

**The chemical landscape of leaves as a driver for microbial community
structure and metabolic interactions**

Dissertation

To fulfill the requirement for the Degree of
“doctor rerum naturalium” (Dr. rer. nat.)

**Submitted to the Council of the Faculty
of Biological Sciences
of the Friedrich Schiller University Jena**

**by MSc. Mariana Murillo Roos
born on 03.08.1989 in San José, Costa Rica**

Reviewers:

- 1. Prof. Erika Kothe, Institut für Mikrobiologie, Friedrich-Schiller-Universität**
- 2. Prof. Kirsten Küsel, Institut für Biodiversität, Friedrich-Schiller-Universität**
- 3. Prof. Klaus Schläppi, Departement Umweltwissenschaften, Universität Basel**

Day of public defense: 06.07.2023

Table of Contents

Summary	8
Zusammenfassung.....	9
Chapter 1: Introduction.....	11
Recruitment and maintenance of the plant microbiome is a dynamic and continuous process	11
Plant immune systems help maintain balance in the microbiome	13
Diverse adaptations help microorganisms thrive in the leaf apoplast	15
Nutrient availability in the leaf is constantly fluctuating.....	18
The plant microbiome may protect plants via competition and priming of the innate immune system	20
Stabilizing microbe-microbe metabolic interactions may arise as a consequence of host nutrient constraints.....	22
Linking microbial community structure and function to plant leaf metabolites to discover key processes shaping host health	23
Chapter 2: Methods.....	25
Propagation of plant material	25
Plant sampling	25
<i>Optimization of infiltration-centrifugation method to obtain leaf apoplast fluid wash (AFW)</i>	25
<i>Growth of multiple Flaveria species for AFW collection</i>	25
<i>Comparison of whole leaf and AFW-based methods for characterization of bacterial diversity in Flaveria</i>	26
<i>Enrichment of leaf microbiomes from Flaveria</i>	26
Leaf apoplast fluid wash (AFW) extraction	27
<i>Recovery of AFW from leaves using the syringe technique</i>	27
<i>Selection of infiltration solution</i>	27
<i>Selection of centrifugation conditions</i>	28
<i>Assessing the dilution effect on amino acid recovery from AFW</i>	28
<i>Recovery of apoplast fluid wash from leaves using a vacuum pump</i>	29
<i>AFW recovery for bacterial community analysis</i>	29
<i>In vitro</i> experiments.....	30
<i>Leaf extract preparation and enrichment of leaf microbiomes</i>	30
<i>Isolation of bacteria and first assessment of carbon preference</i>	30
<i>Preparation of isolates' precultures and testing their growth in different media</i>	31
<i>Production of spent media and evaluation of cross-feeding interactions</i>	31
<i>Tagging of Pseudomonas strains with fluorescent proteins</i>	32
<i>Co-culture experiments using tagged Pseudomonas strains</i>	33
<i>In-planta</i> experiments	34
<i>In-planta testing of Pseudomonas colonization</i>	34

DNA extraction and amplification.....	34
<i>DNA extraction and amplification from whole leaves and AFW</i>	34
<i>DNA extraction and amplification from bacterial isolates</i>	36
<i>DNA extraction and whole genome sequencing of bacterial isolates</i>	36
Metabolomics	36
<i>Ultra-high performance liquid chromatography - high resolution mass spectrometry</i>	36
<i>Metabolomics data analysis</i>	37
Bioinformatics	39
<i>Processing of amplicon sequencing data</i>	39
<i>Processing of Sanger sequencing data</i>	40
<i>Whole genome assembly, annotation, ANI calculation and single nucleotide polymorphism detection</i>	40
Data availability and scripts to generate figures from raw data	40
Chapter 3.1 Parallel characterization of apoplast chemical composition and microbial diversity.....	41
Sodium phosphate buffer allows UHPLC-HRMS analysis of apoplast fluid wash (AFW) from several plant species	41
Gentle centrifugation allows recovery of clean apoplast fluid wash from infiltrated leaves	43
Dilution of samples does not hinder detection of compounds of interest	45
Potential resources for microorganisms in the apoplast differ between <i>Flaveria</i> species	47
Higher bacterial diversity can be recovered from AFW than from whole leaves.	48
Chapter 3.2: The leaf microbiome and chemical landscape simultaneously cycle diurnally	50
Leaf metabolic profiles fluctuate throughout the day in a species-dependent manner	50
Bacterial communities in the leaf apoplast may vary throughout the day	56
Chapter 3.3: Leaf bacterial strain-level diversity is supported by niche separation during cross-feeding ..	58
Nutrient and taxonomic diversity shape microbiome function in <i>Flaveria</i> leaf enrichments	58
Cross-feeding on diverse metabolites sustains <i>Pseudomonas</i> in the absence of a primary carbon source.	59
Diverse plant-derived <i>Pseudomonas</i> strains cross-feed in parallel from <i>Pantoea</i>	62
Niche differentiation among distinct <i>Pseudomonas</i> strains maintains diversity during cross-feeding	64
Chapter 3.4 Cross-feeding is selected for in a leaf nutrient-dependent manner	68
Cross-feeding potential emerged as a convergent trait in bacterial isolates from <i>F. robusta</i>	68
The leaf nutrient landscape might actively select on cross-feeding interactions.....	70
Factors driving selection for cross-feeding depend on the partner strains.....	71
Chapter 4: Discussion	74
4.1 Apoplast fluid washes offer the potential to simultaneously study microbial diversity and the apoplast environment.....	74
4.2 Metabolite and microbial dynamics follow species-specific and distinct diurnal patterns	78
4.3 Leaf bacteria can employ metabolic cross-feeding to survive in nutrient-limited environments... 82	

