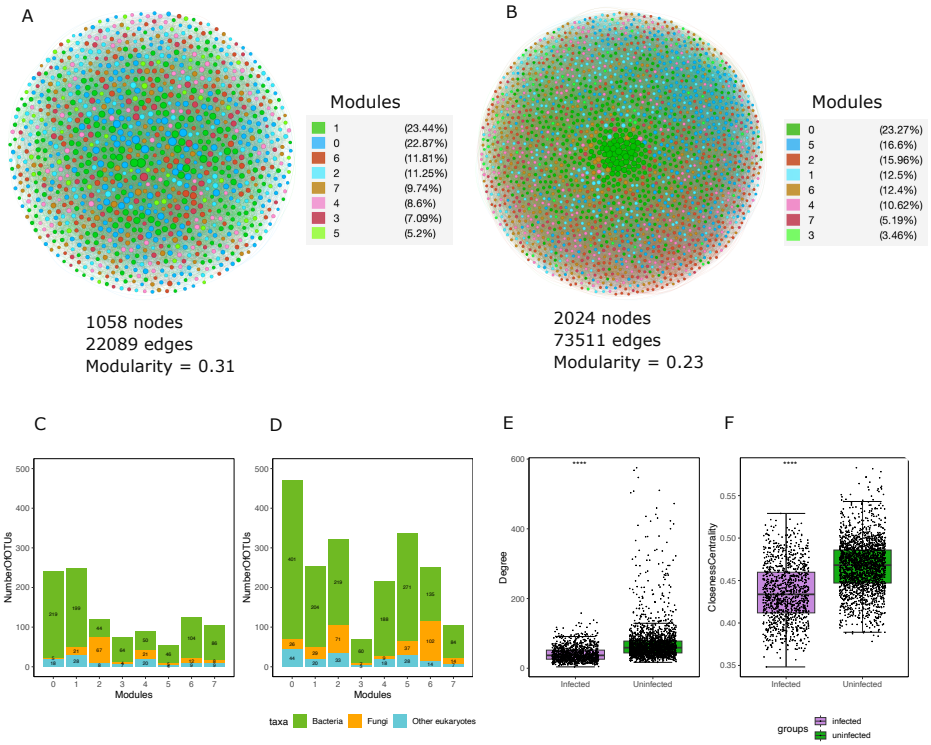


Figure 4. Infected and uninfected leaf changes in microbial co-occurrence networks. We have reconstructed co-occurrence networks for both infected (A) and uninfected (B) samples, where nodes represent OTUs and edges indicate correlations between these OTUs. Nodes are color-coded according to modularity. Histograms (C and D) illustrating the distribution of OTUs within modules for the network of infected and uninfected samples, respectively. These histograms are further color-coded to distinguish microbial taxa, with green representing bacteria, orange representing fungi, and blue representing other eukaryotes. Box plots show features per node degree (E) and closeness centralities (F) for comparison between infected and uninfected networks. Whiskers indicate the dispersion of the data (1.5 x interquartile range). Significance values indicate

differences in microbes between seasons based on the Wilcoxon test (ns: p-value > 0.05, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).

Pairwise interactions between microbial signatures and pathogen *Albugo* through in planta assays

Our study was guided by the hypothesis that specific microbes associated with uninfected plants could have a beneficial effect by reducing *Albugo* pathogenicity, whereas microbes associated with infected plants could have a detrimental effect by either amplifying infection or potentially reducing it by exploiting the pathogen for survival, resulting in increased compositional proliferation. To test this hypothesis, we chose four candidate microbes from bacterial, fungal, and other eukaryotic groups in each category (infected and uninfected) based on the outcomes of recursive feature elimination using all three models (Fig. 3A). The selection criteria included diverse representations of phylogenetic trees, varying relative abundances, laboratory availability, and the use of different machine learning algorithms for feature selection (Fig. 3). SVM and LR provide positive and negative coefficients to determine the microbial importance in categorizing infected versus uninfected leaves. We assigned scores based on these values, with negativity indicating the uninfected class and positivity representing the infected class (see Fig. 3B). However, RF only offers positive scores, necessitating further examination to determine the relationship between the selected OTUs and their respective classes (see Fig. 3B). Among the candidate microbes associated with uninfected plants, MethylobacteriumOTU3 (*Methylobacterium goesingense*) and SphingomonasOTU15 (*Sphingomonas melonis*) represented the bacterial candidates, RhizophydialesOTU126 (*Cystofilobasidium macerans*) for fungi, and RhogostomaOTU3 (*Rhogostoma ephyphila*) for other eukaryotes. In contrast, the microbes identified as candidates linked to infected plants were DuganellaOTU4 (*Duganella zoogloeoides*) and PseudomonasOTU6 (*Pseudomonas viridiflava*) from the bacterial group, PlectosphaerellaOTU16 (*Plectosphaerella niemeijerarum*) as the fungal candidate, and AlbugoOTU5 (*Albugo laibachii*), representing the other-eukaryotic microbes (Fig. 3B).

To investigate the protective effects of the chosen microbes against the pathogen, a mixture of *Albugo* and each of the four microbes associated with uninfected plants were administered to *Arabidopsis* leaves (Fig. 5A). The level of protection was determined by measuring the percentage of infected leaves (Fig. 5B). The findings revealed that all the candidate microbes significantly decreased the infection caused by *Albugo* ($P < 0.001$). *Cystofilobasidium* had the most significant effect, reducing *Albugo* levels by an average of 73%. Sphingomonas came in second with a 65% decrease, followed by *Methylobacterium* and *Rhogostoma* with 39% and 38% decreases, respectively (Fig. 5A). These observations were confirmed by qPCR analysis that showed samples exposed to the uninfected-associated microbes had substantially lower amounts of *Albugo* compared to

control samples (P < 0.05) (Fig. 5C). The candidates reduced the average biomass of *Albugo*, ranging from 87% to 51%. This reduction was achieved with *Cystofilobasidium*, *Methylobacterium*, *Sphingomonas*, and *Rhogostoma*, respectively. We then investigated whether microbes associated with infections were capable of promoting or alleviating the infection. The in planta infection assay demonstrated no significant variations in the infection level of *Plectosphaerella*, in contrast to the *Albugo* control (Fig. 5E). However, *Pseudomonas* showed a decrease of 50% and *Duganella* showed a decrease of 34% in infection (P < 0.05) (Fig. 5E). Likewise, qPCR outcomes substantiated phenotype observations, where the biomass of *Albugo* in the control group did not exhibit significant changes in comparison with *Plectosphaerella*. However, *Pseudomonas* and *Duganella* showed significant differences with a 77% and 65% reduction in *Albugo* biomass, respectively, compared to the control group (Fig. 5F).

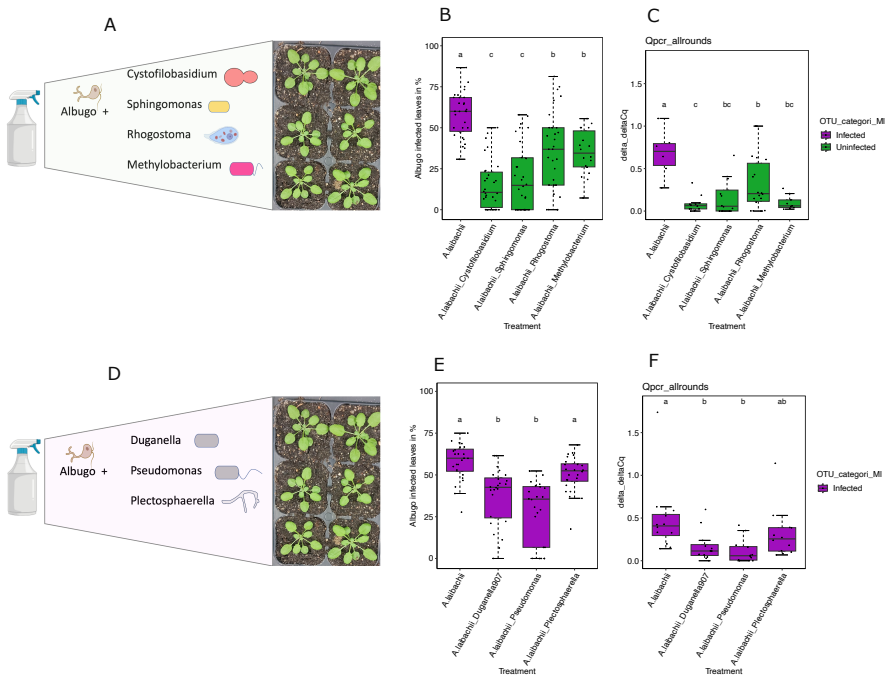


Figure 5 shows the infection assay performed using *A. laibachii* on *A. thaliana* in the presence of both infected and uninfected indicator microorganisms. The selected indicator microorganisms were chosen on the basis of their relevance to both infected and uninfected sample groups, as indicated in Figure 3. For each of these indicator microorganisms, co-inoculation with *Albugo* was performed on *Arabidopsis* plants at 3-4 weeks of age. (B) Illustrates the percentage of leaves infected with *Albugo* in the presence of healthy indicator microorganisms, while (E) presents the percentage of Albugo-infected leaves in the presence of infected indicator microorganisms. Each dot on the graph reflects the percentage of infected leaves in an individual sample. The experiment comprised of six individual replicates, each with three technical replicates. Relative quantification of *A. laibachii* biomass in response to uninfected and infected indicator microbes, respectively, was conducted through qPCR. To allow quantification of the amount of *A. laibachii* DNA in the samples at 14 days post-infection, the oomycete internal transcribed spacer (ITS) 5.8S was

normalized to the A. thaliana EF1-α gene. The relative biomass was then calculated via the ddCT method. The three graphs displayed herein illustrate three independent replicates, each with three technical replicates. Statistically significant differences between groups were evaluated using the Tukey's Honestly Significant Difference (HSD) test, with distinct letters denoting significant discrepancies.

The impact of the individual microbial signatures on the plants

Spray experiments were conducted on gnotobiotic plants to evaluate the impact of individual microbes on plant health and pathogenicity (Fig. S4). Notably, *Duganella* and *Pseudomonas* had a detrimental effect, resulting in the death of all plants within three weeks. *Duganella* visibly deteriorated plant health, leading to complete plant death, while *Pseudomonas* only allowed 9.09% of treated plants to survive. In contrast, *Plectosphaerella* had a less severe but still noteworthy impact, with 70% of the treated plants surviving. Nevertheless, both leaves and roots showed significant harm as evidenced by browning. Although *Cystophilobasidium* is associated with plant health, it exhibited a lower survival rate, with just 33.33% of plants remaining viable. In contrast, the plants that were treated with health-associated microbes (*Sphingomonas*, *Methylobacterium*, *Rhogostoma*) thrived, with over 90% of them surviving and presenting no noticeable negative consequences (Fig. S4).

Discussion

Natural probiotics for managing plant health have gained ecological and economic attention. However, a comprehensive understanding of pathogens in the phyllosphere and the progression of disease in a community context is currently lacking [26]. In natural populations of *Arabidopsis thaliana*, *Albugo laibachii* has been identified as a hub microbe causing white blister rust with significant effects on epiphytic and endophytic bacterial colonization [4]. Albugo's growth patterns, mainly observed during the spring season and found in half of the sampled locations, highlight the influence of geographical locations and environmental conditions on its colonization dynamics [4]. It is unclear what distinguishes the microbiome of infected versus uninfected plants and which specific microbes either promote or suppress the growth of the *Albugo* pathogen. This natural pathosystem presents an opportunity to enhance disease management strategies.

To further explore this matter, an extensive study was performed on the diversity of microbial communities in leaves affected by *Albugo*, both infected and uninfected. Utilizing a detailed microbiome dataset spanning six years, we analyzed microbial taxa at a more resolved level, including bacteria, fungi, and other eukaryotes [25]. The results of the analysis revealed a significant difference between infected and uninfected plants. Infected plants showed lower diversity. In particular, diversity levels within the other eukaryotic groups were found to be lower compared to fungi and bacteria (Fig. 1). This finding aligns with our previous discovery that *Albugo* is a hub microbe with more negative interactions in the microbial network of phyllosphere [4, 13]. The variation in diversity among

different microbial groups, such as bacteria, fungi, and other eukaryotes, emphasizes the importance of these groups in preserving plant health. It has been postulated that reduced diversity in infected plants may be due to the influence of hub microbes, which can indirectly affect diversity through interactions with the host or directly impact the efficiency of other microbe colonization [4]. Our recent research has revealed that *Albugo* plays a direct role in shaping plantassociated microbial communities by releasing proteins and peptides with antimicrobial properties into the apoplast [13]. The reduction in diversity can be attributed to various factors, including the enrichment of pathogenic microorganisms and the proliferation of microorganisms capable of suppressing pathogens [22, 27]. Consistent with this principle, our results provide compelling evidence for a increase in 36% order-level microbial community abundance in both infected and uninfected plants (Fig. 1 and Fig. S1). These results shed light on important microbial dynamics and provide valuable insights into the complex interactions that impact plant health and disease progression.

Our objective was to differentiate microbiome patterns between infected and uninfected leaves, ultimately identifying the microbiome as a potential biomarker for plant health and disease. We utilized machine learning techniques, which have demonstrated successful applications in clinical and plant disease prediction scenarios [21]. Using prediction models, we conducted a systematic analysis of variations in the health and infection stages of plants. Our results revealed highly accurate classifications, achieving a range of 75% to 86% accuracy (Fig. 2), which suggests the existence of predictive patterns among the different groups. To identify the critical microbial players that contribute to these predictive models, we employed feature selection, a popular technique in machine learning. This method enabled the identification of 3.1% of OTUs, including bacteria, fungi, and eukaryotes, that have the potential to act as discriminators (Fig. 3).

Microbial networks have gained popularity as a useful tool for analyzing microbial community structure due to their ability to integrate various types of information and potentially represent system-level behavior [12]. The hypothesis of this study is that pathogen infection can have a significant impact on the complexity of microbial interactions. Our findings indicate that uninfected constructed networks display a heightened level of complexity, which is evident through an increased number of connections and OTUs (Fig. 4). As individuals navigate through an ecological network, they encounter different densities of connections, with regions of dense connections referred to as 'modules' and regions of sparse connections defining their boundaries. Modules suggest stronger links between species within their boundaries than between species in other modules, and this organization is referred to as the 'modularity' of the network. Modularity can indicate diverse habitats, varying selective pressures, and phylogenetic clustering of related species [28]. Our study demonstrates that our networks display modularity (Fig. 4). Additionally, we find that infected networks have higher modularity than uninfected networks, bolstering the notion that increased modularity is connected to greater network stability (Fig. 4) [29].

Our analysis reveals that specific microbes, which have been identified as discriminators between infected and uninfected samples (Fig. 3), are distributed across various modules, indicating their crucial role in the functionality of different community modules within the overall microbial community (Fig. S3). The plant microbiome has a critical role as a defense shield against plant pathogens, utilizing direct and indirect interactions through various mechanisms, including parasitism, competition for resources, and induction of systemic resistance in response to biotic and abiotic stresses [4, 7, 30]. Building upon this understanding, our study aimed to investigate the potential roles of selected microbes in microbe-microbe interactions.

We hypothesized that the plant-associated microbiome, alone or from distinct phylogenies, could provide protection against the pathogen *Albugo laibachii*. Our findings demonstrate that all candidates exhibited protective effects, resulting in infection reductions ranging from 87% to 51%. Specifically, the order of effectiveness was *Cystofilobasidum*, followed by *Methylobacterium*, *Sphingomonas*, and then *Rhogostoma*. The *Cystophilobasidum* fungi exhibited the strongest protective effects among the candidates (Fig. 5A-C). This discovery is particularly intriguing due to the limited information available about the role of *Cystophilobasidium* in protecting *Arabidopsis* leaves from pathogens. Nonetheless, similar studies on other *Basidiomycete* fungi like the epiphytic yeast *Moesziomyces bullatus ex Albugo* have offered valuable insights into their antagonistic behavior against *Albugo laibachii* in *Arabidopsis*. Through a comprehensive approach involving transcriptomics, reverse genetics, and protein characterization, this antagonism was deciphered, highlighting the significance of a GH25 hydrolase with lysozyme activity as a crucial effector [31]. Interestingly, the bacterial candidates we analyzed are well-known core components of the phyllosphere microbiome [6, 32]. The protective capabilities of *Sphingomonas* species, established in prior research as protecting *Arabidopsis* against other pathogens such as *Pseudomonas syringae*, are consistent with our findings [2]. Nevertheless, colonization by *Methylobacterium* genus members yielded varying effects and did not provide plant protection. However, we observed intriguing effects that decreased the pathogenicity of *Albugo*.

The recent focus on protists, specifically *Rhogostoma spp* from the phylum *Cercozoa*, in *Arabidopsis thaliana* is noteworthy [33, 34]. These protists have been identified as influential agents that shape the composition and actions of bacterial and fungal communities through selective predation [33, 34]. Protists act as natural microbial predators. They directly target bacterial and fungal pathogens, resulting in a consuming effect that ultimately leads to the removal of various pathogenic strains [35, 36]. Our findings emphasize the potential of fungi, followed by bacteria, and least of other eukaryotes, in reducing pathogenicity by over 50%. This emphasizes that single microorganisms from various phylogenetic origins demonstrate functional redundancy, which offers favorable prospects for utilizing them in biocontrol methods to efficiently reduce plant infections.

Expanding upon this observation, we propose that in the context of infected plants, the presence of pathogens may lead to a dual response: the pathogen can either establish its unique microbial community, or stimulate the growth of other microbial cohorts. These additional microorganisms can work together to suppress or regulate the pathogen's propagation. The complex interplay of microbial interactions within the plant ecosystem is highlighted through this study, emphasizing the potential for intricate responses to pathogen-induced perturbations.

Our investigation into disease-associated microbes indicated no significant reduction in the infection caused by selected fungi Plectospherella (Fig. 5E and 5F). Interestingly, this species also showed no promotion of infection. The significant reduction in infection by 77% for *Psudomonas* and 65% for *Duganella* deserves attention (Fig. 5E and 5F). According to studies, commensal *Pseudomonas* can activate a host response that specifically inhibits a specific pathogenic lineage, leading to plant protection. This highlights the ability of Pseudomonas to demonstrate a range of pathogenic and commensal behaviors [37]. Indeed, the results of our study reveal that infected-associated microbes have different effects on *Albugo* infection levels, which may be attributed to the phenomenon of mutual infectivity relationship. An instance of such a correlation has been discovered in *Arabidopsis thaliana*, where initial exposure to the pathogen *A. laibachii* generates a setting that promotes growth and spore production of avirulent strains of the *Arabidopsis* downy mildew pathogen *Hyaloperonospora arabidopsidis* on *Arabidopsis* accessions that would otherwise be resistant to this pathogen [38].

A "recolonization experiment" was performed to determine the distinct effects of each candidate [39]. In this experiment, we introduced isolates to sterile plants and evaluated their impact on plant health (Fig. S4). Our observations yielded interesting results: the plants inoculated with health-associated microbes showed a significantly higher level of healthy leaves compared to those associated with disease-associated microbes. Notably, the experiment yielded intriguing results. Inclusion of *Duganella* and *Pseudomonas* strains led to remarkable outcomes. The strains have potential pathogenic properties as nearly all exposed plants died, particularly in the case of *Duganella* and a significant proportion (91%) in the case of *Pseudomonas* (see Fig. S4). The function of these strains in their natural environment poses interesting questions. These strains may strategically modulate their pathogenicity, potentially adjusting their virulence levels or engaging in complex interactions that contribute to their pathogenic behavior. Such behavior presents an intriguing aspect for studying microbial communities and their impact on plant health. In addition, *Plectosphaerella* caused 25% damage to whole plants, extending to the roots as evidenced by brown discoloration. Compared to *Duganella* and *Pseudomonas*, *Plectosphaerella* has a different mode of action. These results underscore the intricate and dynamic interplay between microbial communities and their indispensable role in plant health and growth outcomes. Further investigations

into the interactions between these strains and plants and other microorganisms could yield important knowledge related to plant health management in agricultural or natural settings. It should be highlighted that *A. thaliana* roots inherently contain fungi that boast considerable pathogenic potential, such as *Fusarium* and *Plectosphaerella* [40]. This finding indicates that these fungi could benefit from associations with other microorganisms or may play alternative roles that contribute to their presence within the plant ecosystem.

To summarize, our study employed classification machine learning to analyze the long-term microbiome of *Arabidopsis* in infected and uninfected plants. We effectively identified unique microbiome patterns associated with plant health and infection, including bacteria, fungi, and cercozoa. Among the identified microbes, we discovered evidence of health indicator microbes that had a positive impact on plant resistance against *Albugo* infection. Additionally, we examined the impact of disease-associated microbes and observed that some decreased infection while others had no significant effect. These findings underscore the complexity of microbe-microbe interactions in the plant ecosystem, emphasizing the importance of further investigation to gain a complete understanding of these intricacies.

Conclusion

Our study has provided valuable insights into the diversity, variation, and composition of the leaf microbiome in both infected and uninfected plants, revealing the potential roles of specific microbes in disease dynamics. The discovery of unique microbial signatures associated with plant health and infection has promising implications for developing strategies to maintain plant health and mitigate the effects of plant diseases. However, the findings indicate the complexity of microbe-microbe interactions and the need for further investigation to fully understand the underlying mechanisms involved in plant-pathogenmicrobiome interactions. Comprehending these mechanisms will be vital to developing innovative techniques to efficiently maintain plant health and fight diseases in both agricultural and natural ecosystems. Expanding our understanding of plant-microbe interactions and the role of the microbiome in disease resistance has the potential to develop innovative and sustainable approaches for crop protection and disease management. Future research should prioritize unraveling these interactions and exploring new opportunities to utilize the plant microbiome for improved plant health and productivity.

Method

Sample collection of wild *Arabidopsis thaliana*

Wild *Arabidopsis thaliana* samples were collected from six locations in the Tuebingen area over multiple years, including autumn and spring seasons from 2014 to 2019 (11 distinct time points in total, as presented in Table S1). The methodology for obtaining both epiphytic and endophytic microorganisms from each leaf sample followed the procedures described in Aglar et al. [4]. Leaves were washed gently with water for 30 s, followed by a 1 min wash in 3-5 ml epiphyte solution (consisting of 0.1% Triton X-100 in 1x TE buffer). Epiphytic microorganisms were collected by filtering the solution through a 0.2 mm nitrocellulose membrane filter (Whatman, Piscataway, NJ, USA). The filter was then inserted into a screw-cap tube and frozen on ice for storage. To obtain endophytic fractions, the same leaves underwent surface sterilization, which consisted of a 15-second wash with 80% ethanol, followed by a 30-second treatment with 2% bleach (sodium hypochlorite). Then, the leaves were rinsed three times with sterile autoclaved water for 10 seconds. Afterward, the samples were placed in a screw-cap tube and immediately frozen using dry ice. Phenol-chloroform-based protocol was used for DNA extraction and amplicon sequencing, following the customized procedure stated in Agler et al. [27]. The extracted DNA was utilized in a two-step PCR amplification that targeted the V5-V7 region of bacterial 16S rRNA (employing primers 799F/1192R), the ITS2 region of fungi (utilizing primers fITS7/ITS4), and the 18S rRNA region of eukaryotes (with primers F1422/R1797). Using blocking oligos helped to minimize the amplification of plant DNA. Afterwards, the PCR products were purified and combined in equal proportions before undergoing sequencing in Illumina MiSeq runs (600-cycle) with PhiX control.

Analysis of amplicon sequencing data

The amplicon sequencing data was processed using Mothur (Version=1.42.3) [41], following the guidelines described in Mahmoudi et al. [25]. The initial step involved merging singleend reads to generate paired-end reads, while pairs with less than 5-base overlap between the forward and reverse reads were excluded. Only sequences with lengths ranging from 100 to 600 bases were considered for further analysis. Chimeric sequences were identified using Vsearch within the Mothur environment, with more abundant sequences serving as reference. Cutadapt 2.10 was then utilized to refine the dataset by removing adapter sequences from 16S rRNA and 18S reads. This was especially critical in the case of short amplicons that were fully sequenced, as it permitted retrieval of matching pairs and addressed ”orphan” reads. For fungal reads, we utilized ITSx 1.1b [42] to trim sequences to the ITS2 region. The default parameters were utilized, with the exceptions of preserving sequence headers, performing checks against fungal, oomycete, and plant profiles, allowing single-domain matching with an e-value cutoff of 1e-5.0, allowing matching of only one HMM gene profile, and allowing preservation of partial ITS regions.Subsequently, the sequences

were categorized into Operational Taxonomic Units (OTUs) based on a 97% similarity threshold, utilizing the VSEARCH program incorporated within Mothur. The taxonomic classification of each sequence used the rdp classifier with a consensus confidence threshold of 80 and method=wang. The Silva database (version 138.1) was used for 16S rRNA data, the UNITE public database (version 02 02 2019) for fungal ITS2, and the Pr2 database (version 4.12.0) for eukaryotes. To improve identification of the remaining PhiX reads, the PhiX genome was incorporated into each of these databases. Each operational taxonomic unit (OTU) underwent taxonomic classification with a consensus confidence threshold of 80. For analysis, OTUs lacking known taxonomy at the kingdom level and those with low abundance (¡50 reads) were excluded.

Diversity analysis

To determine the alpha diversity of the samples, the OTU abundance tables were analyzed using Shannon's H diversity index. To evaluate sample alpha-diversity, we analyzed OTU abundance tables with the Shannon's H diversity index using the 'estimate-richness' function in the 'phyloseq' package for R. We assessed normality of data using ShapiroWilk's test before analysis. For mean comparisons, we utilized non-parametric tests for two groups (Wilcoxon Rank Sum and Signed Rank Tests) and for multiple groups (Dunn's test with Bonferroni correction). Significance was determined using a P-value threshold of ¡0.05. Next, we transformed relative abundance tables using a logarithmic function (log10(x+1)) and calculated Bray-Curtis dissimilarities between samples to perform Principal Coordinate Analysis (PCoA). Our analysis utilized the 'ordinate' function from the 'phyloseq' package. To investigate further, we performed a PerMANOVA analysis on the Bray-Curtis dissimilarities to determine the "infection stages" that impact the structure of the leaf microbiome. We used the "Adonis2" function from the "Vegan" package with 10,000 permutations and a significance threshold of P ¡ 0.05. All analyses were performed using R version 4.1.2.

Machine learning analysis

We combined the OTU tables containing bacteria, fungi, and eukaryotes to comprehensively study the interactions between microorganisms. Afterwards, we converted the OTU table into a relative abundance table. To ensure durability, the OTU tables underwent a filtering process which only kept those OTUs that were present in at least 5 samples with an OTU read count over 10. The curated dataset utilized binary labels "infected" and "uninfected" to designate samples in the constructed relative frequency table. Subsequently, machine learning models were trained, including Support Vector Machine (SVM), Random Forest (RF), multilayer perceptron (MLP) and Logistic Regression (LR), using scikit-learn packages in the Python programming environment [43]. To clarify the importance of individual OTUs in distinguishing infected from uninfected samples, we utilized recursive feature elimination, a method known for identifying the most influential

OTUs for the classification task at hand. This process was executed through the RFECV function from the scikit-learn packages in Python [43].

Microbial network analysis and properties

To conduct a comprehensive examination of microbial interactions, the OTU tables, which encompass bacteria, fungi, and other eukaryotes, were merged. The merged tables were utilized to calculate correlations with the SparCC algorithm [44], which utilizes Aitchison's log-ratio analysis and is specifically designed to handle compositional data exhibiting high sparsity. Before computing correlations, the OTU tables were filtered. Only operational taxonomic units (OTUs) with a minimum of 5 sample occurrences, each exhibiting an OTU raw abundance of more than 10 reads, were retained. Afterwards, SparCC correlation scores were computed on the FastSpar platform [45] for these filtered OTU tables with default parameters. 1000 bootstraps generated the pseudo P-values, which allowed for statistical significance assessment. For further analysis, only correlations meeting the criteria of P ¡ 0.01 and an absolute correlation value greater than 0 were included. These chosen correlations served as the basis for subsequent investigations. Modularity analysis was performed using Python's networkx package (v3.1). Community detection was applied using Louvain Community Detection Algorithm [46] before calculating the modularity score. To visualize and analyze the complex microbial interaction network, we used Cytoscape (version 3.7.1) and Gephi (version 0.10.1) software. This platforms enabled the development of a visual rendering of the microbial network, assisting in the exploration and comprehension of intricate associations among diverse microbial entities.

Infections of *Arabidopsis thaliana* leaf and quantification of *Albugo* biomas by qPCR

Overnight liquid cultures of bacteria and yeast were cultivated until they reached an OD_{600} of 0.2. The resulting cultures were then centrifuged at 1200 g for 5 minutes, and the resulting pellets were resuspended in $MgCl_2$. A spore concentration of 25×10^4 spores/mL was meticulously prepared for Plectospherella and Rhogostoma. Approximately 4-5 mL of each sample were carefully combined with 5-6 mL of *Albugo* culture (25×10^4 spores/mL) and evenly applied to 4-5 week old *A. thaliana* seedlings using airbrush guns. Following a two-week period, leaf disease symptoms were assessed, differentiating between infected and non-infected leaves, and quantified as a percentage. The leaves were then stored with great care at -80°C. DNA extraction was performed following the methodology outlined in Lutap et al. (currently under preparation). To measure qPCR, a mixture was prepared consisting of 7.5 µL of SYBR Green supermix, 5 µL of DNA (approximately 50 ng), 1.9 µL of NFW, and 0.3 µL of forward and reverse primers (10 µM each), resulting in a total reaction volume of 15 µL. The sample measurements were performed in triplicate using a Bio-Rad CFX Connect real-time PCR detection system. The quantification of *A. laibachii* was determined. The

DNA was extracted using the following oligonucleotide sequences: *A. thaliana* EF1-α: 5'-AAGGAGGCTGCTGAGATGAA-3', 5'-TGGTGGTCTCGAACTTCCAG-3'; *Albugo* EF1-α: 5'-GTGTTCTGCACATCCACACC-3', 5'-GACCTTGACGGATGAAAGGA-3'. Cq values obtained during the amplification of oomycete DNA were carefully normalized to the DNA amplicon of *A. thaliana*. Afterwards, the ddCq method was used to quantify the contrast between the control group (consisting of only *Albugo*) and the treatment group (*Albugo* + *Arabidopsis*). The relative biomass of *Albugo* was determined using the formula 2^{-ddCq}, and each data point on the graph represents the results of three independent biological replicates.

Availability of data and material

Sequencing data are available under NCBI Bioproject PRJNA961058. OTU tables and scripts are available here
https://github.com/IshtarMM/HealthmarkerMicrobe

Competing interests

Authors declare no competing financial interests in relation to the work.

Author Contributions:

Acknowledgements:

We thank KemenLabSamplingTeam for organizing and participating in several sampling trips and Elke Klenk for helping in MiSeq sequencing preparation. This project has been funded by the European Research Council (ERC) under the DeCoCt research program (grant agreement: ERC-2018-COG 820124), the Cluster of Excellence "Controlling Microbes to Fight Infections" (CMFI; Exc 2124) and the SPP 2125 DECRyPT program from the DFG.

References

[1] St´ephane Hacquard, Ruben Garrido-Oter, Antonio Gonz´alez, Stijn Spaepen, Gail Ackermann, Sarah Lebeis, Alice C McHardy, Jeffrey L Dangl, Rob Knight, Ruth Ley, et al. Microbiota and host nutrition across plant and animal kingdoms. *Cell host & microbe*, 17(5):603–616, 2015.

[2] Gerd Innerebner, Claudia Knief, and Julia A Vorholt. Protection of arabidopsis thaliana against leaf-pathogenic pseudomonas syringae by sphingomonas strains in a controlled model system. *Applied and environmental microbiology*, 77(10):3202–3210, 2011.

[3] Rupali Gupta, Gautam Anand, Rajeeva Gaur, and Dinesh Yadav. Plant–microbiome interactions for sustainable agriculture: a review. *Physiology and Molecular Biology of Plants*, 27:165–179, 2021.

[4] Matthew T Agler, Jonas Ruhe, Samuel Kroll, Constanze Morhenn, Sang-Tae Kim, Detlef Weigel, and Eric M Kemen. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology*, 14(1):e1002352, 2016.

[5] Paul Runge, Freddy Ventura, Eric Kemen, and Remco Stam. Distinct phyllosphere microbiome of wild tomato species in central peru upon dysbiosis. *Microbial ecology*, 85(1):168–183, 2023.

[6] Juliana Almario, Maryam Mahmoudi, Samuel Kroll, Mathew Agler, Aleksandra Placzek, Alfredo Mari, and Eric Kemen. The leaf microbiome of arabidopsis displays reproducible dynamics and patterns throughout the growing season. *Mbio*, 13 (3):e02825–21, 2022.

[7] Vasvi Chaudhry, Paul Runge, Priyamedha Sengupta, Gunther Doehlemann, Jane E Parker, and Eric Kemen. Shaping the leaf microbiota: plant–microbe–microbe interactions. *Journal of Experimental Botany*, 72(1):36–56, 2021.

[8] Paloma Dur´an, Thomas James Ellis, Thorsten Thiergart, Jon Agren, and St´ephane Hacquard. Climate drives rhizosphere microbiome variation and divergent selection between geographically distant arabidopsis populations. *New Phytologist*, 236:608– 621, 2022.

[9] Alexander T Neu, Eric E Allen, and Kaustuv Roy. Defining and quantifying the core microbiome: challenges and prospects. *Proceedings of the National Academy of Sciences*, 118(51):e2104429118, 2021.

[10] Gordon F Custer, Maya Gans, Linda TA van Diepen, Francisco Dini-Andreote, and C Alex Buerkle. Comparative analysis of core microbiome assignments: implications for ecological synthesis. *Msystems*, 8(1):e01066–22, 2023.

[11] Karoline Faust. Open challenges for microbial network construction and analysis. *The ISME Journal*, 15(11):3111–3118, 2021.

[12] Lisa R¨ottjers and Karoline Faust. From hairballs to hypotheses–biological insights from microbial networks. *FEMS microbiology reviews*, 42(6):761–780, 2018.

[13] Daniel G´omez-P´erez, Monja Schmid, Vasvi Chaudhry, Yiheng Hu, Ana Velic, Boris Ma˘cek, Jonas Ruhe, Ariane Kemen, and Eric Kemen. Proteins released into the plant apoplast by the obligate parasitic protist albugo selectively repress phyllosphereassociated bacteria. *New Phytologist*, 2022.

[14] Patricia A Rodriguez, Michael Rothballer, Soumitra Paul Chowdhury, Thomas Nussbaumer, Caroline Gutjahr, and Pascal Falter-Braun. Systems biology of plantmicrobiome interactions. *Molecular plant*, 12(6):804–821, 2019.

[15] Sebastian Pfeilmeier, Anja Werz, Marine Ote, Miriam Bortfeld-Miller, Pascal Kirner, Andreas Keppler, Lucas Hemmerle, Christoph G. G¨abelein, Christine M. Pestalozzi, and Julia A. Vorholt. Dysbiosis of a leaf microbiome is caused

by enzyme secretion of opportunistic xanthomonas strains. *bioRxiv*, 2023. doi: 10.1101/2023.
05.09.539948. URL https://www.biorxiv.org/content/early/2023/05/10/2023.05.09.539948.

[16] Tao Chen, Kinya Nomura, Xiaolin Wang, Reza Sohrabi, Jin Xu, Lingya Yao, Bradley C Paasch, LI Ma, James Kremer, Yuti Cheng, et al. A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature*, 580(7805):653–657, 2020.

[17] Gabriele Berg, Peter Kusstatscher, Ahmed Abdelfattah, Tomislav Cernava, and Kornelia Smalla. Microbiome modulation—toward a better understanding of plant microbiome response to microbial inoculants. *Frontiers in Microbiology*, 12:650610, 2021.

[18] Zhenyan Zhang, Qi Zhang, Hengzheng Cui, Yan Li, Nuohan Xu, Tao Lu, Jian Chen, Josep Penuelas, Baolan Hu, and Haifeng Qian. Composition identification and functional verification of bacterial community in disease-suppressive soils by machine learning. *Environmental Microbiology*, 24(8):3405–3419, 2022.

[19] Anna Bonaterra, Esther Badosa, Nu´ria Daranas, Jesu´s Franc´es, Gemma Rosell´o, and Emilio Montesinos. Bacteria as biological control agents of plant diseases. *Microorganisms*, 10(9):1759, 2022.

[20] Michael J Sweet and Mark T Bulling. On the importance of the microbiome and pathobiome in coral health and disease. *Frontiers in Marine Science*, 4:9, 2017.

[21] Ricardo Hern´andez Medina, Svetlana Kutuzova, Knud Nor Nielsen, Joachim Johansen, Lars Hestbjerg Hansen, Mads Nielsen, and Simon Rasmussen. Machine learning and deep learning applications in microbiome research. *ISME Communications*, 2(1):98, 2022.

[22] Jun Yuan, Tao Wen, He Zhang, Mengli Zhao, C Ryan Penton, Linda S Thomashow, and Qirong Shen. Predicting disease occurrence with high accuracy based on soil macroecological patterns of fusarium wilt. *The ISME Journal*, 14(12):2936–2950, 2020.

[23] Hao-Xun Chang, James S Haudenshield, Charles R Bowen, and Glen L Hartman. Metagenome-wide association study and machine learning prediction of bulk soil microbiome and crop productivity. *Frontiers in Microbiology*, 8:519, 2017.

[24] Duncan B Crosbie, Maryam Mahmoudi, Viviane Radl, Andreas Brachmann, Michael Schloter, Eric Kemen, and Macarena Mar´ın. Microbiome profiling reveals that pseudomonas antagonises parasitic nodule colonisation of cheater rhizobia in lotus. *New Phytologist*, 234(1):242–255, 2022.

[25] Maryam Mahmoudi, Juliana Almario, Katrina Lutap, Kay Nieselt, and Eric Kemen. Seasonality adaptation patterns of the natural arabidopsis leaf microbiome over several plant generations are shaped by environmental factors. *in-preparation*.

[26] Yian Gu, Samiran Banerjee, Francisco Dini-Andreote, Yangchun Xu, Qirong Shen, Alexandre Jousset, and Zhong Wei. Small changes in rhizosphere microbiome composition predict disease outcomes earlier than pathogen density variations. *The ISME Journal*, 16(10):2448–2456, 2022.

[27] Wu Xiong, Rong Li, Yi Ren, Chen Liu, Qingyun Zhao, Huasong Wu, Alexandre Jousset, and Qirong Shen. Distinct roles for soil fungal and bacterial communities associated with the suppression of vanilla fusarium wilt disease. *Soil Biology and Biochemistry*, 107:198–207, 2017.

[28] Jens M Olesen, Jordi Bascompte, Yoko L Dupont, and Pedro Jordano. The modularity of pollination networks. *Proceedings of the National Academy of Sciences*, 104(50): 19891–19896, 2007.

[29] Jacopo Grilli, Tim Rogers, and Stefano Allesina. Modularity and stability in ecological communities. *Nature communications*, 7(1):12031, 2016.

[30] Shalini Rai, Ayman F Omar, Medhat Rehan, Ahmad Al-Turki, Alka Sagar, Noshin Ilyas, RZ Sayyed, and Mirza Hasanuzzaman. Crop microbiome: Their role and advances in molecular and omic techniques for the sustenance of agriculture. *Planta*, 257(2):27, 2023.

[31] Katharina Eitzen, Priyamedha Sengupta, Samuel Kroll, Eric Kemen, and Gunther Doehlemann. A fungal member of the arabidopsis thaliana phyllosphere antagonizes albugo laibachii via a gh25 lysozyme. *Elife*, 10:e65306, 2021.

[32] Davide Bulgarelli, Klaus Schlaeppi, Stijn Spaepen, Emiel Ver Loren Van Themaat, and Paul Schulze-Lefert. Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology*, 64:807–838, 2013.

[33] Kenneth Dumack, Sebastian Flues, Karoline Hermanns, and Michael Bonkowski. Rhogostomidae (cercozoa) from soils, roots and plant leaves (arabidopsis thaliana): Description of rhogostoma epiphylla sp. nov. and r. cylindrica sp. nov. *European journal of protistology*, 60:76–86, 2017.

[34] Michael Bonkowski. Protozoa and plant growth: the microbial loop in soil revisited. *New Phytologist*, 162(3):617–631, 2004.

[35] Kenneth Dumack, Christina Baumann, and Michael Bonkowski. A bowl with marbles: revision of the thecate amoeba genus lecythium (chlamydophryidae, tectofilosida, cercozoa, rhizaria) including a description of four new species and an identification key. *Protist*, 167(5):440–459, 2016.

[36] Bao-Anh Thi Nguyen, Kenneth Dumack, Pankaj Trivedi, Zahra Islam, and Hang-Wei Hu. Plant associated protists—untapped promising candidates for agrifood tools. *Environmental Microbiology*, 25(2):229–240, 2023.

[37] Or Shalev, Talia L Karasov, Derek S Lundberg, Haim Ashkenazy, Pratchaya Pramoj Na Ayutthaya, and Detlef Weigel. Commensal pseudomonas strains facilitate protective response against pathogens in the host plant. *Nature ecology & evolution*, 6(4): 383–396, 2022.

[38] AJ Cooper, AO Latunde-Dada, A Woods-T¨or, J Lynn, JA Lucas, IR Crute, and EB Holub. Basic compatibility of albugo candida in arabidopsis thaliana and brassica juncea causes broad-spectrum suppression of innate immunity. *Molecular PlantMicrobe Interactions*, 21(6):745–756, 2008.

[39] Fantin Mesny, St´ephane Hacquard, and Bart PHJ Thomma. Co-evolution within the plant holobiont drives host performance. *EMBO reports*, page e57455, 2023.

[40] Fantin Mesny, Shingo Miyauchi, Thorsten Thiergart, Brigitte Pickel, Lea Atanasova, Magnus Karlsson, Bruno Hu¨ttel, Kerrie W Barry, Sajeet Haridas, Cindy Chen, et al. Genetic determinants of endophytism in the arabidopsis root mycobiome. *Nature communications*, 12(1):7227, 2021.

[41] Patrick D Schloss. Reintroducing mothur: 10 years later. *Applied and environmental microbiology*, 86(2):e02343–19, 2020.

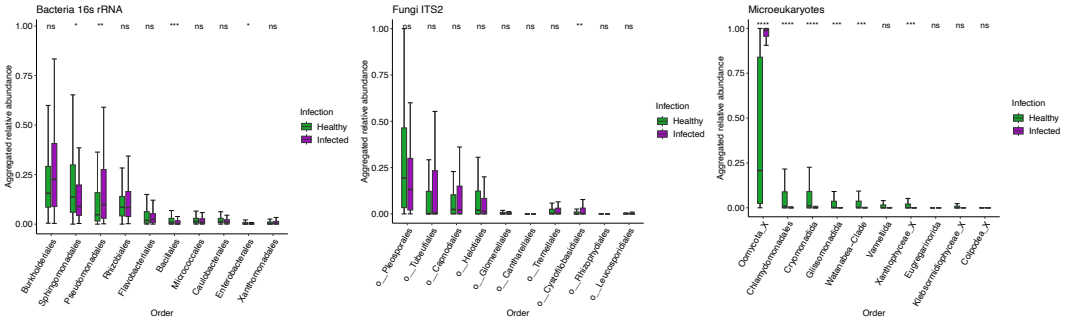
[42] Johan Bengtsson-Palme, Martin Ryberg, Martin Hartmann, Sara Branco, Zheng Wang, Anna Godhe, Pierre De Wit, Marisol S´anchez-Garc´ıa, Ingo Ebersberger, Filipe de Sousa, et al. Improved software detection and extraction of its1 and its 2 from ribosomal its sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in ecology and evolution*, 4(10):914–919, 2013.

[43] Fabian Pedregosa, Ga¨el Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

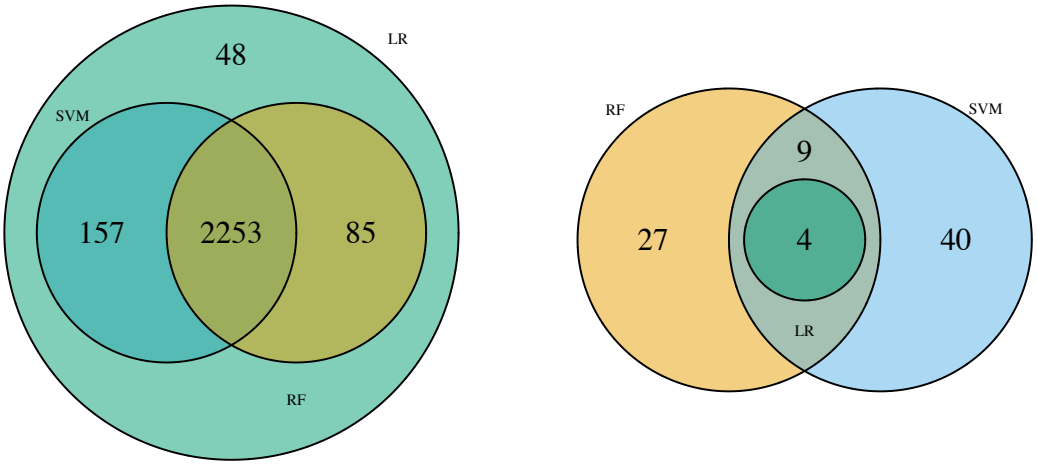
[44] Jonathan Friedman and Eric J Alm. Inferring correlation networks from genomic survey data. *PLoS computational biology*, 8(9):e1002687, 2012.