4.4 Strain-level leaf bacterial diversity is linked to cross-feeding niches.....	83
4.5 Complex and variable leaf nutrient environments can affect selection on metabolic interactions.....	84
Conclusions and outlook	86
Supplementary Information	88
Methods Supp Figures.....	88
Chapter 3.1 Supp. Figures.....	89
Chapter 3.2 Supp. Figures.....	95
Chapter 3.3 Supp. Figures.....	101
Chapter 3.4 Supp. Figures.....	111
Methods Supp.Tables	112
Chapter 3.1 Supp. Tables	118
Chapter 3.2 Supp. Tables	121
Chapter 3.3 Supp. Tables	124
Chapter 3.4 Supp. Tables	125
References	126
Acknowledgments	146
Statement of authorship	148
Curriculum vitae	149

Summary

The intercellular space of leaves (e.g., the apoplast) is critical for plants, functioning in nutrient, water and photosynthate transport as well as cell homeostasis. It is also where important microbiota colonize, presumably by utilizing host resources. This environment is truly dynamic; the fluctuation of nutrients and a large array of defense molecules suggests it is a challenging niche for microbial colonization, where microbe-microbe interactions may play an important role. Thus, linking leaf resources to assembly of the leaf microbial community would have major impacts on understanding plant-microbe interactions. In the frame of this thesis, an infiltration-centrifugation method was optimized to recover apoplastic fluid wash (AFW) from leaves of diverse plant species. Closely related *Flaveria* species could be differentiated based on the AFW metabolomic profile, specifically on the quantities of valuable amino acids that are costly for microorganisms to synthesize. The AFW could also be used to assess bacterial diversity in the leaf apoplast. Amplicon sequencing analyses revealed that more alpha diversity could be recovered from AFW samples than from crushed leaves (the conventional method to assess leaf microbial diversity) and that beta diversity differences were clearer when using AFW. The optimized method was then used to track diurnal oscillations of metabolites and microbial communities in leaves of three *Flaveria* species with different carbon fixing strategies. Although additional work is needed to overcome low bacterial reads at some time points, the results suggested that enrichment of taxa over the day corresponded with increasing metabolite complexity of the AFW in a species-dependent manner. Considering the differences found in levels of amino acids between *F. robusta* and *F. trinervia*, an *in vitro* assay was established. Leaf bacteria from these two species were enriched with either sucrose as a sole carbon source or with addition of amino acids. Enrichments from *F. trinervia* were dominated by the bacteria *Pantoea* sp., while in *F. robusta*, *Pantoea* sp. was co-enriched with *Pseudomonas* sp., regardless of the nutrient regime. *Pseudomonas* sp. was unable to grow on sucrose alone but persisted in the sucrose-only enrichment thanks to cross-feeding diverse metabolites from *Pantoea* sp. Niche partitioning enabled genetically similar *Pseudomonas* strains (>99% average nucleotide identity) to cross-feed in parallel, even though they had different growth rates. *Pantoea* strains were also closely related, but those enriched in *F. robusta* were more efficient cross-feeders. An experimental evolution approach suggested that this could be explained in part by the plant environment, since selection for or against cross-feeding interactions depended on both the environment and the cross-feeding partners. Together, this work provides tools to better link leaf physiology with leaf microbial communities and advances the current knowledge on factors driving the assembly of the leaf microbiome. Specifically, it shows that metabolic interactions and niche partitioning may be relevant in shaping bacterial taxonomic and genetic diversity in plant-associated microbiomes. Moreover, it suggests that environmental conditions relevant in plants can exert selection pressures on cross-feeding interactions in bacterial strain-dependent ways. Whether these interactions are relevant for plant-microbe dynamics and host health remains an open and exciting question.

Zusammenfassung

Der interzelluläre Raum von Blättern (z. B. der Apoplast) ist für Pflanzen von entscheidender Bedeutung, da er für den Transport von Nährstoffen, Wasser und Photosyntheseprodukten sowie für die Zellhomöostase zuständig ist. Hier siedeln sich auch wichtige Mikroorganismen an, da sie vermutlich die Ressourcen des Wirts nutzen. Diese Umgebung ist sehr dynamisch und die Fluktuation von Nährstoffen und einer Vielzahl von Abwehrmolekülen lassen vermuten, dass es sich um eine anspruchsvolle Nische für die mikrobielle Besiedlung handelt und dass Mikroben-Mikroben-Interaktionen eine wichtige Rolle spielen könnten. Die Korrelation von Blattressourcen mit dem Aufbau der mikrobiellen Gemeinschaft des Blattes hätte daher große Auswirkungen auf das Verständnis der Interaktionen zwischen Pflanzen und Mikroben. In der vorliegenden Doktorarbeit wurde daher eine Infiltrations-Zentrifugations-Methode zur Gewinnung von apoplastischer Flüssigkeit (*apoplastic fluid wash*; AFW) aus Blättern verschiedener Pflanzenarten optimiert. Nahverwandte *Flaveria*-Arten konnten anhand des metabolomischen Profils der AFW unterschieden werden, insbesondere anhand der Mengen an wertvollen Aminosäuren, deren Synthese für Mikroorganismen aufwändig ist. Die AFW könnte auch zur Bewertung der bakteriellen Vielfalt im Blattapoplast verwendet werden. Amplikon-Sequenzanalysen ergaben, dass aus AFW-Proben mehr Alpha-Diversität gewonnen werden konnte als aus zerkleinerten Blättern (die herkömmliche Methode zur Bewertung der mikrobiellen Diversität von Blättern) und dass die Unterschiede in der Beta-Diversität bei Verwendung von AFW deutlicher waren. Die optimierte Methode wurde anschließend verwendet, um die täglichen Oszillationen von Metaboliten und mikrobiellen Gemeinschaften in Blättern von drei *Flaveria*-Arten mit unterschiedlichen Strategien zur Kohlenstoffbindung zu verfolgen. Die Ergebnisse deuten darauf hin, dass die Anreicherung von Taxa im Laufe des Tages mit der zunehmenden Metaboliten-Komplexität der AFW in einer spezieabhängigen Weise korrespondiert, wenngleich aufgrund der teilweise geringen bakteriellen Werte weitere Analysen erforderlich sind. In Anbetracht der Unterschiede zwischen *F. robusta* und *F. trinervia* in Bezug auf den Gehalt an Aminosäuren wurde ein *In vitro*-Assay durchgeführt. Die Blattbakterien dieser beiden Arten wurden entweder mit Saccharose als einziger Kohlenstoffquelle oder mit Aminosäuren angereichert. Während bei *F. trinervia* das Bakterium *Pantoea sp.* dominierte, wurde bei *F. robusta* – unabhängig vom Nährstoffregime – *Pantoea sp.* gemeinsam mit *Pseudomonas sp.* angereichert. *Pseudomonas sp.* war dagegen nicht in der Lage, allein auf Saccharose zu wachsen, überlebte aber in der reinen Saccharose-Anreicherung dank der Kreuzfütterung mit verschiedenen Metaboliten von *Pantoea sp.* Die Nischenaufteilung ermöglichte es genetisch ähnlichen *Pseudomonas*-Stämmen (>99 % durchschnittliche Nukleotid-Identität), sich parallel zu ernähren, obwohl sie unterschiedliche Wachstumsraten hatten. Auch die *Pantoea*-Stämme waren eng miteinander verwandt, aber die mit *F. robusta* angereicherten Stämme waren effizientere Kreuzfütterer. Ein experimenteller Evolutionsansatz deutet darauf hin, dass dies zum Teil durch die pflanzliche Umwelt erklärt werden könnte, da die Selektion für oder gegen Kreuzfütterungsinteraktionen sowohl von der Umwelt als auch von den Kreuzfütterungspartnern abhängt. Insgesamt bietet die vorliegende Arbeit Ansätze, um die Blattphysiologie besser mit den mikrobiellen Gemeinschaften der Blätter zu verknüpfen, und erweitert das derzeitige Wissen