[45] Stephen C Watts, Scott C Ritchie, Michael Inouye, and Kathryn E Holt. Fastspar: rapid and scalable correlation estimation for compositional data. *Bioinformatics*, 35 (6):1064–1066, 2019.

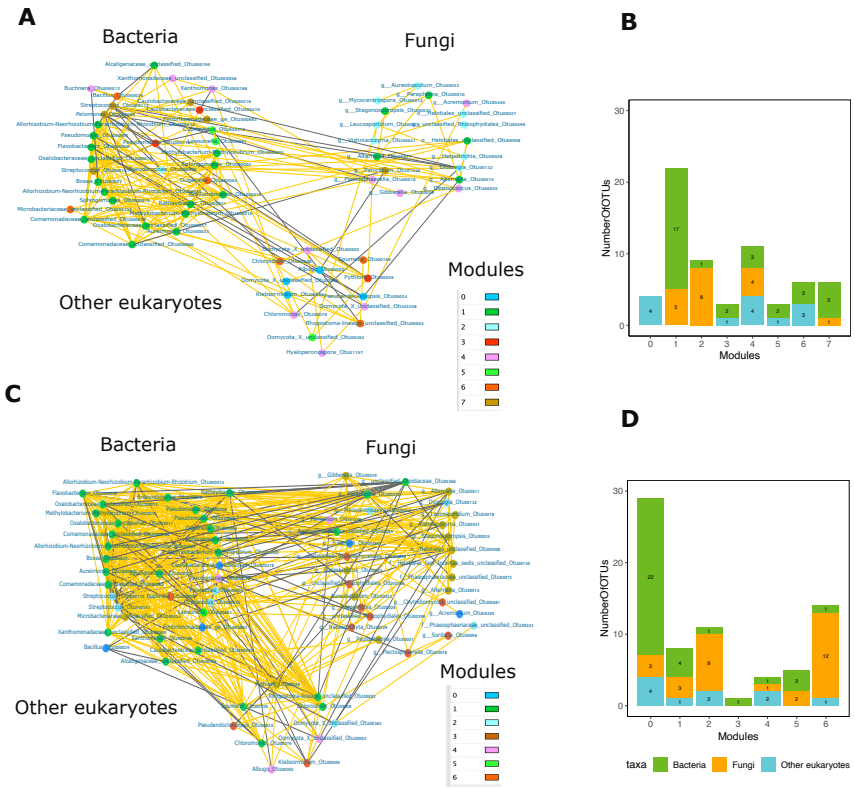
[46] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.



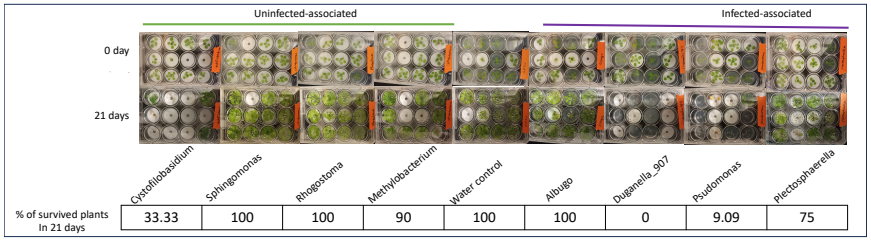
SFigure 1 Changes in high abundance microbial taxa colonising A. thaliana's in infected and uninfected leaves. Box plots (green = uninfected, purple = infected) show the relative abundance of the orders Bacteria, Fungi and Oomycetes in individual samples aggregated by 'infection phenotype'. Whiskers show the dispersion of the data (1.5 x interquartile range). Significance values show the differences of microbes between seasons based on the Wilcoxon test (ns: p-value > 0.05, *p < = 0.05, **p < = 0.01, ***p < = 0.001).



SFigure 2. Comparing the importance of microbes for classifying uninfected leaves from infected samples using machine learning models. To classify infected samples from uninfected samples, four classification models were trained. The Venn diagram on the left represents the number of common microbes with scores greater than 0 between different models in the training model phase. The right Venn diagram represents the number of shared microbes scored higher than 0 after performing recursive feature elimination.



SFigure3. Changes in the microbial co-occurrence intervals of the OTUs selected as discriminators of infected and uninfected leaves by the machine learning classifiers. Interactions between OTUs selected as discriminators of infected and uninfected samples (see Fig. 3) in the network constructed for infected and uninfected samples shown in A and C, respectively. OTUs are in color by module. The edges are colored according to the correlations (positive correlations are colored in orange and negative correlations are colored in gray). Histograms (B and D) showing OTU distribution within modules for the network of infected and uninfected samples. These histograms are further color-coded for distinction of microbial taxa. Green represents bacteria, orange fungi, blue other eukaryotes.



Sfigure 4. Effect of different microbial signatures on four-week-old WS-0 Arabidopsis plants under gnotobiotic conditions. The observed phenotypes were evaluated after a period of three weeks, and the table shows the percentage of surviving plants after infection.

Chapter 5

Pseudomonas antagonises parasitic nodule colonisation of cheater *rhizobia* in *Lotus*

- **Status** Published
- **URL** https://doi.org/10.1111/nph.17988
- **Own contribution**

  Interpreted the data (with co-authors)

  Performed microbial networks and machine learning analysis

  Wrote the manuscript (with co-authors)

Note: This publication has also been included in the doctoral thesis of a co-author.

Research

New Phytologist

Microbiome profiling reveals that *Pseudomonas* antagonises parasitic nodule colonisation of cheater rhizobia in *Lotus*

Duncan B. Crosbie[1], Maryam Mahmoudi[2], Viviane Radl[3], Andreas Brachmann[1], Michael Schloter[3,4], Eric Kemen[2] and Macarena Marín[1]

[1]Genetics, Biocentre, LMU Munich, Martinsried 82152, Germany; [2]Microbial Interactions in Plant Ecosystems, Centre for Plant Molecular Biology, University of Tübingen, Tübingen 72076, Germany; [3]Comparative Microbiome Analysis, Helmholtz Centre for Environmental Health, Oberschleissheim 85764, Germany; [4]Chair for Soil Science, Technical University of Munich, Freising 85354, Germany

Author for correspondence:
Macarena Marín
Email: m.marin@bio.lmu.de

Received: *22 October 2021*
Accepted: *11 January 2022*

New Phytologist (2022) **234:** 242–255
doi: 10.1111/nph.17988

Key words: ineffective nodules, *Lotus*, *Mesorhizobium*, microbiome, *Pseudomonas*, root–nodule symbiosis.

Summary

- Nodule microbiota are dominated by symbiotic nitrogen-fixing rhizobia, however, other non-rhizobial bacteria also colonise this niche. Although many of these bacteria harbour plant-growth-promoting functions, it is not clear whether these less abundant nodule colonisers impact root–nodule symbiosis.
- We assessed the relationship between the nodule microbiome and nodulation as influenced by the soil microbiome, by using a metabarcoding approach to characterise the communities inside nodules of healthy and starved *Lotus* species. A machine learning algorithm and network analyses were used to identify nodule bacteria of interest, which were re-inoculated onto plants in controlled conditions to observe their potential functionality.
- The nodule microbiome of all tested species differed according to inoculum, but only that of *Lotus burttii* varied with plant health. Amplicon sequence variants representative of *Pseudomonas* species were the most indicative non-rhizobial signatures inside healthy *L. burttii* nodules and negatively correlated with *Rhizobium* sequences. A representative *Pseudomonas* isolate co-colonised nodules infected with a beneficial *Mesorhizobium*, but not with an ineffective *Rhizobium* isolate and another even reduced the number of ineffective nodules induced on *Lotus japonicus*.
- Our results show that nodule endophytes influence the overall outcome of the root–nodule symbiosis, albeit in a plant host-specific manner.

Introduction

Leguminous plants have evolved a mutualistic interaction with nitrogen-fixing rhizobia in which the bacteria are hosted and nourished in root organs called nodules in exchange for ammonia. This so-called root–nodule symbiosis is initiated by a two-way signalling between the symbiosis partners, which activates distal cell divisions in the root cortex and culminates in the formation and infection of nodules (Venado *et al.*, 2020). Here the bacteria differentiate into plant-dependent, nitrogen-fixing endosymbiotic bacteroids (Kereszt *et al.*, 2011). The fixation of nitrogen is an energetically expensive process for the host that requires at least 16 ATP molecules per N_2 molecule to fuel the nitrogenase enzyme produced by the rhizobia (Seefeldt *et al.*, 2009). Therefore, to prevent infection of the carbon-rich nodules by pathogens, host plants have evolved complex recognition mechanisms that ensure symbiotic specificity (Wang *et al.*, 2012).

Root–nodule symbiosis is highly species specific and many plants will only form an effective symbiosis with a narrow range of rhizobia (Remigi *et al.*, 2016). Even within these pairings there is variation in nitrogen fixation efficiency (Schumpp & Deakin, 2010). Some bacteria can also nodulate plants and not fix any nitrogen at all (Sachs & Simms, 2008). Examples of ineffective nitrogen fixation have been described after the introduction of crop legumes into areas where native legumes previously grew. For instance, inefficient nitrogen fixation occurs in fields where perennial and annual clovers co-exist (Howieson *et al.*, 2005). Native rhizobial species associated with native legumes can outcompete inoculant strains (Streeter, 1994). In extreme cases, endogenous rhizobia can completely block the nodulation of introduced rhizobia. For example, the nodulation of the pea cultivars Afghanistan and Iran by rhizobial inoculants is suppressed in natural soils by the presence of a non-nodulating strain (Winarno & Lie, 1979). This suggests that interactions of the soil microbiota with the host plant are critical for the establishment of efficient nodules. However, we are far from understanding what factors determine the success of single microbes that compete for resources at the plant soil interface, in particular nodule endophytes and how these affect the outcome of the symbiosis.

There is clear evidence to suggest that the host controls the makeup of the microbiota in its vicinity. *Lotus japonicus* selects for

a broad taxonomic range of bacteria, in addition to the symbiont, within the rhizosphere, endosphere and the nodule. This selectivity filters the diverse soil microbiome into a distinct and taxonomically narrow community within the nodule (Zgadzaj *et al.*, 2016). Despite this selective pressure, nonnodulating bacteria, such as *Pseudomonas* sp., *Klebsiella* sp. and *Rhodococcus* sp. have been isolated from plant nodules (Ibáñez *et al.*, 2009; Ampomah & Huss-Danell, 2011; Martínez-Hidalgo & Hirsch, 2017). Although these isolates do not directly nodulate the plant, they contribute to plant growth in some ways, such as increasing the availability of soluble phosphate and producing plant compounds beneficial for plant growth such as siderophores and indoleacetic acid (Dey *et al.*, 2004; Malik & Sindhu, 2011; Zhao *et al.*, 2013). In addition, non-Rhizobiales microbes found in nodules of *Medicago truncatula* produce antimicrobial compounds that may shape the community and the overall function of the nodule microbiome (Hansen *et al.*, 2020). Microbe–microbe interactions could also impart an effect on the overall functionality of the symbiosis, for instance via antimicrobial activity (Tyc *et al.*, 2014), suppression of plant pathogens (Gu *et al.*, 2020) or by horizontal gene transfer (Cytryn, 2013). Although these complex interactions could dictate the effectiveness and specificity of the symbiosis, little information is known about how rhizobia interact with other members of the nodule microbiota.

In this work, we determined the nodule microbiome of three *Lotus* species upon inoculation with soil suspensions that led to the growth of either starved or healthy plants. We used metabarcoding-based high-throughput sequencing to characterise the microbiome in nodule samples that varied in plant species origin, soil inocula and plant health. Network analyses and machine learning algorithms identified microbiome members specifically associated to nodules of healthy, but not of starved *Lotus burttii* plants. Tripartite interactions between rhizobia, nodule endophytes and the host were further investigated in co-inoculation assays. Our results show that although root–nodule symbiosis is a binary interaction, there are other nodule microbes that modulate this mutualism.

Materials and Methods

Soil collection and inoculum preparation

Soil samples were collected from two neighbouring sites in a semiurban area south west of Munich, Germany. Site 1 (48°06′29.9″N, 11°27′38.9″E) has consistently been home to wild *Lotus corniculatus*, whereas site 2 (48°06′33.2″N, 11°27′41.4″ has been subjected to tilling and physical disturbance and did not contain *Lotus* plants at the time of collection. Soil samples were taken from the top layer (0–20 cm deep) after plant material was removed from the site in May 2019 and October 2018. Physicochemical property measurements of each soil were performed by AGROLAB Agrarzentrum GmbH (Landshut, Germany). Soil samples were sieved to remove stones and plant material with a 2 mm sieve, mixed 1 : 5 with a nitrogen-limiting Fabaceae (FAB) liquid medium, and stirred for 2 h. Soil particulate matter was removed by centrifugation at 1000 ***g*** for 5 min. Soil suspensions were used as inputs and a quantitative PCR (qPCR) was run to compare the quantity of soil bacteria present in both soil suspensions inputs.

Plant growth and inoculation conditions

Lotus burttii B-303 (seed bag no. 91105), *L. japonicus* Gifu B-129 (seed bag no. 110913) and *L. corniculatus* cv Leo (Andreae Saaten, Regensburg, Germany) seeds were scarified and then sterilised by incubation in a sterilising solution (1.2% NaOCl, 1% SDS) for 8 min before being washed three times with sterile water. Seeds were then soaked in sterile water for 2–3 h and germinated on 0.5 B5 agar medium (Gamborg *et al.*, 1968) for 3 d in dark followed by 3 d in a long-day photoperiod (16 h : 8 h, light : dark) at 24°C. Seedlings were then transferred into sterilised tulip-shaped Weck jars (10 seedlings per jar) containing 300 ml of a sand : vermiculite mix (1 : 2) and supplemented with 40 ml of a low nitrogen FAB medium, to create nitrogen-limiting growth conditions as mentioned above (Liang *et al.*, 2019). Jars were sealed with micropore tape to create a closed system. Seedlings were left to recover for 2 d in a long-day photoperiod. After the 2-d recovery, each seedling was inoculated with 1 ml of soil suspension. *Lotus burttii* and *L. japonicus* treatments consisted of 150 plants from three independent experiments, and *L. corniculatus* treatments consisted of 50 plants per condition from one independent experiment.

Harvesting, phenotyping and nodule surface sterilisation

Plants were harvested and phenotyped 5 wk post inoculation across five independent experiments. Shoot length, shoot dry weight, nodule number, nodule colour and plant health were recorded. Nodules were classified as pink or white, which indicated the presence or absence of leghaemoglobin, respectively, a prerequisite for, but not a guarantee of, nitrogen fixation (Downie, 2005). Roots were removed from shoots and sonicated using the Bioruptor® (Diagenode, Seraing, Belgium) twice for 15 min. Nodules from three or four plants were excised and pooled based on similarity of plant shoot and nodule phenotype. Pink and white nodules were collected separately. Pooled nodules were treated with 70% ethanol for 1 min followed by 2% NaOCl for 2.5 min. Nodules were then washed with sterile water eight times and after the removal of the final water wash, samples were snap frozen in liquid nitrogen. The final wash was plated onto 20Q agar supplemented with 3.8% w/v mannitol (modified from Werner *et al.*, 1975) to assess sterilisation.

DNA extraction

Nodule samples were homogenised six times in a Mixer Mill 400 (Retsch, Haan, Germany) at a frequency of 30 s^{-1} for 1 min. DNA was then extracted according to a modified protocol from Töwe *et al.* (2011). For extraction of DNA from the inputs, soil suspensions were centrifuged at 5000 ***g*** and DNA from pellets was extracted according to the CTAB method described by the Doe Joint Genomics Institute (William *et al.*, 2012). The

concentrations of extracted DNA samples were quantified using a Qubit 2.0® fluorometer (Invitrogen, Carlsbad, CA, USA).

Quantitative PCR

Quantitative PCR was performed using the forward primer (FP) *16S rDNA* (5′-GGTAGTCYAYGCMSTAAACG-3′) and reverse primer (RP) *16S rDNA* (5′-GACARCCATGCASCACCTG-3′) primers (Bach *et al.*, 2002). The 25-µl PCR mixture contained 12.5 µl SYBR Green, 2 µl template DNA, 7.5 µl Milli-Q water, 1 µl 10 µM of primers and 1 µl 15% bovine serum albumin (BSA). The mixture was amplified using a CFX96 Real-Time System (Bio-Rad, Hercules, CA, USA) under the following conditions: template was denatured at 94°C for 10 min before 40 cycles of 95°C for 20 s, 57°C for 30 s and 72°C for 45 s, followed by dissociation curve steps of 95°C for 15 s, 60°C for 30 s and 95°C for 15 s. Quantification of the *16S* rRNA gene molecules was correlated with a calibration curve constructed with known amounts of a *16S* rRNA gene standard plasmid constructed of a *Mesorhizobium septentrionale 16S* rRNA gene sequence cloned into a pUC57 plasmid.

Amplification, library preparation and sequencing

To determine bacterial diversity, a metabarcoding approach was utilised. The hypervariable region V3–V4 of the *16S* rRNA gene was amplified using universal bacterial primers 335F (5′-CADACTCCTACGGGAGGC-3′) and 769R (5′-ATCCTGTTTGMTMCCCVCRC-3′) fused to Illumina adapters. The primers were specific for bacterial DNA and did not amplify plastidial and mitochondrial plant DNA (Dorn-In *et al.*, 2015). Amplification reaction volumes were 25 µl using 1 unit of Phusion polymerase, 5 µl 5× High-Fidelity Phusion buffer, 7.5 µl of 1% BSA, 0.5 µl of 10 mM dNTPs, 0.5 µl of 50 mM $MgCl_2$, 0.5 µl of 10 pmol μl^{-1} primer and 5 ng of template DNA. The assay was conducted in triplicate under the following conditions: template was denatured at 98°C for 1 min, then 25 cycles of 98°C for 10 s, 55°C for 30 s and 72°C for 30 s, followed by a final step at 72°C for 5 min. PCR products were verified via gel electrophoresis, pooled, and cleaned using CleanPCR beads (CleanNA, Waddinxveen, the Netherlands). Fragments were then indexed with 10 nucleotide barcode sequences using the Nextera XT Index Kit v.2 Set D primers (Illumina, San Diego, CA, USA). Indexing PCR reactions were run in triplicate with a volume of 25 µl using 12.5 µl NEB Next High-Fidelity Master Mix, 2.5 µl of each delegated primer and 20 ng of amplicon under the following conditions: template was denatured at 98°C for 30 s, then eight cycles of 98°C for 10 s, 55°C for 30 s and 72°C for 30 s, followed by a final step at 72°C for 5 min. PCR products were pooled and cleaned with CleanPCR beads (CleanNA). Quantification and quality control were conducted using an AATI Fragment Analyser (Santa Clara, CA, USA). All samples were pooled at an equimolar concentration for paired-end 2 × 300-bp sequencing via the MiSeq system (Illumina) using the MiSeq Reagent Kit v.3 (600 cycles), as per the manufacturer's recommendation.

Sequence and statistical analysis

An average of *c.* 141 000 raw Illumina reads per sample were obtained, which were then demultiplexed and had adapter and barcode sequences removed using CUTADAPT v.3.1 (Martin, 2011). Reads were then trimmed, merged and filtered using DADA2 (Callahan *et al.*, 2016) in R. The criteria for filtering were minimum lengths of 280 bp for the forward reads and 160 bp for the reverse, as these lengths corresponded to a minimum quality score of 25. Merged sequences had chimeras and chloroplastic and mitochondrial sequences removed. Amplicon sequence variants (ASVs) were assigned in R using the Silva database v.132 (Quast *et al.*, 2012).

The PHYLOSEQ v.1.26.1 package in the R (McMurdie & Holmes, 2013) pipeline was used to infer alpha diversity of ASVs rarefied corresponding to the sample with the lowest number of reads. Multidimensional Scaling using Bray–Curtis (Bray & Curtis, 1957) distance was performed using the PHYLOSEQ v.1.26.1 package in R (McMurdie & Holmes, 2013) to assess the beta diversity of microbial communities. Comparisons were visualised using GGPLOT2 (Wickham, 2009) in R and tested for statistical significance (*adonis* test, $P < 0.01$) via permutational multivariate analysis of variance (PERMANOVA) utilising 999 permutations in the VEGAN package (Oksanen *et al.*, 2018). Relative abundance of each genera per sample was calculated using transformed count data. To further specify the composition of the sample microbiome the relative abundance of the most prevalent ASVs (abundance > 0.1%) was calculated for each sample. All abundance levels were calculated using the PHYLOSEQ v.1.26.1 package (McMurdie & Holmes, 2013) in R.

Machine learning model

A support vector machine learning model by svm.SVC (kernel=linear) in PYTHON SCIKIT-LEARN (Pedregosa *et al.*, 2011) was used to discriminate between starved and healthy plant samples of *L. burttii* on relative abundance filtered ASVs using five-fold cross-validation. The ASV tables were filtered to ASVs present at ≥ 50 reads in soil suspension 2 inoculated *L. burttii* nodule samples. The *svm.SVC.coef* function was used to calculate the coefficient value of the ASVs. These values were then used to identify signature ASVs characteristic of certain sample types. The model was trained with 70% of the data and evaluated by 30% of the data five times (mean of accuracy = 0.89) with the average coefficient value of each ASV being used to select for important features.

Microbial correlation networks

Filtered ASV tables comprised of samples of *L. burttii* inoculated with soil suspension 2 (ASV raw abundances) were used to calculate microbial correlation networks among ASVs using the SPARCC (Friedman & Alm, 2012) algorithm in FASTSPAR (Watts *et al.*, 2019). This algorithm uses log-ratio variances of ASV fractions to calculate pairwise correlations between ASVs in an iterative manner. The filtered tables were used to calculate the

correlation between ASVs using the FastSpar implementation and the default parameters. Pseudo *P*-values were inferred from 1000 bootstraps. Only correlations with $P < 0.01$ were kept for further analyses. Network visualisation was performed in Cytoscope v.3.8.2 (Shannon *et al.*, 2003). Analysis of interactions between and within ASVs of different genera were carried out using the same methods.

Isolation of strains

Strains were isolated from crushed nodules on a variety of media. Nodules from either *L. burttii*, *L. corniculatus* or *L. japonicus* inoculated with either soil suspension 1 or 2 were sterilised as described above (see 'Harvesting and nodule sterilisation' in the Materials and Methods section). Individual nodules were crushed in 10 mM $MgSO_4$ and the content was then spread onto 20Q agar plates supplemented with mannitol, lysogeny broth (LB) (Bertani, 1951), yeast mannitol (YM) (Vincent, 1970), *Pseudomonas* minimal medium (PMM) (Sandman & Ecker, 2014), and tryptone soy (TS) (Sigma, Darmstadt, Germany). Plates were incubated at 28°C for up to 3 wk and further isolation of single colonies was carried out 7–9 times until pure cultures were attained. The taxonomy of each strain was determined by amplifying the *16S* rRNA gene using primers 41f (5′-GCTCAGATTGAACGCTGGCG-3′) and 1488r (5′-CGGTTACCTTGTTACGACTTCACC-3′) (Herrera-Cervera *et al.*, 1999) and Phusion polymerase. Amplicons were purified using a 1 : 0.8 ratio of PCR product to CleanPCR beads (CleanNA) and sequenced using *16S* rRNA gene-specific primers, 41f and 1488r (Herrera-Cervera *et al.*, 1999), by Sanger sequencing. Sequences were aligned to DNA sequences from the NCBI Nucleotide collection online database using Blastn (Altschul *et al.*, 1990). The sequences of the isolates were aligned with ASV sequences using CLC Main Workbench 7 (Qiagen, Hilden, Germany). Strains were stored in 40% glycerol at −80°C.

Isolate inoculations

Bacteria were streaked and grown until single colonies formed. Single colonies were inoculated into 2 ml of the appropriate liquid medium and grown at 28°C for 2, 3 or 5 d for *Pseudomonas* sp. strains, *Rhizobium* sp. BW8-2 and *Mesorhizobium* sp. Qb1E3-1, respectively. Bacteria were then washed in sterile water and resuspended in FAB medium to allow for a final OD_{600} of 0.005. Each plant was inoculated with 1 ml of bacterial suspension. Plants were prepared as described above.

Conjugation

Strains used in this work are listed in Supporting Information Table S1. *Escherichia coli* ST18 transformed with pFAJ-GFP and pFAJ-*Ds*Red plasmids (Kelly *et al.*, 2013) and *E. coli* S17.1 transformed with pABC-Cerulean were grown at 37°C overnight in LB supplemented with appropriate antibiotics. *Pseudomonas* sp. Lb2C2, *Rhizobium* sp. BW8-2 and *Mesorhizobium* sp. Qb1E3-1 were grown at 28°C in 20Q liquid with the appropriate antibiotics for 2, 3 and 5 d, respectively. Conjugations were conducted as in Liang *et al.* (2019). Successful conjugation was confirmed via fluorescence microscopy and *16S* rRNA gene sequencing.

Section preparation and microscopy

Lotus burttii seeds were sterilised, germinated, potted and inoculated with fluorescent strains as described above (see 'Plant growth and inoculation conditions' and 'Isolate inoculations' in the Materials and Methods section). Plants were harvested 2 wk after inoculation and the nodules were excised and embedded in 6% low melting agarose. The nodules were then sliced into 100 µm-thick sections using a VT1000S vibratome (Leica Biosystems, Wetzlar, Germany) and visualised with a TCS SP5 confocal microscope (Leica Microsystems, Wetzlar, Germany) equipped with a ×20 HCX PL APO water immersion lens. GFP and Cerulean were excited with an argon laser line at 488 and 433 nm, and the emissions were detected at 492–515 and 455–474 nm, respectively. *Ds*Red was excited with a diode pumped solid-state laser at 561 nm and detected at 580–620 nm.

Statistical analyses

The nodule and root phenotype of plants inoculated with nodule isolates were recorded. Statistical significance was assessed using analysis of variance (ANOVA) and Tukey honestly significant difference (HSD) tests in R (Graves *et al.*, 2015).

Results

Species-specific effect of soil inoculum on *Lotus* plant growth

Two different soil suspensions were used to inoculate *L. burttii*, *L. japonicus*, and *L. corniculatus* plants. These *Lotus* species were selected as they all belonged to the *L. corniculatus* clade (Kramina *et al.*, 2016), but nodulated with a different range of microsymbionts (Gossmann *et al.*, 2012; Sandal *et al.*, 2012). The first soil (soil 1) was collected at a site that contained healthy wild growing *L. corniculatus* plants, while the second soil (soil 2) site contained no leguminous plants at all. The soils had minor differences in mineral content and grain size (Table S2). The quantity of soil bacteria present in the soil suspensions used as inputs was compared by qPCR. Soil suspension inputs 1 and 2 contained 1.62×10^5 and 2.28×10^5 molecules of the *16S* rRNA gene per nanogram of extracted DNA, respectively.

Lotus japonicus, *L. burttii* and, to a lesser extent, *L. corniculatus*, produced exclusively healthy plants (green leaves, elongated shoots) when inoculated with soil 1 suspension (Soil S1; Fig. 1). Contrastingly, there was marked variation in the shoot growth phenotype seen in all species when inoculated with the soil 2 suspension (Soil S2). Growing alongside the healthy plants was a large contingent of starved plants presenting with shorter shoots and yellow leaves (Fig. 1). Similar results were observed across five independent experiments (Fig. S1). Nodule number also varied dependent on soil suspension inoculum. Plants inoculated with soil 1 suspension

New Phytologist

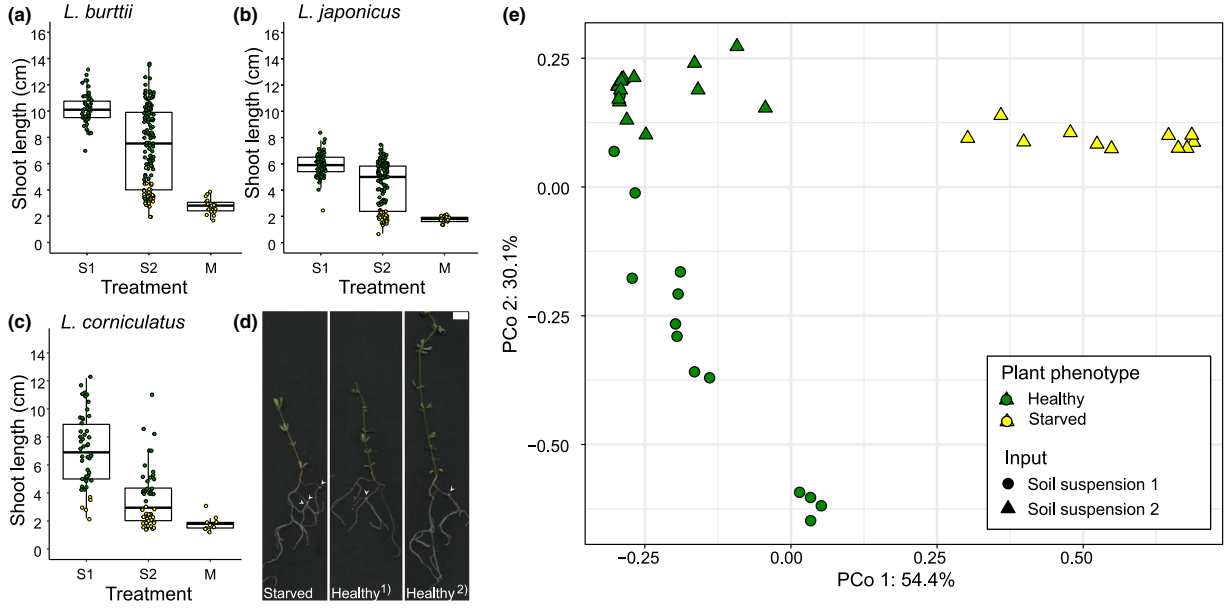


Fig. 1 Shoot growth phenotype of *Lotus* plants inoculated with Munich soil suspensions. Shoot growth quantification of *Lotus burttii* (a), *Lotus japonicus* (b) and *Lotus corniculatus* (c) plants 5 wk post inoculation with soil suspensions 1 (S1) and 2 (S2) and a mock (M) treatment. Green and yellow dots indicate plants with healthy and starved phenotypes, respectively. Box plots display the results of 50–150 plants per condition. The bold black line and the box depict the median and the interquartile range, respectively. In total, 49 mock treated plants were included. (d) Scanned images of *L. burttii* 5-wk post inoculation with soil suspension 2. Starved plants exhibited pale green leaves despite having nodules on their roots. The shoots of healthy dark green plants varied in length. Phenotypic variation is depicted in [1)] and [2)]. White arrowheads indicate the position of nodules on plant roots. Plots show the results from one representative experiment. Bar, 1 cm. (e) Principal coordinates analysis plot of *L. burttii* nodules based on beta diversity calculated using the Bray–Curtis dissimilarity index (Bray & Curtis, 1957) revealed a clustering of common sample types and a separation of dissimilar sample types.

consistently developed a higher number of nodules per plant across all species (Fig. S2a–c). Starved plants inoculated with the soil 2 suspension exhibited roots either with or without nodules. In *L. burttii*, 73.8% of starved plants contained nodules, while in *L. japonicus* and *L. corniculatus*, 45.2% and 59.3% exhibited nodules, respectively (Fig. S2d). However, the most striking difference was that in *L. burttii* 88.4% of the nodules on starved-nodulated plants were white, whereas in the other species most of the nodules were pink (Fig. S2e). These results showed that the microbiota of the soil 2 suspension is capable of mediating both effective and ineffective symbiosis, although the frequency at which each plant succumbs to an ineffective nodulation differs.

Richness, diversity and community structure of the *Lotus* nodule microbiome

The microbiome of an effective plant nodule is typically dominated by the respective symbiont, although there can also be colonisation by other microbes (Martínez-Hidalgo & Hirsch, 2017). To investigate how the nodule microbiota varied depending on the plant host, inoculum, and nodule phenotype we sequenced the microbiome of nodules collected from healthy and starved *Lotus* of different species inoculated with different soil suspensions. A variable region of the *16S* rRNA gene was sequenced and the output reads were processed, sorted into ASVs and assigned a taxonomy. ASVs were used as they provide a finer resolution than Operational Taxonomic Units (Callahan *et al.*, 2017), which is important as the 16S rRNA gene of some rhizobia, such as *Mesorhizobium*, can be more than 99% identical between different species (Marcos-García *et al.*, 2015). Sequencing produced 13 989 700 paired-end reads after quality filtering, which clustered into 67 442 unique ASVs. Sequence coverage varied between sample types with the nodule samples having an average of 148 679 reads per sample and the soil suspension input samples having an average of 67 618 reads per sample (Dataset S1). All rarefaction curves reached a saturation plateau (Fig. S3).

To assess the effect of the host genotype and the inoculum on the nodule microbiome diversity, the alpha and beta diversities of the different nodule samples from all three species were determined. Within sample variation (alpha diversity) was calculated using the Shannon diversity index, which was found to be much higher in the soil suspension input samples compared with the nodule samples (Fig. S4a). The soil suspensions 1 and 2 did not significantly vary in their alpha diversities (Welch two sample *t*-test, $P = 0.749$), although it was found that plants inoculated with soil 1 suspension produced nodules with a much higher alpha diversity compared with those inoculated with soil 2 suspension. This observation was most pronounced in *L. japonicus* and *L. corniculatus* (Fig. S4). A similar trend in alpha diversity was seen when considering observed ASVs (Fig. S4b).

To analyse the diversity between sample types (beta diversity), principal coordinate and PERMANOVA analyses were conducted using the Bray–Curtis dissimilarity. A global comparison of the nodule diversity showed an overall separation based on soil suspension input (Fig. S5; Soil S1 vs Soil S2, $Pr(>F) = 0.001$; Table S3), despite the two soil suspension inputs showing insignificant differences between one another (Soil S1 vs Soil S2 (input suspension), $Pr(>F) = 0.072$; Table S3). The most pronounced difference was between nodules of *L. burttii* plants. At the species level, *L. burttii* and *L. japonicus* showed a significant difference in beta diversity based on soil suspension input (*Lb* healthy plants – Soil S1 vs Soil S2, $Pr(>F) = 0.001$; *Lj* healthy plants – Soil S1 vs Soil S2, $Pr(>F) = 0.002$; Table S3). *Lotus burttii* nodules showed a significant difference in beta diversity based on plant health (*Lb* Soil S2 – healthy vs starved plants, $Pr(>F) = 0.001$; Fig. 1d; Table S3), however this was not the case in *L. japonicus* or *L. corniculatus* (*Lj* Soil S2 – healthy vs starved plants, $Pr(>F) = 0.097$; *Lc* Soil S2 – healthy vs starved plants, $Pr(>F) = 0.742$; Table S3). As a control, we compared the microbiome of laboratory grown *L. corniculatus* plants to the microbiome of nodules collected from *L. corniculatus* plants growing on site 1 (*Lc* Soil S1 – laboratory grown vs wild plants, $Pr(>F) = 0.342$) (Table S3). These did not significantly differ, supporting that nodules produced in this growth/inoculation system are representative of nodules grown in the wild.

Bacterial composition of the nodule microbiome

Both soil suspension types were dominated by Alphaproteobacteria and Gammaproteobacteria. To determine the bacterial composition of the nodule microbiome we estimated the relative abundance at an ASV level. The nodule microbiome of all *Lotus* species was dominated by ASVs belonging to the order Rhizobiales. Nodules of healthy plants were largely dominated by *Mesorhizobium*, independent of the host and soil suspension input. However, while nodules from plants inoculated with soil 1 suspension were colonised with a variety of different *Mesorhizobium* ASVs, the nodules of healthy plants inoculated with soil 2 suspension were almost exclusively colonised by *Mesorhizobium* ASV1 (Fig. 2). This disparity in *Mesorhizobium* ASV presence is despite the observation that there is no significant difference between the *Mesorhizobium* ASVs present in the two suspensions (*Meso.* Soil S1 vs *Meso.* Soil S2, $P(>F) = 0.479$). Nodules of starved *L. burttii* plants were largely colonised by bacteria belonging to what was taxonomically defined as *Allorhizobium–Neorhizobium–Pararhizobium–Rhizobium* and will be referred to as *Rhizobium* (Fig. 2a). This suggests that *L. burttii* plants are less selective compared with *L. corniculatus* and *L. japonicus* and develop an ineffective symbiosis with *Rhizobium* strains.

Pseudomonas are more prevalent in healthy plant nodules and negatively correlate with ineffective *Rhizobium*

Support Vector Machine (SVM) is a machine learning method used to separate a data set using a linear or nonlinear surface (Noble, 2006). In this instance we used a linear-kernel to transform the data and then based on this transformation defined a boundary separating data points, ASVs, based on the nodule phenotype of *L. burttii* plants inoculated with soil 2 suspension. The SVM model revealed that *Mesorhizobium* ASV 1 (M.1) was by far the most dominant indicator of a healthy nodule (Fig. 3), which is not surprising as *Mesorhizobium* is the typical symbiont of *L. burttii* (Rodpothong *et al.*, 2009). The second two most influential indicators of a healthy microbiome were *Pseudomonas* ASVs 28 and 57 (P.28 and P.57), which were present in both soil suspension inputs. The three ASVs most indicative of a starved *L. burttii* nodule microbiome were *Rhizobium* ASVs. Once we had identified the genera most characteristic of healthy and starved *L. burttii* nodules we wanted to predict how they interacted. A microbial network was constructed with soil suspension 2 inoculated *L. burttii* samples by using SparCC (Friedman & Alm, 2012) which analysed interactions between and within ASVs from different genera (Figs 4, S6). The ratios of negative to positive interactions within *Rhizobium* and *Mesorhizobium* ASVs were 1.13 and 1.05, respectively. *Pseudomonas* ASVs all correlated positively with one another (number of edges = 8). The ratio of negative to positive interactions between *Rhizobium* and *Mesorhizobium* (ratio = 1.64) was higher compared with this ratio among *Pseudomonas* and *Mesorhizobium* (ratio = 0.77), indicating that symbiotically beneficial *Mesorhizobium* ASVs co-occur with *Pseudomonas*. Strikingly, between *Pseudomonas* and *Rhizobium* ASVs, all correlations were negative (number of edges = 37) also supporting the SVM analysis, which showed that these ASVs were characteristic of healthy and starved *L. burttii* nodules, respectively.

Pseudomonas isolate co-colonises *Mesorhizobium* but not *Rhizobium*-induced nodules

To validate sequencing data, we inoculated *Lotus* plants with bacterial strains isolated from *Lotus* nodules (Table S1). To determine the nodule colonisation pattern of *Pseudomonas* sp. PLb11B, we co-inoculated a fluorescently tagged strain with either *Mesorhizobium* sp. Qb1E3-1, which induces effective nodules or *Rhizobium* sp. BW8-2, which induces ineffective nodules, onto *L. burttii*. Fluorescence microscopy revealed that 32.5% (14/43) of nodules induced by *Mesorhizobium* sp. Qb1E3-1 contained *Pseudomonas* sp. PLb11B. The *Pseudomonas* nodule colonisation was intracellular and was confined to particular areas of the nodule, only infecting a minority of nodule cells (Fig. 5). Conversely no plant nodules (0/22) induced by *Rhizobium* sp. BW8-2 contained *Pseudomonas* sp. PLb11B after a co-inoculation (Fig. 5).

Co-inoculation of a *Pseudomonas* isolate decreases ineffective nodulation by a *Rhizobium* but not a *Mesorhizobium* isolate in a species-specific manner

To investigate if the negative correlation between *Pseudomonas* and *Rhizobium* ASVs in nodules underlay an antagonistic interaction, *Lotus* plants were co-inoculated with nodule isolates *Pseudomonas* sp. Lb2C2 and *Rhizobium* sp. BW8-2 representing the

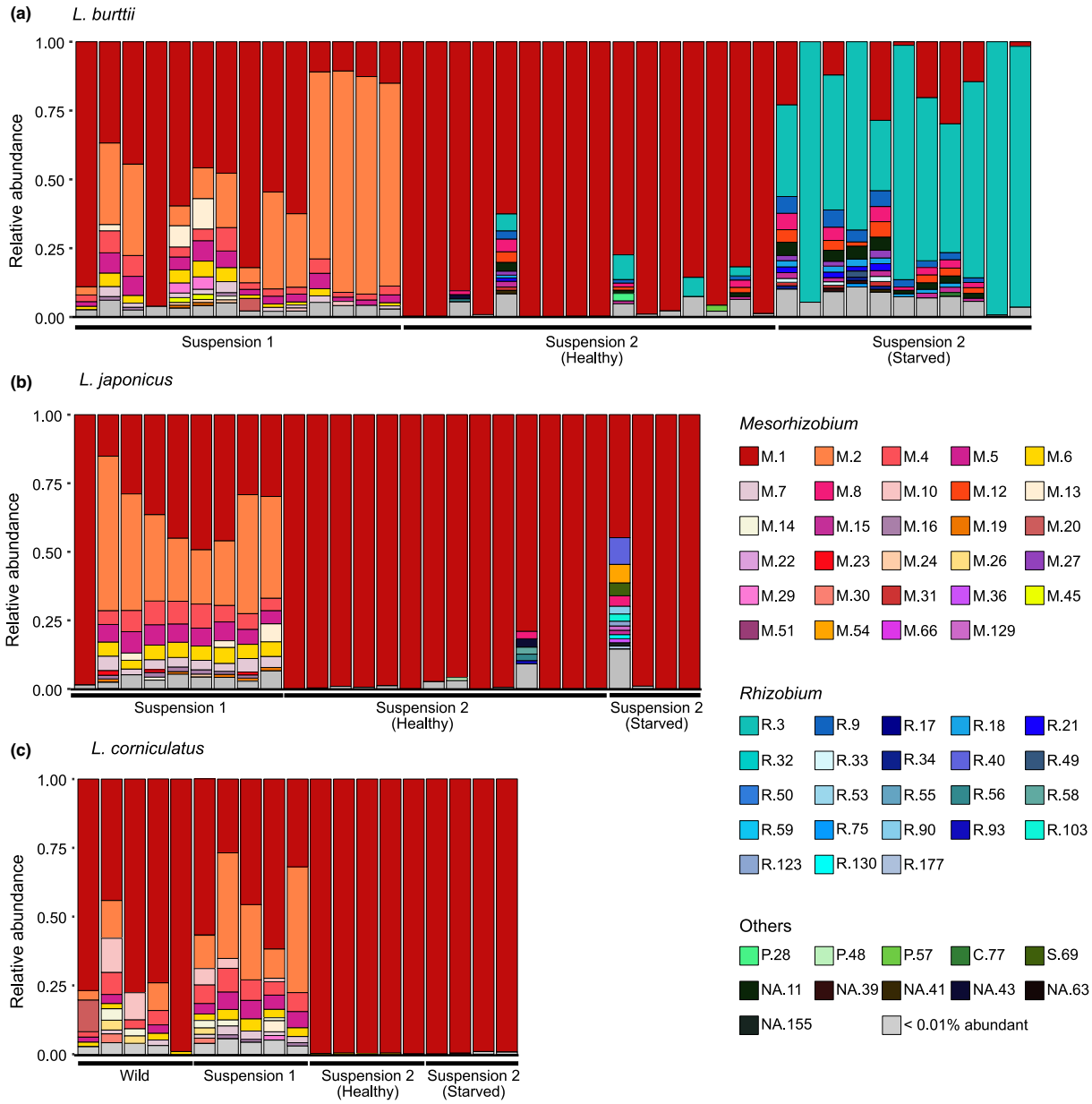


Fig. 2 Community profile showing the relative abundance of amplicon sequence variants (ASVs) present in *Lotus* nodules. The relative abundance of ASVs was estimated for *Lotus burttii* (a), *Lotus japonicus* (b) and *Lotus corniculatus* (c) using transformed data and the PHYLOSEQ v.1.26.1 package in R (McMurdie & Holmes, 2013). *Mesorhizobium* (M) ASVs are depicted in red, yellow, orange, pink and purple shades, *Rhizobium* (R) ASVs are depicted in cyan and blue shades. Other ASVs are depicted in green and black. Amplicon sequence variants < 0.01% abundant are coloured grey. NA, not assigned (taxonomy could only be defined to a Family level).