über die Faktoren, die den Aufbau des Blattmikrobioms bestimmen. Insbesondere legt sie nahe, dass metabolische Interaktionen und Nischenaufteilung für die Gestaltung der bakteriellen taxonomischen und genetischen Vielfalt in pflanzenassoziierten Mikrobiomen von Bedeutung sein können. Darüber hinaus deuten die Ergebnisse darauf hin, dass pflanzenrelevante Umweltbedingungen einen Selektionsdruck auf stammesabhängige Wechselwirkungen zwischen Bakterien ausüben können. Ob diese Interaktionen für die Dynamik von Pflanzen-Mikroben und die Gesundheit des Wirts relevant sind, bleibt eine offene und spannende Frage.

Chapter 1: Introduction

One of the main challenges in agriculture is how to keep the balance between profitable crops and environmental sustainability. For example, the consumption of pesticides in the past decades has steadily increased (FAO, 2021), partly as a response to the appearance of new pests (Wang et al., 2022). While these measures are effective, long-term use of pesticides can lead to pathogen resistance (Corkley et al., 2022) and a loss of biodiversity (Brühl and Zaller, 2019). In addition, the cost of agrochemicals is highly susceptible to raw material and energy availability and even to political affairs (FAO, 2022). Therefore, it is extremely risky for an agroecosystem to fully depend on the acquisition and application of pesticides. One way to reduce reliance on chemicals is to integrate strategies that contemplate the relationship of plants with their environment in production systems. This approach will require development of resilient crops, and this can be achieved in part by providing plants with the conditions to build and maintain a healthy and stable microbiome (Nerva et al., 2022). Thus, it is crucial to understand what keeps microbial communities balanced in response to fluctuating environmental conditions.

Plants are naturally colonized by diverse communities of microorganisms, including bacteria, fungi, protists, viruses and even algae (Zhu et al., 2018; Chaudhry et al., 2021). However, only a small subset of the microbes found in inoculation sources like soil and air can efficiently colonize plants, especially the phyllosphere, which includes aboveground tissues like the leaf (Fitzpatrick et al., 2020). This strong filtering is largely explained by host plant barriers like the immune system (Vannier et al., 2019) and a tight regulation of nutrients. While a few microorganisms (pathogens) have developed large toolboxes of special adaptations to tackle these barriers (Levy et al., 2018), most plant-associated microbial diversity are non-pathogenic or commensals and lack these strategies. Despite their lack of specialized adaptations, commensals have the potential to serve important functions for the plant (Morris and Moury, 2019). In the last decades, a lot of effort has been placed in studying how commensals colonize and survive on roots, but less focus has been placed on leaf microbial communities (Koskella, 2020). Since leaves represent about 60% of the total biomass on Earth (Bar-On et al., 2018), it is necessary to study the dynamics of their microbiomes and their impact on the plant's health.

Recruitment and maintenance of the plant microbiome is a dynamic and continuous process

Plant assembly of microbial communities can generally be divided into vertical and horizontal processes. A small fraction is transferred vertically from the mother plant to the seeds. Despite its relatively small size, this fraction may play important roles for plant protection and nutrition due to, for example, enrichment for protective chitinolytic and phosphorus-solubilizing activities (Bziuk et al., 2021). If so, this would suggest that well-known protective vertical transfer of microbes like in vaginal birth of mammals (Mortensen et al., 2021), is a widespread strategy in host-associated microbiomes. Regardless, vertical transmission in plants has been proven for some plant species by inoculating microbes in the flowers of parent plants and following

their journey to mature seeds (Dutta et al., 2015; Mitter et al., 2017). In this case, bacteria that colonize the stigma of the flower seem more likely to enter the seeds (Dutta et al., 2015). However, microbes can also be transferred from pollen grains or through the carbohydrate transport route (Berg and Raaijmakers, 2018). The latter might be especially important for clonal plants that are propagated via vegetative tissue (Vannier et al., 2018).

A large portion of the plant microbiome is thought to be horizontally acquired. In the early stages of development, soil and air are the main inocula (Maignien et al., 2014; Rochefort et al., 2021). As plants continue to grow and develop above-ground structures (stem, leaves, flowers, and fruits), potential colonizers continue to arrive by other mechanisms like aerial dispersion (Schirotto et al., 2018), soil splashing (Penet et al., 2014), transfer of microbes in arthropod oral secretions or feces (Rashed et al., 2012; Dutta et al., 2014), irrigation (Truchado et al., 2019) or water dripping through the leaves (Bittar et al., 2018). The environment in which the plant develops thus exerts a significant impact on the composition of the microbiome, but its relevance varies across tissues. In rhizosphere communities, for example, the type of soil and crop management practices are among the strongest determining factors (Caradonia et al., 2019; Zhang et al., 2022), while in leaf or phyllosphere communities, radiation, desiccation and foliar applications have a major impact (Kamo et al., 2018; McGarvey et al., 2019; Carvalho et al., 2020). The influence of environmental factors may also vary across spatial scales and for different organisms. Recent data suggests that across different European sampling sites, climate had a larger influence on root-associated communities of filamentous eukaryotes, while soil properties were more important for bacterial communities (Thiergart et al., 2020).

Studies conducted in common environments, where all plants received homogenous inoculum and environmental conditions, have shown differences in the microbiome assembly depending on the host identity (Wagner et al., 2016; Li et al., 2018; Brown et al., 2020). Interestingly, the host effect is apparently stronger in the leaves compared to the roots (Coleman-Derr et al., 2016; Wagner et al., 2016; Li et al., 2018) and is especially important for bacteria, compared to fungi (Coleman-Derr et al., 2016). Leaf characteristics, like stomatal conductance and leaf chemistry (Li et al., 2018) or plant functional traits like growth and mortality rates (Kembel et al., 2014) have important effects in colonization. In addition, the use of *Arabidopsis* mutants showed that changes in the permeability of the leaf cuticle may lead to a different assembly of the community (Bodenhausen et al., 2014). In some species, the effect of the host can be so determinant, that leaf bacterial diversity remains stable across geographical gradients (Redford et al., 2010). For other species however, it is the combination of host genotype and environment that determines the diversity of the microbiome (Wagner et al., 2016).

Plants can also actively recruit microbes, and this can occur from a very early stage. Seed exudates of common bean for example, were shown to attract a beneficial *Bacillus* strain that conferred protection against drought (Martins et al., 2018). Root exudates are also known to play a key role in this recruitment, whereby their chemical composition changes throughout development stages, directly influencing the

diversity of the rhizosphere community (Chaparro et al., 2014). In early stages, *Arabidopsis* exudes more sugars and sugar alcohols, while it later produces more amino acids and phenolic compounds (Chaparro et al., 2013). This recruitment may also vary under specific conditions. Salt-stressed plants, for example, may exude different organic acids which attract microbes that help regulating their ion balance (Liu et al., 2021). Under phosphorus depletion, root exudates of maize were shown to alter the expression of motility genes in a rhizobacteria that produces phytases but is unable to fix nitrogen. Under nitrogen deficiency, the same function was repressed, suggesting that the plant recruited the microbe when phosphorus levels were low, but kept it away under nitrogen depletion to reduce competition for nitrogen sources (Carvalhais et al., 2013). Similarly, root-derived flavones by maize under nitrogen deprivation were shown to recruit members of Oxalobacteraceae, which help with nitrogen acquisition (Yu et al., 2021).

Deterministic factors like environmental conditions and host traits and are good examples of the niche-based theory used to describe community structure (Zhou and Ning, 2017). However, they do not fully explain the diversity and variation observed in plant microbial communities (Mayer et al., 2022). Neutral theory is increasingly used as a complement to explain the ecology of microbial communities (Maignien et al., 2014). This concept proposes that community dynamics are modulated by stochastic events such as drift (birth, death and reproduction rates of each microbe), dispersal or migration rates and speciation generated through genetic variation (Chave, 2004; Hanson et al., 2012). Maignien et al. (2014) for example, observed that although there was a strong effect of host-filtering selecting for the establishment of certain taxa in *Arabidopsis* leaves, there was huge variability in the relative abundance of the taxa, suggestive of stochastic processes in the early stages of colonization. Some of these stochastic, early arrivers can have a substantial impact in structuring the leaf microbial community Carlström et al (2019). These microbes are possibly better suited to overcome the significant barriers to colonization in the plant environment. Thus, their early colonization may facilitate arrival and establishment of following microbes.

Plant immune systems help maintain balance in the microbiome

Plants have complex immune systems that help regulate their interactions with microorganisms. The immunity components that directly interact with microorganisms can be roughly divided into three groups: 1) pre-formed barriers to colonization like the epidermal cell layer, cuticular wax and some phytoalexins, 2) extracellular pattern recognition receptors (PRRs) and 3) intracellular effector receptors like NB-LRRs. PRRs located in the cell membrane of plants recognize microorganisms and their associated molecular patterns (PAMPs or MAMPS) such as flagellin, peptidoglycan, chitin, and others. This results in a cascade of reactions known as PAMP-triggered immunity (PTI) (Fig. 1, upper left). PTI can be partly overcome when pathogens inject effectors into cells, which renders the plant susceptible to infection. NB-LRR proteins, located in the cytoplasm of the plant cell, in turn can recognize these effectors and in concert with PTI activate an effector-triggered immunity (ETI) (Fig. 1, lower left). If the ETI response is strong enough, it can lead to a hypersensitive response (HR), characterized by a burst of reactive oxygen species and cell death to prevent pathogens from spreading further (Jones and Dangl, 2006). These defense networks depend on

the action of phytohormones. Which compounds are specifically synthesized depends on the attacker: generally, biotrophs, or pathogens that rely on live tissue, trigger a response mediated by high levels of salicylic acid, whereas necrotrophs, which destroy the host cells, are counterattacked by a response which involves jasmonic acid and ethylene signaling (Pieterse et al., 2009).