ASVs in question. *Rhizobium* sp. BW8-2 induced a large number of ineffective nodules and nodule primordia on the roots of *L. japonicus*. Co-inoculation with Lb2C2 significantly decreased the number of nodule structures (Fig. 6). No significant difference was observed regarding root weight and shoot length (Fig. S7). This inhibitory effect was host specific, as no variation in nodule number was observed in *L. burttii* upon co-inoculation with BW8-2 and Lb2C2. By contrast, co-inoculation of Lb2C2 with the effective symbiont *Mesorhizobium* sp. Qb1E3-1 saw no reduction in the nodulation of *L. burttii* or *L. japonicus* (Fig. 6) and only minimal variation in shoot length and root weight (Fig. S7). Inoculation with all three strains did not induce a

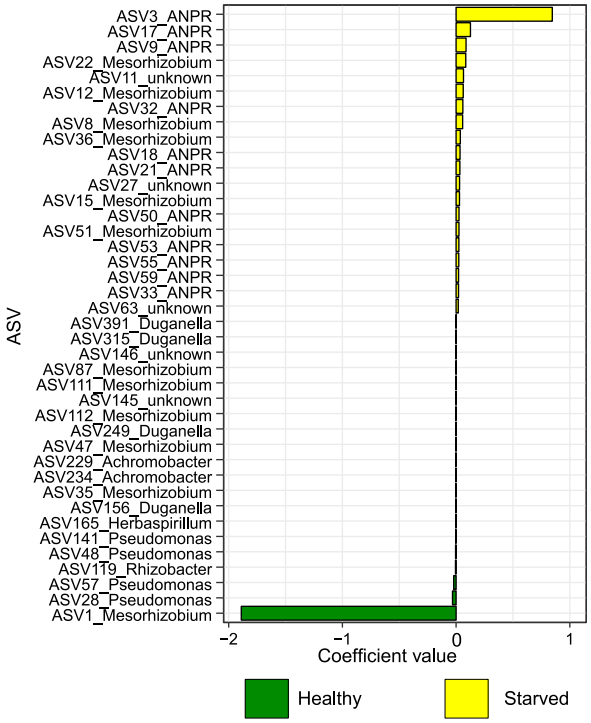


Fig. 3 Indicator amplicon sequence variants (ASVs) of samples. Support Vector Machine linear model from SCIKIT-LEARN packages were used to identify separator ASVs between healthy and starved *Lotus burttii* plants inoculated with soil 2 suspension. Histogram represents the coefficient scores of top 20 ASVs form healthy and starved plants. Negative coefficient values (green bars) represent indicator ASVs in healthy plants while positive values (yellow bars) show indicator ASVs in starved samples. The family of representative ASVs are shown after the ASV number with 'ANPR' indicating *Rhizobium* and 'unknown' indicating that no taxonomy could be assigned at a species level (NA).

different phenotype compared with co-inoculations with *Mesorhizobium* sp. Qb1E3-1 and *Rhizobium* sp. BW8-2 (Fig. S8).

Discussion

Nodules of legumes are not only colonised by rhizobia. Despite this, little information is known about how microbes other than rhizobia affect the root–nodule symbiosis, in particular nodule function and plant health. Here, we characterised variation in the bacterial microbiome of nodules dependent on plant species and soil suspension inoculum and determined correlations between the microbiome makeup and plant health using *16S* rRNA gene amplicon sequencing. Our study revealed that (1) the nodule microbiome of *L. japonicus*, *L. corniculatus* and *L. burttii* is dependent on soil suspension inoculum, (2) the nodule microbiome of starved *L. burttii* plants differs from that of the healthy plants, (3) *Pseudomonas* strains are more prevalent in healthy plant nodules than in starved-plant nodules; co-colonise effective nodules; and reduce the formation of ineffective nodules in a host-specific manner.

Soil suspension input influences *Lotus* spp. nodule microbiome

The nodule microbiome of *Lotus* plants is dependent on the soil suspension inoculum (Fig. S5; Table S3). Soil is the main influencing factor on the rhizosphere, root or nodule microbiomes in nonlegumes (Simonin *et al.*, 2020; Thiergart *et al.*, 2020) and legumes such as *M. truncatula* (Brown *et al.*, 2020) and soybean (Liu *et al.*, 2019; Han *et al.*, 2020). However, many of these studies cite the vast differences in the diversity of the microbial communities or the physicochemical properties of the soil suspension inputs as the reason for the disparity in plant microbiomes (Brown *et al.*, 2020; Han *et al.*, 2020; Simonin *et al.*, 2020). Our results showed that the nodule microbiomes of plants inoculated with different soil suspensions varied significantly (Table S3). This difference is highlighted by soil 1 suspension-inoculated nodules being colonised by a range of *Mesorhizobium* ASVs and soil 2 suspension nodules almost colonised exclusively by *Mesorhizobium* ASV M.1 (Fig. 2). Also, plants grown in soil 1 suspension produced, on average, more nodules and had a broader range of shoot growth than those inoculated with soil 2 suspension (Figs 1, S2). However, the original soil suspensions inoculated onto the plants showed no differences in alpha diversity and only slight, although not significant, differences in beta diversity (Fig. S4; Table S3). The soils from which the suspensions were produced also had no noteworthy differences in their microbiome diversity or physicochemical properties (Fig. S4; Table S2). This suggests that lowly abundant soil microbes that do not sway diversity measures, may play a pivotal role in how the microbiome functions as a whole. Such a phenomenon has been described in peat soil, where a *Desulfosporosinus* sp., which comprised only 0.006% of the total microbiome, acted as an important sulphate reducer in the biogeochemical process that diverts carbon flow from methane to CO_2 (Pester *et al.*, 2010). Also, *Bacillus* species, typically found at a low abundance in the rhizosphere compared with rhizobia, increase the number of nodules and/or the size of nodules in legumes (Rajendran *et al.*, 2008; Mishra *et al.*, 2009; Schwartz *et al.*, 2013).

Starved *L. burttii* plant nodules harbour a microbiome different to that of healthy plants

Lotus burttii is the only species that we tested that showed a significant difference between the nodule microbiome of heathy and starved plants. Nodules of starved *L. burttii* plants were dominated by *Rhizobium* ASVs, while the nodules of healthy plants were predominantly colonised by *Mesorhizobium* ASVs. *Lotus burttii* is known to form infected but ineffective nodules upon inoculation with *Rhizobium leguminosarum* Norway, however this does not form nodules on *L. japonicus* or *L. corniculatus* (Gossmann *et al.*, 2012). This correlates with the observation that starved *L. japonicus* and *L. corniculatus* harboured nodules that were not dominated by *Rhizobium*, but rather by *Mesorhizobium*

New Phytologist

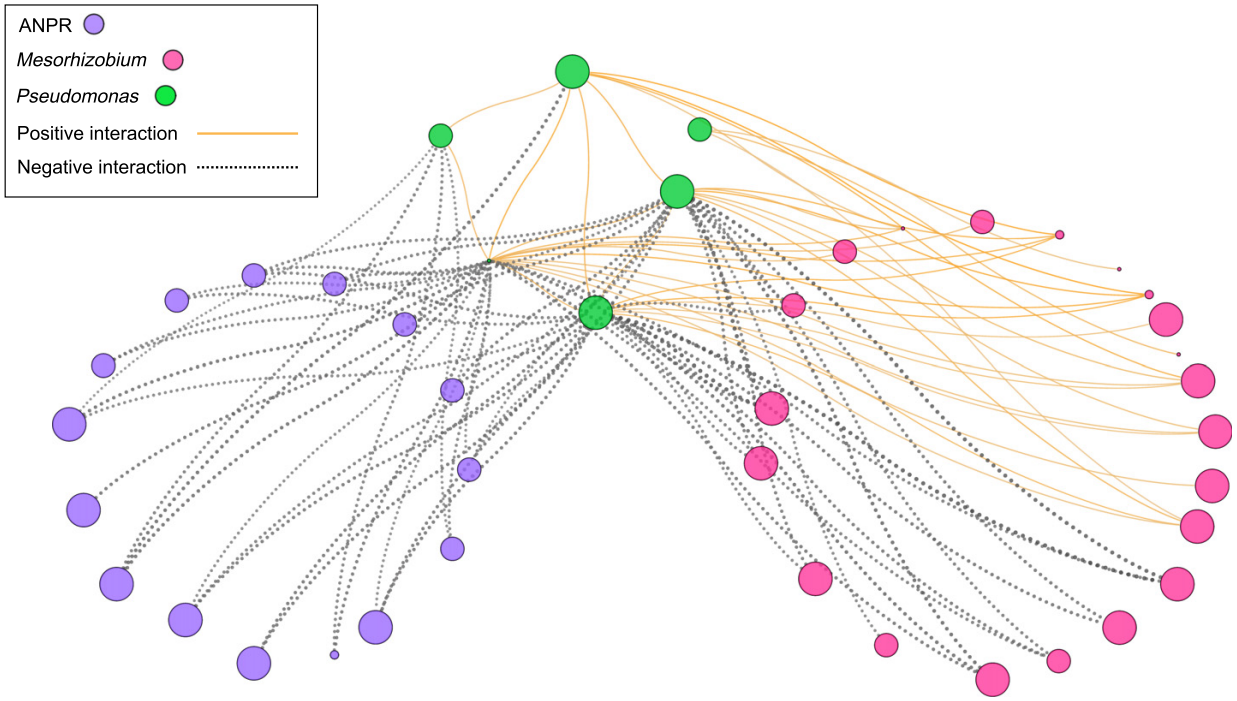


Fig. 4 Microbial co-occurrence network of *Lotus burttii*. An amplicon sequence variant (ASV) table of soil suspension 2-inoculated *L. burttii* samples was used to infer a correlation network SparCC (Friedman & Alm, 2012) algorithm implemented using the FastSpar (Watts *et al*., 2019) tool. The nodes (dots) of this network corresponding to ASVs are grouped and coloured by genus. Node size indicates the relative abundance. Each edge (line) between two ASVs represents either a positive (orange line) or negative (grey-dashed line) correlation. Only significant correlations ($|R| \geq 0.2$, $P < 0.01$) between Rhizobiaceae and Pseudomonadaceae families and first neighbours of Pseudomonadaceae are shown in the network.

ASVs, similar to the microbiome of healthy plants (Fig. 2). The variation in the starved-plant nodule microbiomes of *Lotus* species may be explained by how readily each plant is nodulated. Liang *et al.* (2019) described that ineffective *R. leguminosarum* Norway colonises nodules of *L. burttii* via cracks in the epidermis. *Lotus burttii* is more susceptible to less-specific infections (Zarrabian *et al.*, 2021), which is likely to increase its vulnerability to forming an ineffective symbiosis. This reduced specificity by *L. burttii* is also highlighted in the number of starved plants that contained nodules. In total, 73.8% of starved *L. burttii* plants grew nodules, much more than in *L. japonicus* and *L. corniculatus* (Fig. S2). The higher frequency of nodulation coupled with the reduced specificity that *L. burttii* exhibits in choosing a nodulation partner might leave the plant susceptible to expending energy on ineffective symbiotic processes, resulting in the starvation of the plant. Conversely, *L. corniculatus* and *L. japonicus* do not exhibit this same level of promiscuity, which is evidenced by their nodules being dominated by *Mesorhizobium* in all sample types. The reason as to why a starved plant would harbour a nodule microbiome similar to that of a healthy plant remains to be elucidated. We postulate that it may be simply a delay in the establishment of a successful symbiosis or due to being colonised by nonnitrogen-fixing *Mesorhizobium* strains. Rodpothong *et al.*, came to similar conclusions when inoculating *Mesorhizobium loti* Nod factor synthesis mutants onto different *Lotus* species. Nodulation of *L. burttii* was unaffected, while *L. japonicus* and *L. corniculatus* exhibited delayed nodulation and reduced infection (Rodpothong *et al.*, 2009). Taken together, our results support the idea that the reduced specificity exhibited by *L. burttii* during root–nodule symbiosis allows for a broader range of bacteria to colonise its nodules.

Pseudomonas ASVs are more prevalent in healthy *L. burttii* nodules and can reduce ineffective nodulation in *L. japonicus*

Although the microbiota of all nodule types were dominated by Rhizobiales bacteria, there was a small contingent of non-Rhizobiales ASVs detected as well (Fig. 2). This is not uncommon in legume nodules, as non-Rhizobiales bacteria are often isolated from nodules. Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria and Actinobacteria have all been found in various legumes nodules (Benhizia *et al.*, 2004; Dey *et al.*, 2004; Cummings *et al.*, 2009; Ibáñez *et al.*, 2009; Ampomah & Huss-Danell, 2011; Zhao *et al.*, 2013; Dobritsa & Samadpour, 2016; Ferchichi *et al.*, 2019). Of the non-rhizobia that were present in *Lotus* nodules, *Pseudomonas* was the most prevalent. We found that *Pseudomonas* ASVs were characteristic of healthy, but not of starved, *L. burttii* nodules (Fig. 3) suggesting that they have the potential to support plant health. Previous studies have shown that

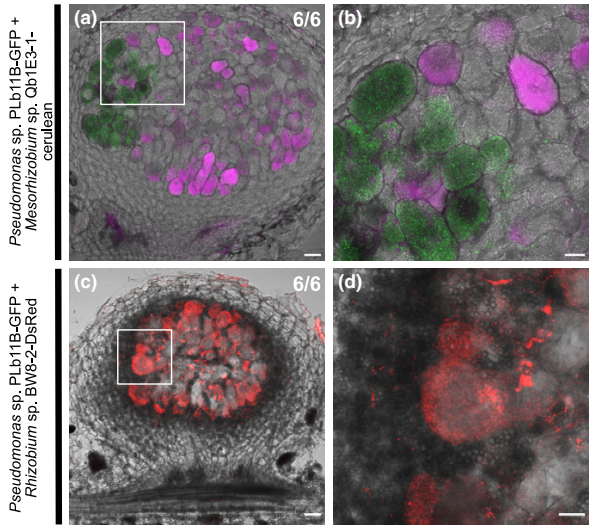


Fig. 5 *Pseudomonas* sp. PLb11B-GFP colonisation in nodules induced by *Mesorhizobium* sp. Qb1E3-1-cerulean and *Rhizobium* sp. BW8-2-DsRed. Nodules were prescreened for signs of fluorescence before sectioning. (a) Overview and (b) zoomed-in confocal microscopy image of *Pseudomonas* sp. PLb11B-GFP (green) and *Mesorhizobium* sp. Qb1E3-1-cerulean (magenta) colonisation in an effective *Lotus burttii* nodule. Overview image bar, 50 μm. Zoomed image bar, 20 μm. (c) Overview and (d) zoomed-in confocal microscopy image of *Pseudomonas* sp. PLb11B (green) and *Rhizobium* sp. BW8-2 (red) colonisation in an ineffective *L. burttii* nodule. Overview image bar, 50 μm. Zoomed image bar, 20 μm.

Pseudomonas can influence plant growth directly by producing siderophores, solubilising phosphate and producing indoleacetic acid (Dey *et al.*, 2004; Ibáñez *et al.*, 2009; Zhao *et al.*, 2013; Ferchichi *et al.*, 2019) or indirectly via antagonistic behaviour towards phytopathogenic fungi (Sindhu & Dadarwal, 2001; Chandra *et al.*, 2020). A *Pseudomonas* strain isolated from *Sophora alopecuroides* also promotes plant growth upon reinoculation with *Mesorhizobium* (Zhao *et al.*, 2013). We posit that potential microbe–microbe interactions involving *Pseudomonas* also influence the outcome of the root–nodule symbiosis. To analyse any potential microbe–microbe interactions within the nodules we looked for interactions between nodule ASVs. Network analysis comparing the nodule microbiome of healthy and starved *L. burttii* plants revealed significant negative correlations between *Pseudomonas* ASVs and multiple *Rhizobium* ASVs, as well as positive and negative interactions with *Mesorhizobium* ASVs (Fig. 4). These predicted interactions were supported by co-inoculating either an ineffective symbiont, *Rhizobium* sp. BW8-2, or an effective symbiont, *Mesorhizobium* sp. Qb1E3-1 with *Pseudomonas* sp. PLb11B. Each isolate had been previously isolated from *L. burttii* nodules, however it was found that *Pseudomonas* sp. PLb11B was only present in *Mesorhizobium*-induced nodules. Using fluorescently tagged strains and microscopy we found that 32.5% of nodules formed by the *Mesorhizobium* and 0% of nodules formed by *Rhizobium* contained *Pseudomonas*. *Pseudomonas* bacteria have been shown to colonise root hairs (Berggren *et al.*, 2005) or nodules intercellularly (Pastor-Bueis *et al.*, 2021). But in contrast, we found that *Pseudomonas* sp. PLb11B infection was intracellular and typically confined to small regions of each nodule with only a small number of cells showing extensive colonisation (Fig. 5). The lack of *Pseudomonas* sp. PLb11B in *Rhizobium* sp. BW8-2 induced nodules aligned with the sequencing data and interaction network observations. This negative interaction was further highlighted when observing another *Pseudomonas* isolate, Lb2C2, co-inoculated with *Rhizobium* sp. BW8-2 on *L. japonicus*. There was a significant reduction in the number of nodules and nodule primordia in *L. japonicus* compared with the single inoculation with the *Rhizobium* sp. BW8-2. Noticeably, this effect was host and inoculum specific, as no reduction in nodule number was observed in *L. burttii* or in co-inoculations of *Pseudomonas* with *Mesorhizobium* (Figs 6, S8). This contrasts with publications that suggested that *Pseudomonas* and *Rhizobium* strains interact synergistically (Tilak *et al.*, 2006; Egamberdieva *et al.*, 2010, 2013; Sánchez *et al.*, 2014). Interactions can be direct, for example filtrates from *Rhizobium* sp. increasing the cell density of *Pseudomonas fluorescens* (Samavat *et al.*, 2011), or mediated via the plant, for example indoleacetic acid produced by *Pseudomonas* sp. resulted in a more extensive root system in *Galega officinalis* and an increased number of potential infection sites for the compatible *Rhizobium* sp. (Egamberdieva *et al.*, 2013). The negative correlation we observed between *Pseudomonas* and *Rhizobium* ASVs in *L. burttii* nodules may also have been due to an indirect effect mediated by *Mesorhizobium*. Negative correlations were also seen between *Mesorhizobium* ASVs and *Rhizobium* ASVs. This can be explained by both bacteria competing for nodule colonisation. Significant positive correlations were apparent between *Pseudomonas* and *Mesorhizobium* ASV M.1, which was dominant in the healthy nodules of plants inoculated with soil suspension 2 (Fig. 2). Positive interactions have already been seen after the co-inoculation of *Pseudomonas* sp. isolates with a *Mesorhizobium* sp., which led to an increase in nodule number in chickpea (Malik & Sindhu, 2011). Positive correlations between *Mesorhizobium* and *Pseudomonas* coupled with the reduction in ineffective nodulation by cheater rhizobia upon co-inoculation with *Pseudomonas* supported the hypothesis that these *Pseudomonas* ASVs have a beneficial role in root–nodule symbiosis.

Our results add to the growing assertion that the soil microbiome, including non-Rhizobiales bacteria, greatly shape the overall functionality of root–nodule symbiosis and healthy plant growth (Martínez-Hidalgo & Hirsch, 2017). The ability for *Pseudomonas* to selectively colonise healthy plant nodules and reduce the number of ineffective nodules in *L. japonicus* indicated that root–nodule symbiosis is influenced by the broader soil microbiota. This research will aid the construction of synthetic communities capable of recreating observed patterns in a bid to narrow down which soil microbes and which microbe–microbe interactions are pivotal in forming the ideal microbiome to maximise plant growth.

Acknowledgements

We thank Juan Liang for help to isolate rhizobia strains, Anke Becker and Doreen Meier for providing the pABC-cerulean

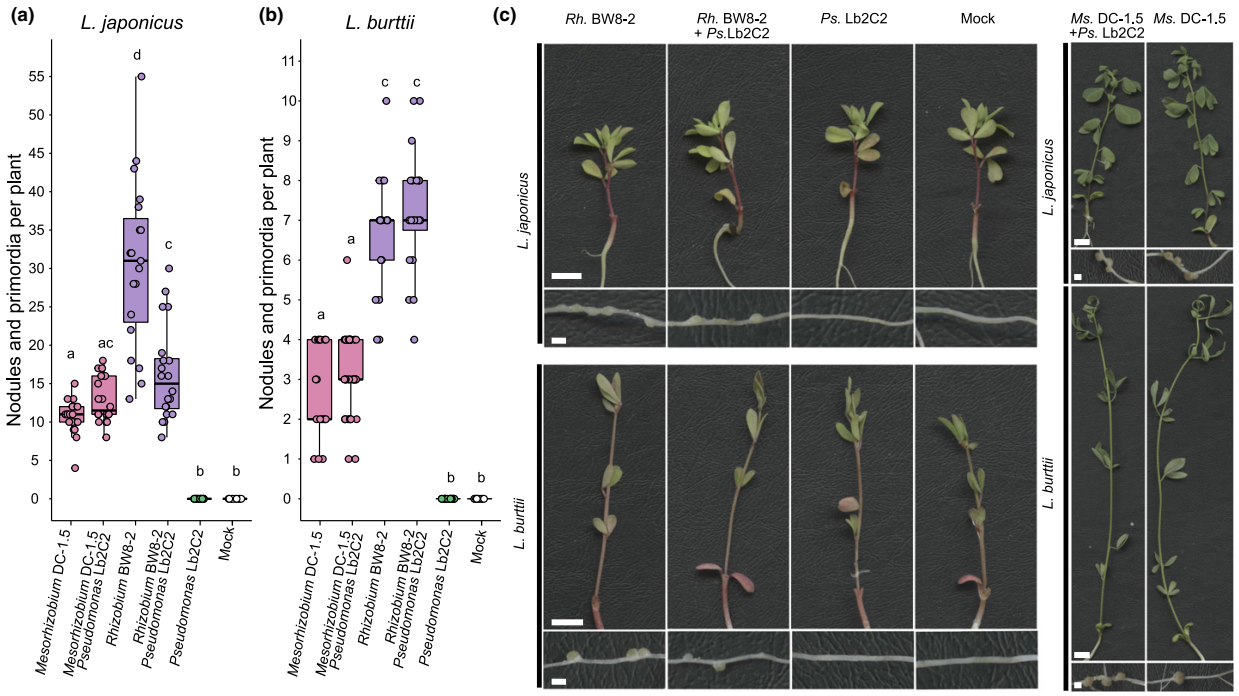


Fig. 6 Nodule organogenesis phenotype of *Lotus* plants inoculated with *Rhizobium* sp. BW8-2, *Mesorhizobium* sp. DC-1.5, and *Pseudomonas* sp. Lb2C2. Box plots of the number of nodules and nodule primordia formed on *Lotus japonicus* (a) and *Lotus burttii* (b) roots. In total, 20 plants were inoculated with *Rhizobium* sp. BW8-2, *Mesorhizobium* sp. DC-1.5 or *Pseudomonas* sp. Lb2C2 nodule isolates. *Lotus burttii* and *L. japonicus* were harvested at 4 and 5 wk post inoculation, respectively. Significance calculated using ANOVA and Tukey HSD is indicated as lowercase letters. Each point represents the number of nodules in one plant. The bold black line and the box depict the median and the interquartile range, respectively. (c) Representative images of root and shoot phenotypes of each inoculation treatment. Shoots bar, 5 mm. Nodules bar, 1 mm.

plasmid as well as Anna Beatrice Preinsberger and Oscar Gonzalez Puiggené for assisting in the construction of conjugated isolates. This work is supported by the German Research Foundation grants MA 7269/2-1 and Schl446/38-1. Open access funding enabled and organised by ProjektDEAL.