Although plant immunity is often considered a tool to regulate pathogens, it plays important roles in keeping all microorganisms in check. *Arabidopsis thaliana* was recently shown to exhibit a general non-self response to diverse leaf-colonizing bacterial strains, regardless of their pathogenicity. The response consists of the expression of 24 core genes, of which several are involved in cell-wall strengthening and tryptophan-derived metabolism, a hallmark for defense (Maier et al., 2021). Nonpathogenic bacteria also activate PTI, keeping their population at low numbers without fully eliminating them (Velásquez et al., 2022). Indeed, whether and how plants discriminate pathogenic from commensal microbes is currently debated, but there appear to be several mechanisms. *A. thaliana* PRRs, for example, can differentiate types of the flagellin peptide flg22, effectively allowing colonization of bacteria that harbor immune-evading types (Colaiani et al., 2021). In rice, symbiotic fungi are distinguished because they secrete shorter chitooligosaccharides than pathogens (Zhang et al., 2021). In general, non-pathogenic microbes seem to avoid triggering the onset of antimicrobial responses by carefully colonizing roots and preventing wounding or cellular damage (Zhou et al., 2020).

Taken together, the plant immune system can be considered as a multi-functional tool that keeps microbial colonizers in balance. Pathogens, then, are extremely well-adapted colonizers that can overcome this system, becoming detrimental to plant. Surprisingly, however, pathogens still often colonize plants without causing disease (Malcolm et al., 2013). Understanding how this is possible may require looking toward the rest of the microbiome.

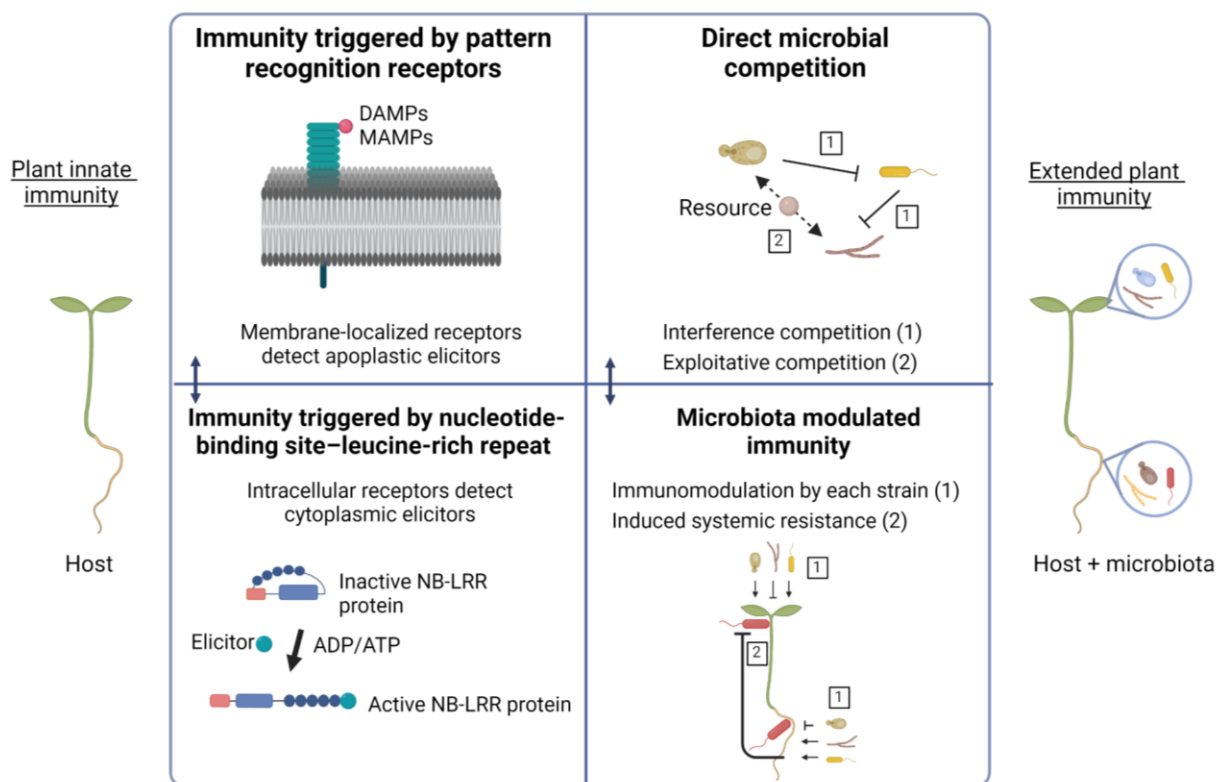


Figure 1. A) Microbiota-mediated extension of the plant immune system. Adapted from Vannier et al (2019).

Diverse adaptations help microorganisms thrive in the leaf apoplast

A small fraction of plant colonizers (i.e., endophytes) can enter the inner tissues of the plant through wounds or natural openings. Microbes that enter the plant are protected from direct sunlight and dehydration but are possibly more prone to changes in the physiological state of the plant. Thus, understanding how endophytes colonize and survive can give important insights into host-microbe interactions. Endophytic colonization of leaves occurs mainly through stomata. These allow gas exchange and transpiration, and their opening and closing is generally regulated by abiotic conditions. High atmospheric humidity, for example, promotes stomatal opening, which can allow more microbial cells to enter the leaf (Panchal et al., 2016). Some colonizers, especially pathogens, are adapted to manipulate this process to suppress closure, rendering the plant more susceptible to infection (Melotto et al., 2006; Xiong et al., 2021).