Author contributions

DBC and MMarín conceived and designed the study. DBC performed the experiments. DBC and MMahmoudi analysed the data. AB and VR performed MiSeq sequencing. EK and MS contributed with reagents and materials. DBC, MMarín and MMahmoudi contributed to writing and preparation of the manuscript. All authors read and approved the final manuscript.

ORCID

Andreas Brachmann https://orcid.org/0000-0001-7980-8173
Duncan B. Crosbie https://orcid.org/0000-0001-9917-1528
Eric Kemen https://orcid.org/0000-0002-7924-116X
Macarena Marín https://orcid.org/0000-0002-6966-2446
Viviane Radl https://orcid.org/0000-0001-6898-5430
Michael Schloter https://orcid.org/0000-0003-1671-1125

Data availability

All *16S* rRNA gene sequencing data were submitted to the Short Read Archive of NCBI and can be found under BioProject accession no. PRJNA731628. The data and scripts used for the machine learning and network analysis are available at https://github.com/IshtarMM/LotusNodules.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

Ampomah OY, Huss-Danell K. 2011. Genetic diversity of root nodule bacteria nodulating *Lotus corniculatus* and *Anthyllis vulneraria* in Sweden. *Systematic and Applied Microbiology* **34**: 267–275.

Bach HJ, Tomanova J, Schloter M, Munch JC. 2002. Enumeration of total bacteria and bacteria with genes for proteolytic activity in pure cultures and in environmental samples by quantitative PCR mediated amplification. *Journal of Microbiological Methods* **49**: 235–245.

Benhizia Y, Benhizia H, Benguedouar A, Muresu R, Giacomini A, Squartini A. 2004. Gamma proteobacteria can nodulate legumes of the genus *Hedysarum*. *Systematic and Applied Microbiology* **27**: 462–468.

Berggren I, Alström S, Van Vuurde J, Mårtensson AM. 2005. Rhizoplane colonisation of peas by *Rhizobium leguminosarum* bv. *viceae* and a deleterious *Pseudomonas putida*. *FEMS Microbiology Ecology* **52**: 71–78.

Bertani G. 1951. Studies on lysogenesis I: the mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology* **62**: 293–300.

Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**: 326–349.

Brown SP, Grillo MA, Podowski JC, Heath KD. 2020. Soil origin and plant genotype structure distinct microbiome compartments in the model legume *Medicago truncatula*. *Microbiome* **8**: 1–17.

Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* **11**: 2639–2643.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581–583.

Chandra H, Kumari P, Bisht R, Prasad R, Yadav S. 2020. Plant growth promoting *Pseudomonas aeruginosa* from *Valeriana wallichii* displays antagonistic potential against three phytopathogenic fungi. *Molecular Biology Reports* **47**: 6015–6026.

Cummings SP, Gyaneshwar P, Vinuesa P, Farruggia FT, Andrews M, Humphry D, Elliott GN, Nelson A, Orr C, Pettitt D. 2009. Nodulation of *Sesbania* species by *Rhizobium* (*Agrobacterium*) strain IRBG74 and other rhizobia. *Environmental Microbiology* **11**: 2510–2525.

Cytryn E. 2013. The soil resistome: the anthropogenic, the native, and the unknown. *Soil Biology and Biochemistry* **63**: 18–23.

Dey R, Pal K, Bhatt D, Chauhan S. 2004. Growth promotion and yield enhancement of peanut (*Arachis hypogaea* L.) by application of plant growth-promoting rhizobacteria. *Microbiological Research* **159**: 371–394.

Dobritsa AP, Samadpour M. 2016. Transfer of eleven species of the genus *Burkholderia* to the genus *Paraburkholderia* and proposal of *Caballeronia* gen. nov. to accommodate twelve species of the genera *Burkholderia* and *Paraburkholderia*. *International Journal of Systematic and Evolutionary Microbiology* **66**: 2836–2846.

Dorn-In S, Bassitta R, Schwaiger K, Bauer J, Hölzel CS. 2015. Specific amplification of bacterial DNA by optimized so-called universal bacterial primers in samples rich of plant DNA. *Journal of Microbiological Methods* **113**: 50–56.

Downie JA. 2005. Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules. *Current Biology* **15**: R196–R198.

Egamberdieva D, Berg G, Lindström K, Räsänen LA. 2010. Co-inoculation of *Pseudomonas* spp. with *Rhizobium* improves growth and symbiotic performance of fodder galega (*Galega orientalis* Lam.). *European Journal of Soil Biology* **46**: 269–272.

Egamberdieva D, Berg G, Lindström K, Räsänen LA. 2013. Alleviation of salt stress of symbiotic *Galega officinalis* L. (goat's rue) by co-inoculation of *Rhizobium* with root-colonizing *Pseudomonas*. *Plant and Soil* **369**: 453–465.

Ferchichi N, Toukabri W, Boularess M, Smaoui A, Mhamdi R, Trabelsi D. 2019. Isolation, identification and plant growth promotion ability of endophytic bacteria associated with lupine root nodule grown in Tunisian soil. *Archives of Microbiology* **201**: 1333–1349.

Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. *PLoS Computational Biology* **8**: e1002687.

Gamborg OL, Miller RA, Ojima K. 1968. Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research* **50**: 151–158.

Gossmann JA, Markmann K, Brachmann A, Rose LE, Parniske M. 2012. Polymorphic infection and organogenesis patterns induced by a *Rhizobium leguminosarum* isolate from *Lotus* root nodules are determined by the host genotype. *New Phytologist* **196**: 561–573.

Graves S, Piepho H-P, Selzer ML. 2015. *Package 'MULTCOMPVIEW'. Visualizations of paired comparisons*. [WWW document] URL https://cran.r-project.org/web/packages/multcompView/multcompView.pdf [accessed 14 July 2021].

Gu S, Wei Z, Shao Z, Friman V-P, Cao K, Yang T, Kramer J, Wang X, Li M, Mei X *et al.* 2020. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. *Nature Microbiology* **5**: 1002–1010.

Han Q, Ma Q, Chen Y, Tian B, Xu L, Bai Y, Chen W, Li X. 2020. Variation in rhizosphere microbial communities and its association with the symbiotic efficiency of rhizobia in soybean. *The ISME Journal* **14**: 1915–1928.

Hansen BL, Pessotti RdC, Fischer MS, Collins A, El-Hifnawi L, Liu MD, Traxler MF. 2020. Cooperation, competition, and specialized metabolism in a simplified root nodule microbiome. *mBio* **11**: e01917-20.

Herrera-Cervera JA, Caballero-Mellado J, Laguerre G, Tichy H-V, Requena N, Amarger N, Martínez-Romero E, Olivares J, Sanjuan J. 1999. At least five rhizobial species nodulate *Phaseolus vulgaris* in a Spanish soil. *FEMS Microbiology Ecology* **30**: 87–97.

Howieson JG, Yates RJ, O'Hara GW, Ryder M, Real D. 2005. The interactions of *Rhizobium leguminosarum* biovar *trifolii* in nodulation of annual and perennial *Trifolium* spp. from diverse centres of origin. *Australian Journal of Experimental Agriculture* **45**: 199–207.

Ibáñez F, Angelini J, Taurian T, Tonelli ML, Fabra A. 2009. Endophytic occupation of peanut root nodules by opportunistic *Gammaproteobacteria*. *Systematic and Applied Microbiology* **32**: 49–55.

Kelly SJ, Muszyński A, Kawaharada Y, Hubber AM, Sullivan JT, Sandal N, Carlson RW, Stougaard J, Ronson CW. 2013. Conditional requirement for exopolysaccharide in the *Mesorhizobium-Lotus* symbiosis. *Molecular Plant–Microbe Interactions* **26**: 319–329.

Kereszt A, Mergaert P, Kondorosi E. 2011. Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? *Molecular Plant–Microbe Interactions* **24**: 1300–1309.

Kramina TE, Degtjareva GV, Samigullin TH, Valiejo-Roman CM, Kirkbride J, Joseph H, Volis S, Deng T, Sokoloff DD. 2016. Phylogeny of *Lotus* (Leguminosae: Loteae): partial incongruence between nrITS, nrETS and plastid markers and biogeographic implications. *Taxon* **65**: 997–1018.

Liang J, Klingl A, Lin Y-Y, Boul E, Thomas-Oates J, Marín M. 2019. A subcompatible rhizobium strain reveals infection duality in *Lotus*. *Journal of Experimental Botany* **70**: 1903–1913.

Liu F, Hewezi T, Lebeis SL, Pantalone V, Grewal PS, Staton ME. 2019. Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiology* **19**: 1–19.

Malik DK, Sindhu SS. 2011. Production of indole acetic acid by *Pseudomonas* sp.: effect of coinoculation with *Mesorhizobium* sp. *Cicer* on nodulation and plant growth of chickpea (*Cicer arietinum*). *Physiology and Molecular Biology of Plants* **17**: 25–32.

Marcos-García M, Menéndez E, Cruz-González X, Velázquez E, Mateos PF, Rivas R. 2015. The high diversity of *Lotus corniculatus* endosymbionts in soils of northwest Spain. *Symbiosis* **67**: 11–20.

Martin M. 2011. CUTADAPT removes adapter sequences from high-throughput sequencing reads. *Embnet.Journal* **17**: 10–12.

Martínez-Hidalgo P, Hirsch AM. 2017. The nodule microbiome: N_2-fixing rhizobia do not live alone. *Phytobiomes* **1**: 70–82.

McMurdie PJ, Holmes S. 2013. PHYLOSEQ: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**: e61217.

Mishra PK, Mishra S, Selvakumar G, Kundu S, Shankar GH. 2009. Enhanced soybean (*Glycine max* L.) plant growth and nodulation by *Bradyrhizobium japonicum-SB1* in presence of *Bacillus thuringiensis-KR1*. *Acta Agriculturae Scandinavica Section B – Soil and Plant Science* **59**: 189–196.

Noble WS. 2006. What is a support vector machine? *Nature Biotechnology* **24**: 1565–1567.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'hara R, Simpson G, Solymos P, Stevens MHH, Wagner H. 2018. *Community ecology package*. R package v.2.5-2. [WWW document] URL https://CRAN.R-project.org/package=vegan [accessed 22 April 2020].

Pastor-Bueis R, Jimenez-Gomez A, Barquero M, Mateos PF, González-Andrés F. 2021. Yield response of common bean to co-inoculation with *Rhizobium* and *Pseudomonas* endophytes and microscopic evidence of different colonised spaces inside the nodule. *European Journal of Agronomy* **122**: 126187.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. 2011. SCIKIT-LEARN: machine learning in PYTHON. *The Journal of Machine Learning Research* **12**: 2825–2830.

Pester M, Bittner N, Deevong P, Wagner M, Loy A. 2010. A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *The ISME Journal* **4**: 1591–1602.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**: D590–D596.

Rajendran G, Sing F, Desai AJ, Archana G. 2008. Enhanced growth and nodulation of pigeon pea by co-inoculation of *Bacillus* strains with *Rhizobium* spp. *Bioresource Technology* **99**: 4544–4550.

Remigi P, Zhu J, Young JPW, Masson-Boivin C. 2016. Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends in Microbiology* **24**: 63–75.

Rodpothong P, Sullivan JT, Songsrirote K, Sumpton D, Cheung KW-T, Thomas-Oates J, Radutoiu S, Stougaard J, Ronson CW. 2009. Nodulation gene mutants of *Mesorhizobium loti* R7A—*nodZ* and *nolL* mutants have host-specific phenotypes on *Lotus* spp. *Molecular Plant–Microbe Interactions* **22**: 1546–1554.

Sachs JL, Simms EL. 2008. The origins of uncooperative rhizobia. *Oikos* **117**: 961–966.

Samavat S, Besharati H, Behboudi K. 2011. Interactions of rhizobia cultural filtrates with *Pseudomonas fluorescens* on bean damping-off control. *Journal of Agricultural Science and Technology* **13**: 965–976.

Sánchez AC, Gutiérrez RT, Santana RC, Urrutia AR, Fauvart M, Michiels J, Vanderleyden J. 2014. Effects of co-inoculation of native *Rhizobium* and *Pseudomonas* strains on growth parameters and yield of two contrasting *Phaseolus vulgaris* L. genotypes under Cuban soil conditions. *European Journal of Soil Biology* **62**: 105–112.

Sandal N, Jin H, Rodriguez-Navarro DN, Temprano F, Cvitanich C, Brachmann A, Sato S, Kawaguchi M, Tabata S, Parniske M. 2012. A set of *Lotus japonicus* Gifu× *Lotus burttii* recombinant inbred lines facilitates map-based cloning and QTL mapping. *DNA Research* **19**: 317–323.

Sandman K, Ecker C. 2014. *Pseudomonas* isolation and identification: an introduction to the challenges of polyphasic taxonomy. *Journal of Microbiology & Biology Education* **15**: 287–291.

Schumpp O, Deakin WJ. 2010. How inefficient rhizobia prolong their existence within nodules. *Trends in Plant Science* **15**: 189–195.

Schwartz A, Ortiz I, Maymon M, Herbold C, Fujishige N, Vijanderan J, Villella W, Hanamoto K, Diener A, Sanders E *et al.* 2013. *Bacillus simplex*—a little known PGPB with anti-fungal activity—alters pea legume root architecture and nodule morphology when coinoculated with *Rhizobium leguminosarum* bv. *viciae*. *Agronomy* **3**: 595–620.

Seefeldt LC, Hoffman BM, Dean DR. 2009. Mechanism of Mo-dependent nitrogenase. *Annual Review of Biochemistry* **78**: 701–722.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. CYTOSCAPE: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**: 2498–2504.

Simonin M, Dasilva C, Terzi V, Ngonkeu EL, Diouf D, Kane A, Béna G, Moulin L. 2020. Influence of plant genotype and soil on the wheat rhizosphere microbiome: evidences for a core microbiome across eight African and European soils. *FEMS Microbiology Ecology* **96**: fiaa067.

Sindhu S, Dadarwal K. 2001. Chitinolytic and cellulolytic *Pseudomonas* sp. antagonistic to fungal pathogens enhances nodulation by *Mesorhizobium* sp. *Cicer* in chickpea. *Microbiological Research* **156**: 353–358.

Streeter JG. 1994. Failure of inoculant rhizobia to overcome the dominance of indigenous strains for nodule formation. *Canadian Journal of Microbiology* **40**: 513–522.

Thiergart T, Durán P, Ellis T, Vannier N, Garrido-Oter R, Kemen E, Roux F, Alonso-Blanco C, Ågren J, Schulze-Lefert P *et al.* 2020. Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations. *Nature Ecology & Evolution* **4**: 122–131.

Tilak K, Ranganayaki N, Manoharachari C. 2006. Synergistic effects of plant-growth promoting rhizobacteria and *Rhizobium* on nodulation and nitrogen fixation by pigeonpea (*Cajanus cajan*). *European Journal of Soil Science* **57**: 67–71.

Töwe S, Wallisch S, Bannert A, Fischer D, Hai B, Haesler F, Kleineidam K, Schloter M. 2011. Improved protocol for the simultaneous extraction and column-based separation of DNA and RNA from different soils. *Journal of Microbiological Methods* **84**: 406–412.

Tyc O, van den Berg M, Gerards S, van Veen JA, Raaijmakers JM, De Boer W, Garbeva P. 2014. Impact of interspecific interactions on antimicrobial activity among soil bacteria. *Frontiers in Microbiology* **5**: 567.

Venado RE, Liang J, Marín M. 2020. Rhizobia infection, a journey to the inside of plant cells. *Advances in Botanical Research* **94**: 97–118.

Vincent JM. 1970. *A manual for the practical study of the root-nodule bacteria.* Oxford, UK: Blackwell Scientific.

Wang D, Yang S, Tang F, Zhu H. 2012. Symbiosis specificity in the legume–rhizobial mutualism. *Cellular Microbiology* **14**: 334–342.

Watts SC, Ritchie SC, Inouye M, Holt KE. 2019. FASTSPAR: rapid and scalable correlation estimation for compositional data. *Bioinformatics* **35**: 1064–1066.

Werner D, Wilcockson J, Kalkowski B. 1975. Nitrogen fixing activity in separated from plant cell cultures. *Zeitschrift für Naturforschung C* **30**: 687–688.

Wickham H. 2009. Elegant graphics for data analysis. *Media* **35**: 10.1007.

William S, Feil H, Copeland A. 2012. Bacterial genomic DNA isolation using CTAB. *Sigma* **50**: 6876.

Winarno R, Lie T. 1979. Competition between *Rhizobium* strains in nodule formation: interaction between nodulating and non-nodulating strains. *Plant and Soil* **51**: 135–142.

Zarrabian M, Montiel J, Sandal N, Jin H, Lin YY, Klingl V, Marín M, James E, Parniske M, Stougaard J *et al.* 2021. A promiscuity locus confers *Lotus burttii* nodulation with rhizobia from five different genera. *bioRxiv*. doi: 10.1101/2021.08.26.457880.

Zgadzaj R, Garrido-Oter R, Jensen DB, Koprivova A, Schulze-Lefert P, Radutoiu S. 2016. Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *Proceedings of the National Academy of Sciences, USA* **113**: E7996–E8005.

Zhao LF, Xu YJ, Ma ZQ, Deng ZS, Shan CJ, Wei GH. 2013. Colonization and plant growth promoting characterization of endophytic *Pseudomonas chlororaphis* strain Zong1 isolated from *Sophora alopecuroides* root nodules. *Brazilian Journal of Microbiology* **44**: 629–637.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 Metadata of all samples.

Fig. S1 Reproducibility of plant growth experiments.

Fig. S2 Number of nodules per plant after inoculation with soil suspensions.

Fig. S3 Rarefaction curves of sequencing data.

Fig. S4 Nodule microbiome alpha diversity plotted by species and soil suspension input.

Fig. S5 Global principal coordinate analysis of all samples.

Fig. S6 Overview network analysis.

Fig. S7 Root weight and shoot length phenotype of *Lotus* plants inoculated with *Rhizobium* sp. BW8-2, *Mesorhizobium* sp. DC-1.5, and *Pseudomonas* sp. Lb2C2.

Fig. S8 Nodule organogenesis phenotype of *Lotus* plants inoculated with *Rhizobium* sp. BW8-2, *Mesorhizobium* sp. DC-1.5, and *Pseudomonas* sp. Lb2C2.

Table S1 Strains and plasmids used.

Chapter 6

General discussion

6.1 Decoding the influences of environmental factors on leaf microbiome dynamics: a long-term study of *Arabidopsis thaliana* populations

The leaf microbiome is a dynamic and complex ecosystem structurally altered by various abiotic and biotic factors. Among these factors, critical determinants in shaping microbial communities within natural plants include geographic location, plant compartment, and growing season (Runge et al., 2023, Durán et al., 2018, Ou et al., 2019, Ware et al., 2021, Edwards et al., 2015). However, there remains a significant gap in our understanding of the role of environmental factors (e.g., temperature and precipitation) in shaping microbial communities across multiple plant generations and sampling sites. Furthermore, the extent to which the stability of biotic microbial interaction networks is vulnerable to environmental factors remains an unanswered question. To address these fundamental questions, in Chapter 2 of my thesis, I conducted a comprehensive study focusing on the *Arabidopsis thaliana* leaf microbiome. Collected over five years from natural populations from six geographical locations in the Tübingen region of Germany (Agler et al., 2016). The results of this study are discussed in the following sections.

6.1.1 Most effective environmental factors shaping leaf microbiome's structure and microbial interactions

Our study in Chapter 2 examined the impact of environmental factors on the leaf microbiome, revealing a significant influence on associated microbial communities. Specifically, our investigation showed that solar radiation and humidity-related factors are the most critical factors shaping the leaf microbiome. Here, we identified that solar radiation negatively correlates with microbial alpha-diversity. This effect could be directly caused by damage to microbial DNA, particularly on the leaf surface. Alternatively, it could indirectly affect diversity by promoting the production of reactive oxygen species (ROS) that inhibit the growth and diversity of sensitive species. Long-term low-dose ionizing radiation has been shown to affect soil microbial communities by inhibiting predatory or

parasitic fungi (Cheng et al., 2023). While analyzing microbial interaction networks, solar radiation emerged as an essential factor positively correlated with their connectivity. This could represent a survival strategy in which the microorganisms establish a stronger bond or dependency to cope with the environmental burden of radiation exposure and thus become more resistant to outside disturbances.

Factors related to humidity, such as precipitation and vapor pressure, significantly contributed to higher microbial alpha-diversity. This pattern could be related to modifying substrate diffusion (Yang et al., 2021) and facilitating microbial dispersal via rainfall. However, we found negative correlations between the parameters of high humidity and the connectivity of the microbial network. This suggests that under conditions such as precipitation, microbes may choose strategies such as adhesion over motility. This may reduce the connectivity of microbial interactions. In addition, intense precipitation may physically disrupt microbial habitats and structures, such as biofilms or microbial aggregates, leading to temporary disintegration of microbial networks and reduced connectivity. Although microbial variations have been identified in the plant compartments under seasonal variations of environmental factors, there is still a gap in our understanding of how the microbiome changes during the growth seasons of the host. I addressed this research question in more detail in Chapter 3 of my thesis.

6.2 Capturing temporal dynamics patterns of leaf microbiome using time series data

In Chapter 3 of this thesis, we have set up an experiment to monitor the monthly changes of the microbiome of *Arabidopsis* leaves during the natural growth season in the Cologne area. This season ranges from November, which marks seedling emergence, to March, which marks flowering. The experiment was repeated for three consecutive years. This experiment aimed to answer these questions: Are there predictable patterns of microbe-microbe interactions over time? How can we identify the most critical persistent microbes and their role within microbial interaction networks? We investigated these aspects in our second paper (Almario et al., 2022), which I will discuss in the following.