After entering the leaf, endophytes can establish in the apoplast, the space between cells (Geilfus et al., 2020). Evidence gathered from different studies suggest the apoplast is a harsh environment and thus survival of microbes there requires special adaptations (Hu et al., 2022). In general, endophytes must regulate several processes to avoid triggering the immune response of the plant. Most research on

adaptations has been so far conducted in endophytic pathogens (Naseem et al., 2017), but in the last years the interest to understand how commensals thrive has also increased. A recent study showed for example, that flagellum biosynthesis and chemotaxis were downregulated in commensals once they entered the apoplast and that this effect was intensified as weeks after colonization went by (Velásquez et al., 2022). This underscores the intense pressure colonizers are under to avoid flagellar detection in plant leaves (Vorholt, 2012; Colaianni et al., 2021). Another study showed suppressed genes encoding ribosomal subunits and proton ATPases, suggesting that commensals reduce metabolic activity upon entry in the leaf (Nobori et al., 2022). While these examples provide some of the first insights into commensal gene regulation in the apoplast, it is important to note that in all of them bacteria was injected into the apoplast, which may not be completely comparable to natural colonization.

One major constraint for bacterial growth in the apoplast appears to be water availability, which constantly fluctuates due to environmental conditions and high solute concentration (Beattie, 2011). Plants also further limit water availability to pathogens by increasing the solute concentration as part of the hypersensitive response (Wright and Beattie, 2004). To deal with osmotic stress, some bacteria can increase trehalose synthesis (Freeman et al., 2010; Yu et al., 2013). Indeed, trehalose accumulation is observed with a variety of pathogen infections, including eukaryotes (Brodmann et al., 2002; Foster et al., 2003). Other microbes may form aggregates via exopolysaccharide substances which help maintain hydration (Monier and Lindow, 2003). Furthermore, pathogens such as *P. syringae* cause wetting of the apoplastic space by disturbing *Arabidopsis* cell membranes via the HopM1 and AvrE1 effectors. This may in addition favor the flow of nutrients and promote the spread of the bacterial cells (Xin et al., 2016).

Another major limitation for growth is nutrient availability. In fact, pathogens that establish in the leaf apoplast are in a constant battle with the plant to gain access to resources (Chen et al., 2010; Xin et al., 2016). For example, both *P. syringae* DC3000 (Yamada et al., 2016) and the ascomycete fungal pathogen *Botrytis* (Veillet et al., 2016) induce the production of cell wall invertases in *Arabidopsis* to facilitate the cleavage of sucrose, while *Xanthomonas* uses effectors to activate SWEET transporters and increase sucrose levels in the apoplast (Streubel et al., 2013). Furthermore, production of auxins by some pathogens may loosen the plant cell wall, releasing saccharides (Vorholt, 2012). Commensals presumably have fewer adaptations to manipulate plant nutrient availability and therefore are probably more directly reliant on scavenging nutrients (Paulsen et al., 2005). In fact, their maximum population size may correlate to the amount of available resources (Mercier and Lindow, 2000; Ryffel et al., 2016).

Direct evidence of nutrient availability in plants is limited, but the identification of microbial genes important for survival in the plant environment have shed light on what resources might be most important and the limitations in finding them. For example, fast-growing bacteria like Alphaproteobacteria and Gammaproteobacteria, which are commonly found in plant microbiomes, have disproportionately many low-affinity transporters (Levy et al., 2018), which may enable them to grow quickly even when the nutrient

availability is fluctuating or limiting (Trivedi et al., 2020). Likewise, epiphytic bacteria were shown to activate ABC and TonB transporters to possibly make use of amino acids and carbohydrates in the leaf surface (Delmotte et al., 2009). Genes involved in carbohydrate and amino acid metabolism were also found to be especially important for *Arabidopsis* root colonization by the commensal *Pseudomonas simiae* (Cole et al., 2017). In fact, amino acid availability may be a general limitation inside plants since their synthesis was shown to be relevant for survival of the pathogen *Pantoea stewartii* in the xylem of maize (Duong et al., 2018). This concept is also supported by the reduction of in-plant fitness of *P. syringae* mutants in which the biosynthetic pathways of tryptophan, proline and isoleucine/leucine/valine were disrupted, suggesting they were not able to obtain sufficient supplies from the host (Helmann, Deutschbauer, and Lindow 2019).

Different commensal colonizers may adapt to the plant environment by occupying unique niches. A recent study analyzed the transcriptome of nine diverse commensal strains upon injection into *Arabidopsis* leaves and showed that nutrition-related genes were differentially expressed among taxa (Nobori et al., 2022); genes connected to glycan degradation were induced in Bacteroidetes strains, while genes connected to sulfur-metabolism were induced in Proteobacteria strains, which agrees with previous work (Delmotte et al., 2009). Similarly, transporters for urea, arabinose, xylose and ribose were activated in Proteobacteria strains, suggesting that these commensals were possibly limited for nitrogen and carbon (Nobori et al., 2022). Differences on resource utilization or niche separation may allow co-existence of different microbes and thus could be a major driver in structuring the leaf microbiome and maintaining diversity.

Indeed, the leaf microbiome harbors astounding genetic diversity even within single taxa (i.e. pangenomic diversity) (Karasov et al., 2018; Ramírez-Sánchez et al., 2022) which may also play important roles in adaptation. For example, diversity within the same 16S amplicon sequencing variant (ASV) of both leaf-associated *Pseudomonas* and *Sphingomonas* was shown to be relevant for host colonization (Karasov et al., 2018; Lundberg et al., 2022). Similarly, continuous transfers of *Bacillus subtilis* on *A. thaliana* roots, led to mutations that improved its ability to colonize the plant (Nordgaard et al., 2022). The arisal of genetic diversity within taxa that is relevant for interaction with hosts is not surprising since functional diversity is generally known to be very advantageous in adaptation to different environments. For example, passaging of an isogenic *Curtobacterium* soil bacterium in distinct environments along an elevation gradient led to the emergence of mutations related to nutrient acquisition, stress response and exopolysaccharide production (Chase et al., 2021). Such diversity can make communities more resilient as well. In leaf microbiomes, reshuffling of the microbiome and occurrence of specific functional traits provided better responses to pathogen disturbances (Gao et al., 2021; Li et al., 2022).

Overall, microorganisms deploy a large diversity of strategies to survive in leaves. Many of these traits are involved in obtaining and partitioning resources, suggesting leaf survival depends on overcoming dramatic resource limitations in the host. However, while massive pangenomic diversity is clearly important for host interactions, it is not clear whether this also derives from the pressure to diversify to occupy distinct nutritional niches.

Nutrient availability in the leaf is constantly fluctuating

Despite the clear importance of nutrition in shaping bacterial adaptations required for leaf survival, very little is actually known about nutrient availability in colonizable leaf spaces. A range of carbon sources, including carbohydrates, amino acids, organic acids, and sugar alcohols have been detected on leaf surfaces and in the apoplast that could serve as nutrients for phyllosphere microorganisms (Lohaus et al., 2001; Vorholt, 2012). The presence of these compounds on the surfaces of leaves is mainly thought to be due to the process of leaching (Tukey, 1966), but their distribution is very heterogeneous (Leveau and Lindow, 2001; Wang et al., 2014). Following the fate of single bacterial cells in the phyllosphere showed that there are nutrient hotspots on leaf surfaces that can sustain large densities of cells with relatively high reproduction rates (Remus-Emsermann et al., 2012).

The presence of sugars and amino acids in the apoplast most likely derives from transport between photosynthetic cells and sink organs. Sucrose, the primary sugar produced in leaves, is transported from the source cells to the phloem parenchyma via plasmodesmata. There, SWEET transporters transfer it to the apoplast, from which it is translocated into the sieve element-companion cell complex (SE-CC) by sucrose transporters (SUTs). Once it reaches its destiny (e.g., a young leaf or other sink), it is unloaded from the sieve elements to the companion cells. From these cells it is either released as such to the apoplast or cleaved by cell wall invertases, releasing hexoses into the apoplast. In either case, sucrose or hexoses are transported into the phloem parenchyma cells via SUTs and distributed to other cells (Ruan et al., 2010; Ma et al., 2019). Amino acids, on the other hand, can be taken up by roots and transported directly via xylem, or they can be directly produced in photosynthetic leaves. To reach the sink tissues, they are first passively released into the leaf apoplast of the source leaf and then actively re-imported into the sieve element-companion cells. The unloading process into the sink cells is thought to be symplastic or independent of an apoplast step (Tegeeder and Hammes, 2018).

Loading and unloading processes are tightly regulated by the plant, avoiding uncontrolled leakage of sugars into the apoplast, and ensuring their import back into the cytosol via transporter proteins located in the plasma membrane (Ayre, 2011). These proteins are specially important to avoid excessive leakage under environmental stress like wounding (Meyer et al., 2004) or pathogen attack (Veillet et al., 2016; Yamada et al., 2016). Transcriptomic studies have shown that urea and sugar transporters are activated in *Arabidopsis* leaves upon infiltration with bacteria, possibly to exert control over these compounds (Nobori et al., 2022). Some very well-adapted pathogens, however, have higher affinity transporters than the plant, which may help them to recover carbon sources like sucrose (Wahl et al., 2010).

In addition to the tight regulation of nutrient leakage, availability of resources in the leaf may be hindered by diurnal fluctuations in the plant metabolism. Carbohydrate accumulation in leaves, for example, exhibits diurnal cycles. During the day sucrose is directed to sink tissues and starch is stored as a reserve compound in the chloroplasts; at night starch is broken down to sucrose to provide the plant's needs. However, this may vary depending on the length of the day and the developmental stages of the sink tissues. For example,

during short days, *Arabidopsis* leaves store more starch than during long days (Taiz et al., 2015). Organic acids like fumarate and malate are also accumulated during the day and depleted during the night in *Arabidopsis* leaves (Pracharoenwattana et al., 2010). Nitrogen levels show diurnal fluctuation as well; in tobacco, nitrate levels decrease during the day and increase at night, while ammonium, glutamine, glycine and serine behave oppositely. This is partly because during the first part of the day nitrate assimilation doubles the nitrogen uptake; but later in the day, when the activity of the nitrate reductase decreases, the levels are reversed. This reduces the de-novo synthesis of amino acids during the night, while allowing replenishment of the nitrate pool in the leaf (Matt et al., 2001). Respiration during the night also lowers the levels of certain amino acids, as observed by O'Leary et al. (2017). All of these studies, however, have looked at the overall leaf composition, but few reports have focused on the diurnal fluctuation of metabolites in the apoplast, which would better reflect the conditions affecting microbe nutrition. Differences can be expected since concentrations in the apoplast partly depend on the transpiration rate of the plant (Sattelmacher, 2001) and on the apoplastic pH, which changes change with light exposure (Mühling and Läubli, 2000). Lohaus et al. (2001) for example, found that sucrose, amino acids, malate, as well as chloride accumulated in barley leaves throughout the day reaching the highest values at the end of the light period. Similarly, Gabriel and Kesselmeier (1999) found that the concentration of organic acids and inorganic ions in the apoplast fluid of *Quercus ilex* was considerably higher at night. The authors correlated this observation with a decrease in transpiration and a higher leaf water potential.

Differences in nutrient landscapes should also be expected between plant species. One major source of this variation could be the photosynthesis strategy of the host. Plants are divided into three major categories based on their carbon fixing strategy (C_3 , C_4 and crassulacean acid metabolism -CAM). The C_4 and CAM pathways have evolved from C_3 photosynthesis many times independently in the plant kingdom in response to decreasing levels of CO_2 and warming temperatures in the atmosphere (Sage et al., 2012). This has been generally accompanied by complex reorganizations of the leaf anatomy; i.e. more tightly packed tissues and a complex network of minor veins (McKown and Dengler, 2010). Interestingly, C_4 photosynthesis has arisen in particular genera, leaving multiple species from the same genera that perform different types of photosynthesis (Sage et al., 2012). Both C_4 and CAM pathways reduce photorespiration and enhance C-fixing efficiency by decreasing O_2 availability to RuBisCo. This is because in the presence of O_2 , RuBisCo oxygenates RuBP instead of carboxylating it, producing the toxic intermediate 2-phosphoglycolate. This compound is converted through a series of steps first into glycine, then into serine, then into glycerate and finally into 3-phosphoglycerate, which reenters the Calvin-Benson Cycle. In this energy-costing process, fixed carbon is lost and released ammonia must be reassimilated (Taiz et al., 2015).