6.2.1 The variability of the leaf microbiome over time and the forces involved in this process

In Chapter 3, we showed that the composition of leaf-associated bacterial, fungal, and oomycete communities showed a high degree of stochasticity, as expected for dynamic ecological systems (Shoemaker et al., 2020). Only half of the observed variability was explained by the examined factors, such as sampling time and plant ecotype. We consistently observed variations in key microbial groups across months. In particular, we observed the prevalence of *Hyaloperonospora*, the downy mildew pathogen of *Arabidop-*

sis leaves. This aligns with findings from different regions (Lundberg et al., 2022). The abundance of *Peronosporales* increased continuously during the growing season. It peaked in March, even though our plants showed no symptoms of downy mildew. This reflects the dynamics of downy mildew in *Brassicaceae*, which thrives in cold, wet weather. It is suggested that for *Arabidopsis* populations in the Cologne area, pathogenic pressure on the plant increases during the early growth stages, especially during the winter months.

6.2.2 Persistence core microbes across three years

Highlighting the remarkable variation within leaf microbial communities, the presence of stable microbes in *Arabidopsis* leaves, referred to as core taxa, represented only 0.62% of all leaf-associated taxa detected. Interestingly, most microbes considered core in one experiment did not retain this status in subsequent experiments. This suggests a tendency for dominant lineages to shift from year to year, consistent with the observed prioritized structuring of leaf microbiomes during the initial stages of colonization. This suggests that communities may be dominated by different microorganisms or core taxa that are interchangeable with each other (Maignien et al., 2014). Our study showed that these core taxa include potential plant pathogens, such as *Hyaloperonospora* and *Cladosporium* (Coates and Beynon, 2010, Thomma et al., 2005), and microorganisms that are beneficial to plants, such as *Sphingomonas* and *Variovorax*. These findings may explain the reason why the sampled plants remained asymptomatic. *Sphingomonas* bacteria residing on leaves have previously been shown to protect *Arabidopsis* from bacterial pathogens (Ritpitakphong et al., 2016) and have been hypothesized to play a role in enhancing plant disease resistance to root fungal pathogens. Meanwhile, *Variovorax* strains have been shown to modulate the hormonal balance of the plant by breaking down auxins. This promotes plant growth, especially under stressful conditions (Finkel et al., 2020). It is worth noting that bacteria and yeasts are found in *A. thaliana* have been reported to affect plant hormone levels by producing auxin-like indolic compounds (Wang et al., 2016).

6.2.3 Highly connected hub component in microbial networks and associations of them with core microbes

We initially hypothesized that the persistence of the identified core taxa would be due to high levels of connectivity within leaf microbial networks. However, contrary to our hypothesis, we found substantial month-to-month variability in individual core taxa's connectivity levels or hubness. In particular, no taxon maintained a consistently high level of connectivity throughout the growing season. This suggests that high connectivity is not required for a taxon to exhibit high prevalence within the leaf microbiome. Therefore, core taxa do not necessarily act as hubs in the network (Stopnisek and Shade,

2021). In particular, two microbes from the *Bacillus* and *Massilia* lineages bucked this trend by being identified as hubs. They reached peak connectivity within leaf microbial networks in February, a month of reduced community complexity, yet their abundances remained stable. This suggests a potential functional link between these hubs and community stability. Highly interconnected microbes may be reliable predictors of microbial community stability, as indicated by previous studies (Wieczorek et al., 2019).

To summarize, our results suggest that although microbial communities exhibit high variability, they still show identifiable temporal trends containing both stable and unstable elements. Future experimental investigations will be crucial to refine our predictions and understand whether and how the presence or removal of central hubs or core components affects the stability of microbial communities over time.

6.3 Exploring pathogen dynamics and probiotic solutions in plant health

The increasing frequency and intensity of plant infections are a significant threat to primary productivity, food security, and biodiversity in susceptible regions of the world. Massive yield losses and ecological disruption result from such epidemics. Plants can be vulnerable to a wide range of pathogens, from biotrophs that obtain nutrients from living cells to necrotrophs that obtain nutrients from dead cells (Singh et al., 2023). The development of effective methods for the control of these pathogens is therefore of vital importance. One method that has gained attention from the ecological and economic perspectives is implementing natural probiotics to maintain plant health (Gu et al., 2022). In natural populations of *Arabidopsis thaliana*, *Albugo laibachii* has been identified as a significant microbe acting as a hub in microbial networks, and we also identified it as a potential core microbe (Agler et al., 2016, Almario et al., 2022). *Albugo* causing white blister rust. This considerably impacts epiphytic and endophytic bacterial colonization (Agler et al., 2016). However, the factors that distinguish the microbiome of infected from uninfected plants and the specific microbes that promote or inhibit the growth of *Albugo* remain unknown. I delve deeper into these aspects in Chapter 4 of my thesis. I will provide a more detailed discussion of our findings in the following sections.

6.3.1 Microbial community differences in *Albugo*-infected and uninfected plants

We conducted a detailed study to compare the microbial communities in plants infected with *Albugo* and those that are not infected. We analyzed a large microbiome dataset collected over five years as described in Chapter 3 (Mahmoudi et al., 2024). We found that infected and uninfected plants differed significantly in diversity and composition. Except for fungi, infected plants had lower diversity, especially within non-fungal eukaryotic

groups. This finding is consistent with our previous research. It suggests that *Albugo* plays a central role in the microbial community of the phyllosphere, resulting in more negative interactions with other microbes (Agler et al., 2016). It has been suggested that hub microbes may be responsible for the reduced diversity observed in infected plants by exerting an indirect effect on diversity through host interactions or by a direct effect on the colonization efficiency of other microbes (Agler et al., 2016). Another study has shown that *Albugo* actively shapes the microbial communities associated with plants by releasing antimicrobial proteins and peptides into the apoplast. This suggests that *Albugo* has a significant impact on plant microbiome diversity (Gómez-Pérez et al., 2022).

6.3.2 Utilizing machine learning to uncover microbiome patterns as biomarkers for disease prediction

This study aimed to establish the microbiome as a potential biomarker for plant health and disease. To accomplish this, we used machine learning techniques that have previously proven effective in predicting clinical and plant disease (Hernández Medina et al., 2022). We distinguished between the health and infection stages of plants using predictive models. The results showed a high level of accuracy, ranging from 75% to 86%. These findings provide robust evidence that predictive patterns exist between different groups, highlighting the potential of the microbiome as a valuable indicator of plant health and disease status. We used feature selection, a commonly applied machine learning technique, to identify the microbial contributors that support these predictive models. This approach allowed us to identify a subset of 3.1% of OTUs, which include bacteria, fungi, and non-fungal eukaryotes, that have the potential to act as effective discriminators. These discriminators are crucial in distinguishing between infected and uninfected plants. They reveal the specific components of the microbiome that provide valuable information for assessing plant health and predicting disease.

6.3.3 Microbial network complexity in the face of pathogen invasion: exploring modularity and interactions in infected communities

Microbial networks are a valuable tool for studying the structure of microbial communities because they can combine different types of information and potentially capture system-level behaviors (Röttjers and Faust, 2018). Our study aims to investigate the impact of infection by *Albugo* on the complexity of microbial interaction networks. Our research indicates that infection by *Albugo* alters microbial interactions and community dynamics by reducing complexity. On the other hand, the uninfected plant's constructed network exhibited greater complexity, as evidenced by increased connections and OTUs. Ecological networks consist of regions with varying levels of connectivity. Regions with dense connections are called modules, while areas with fewer connections define the boundaries between modules. The presence of modules indicates that connections within

the same boundary are stronger than those between different modules. This structure is known as modularity in the context of networks. This study highlights the presence of modularity in these networks. Furthermore, we showed that infected plants' constructed networks exhibit slightly greater modularity than uninfected networks. This finding supports the notion that higher modularity is associated with increased network stability (Grilli et al., 2016).

6.3.4 Prediction and validation of microbiome involved in plant health and disease

Our analysis showed that certain microbes that can differentiate between infected and uninfected plants are distributed across different modules within the microbial community. This distribution highlights their essential functions in the diversity of the community modules within the broader microbial community. The plant microbiome serves as a crucial defense against plant pathogens. It is involved in direct and indirect interactions through various mechanisms, including parasitism, resource competition, and induction of systemic resistance in response to biotic and abiotic stresses (Agler et al., 2016, Chaudhry et al., 2021, Rai et al., 2023). Based on this understanding, our study investigates the potential role of specific microbes in microbe-microbe interactions within the plant microbiome.

6.3.5 Examining the protective roles of diverse plant microbiome components against *Albugo laibachii* infection in planta

Our study was based on the hypothesis that the plant microbiome, regardless of its phylogenetic origin, could help protect against *Albugo laibachii*. The data demonstrate that all of the selected candidates were protective, significantly reducing infection rates, ranging from 87% to 51%. The efficacy of these candidates is ranked in the following order: Among all candidates, *Cystofilobasidum* fungi exhibited the highest level of protection, followed by *Methylobacterium*, *Sphingomonas*, and *Rhogostoma*. Notably, *Cystofilobasidum* fungi demonstrated the most potent protective effects. This finding is exciting due to the limited existing knowledge of their ability to protect *Arabidopsis* leaves from pathogens. However, studies analyzing *Basidiomycete* fungi, including the epiphytic yeast *Moesziomyces*, have provided important insights into their antagonistic behavior towards *Albugo laibachii* in *Arabidopsis*. Using transcriptomic and proteomic approaches, GH25, a hydrolase with lysozyme activity, was a critical effector against *Albugo* (Eitzen et al., 2021).

The bacterial candidates we analyzed have been identified in Chapter 3 as core components of the phyllosphere microbiome (Almario et al., 2022, Bulgarelli et al., 2013). Our results are consistent with the documented protective abilities of *Sphingomonas* species against pathogens such as *Pseudomonas syringae* in *Arabidopsis*. However, colonization

by members of the genus *Methylobacterium* reported inconsistent effects and did not provide plant protection (Innerebner et al., 2011). In contrast to this study, our research revealed a reduction in the infection level of *Albugo*. In addition, our study highlights the recent focus on protists, particularly *Rhogostoma spp* from the phylum *Cercozoa*, in *Arabidopsis thaliana* (Dumack et al., 2017, Bonkowski, 2004). These protists have been identified as influential agents that shape the composition and behavior of bacterial and fungal communities through selective predation. Protists act as natural microbial predators. They directly target bacterial and fungal pathogens, resulting in a consuming effect that ultimately leads to removing various pathogenic strains (Dumack et al., 2016, Nguyen et al., 2023). Our results highlight the potential of fungi, followed by bacteria and, to a lesser extent, non-fungal eukaryotes, to reduce infection of *Albugo* by over 50%. This highlights the idea of functional redundancy among individual microorganisms from different phylogenetic backgrounds and offers promising opportunities for their use in biocontrol methods to control plant infections effectively.

6.3.6 In a planta experiment, the pairwise interactions between disease-associated microbes and *Albugo*

We have experimentally tested the effect of the disease-associated microbes on the level of infection of the *Albugo*. Compared to the control group, one of the candidate *Plectosphaerella* fungi caused the same level of infection together with *Albugo*. Our bacterial candidates, *Pseudomonas* and *Duganella*, decreased by 77% and 65%, respectively. This highlights the adaptability of *Pseudomonas*, which can exhibit a range of behaviors from pathogenic to commensal (Shalev et al., 2022). Our study highlights the diverse effects of disease-associated microbes on *Albugo* infection levels. This may be related to cross-infection relationships. An example of such a relationship has been found in *Arabidopsis thaliana*: Initial exposure to the pathogen *Albugo laibachii* creates an environment that stimulates the growth and spore production of avirulent strains of *Hyaloperonospora arabidopsidis*, the cause of *Arabidopsis* downy mildew, on *Arabidopsis* accessions that would otherwise be resistant to this pathogen (Cooper et al., 2008). This highlights the complex and diverse microbial interactions in plant ecosystems responding to pathogens.

Together, we showed that microbes from healthy plants could reduce the infection level of the pathogens, while the infected-associated microbes play dual roles. They can act together to suppress the pathogen; another role is to regulate its ability to spread.

6.4 Microbial dynamics in the rhizosphere of leguminous *Lotus*

Leguminous plants have a mutualistic relationship with nitrogen-fixing *Rhizobia*. Bacteria are hosted in root nodules in exchange for ammonia. *Rhizobia* do not solely colonize

nodules of legumes. However, little is known about the impact of non-rhizobia microbes on root-nodule symbiosis, particularly nodule function and plant health (Crosbie et al., 2022). In Chapter 5, we investigated the variation in the bacterial microbiome of nodules as a function of plant species and soil suspension inoculum. We correlated microbiome composition with plant health using 16S rRNA gene amplicon sequencing. We aimed to answer the following questions in our study: How does the microbiome of different lotus species (*L. japonicus*, *L. corniculatus*, and *L. burttii*) vary with soil suspension inoculum? Can we distinguish the nodule microbiome of starved *L. Burttii* plants from those of healthy plants? Which microbial groups are involved in the formation of healthy vs. starved nodules? We have comprehensively investigated these questions in our paper (Crosbie et al., 2022). The results of this work are discussed below.

6.4.1 Nodule microbiome composition in different *Lotus* species under different soil suspensions

Soil properties significantly influence nodule microbiome composition in non-legume and legume plants. Previous research has identified differences in microbial community diversity and soil suspension properties as key factors driving variation in plant microbiomes (Simonin et al., 2020, Brown et al., 2020, Liu et al., 2019). In our study, plants inoculated with different soil suspensions show significant differences in nodule microbiomes. Specifically, nodules inoculated with soil 1 showed colonization by different *Mesorhizobium* ASVs, whereas nodules inoculated with soil two were predominantly colonized by *Mesorhizobium* ASV M.1. In comparison to plants inoculated with soil two suspension, plants inoculated with soil one suspension developed more nodules and exhibited a wider range of shoot growth. Surprisingly, the original soil suspensions did not exhibit any significant differences in alpha diversity. They showed only minor variations in beta diversity, which were not statistically significant. Moreover, there were no noticeable differences in microbiome diversity or physicochemical properties between the source soils for these suspensions. This suggests that low abundance of soil microbes, which may not significantly impact diversity metrics, may be crucial for the overall functionality of the microbiome.

6.4.2 Distinct nodule microbiome in starved compared to healthy *L. burttii* plants

Among the species we studied, *L. burttii* was unique. It showed a significant difference in its nodule microbiome between healthy and starved plants. The nodules of starved *L. burttii* plants were mainly colonized by *Rhizobium* ASVs, whereas *Mesorhizobium* ASVs predominantly colonized the nodules of the healthy plants. Notably, when inoculated with *Rhizobium leguminosarum* Norway, *L. burttii* forms infected but ineffective nodules. This result was not observed in other *Lotus* species (Gossmann et al., 2012).

Despite starvation, both *L. japonicus* and *L. corniculatus* maintained nodules dominated by *Mesorhizobium* ASVs, similar to healthy plants. In contrast, *L. burttii* seemed more prone to form inefficient symbiotic relationships, possibly due to its higher susceptibility to colonization by non-host-specific symbionts (Zarrabian et al., 2022). This is evidenced by the significantly higher nodule formation rate in starved *L. burttii*. Such increased nodule frequency and reduced specificity in selecting a nodule partner could lead to starvation. Our results support the idea that the reduced specificity of *L. burttii* during root nodule symbiosis allows a wider range of bacteria to colonize its nodules.

6.4.3 Prediction of biomarker microbes in starved and healthy nodules

In our study, we observed that while *Rhizobiales* bacteria dominated the microbiota of all types of nodules, a subset of ASVs that did not belong to *Rhizobiales* was also detected. Such diversity is common in legume nodules: Various non-*Rhizobiales* bacteria such as *Proteobacteria*, along with *Actinobacteria* are often isolated from nodules (Benhizia et al., 2004, Dey et al., 2004). Notably, *Pseudomonas* was the predominant non-*rhizobial* genus found in the nodules of *Lotus*. Using machine learning we found that *Pseudomonas* ASVs were characteristic of healthy *L. burttii* nodules. This suggests a potential role in supporting plant health. Previous studies have shown that *Pseudomonas* can directly affect plant growth through siderophore production (Dey et al., 2004). Our results support the idea that *Pseudomonas* may play a crucial role in the outcome of the root nodule symbiosis, possibly through microbe-microbe interactions.

6.4.4 *Pseudomonas* and *Rhizobium* interactions in microbial networks and in experimental legume symbiosis

Examining the nodule microbiomes of both healthy and starved *L. burttii* plants using network analysis revealed notable associations. In particular, significant negative correlations were found between *Pseudomonas* and *Rhizobium*, as well as significant positive and negative interactions involving *Mesorhizobium*. Subsequent co-inoculation experiments were used to validate these expected interactions. The negative interaction between these microbes was highlighted when *Pseudomonas* isolates were co-inoculated with *Rhizobium sp.*, resulting in a significant reduction in both nodule numbers and nodule primordia compared to single *Rhizobium sp.* inoculations. Furthermore, the observed positive interactions between *Mesorhizobium* and *Pseudomonas*, along with the reduction of ineffective nodulation by deceiving *Rhizobia* in the presence of *Pseudomonas*, support the notion that these *Pseudomonas* ASVs make beneficial contributions to the root nodule symbiosis.

In summary, this work highlights the critical role of the soil microbiome and nodule colonization by *Pseudomonas* in healthy plants. Its role in reducing ineffective nodules

underscores the influence of the broader soil microbiota on root-nodule symbiosis. Additionally, the application of machine learning, followed by microbial network analysis, to identify key specific interactions within complex microbial communities that contribute to plant health.

6.5 Conclusions

In this dissertation, we collected and analyzed the leaf microbiome of *Arabidopsis thaliana* from different geographical locations and two seasons during a five-year study. The diversity, composition, and microbe-microbe interactions of the leaf microbiome in response to environmental factors were conducted in Chapter 2 and further discussed in Section 6.1. Temporal dynamics of the leaf microbiome were monitored over a five-month growing season from November to March. Furthermore, core persistent microbes and their connectivity patterns as hubs in microbial interaction networks were investigated in Chapter 3 and further discussed in Section 6.2. In Chapter 4, the potential core microbe *Albugo* was further investigated to identify the patterns of plants infected with this pathogen and uninfected plants. Machine learning models were used to identify discriminant microbes for plant health and infection from different microbial groups of bacteria and eukaryotes. These identified microbes were further validated experimentally to improve plant health in coping with the pathogen. The results of this part are discussed in Section 6.3. The roles of the microbiome in plant health were investigated in three *Lotus* species. As detailed in Chapter 5 and further discussed in section 6.4, nodule microbiomes of healthy and starved plants grown in different soils were analyzed using machine learning models. Microbial interaction networks were used to identify interactions between predicted microbes that discriminate between healthy and starved plants. In addition, interactions between selected nitrogen-fixing candidates such as *Rhizobia* and *Mesorhizobium* with *Pseudomonas* were experimentally confirmed.

In conclusion, based on long-term collected microbial data, this dissertation provides in-depth insights into identifying specific plant-microbe-microbe interactions using machine learning and microbial networks. These findings enhance our understanding of microbial responses to various biotic and abiotic perturbations and contribute to developing probiotics to help plants cope with these stresses.

Bibliography

M. T. Agler, J. Ruhe, S. Kroll, C. Morhenn, S.-T. Kim, D. Weigel, and E. M. Kemen. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology*, 14(1):e1002352, 2016.

J. Almario, M. Mahmoudi, S. Kroll, M. Agler, A. Placzek, A. Mari, and E. Kemen. The leaf microbiome of arabidopsis displays reproducible dynamics and patterns throughout the growing season. *Mbio*, 13(3):e02825–21, 2022.

E. L. Aydogan, G. Moser, C. Müller, P. Kämpfer, and S. P. Glaeser. Long-term warming shifts the composition of bacterial communities in the phyllosphere of galium album in a permanent grassland field-experiment. *Frontiers in Microbiology*, 9:144, 2018.

Y. Benhizia, H. Benhizia, A. Benguedouar, R. Muresu, A. Giacomini, and A. Squartini. Gamma proteobacteria can nodulate legumes of the genus hedysarum. *Systematic and applied microbiology*, 27(4):462–468, 2004.

G. Berg, D. Rybakova, M. Grube, and M. Köberl. The plant microbiome explored: implications for experimental botany. *Journal of experimental botany*, 67(4):995–1002, 2016.

S. Bhalla, K. Chaudhary, R. Kumar, M. Sehgal, H. Kaur, S. Sharma, and G. P. Raghava. Gene expression-based biomarkers for discriminating early and late stage of clear cell renal cancer. *Scientific reports*, 7(1):44997, 2017.

M. Bonkowski. Protozoa and plant growth: the microbial loop in soil revisited. *New Phytologist*, 162(3):617–631, 2004.

A. W. Bowsher, G. M. N. Benucci, G. Bonito, and A. Shade. Seasonal dynamics of core fungi in the switchgrass phyllosphere, and co-occurrence with leaf bacteria. *Phytobiomes Journal*, 5(1):60–68, 2021.

B. Brachi, D. Filiault, H. Whitehurst, P. Darme, P. Le Gars, M. Le Mentec, T. C. Morton, E. Kerdaffrec, F. Rabanal, A. Anastasio, et al. Plant genetic effects on microbial hubs impact host fitness in repeated field trials. *Proceedings of the National Academy of Sciences*, 119(30):e2201285119, 2022.

G. Brader, S. Compant, K. Vescio, B. Mitter, F. Trognitz, L.-J. Ma, and A. Sessitsch. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annual Review of Phytopathology*, 55:61–83, 2017.

R. M. Braga, M. N. Dourado, and W. L. Araújo. Microbial interactions: ecology in a molecular perspective. *brazilian journal of microbiology*, 47:86–98, 2016.

S. P. Brown, M. A. Grillo, J. C. Podowski, and K. D. Heath. Soil origin and plant genotype structure distinct microbiome compartments in the model legume medicago truncatula. *Microbiome*, 8(1):1–17, 2020.

D. Bulgarelli, K. Schlaeppi, S. Spaepen, E. V. L. Van Themaat, and P. Schulze-Lefert. Structure and functions of the bacterial microbiota of plants. *Annual review of plant biology*, 64:807–838, 2013.

C. I. Carlström, C. M. Field, M. Bortfeld-Miller, B. Müller, S. Sunagawa, and J. A. Vorholt. Synthetic microbiota reveal priority effects and keystone strains in the arabidopsis phyllosphere. *Nature Ecology & Evolution*, 3(10):1445–1454, 2019.

H.-X. Chang, J. S. Haudenshield, C. R. Bowen, and G. L. Hartman. Metagenome-wide association study and machine learning prediction of bulk soil microbiome and crop productivity. *Frontiers in Microbiology*, 8:519, 2017.

V. Chaudhry, P. Runge, P. Sengupta, G. Doehlemann, J. E. Parker, and E. Kemen. Shaping the leaf microbiota: plant–microbe–microbe interactions. *Journal of Experimental Botany*, 72(1):36–56, 2021.

W. Chen, K. Ren, A. Isabwe, H. Chen, M. Liu, and J. Yang. Stochastic processes shape microeukaryotic community assembly in a subtropical river across wet and dry seasons. *Microbiome*, 7:1–16, 2019.

F. Cheng, X. Huang, Q. Qin, Z. Chen, F. Li, and W. Song. The effect of aboveground long-term low-dose ionizing radiation on soil microbial diversity and structure. *Frontiers in Ecology and Evolution*, 11:1184582, 2023.

M. E. Coates and J. L. Beynon. Hyaloperonospora arabidopsidis as a pathogen model. *Annual review of phytopathology*, 48:329–345, 2010.

S. Compant, M. C. Cambon, C. Vacher, B. Mitter, A. Samad, and A. Sessitsch. The plant endosphere world–bacterial life within plants. *Environmental Microbiology*, 23 (4):1812–1829, 2021.

A. Cooper, A. Latunde-Dada, A. Woods-Tör, J. Lynn, J. Lucas, I. Crute, and E. Holub. Basic compatibility of albugo candida in arabidopsis thaliana and brassica juncea causes broad-spectrum suppression of innate immunity. *Molecular Plant-Microbe Interactions*, 21(6):745–756, 2008.

J. K. Copeland, L. Yuan, M. Layeghifard, P. W. Wang, and D. S. Guttman. Seasonal community succession of the phyllosphere microbiome. *Molecular Plant-Microbe Interactions*, 28(3):274–285, 2015.

D. B. Crosbie, M. Mahmoudi, V. Radl, A. Brachmann, M. Schloter, E. Kemen, and M. Marín. Microbiome profiling reveals that pseudomonas antagonises parasitic nodule colonisation of cheater rhizobia in lotus. *New Phytologist*, 234(1):242–255, 2022.

K. M. Dastogeer, F. H. Tumpa, A. Sultana, M. A. Akter, and A. Chakraborty. Plant microbiome–an account of the factors that shape community composition and diversity. *Current Plant Biology*, 23:100161, 2020.

F. T. De Vries, P. Manning, J. R. Tallowin, S. R. Mortimer, E. S. Pilgrim, K. A. Harrison, P. J. Hobbs, H. Quirk, B. Shipley, J. H. Cornelissen, et al. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology letters*, 15(11):1230–1239, 2012.

H. Desaint, N. Aoun, L. Deslandes, F. Vailleau, F. Roux, and R. Berthomé. Fight hard or die trying: when plants face pathogens under heat stress. *New Phytologist*, 229(2): 712–734, 2021.

R. Dey, K. Pal, D. Bhatt, and S. Chauhan. Growth promotion and yield enhancement of peanut (arachis hypogaea l.) by application of plant growth-promoting rhizobacteria. *Microbiological research*, 159(4):371–394, 2004.

K. Dumack, C. Baumann, and M. Bonkowski. A bowl with marbles: revision of the thecate amoeba genus lecythium (chlamydophryidae, tectofilosida, cercozoa, rhizaria) including a description of four new species and an identification key. *Protist*, 167(5): 440–459, 2016.

K. Dumack, S. Flues, K. Hermanns, and M. Bonkowski. Rhogostomidae (cercozoa) from soils, roots and plant leaves (arabidopsis thaliana): Description of rhogostoma epiphylla sp. nov. and r. cylindrica sp. nov. *European journal of protistology*, 60: 76–86, 2017.

P. Durán, T. Thiergart, R. Garrido-Oter, M. Agler, E. Kemen, P. Schulze-Lefert, and S. Hacquard. Microbial interkingdom interactions in roots promote arabidopsis survival. *Cell*, 175(4):973–983, 2018.

J. Edwards, C. Johnson, C. Santos-Medellín, E. Lurie, N. K. Podishetty, S. Bhatnagar, J. A. Eisen, and V. Sundaresan. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences*, 112(8):E911–E920, 2015.

K. Eitzen, P. Sengupta, S. Kroll, E. Kemen, and G. Doehlemann. A fungal member of the arabidopsis thaliana phyllosphere antagonizes albugo laibachii via a gh25 lysozyme. *Elife*, 10:e65306, 2021.

O. M. Finkel, I. Salas-González, G. Castrillo, J. M. Conway, T. F. Law, P. J. P. L. Teixeira, E. D. Wilson, C. R. Fitzpatrick, C. D. Jones, and J. L. Dangl. A single bacterial genus maintains root growth in a complex microbiome. *Nature*, 587(7832):103–108, 2020.

J. Friedman and E. J. Alm. Inferring correlation networks from genomic survey data. *PLoS computational biology*, 8(9):e1002687, 2012.

H. Fujita, M. Ushio, K. Suzuki, M. S. Abe, M. Yamamichi, K. Iwayama, A. Canarini, I. Hayashi, K. Fukushima, S. Fukuda, et al. Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome*, 11(1):63, 2023.

G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. Microbiome datasets are compositional: and this is not optional. *Frontiers in microbiology*, 8: 2224, 2017.

C.-H. Goh, D. F. Veliz Vallejos, A. B. Nicotra, and U. Mathesius. The impact of beneficial plant-associated microbes on plant phenotypic plasticity. *Journal of chemical ecology*, 39:826–839, 2013.

D. Gómez-Pérez, M. Schmid, V. Chaudhry, Y. Hu, A. Velic, B. Maček, J. Ruhe, A. Kemen, and E. Kemen. Proteins released into the plant apoplast by the obligate parasitic protist albugo selectively repress phyllosphere-associated bacteria. *New Phytologist*, 2022.

J. A. Gossmann, K. Markmann, A. Brachmann, L. E. Rose, and M. Parniske. Polymorphic infection and organogenesis patterns induced by a rhizobium leguminosarum isolate from lotus root nodules are determined by the host genotype. *New Phytologist*, 196(2):561–573, 2012.

K. L. Grady, J. W. Sorensen, N. Stopnisek, J. Guittar, and A. Shade. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nature communications*, 10(1):4135, 2019.

R. I. Griffiths, B. C. Thomson, P. James, T. Bell, M. Bailey, and A. S. Whiteley. The bacterial biogeography of british soils. *Environmental microbiology*, 13(6):1642–1654, 2011.

J. Grilli, T. Rogers, and S. Allesina. Modularity and stability in ecological communities. *Nature communications*, 7(1):12031, 2016.

T. Großkopf and O. S. Soyer. Synthetic microbial communities. *Current opinion in microbiology*, 18:72–77, 2014.

Y. Gu, S. Banerjee, F. Dini-Andreote, Y. Xu, Q. Shen, A. Jousset, and Z. Wei. Small changes in rhizosphere microbiome composition predict disease outcomes earlier than pathogen density variations. *The ISME Journal*, 16(10):2448–2456, 2022.

R. Gupta, G. Anand, R. Gaur, and D. Yadav. Plant–microbiome interactions for sustainable agriculture: a review. *Physiology and Molecular Biology of Plants*, 27:165–179, 2021.

K. Hamonts, P. Trivedi, A. Garg, C. Janitz, J. Grinyer, P. Holford, F. C. Botha, I. C. Anderson, and B. K. Singh. Field study reveals core plant microbiota and relative importance of their drivers. *Environmental Microbiology*, 20(1):124–140, 2018.

M. A. Hassani, P. Durán, and S. Hacquard. Microbial interactions within the plant holobiont. *Microbiome*, 6:1–17, 2018.

R. Hernández Medina, S. Kutuzova, K. N. Nielsen, J. Johansen, L. H. Hansen, M. Nielsen, and S. Rasmussen. Machine learning and deep learning applications in microbiome research. *ISME Communications*, 2(1):98, 2022.

G. Innerebner, C. Knief, and J. A. Vorholt. Protection of arabidopsis thaliana against leaf-pathogenic pseudomonas syringae by sphingomonas strains in a controlled model system. *Applied and environmental microbiology*, 77(10):3202–3210, 2011.

J. K. Jansson, R. McClure, and R. G. Egbert. Soil microbiome engineering for sustainability in a changing environment. *Nature Biotechnology*, pages 1–13, 2023.

H. Kadivar and A. E. Stapleton. Ultraviolet radiation alters maize phyllosphere bacterial diversity. *Microbial Ecology*, pages 353–361, 2003.

T. L. Karasov, J. Almario, C. Friedemann, W. Ding, M. Giolai, D. Heavens, S. Kersten, D. S. Lundberg, M. Neumann, J. Regalado, et al. Arabidopsis thaliana and pseudomonas pathogens exhibit stable associations over evolutionary timescales. *Cell host & microbe*, 24(1):168–179, 2018.

T. L. Karasov, M. Neumann, G. Shirsekar, G. Monroe, P. Team, D. Weigel, and R. Schwab. Drought selection on arabidopsis populations and their microbiomes. *bioRxiv*, pages 2022–04, 2022.

D. Kishore, G. Birzu, Z. Hu, C. DeLisi, K. S. Korolev, and D. Segrè. Inferring microbial co-occurrence networks from amplicon data: a systematic evaluation. *Msystems*, pages e00961–22, 2023.

S. Kroll, M. T. Agler, and E. Kemen. Genomic dissection of host–microbe and microbe–microbe interactions for advanced plant breeding. *Current Opinion in Plant Biology*, 36:71–78, 2017.

Z. D. Kurtz, C. L. Müller, E. R. Miraldi, D. R. Littman, M. J. Blaser, and R. A. Bonneau. Sparse and compositionally robust inference of microbial ecological networks. *PLoS computational biology*, 11(5):e1004226, 2015.

F. Liu, T. Hewezi, S. L. Lebeis, V. Pantalone, P. S. Grewal, and M. E. Staton. Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC microbiology*, 19(1):1–19, 2019.

H. Liu, L. C. Carvalhais, M. Crawford, E. Singh, P. G. Dennis, C. M. Pieterse, and P. M. Schenk. Inner plant values: diversity, colonization and benefits from endophytic bacteria. *Frontiers in microbiology*, 8:2552, 2017.

D. S. Lundberg, R. de Pedro Jové, P. Pramoj Na Ayutthaya, T. L. Karasov, O. Shalev, K. Poersch, W. Ding, A. Bollmann-Giolai, I. Bezrukov, and D. Weigel. Contrasting patterns of microbial dominance in the arabidopsis thaliana phyllosphere. *Proceedings of the National Academy of Sciences*, 119(52):e2211881119, 2022.

M. Mahmoudi, J. Almario, K. Lutap, K. Nieselt, and E. Kemen. Microbial communities living inside plant leaves or on the leaf surface are differently shaped by environmental cues. *ISME communications*, 4(1):ycae103, 2024.

L. Maignien, E. A. DeForce, M. E. Chafee, A. M. Eren, and S. L. Simmons. Ecological succession and stochastic variation in the assembly of arabidopsis thaliana phyllosphere communities. *MBio*, 5(1):10–1128, 2014.

K. M. Meyer, R. Porch, I. E. Muscettola, A. L. S. Vasconcelos, J. K. Sherman, C. J. E. Metcalf, S. E. Lindow, and B. Koskella. Plant neighborhood shapes diversity and reduces interspecific variation of the phyllosphere microbiome. *The ISME journal*, 16 (5):1376–1387, 2022.

B. E. Morris, R. Henneberger, H. Huber, and C. Moissl-Eichinger. Microbial syntrophy: interaction for the common good. *FEMS microbiology reviews*, 37(3):384–406, 2013.

J. Namkung. Machine learning methods for microbiome studies. *Journal of Microbiology*, 58:206–216, 2020.

J. Namkung, W. Kwon, Y. Choi, S. G. Yi, S. Han, M. J. Kang, S.-W. Kim, T. Park, and J.-Y. Jang. Molecular subtypes of pancreatic cancer based on mirna expression profiles have independent prognostic value. *Journal of gastroenterology and hepatology*, 31 (6):1160–1167, 2016.

D. Naylor, S. DeGraaf, E. Purdom, and D. Coleman-Derr. Drought and host selection influence bacterial community dynamics in the grass root microbiome. *The ISME journal*, 11(12):2691–2704, 2017.

A. T. Neu, E. E. Allen, and K. Roy. Defining and quantifying the core microbiome: challenges and prospects. *Proceedings of the National Academy of Sciences*, 118(51): e2104429118, 2021.

B.-A. T. Nguyen, K. Dumack, P. Trivedi, Z. Islam, and H.-W. Hu. Plant associated protists—untapped promising candidates for agrifood tools. *Environmental Microbiology*, 25(2):229–240, 2023.

T. Ou, W.-f. Xu, F. Wang, G. Strobel, Z.-y. Zhou, Z.-h. Xiang, J. Liu, and J. Xie. A microbiome study reveals seasonal variation in endophytic bacteria among different mulberry cultivars. *Computational and structural biotechnology journal*, 17:1091–1100, 2019.

J. M. Raaijmakers and M. Mazzola. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annual review of phytopathology*, 50:403–424, 2012.

S. Rai, A. F. Omar, M. Rehan, A. Al-Turki, A. Sagar, N. Ilyas, R. Sayyed, and M. Hasanuzzaman. Crop microbiome: Their role and advances in molecular and omic techniques for the sustenance of agriculture. *Planta*, 257(2):27, 2023.

G. Rastogi, A. Sbodio, J. J. Tech, T. V. Suslow, G. L. Coaker, and J. H. Leveau. Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *The ISME journal*, 6(10):1812–1822, 2012.

U. Ritpitakphong, L. Falquet, A. Vimoltust, A. Berger, J.-P. Métraux, and F. L'Haridon. The microbiome of the leaf surface of arabidopsis protects against a fungal pathogen. *New Phytologist*, 210(3):1033–1043, 2016.

F. Romero, S. Cazzato, F. Walder, S. Vogelgsang, S. F. Bender, and M. G. van der Heijden. Humidity and high temperature are important for predicting fungal disease outbreaks worldwide. *New Phytologist*, 234(5):1553–1556, 2022.

L. Röttjers and K. Faust. From hairballs to hypotheses–biological insights from microbial networks. *FEMS microbiology reviews*, 42(6):761–780, 2018.

P. Runge, F. Ventura, E. Kemen, and R. Stam. Distinct phyllosphere microbiome of wild tomato species in central peru upon dysbiosis. *Microbial ecology*, 85(1):168–183, 2023.

J. Sardans, J. PeNUuelas, M. Estiarte, and P. Prieto. Warming and drought alter c and n concentration, allocation and accumulation in a mediterranean shrubland. *Global Change Biology*, 14(10):2304–2316, 2008.

O. Shalev, T. L. Karasov, D. S. Lundberg, H. Ashkenazy, P. Pramoj Na Ayutthaya, and D. Weigel. Commensal pseudomonas strains facilitate protective response against pathogens in the host plant. *Nature ecology & evolution*, 6(4):383–396, 2022.

L. G. Shoemaker, L. L. Sullivan, I. Donohue, J. S. Cabral, R. J. Williams, M. M. Mayfield, J. M. Chase, C. Chu, W. S. Harpole, A. Huth, et al. Integrating the underlying structure of stochasticity into community ecology. *Ecology*, 101(2):e02922, 2020.

M. Simonin, C. Dasilva, V. Terzi, E. L. Ngonkeu, D. Diouf, A. Kane, G. Béna, and L. Moulin. Influence of plant genotype and soil on the wheat rhizosphere microbiome: evidences for a core microbiome across eight african and european soils. *FEMS microbiology ecology*, 96(6):fiaa067, 2020.

B. K. Singh, M. Delgado-Baquerizo, E. Egidi, E. Guirado, J. E. Leach, H. Liu, and P. Trivedi. Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology*, pages 1–17, 2023.

W. T. Sloan, M. Lunn, S. Woodcock, I. M. Head, S. Nee, and T. P. Curtis. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental microbiology*, 8(4):732–740, 2006.

N. Stopnisek and A. Shade. Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. *The ISME Journal*, 15(9):2708–2722, 2021.

M. J. Sweet and M. T. Bulling. On the importance of the microbiome and pathobiome in coral health and disease. *Frontiers in Marine Science*, 4:9, 2017.

M. Thieme, A. Brêchet, Y. Bourgeois, B. Keller, E. Bucher, and A. C. Roulin. Experimentally heat-induced transposition increases drought tolerance in arabidopsis thaliana. *New Phytologist*, 236(1):182–194, 2022.

B. P. Thomma, H. P. Van Esse, P. W. Crous, and P. J. De Wit. Cladosporium fulvum (syn. passalora fulva), a highly specialized plant pathogen as a model for functional studies on plant pathogenic mycosphaerellaceae. *Molecular plant pathology*, 6(4):379–393, 2005.

H. Toju, S. Yamamoto, A. S. Tanabe, T. Hayakawa, and H. S. Ishii. Network modules and hubs in plant-root fungal biomes. *Journal of The Royal Society Interface*, 13(116): 20151097, 2016.

B. D. Topçuoğlu, N. A. Lesniak, M. T. Ruffin IV, J. Wiens, and P. D. Schloss. A framework for effective application of machine learning to microbiome-based classification problems. *MBio*, 11(3):10–1128, 2020.

P. Trivedi, B. D. Batista, K. E. Bazany, and B. K. Singh. Plant–microbiome interactions under a changing world: Responses, consequences and perspectives. *New Phytologist*, 234(6):1951–1959, 2022.

M. G. Van Der Heijden and M. Hartmann. Networking in the plant microbiome. *PLoS biology*, 14(2):e1002378, 2016.

M. G. Van Der Heijden, J. N. Klironomos, M. Ursic, P. Moutoglis, R. Streitwolf-Engel, T. Boller, A. Wiemken, and I. R. Sanders. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*, 396(6706):69–72, 1998.

R. E. Venado, J. Liang, and M. Marín. Rhizobia infection, a journey to the inside of plant cells. In *Advances in Botanical Research*, volume 94, pages 97–118. Elsevier, 2020.

J. A. Vorholt. Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10(12): 828–840, 2012.

W. A. Walters, Z. Jin, N. Youngblut, J. G. Wallace, J. Sutter, W. Zhang, A. González-Peña, J. Peiffer, O. Koren, Q. Shi, et al. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proceedings of the National Academy of Sciences*, 115(28):7368–7373, 2018.

K. Wang, T. P. Sipilä, and K. Overmyer. The isolation and characterization of resident yeasts from the phylloplane of arabidopsis thaliana. *Scientific reports*, 6(1):39403, 2016.

L. Wang, T. S. George, and G. Feng. Concepts and consequences of the hyphosphere core microbiome for arbuscular mycorrhizal fungal fitness and function. *New Phytologist*, 2023a.

Z. Wang, X. Hu, M. K. Solanki, and F. Pang. A synthetic microbial community of plant core microbiome can be a potential biocontrol tool. *Journal of Agricultural and Food Chemistry*, 71(13):5030–5041, 2023b.

I. M. Ware, M. E. Van Nuland, Z. K. Yang, C. W. Schadt, J. A. Schweitzer, and J. K. Bailey. Climate-driven divergence in plant-microbiome interactions generates range-wide variation in bud break phenology. *Communications Biology*, 4(1):748, 2021.

S. Weiss, W. Van Treuren, C. Lozupone, K. Faust, J. Friedman, Y. Deng, L. C. Xia, Z. Z. Xu, L. Ursell, E. J. Alm, et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *The ISME journal*, 10(7):1669–1681, 2016.

A. S. Wieczorek, O. Schmidt, A. Chatzinotas, M. Von Bergen, A. Gorissen, and S. Kolb. Ecological functions of agricultural soil bacteria and microeukaryotes in chitin degradation: a case study. *Frontiers in microbiology*, 10:1293, 2019.

E. H. Wintermute and P. A. Silver. Emergent cooperation in microbial metabolism. *Molecular systems biology*, 6(1):407, 2010.

K. Wippel, K. Tao, Y. Niu, R. Zgadzaj, N. Kiel, R. Guan, E. Dahms, P. Zhang, D. B. Jensen, E. Logemann, et al. Host preference and invasiveness of commensal bacteria in the lotus and arabidopsis root microbiota. *Nature microbiology*, 6(9):1150–1162, 2021.

L. Xu, Z. Dong, D. Chiniquy, G. Pierroz, S. Deng, C. Gao, S. Diamond, T. Simmons, H. M.-L. Wipf, D. Caddell, et al. Genome-resolved metagenomics reveals role of iron metabolism in drought-induced rhizosphere microbiome dynamics. *Nature communications*, 12(1):3209, 2021.

X. Yang, K. Zhu, M. E. Loik, and W. Sun. Differential responses of soil bacteria and fungi to altered precipitation in a meadow steppe. *Geoderma*, 384:114812, 2021.

H. Yuan, R. Mei, J. Liao, and W.-T. Liu. Nexus of stochastic and deterministic processes on microbial community assembly in biological systems. *Frontiers in Microbiology*, 10:1536, 2019.

J. Yuan, T. Wen, H. Zhang, M. Zhao, C. R. Penton, L. S. Thomashow, and Q. Shen. Predicting disease occurrence with high accuracy based on soil macroecological patterns of fusarium wilt. *The ISME Journal*, 14(12):2936–2950, 2020.

M. Zarrabian, J. Montiel, N. Sandal, S. Ferguson, H. Jin, Y.-Y. Lin, V. Klingl, M. Marín, E. K. James, M. Parniske, et al. A promiscuity locus confers lotus burttii nodulation with rhizobia from five different genera. *Molecular Plant-Microbe Interactions*, 35 (11):1006–1017, 2022.

Z. Zhang, Q. Zhang, H. Cui, Y. Li, N. Xu, T. Lu, J. Chen, J. Penuelas, B. Hu, and H. Qian. Composition identification and functional verification of bacterial community in disease-suppressive soils by machine learning. *Environmental Microbiology*, 24(8):3405–3419, 2022.

J. Zhou and D. Ning. Stochastic community assembly: does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews*, 81(4):10–1128, 2017.

Y.-G. Zhu, C. Xiong, Z. Wei, Q.-L. Chen, B. Ma, S.-Y.-D. Zhou, J. Tan, L.-M. Zhang, H.-L. Cui, and G.-L. Duan. Impacts of global change on the phyllosphere microbiome. *New Phytologist*, 234(6):1977–1986, 2022.

I. Zilber-Rosenberg and E. Rosenberg. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS microbiology reviews*, 32(5):723–735, 2008.