In contrast to C_3 plants, where carbon fixing occurs altogether in both mesophyll and bundle sheath cells, C_4 plants divide the process into the two types of cells. In the mesophyll, CO_2 is converted to a four-C compound (oxaloacetate), which is then converted into malate or aspartate and exported to the bundle

sheath for decarboxylation (Sage et al., 2012). Decarboxylation can occur via different enzymes, which helps define different subtypes of C₄ (NADP-ME, NAD-ME and PEPCK) (Hatch et al., 1975). The CO₂ that is released is concentrated around the RuBisCo enzyme, limiting its oxygenation and therefore, increasing the efficiency of C fixation (Sage et al., 2012). CAM plants on the other hand do not separate the process between cells, but rather in time. During the night, these plants open their stomata, CO₂ is fixed into oxaloacetate, which is then converted to malate or other organic acids and stored in vacuoles. During the day, CO₂ is re-released and enters the Calvin Cycle to fix sugars (Taiz et al., 2015). In addition to these three major groups, there are also plants with a C₃-C₄ intermediate photosynthesis. These plants have evolved from C₃ by progressively relocating the decarboxylation of glycine produced in photorespiration to bundle sheath cells. The released CO₂ is then used by Rubisco, increasing photosynthesis efficiency (Sage et al., 2012).

The intermediate compounds that are transferred between cells differ with the photosynthesis type of the plant, which may in turn impact the nutrient landscape of the leaves. While in C₃ and C₃-C₄ plants, there is a recirculation of NH₃, glycine and serine between organelles due to photorespiration, in C₄ plants, it is mostly malate, aspartate and alanine that are transferred (Sage et al., 2012; Taiz et al., 2015). In fact, metabolomic analyses of *Flaveria* leaves, a model genus to study photosynthesis evolution, showed higher concentrations of serine and glycine in C₃ and C₃-C₄ species and higher levels of alanine in C₄ species (Gowik et al., 2011). Given the drastic differences in metabolic processes and transport across different photosynthesis types, it seems plausible that they will affect nutrients available to microorganisms in the apoplast and elsewhere in leaves.

Microbes that manage to cope with the constraints of the apoplast are still likely exposed to nutrient fluctuations. This harsh environment may have direct implications on the interactions across members of the microbiome, and this in turn may affect the communication with the host.

The plant microbiome may protect plants via competition and priming of the innate immune system

Plant immune systems have evolved in the presence of microorganisms. Thus, it is thought that plant immunity is extended by a balanced microbial community, which can work as a barrier against biotic stress, helping to sustain plant health. Indeed, failing to establish a normal microbiota can have devastating effects (Durán et al., 2018; Chen et al., 2020; Runge et al., 2022). This extended immunity is composed of microbe-microbe interactions that play important roles in structuring the microbiomes and in the health of the host (Agler et al., 2016; Durán et al., 2018). Interactions mediated by members of the commensal microbiome that protect plants appear to be mostly competitive (Schäfer et al., 2022) and mechanistically can essentially be divided into interference competition, exploitative competition and immune modulation.

In interference competition, protection from pathogens is provided by producing compounds like antibiotics and hydrolytic enzymes (Fig. 1, upper right). Such is the case of biological control agents like the fungus

Trichoderma which can degrade the cell wall of other fungi (Ting and Chai, 2015) or several *Bacillus* and *Pseudomonas* strains that produce inhibitory compounds affecting a wide range of pathogens (Raaijmakers and Mazzola, 2012). *Pseudomonas piscium*, for example, was shown to protect wheat heads from *Fusarium graminearum* by producing pyocyanins. These molecules inhibit histone acetylation and consequently repress gene expression, fungal growth and pathogenicity (Chen et al., 2018). Interference competition appears to be an important mechanism behind suppressive soils, or soils that result in resistance to a certain pathogen for the plants grown in them. Cha et al. (2016), for example, found that the suppressiveness observed towards *Fusarium oxysporum* causing wilt in strawberry plants was mainly due to an antifungal thiopeptide produced by *Streptomyces*. Likewise, Mendes et al. (2011) found that a chlorinated lipopeptide produced by a Pseudomonadaceae isolate from the soil could inhibit *Rhizoctonia solani*. Furthermore, Carrión et al. (2019) investigated *R. solani*-suppressive soils and determined that the phenomenon could be explained by several *Flavobacterium* from the root endosphere that produce antifungal enzymes and metabolites.

Antibiotic interactions have been rarely observed in leaves, possibly because of the spatial segregation that may hinder one strain encountering another one (Schäfer et al., 2022). On the other hand, the extensive adaptations of leaf colonizers for interference competition suggests that it may be important. For example, the endopeptidase NlaP produced by an *Aeromicrobium* leaf isolate was shown to have strong antimicrobial effects against *Nocardioide*s bacteria *in planta* (Schäfer et al., 2022). In addition, the genome of a *Rhizobium* isolated from the *Arabidopsis* phyllosphere, that confers protection against *P. syringae* pv. tomato DC3000, was enriched in the type-6 secretion system which has been linked to inter-microbial competition among other roles (Vogel et al., 2021).

The second mode of direct protection by commensals is exploitative competition. Here, the protective microorganisms prevent growth of competitors by monopolizing resources (Trivedi et al., 2020). Clear examples of this can be found in biological control agents and probiotics. *Trichoderma*, for instance, gains control over iron resources by producing siderophores and this limits the growth of competing fungi (Segarra et al., 2010). Natural resistance to pathogens in plants could also possibly be explained by resource competition. Kwak et al. (2018) observed that the genus *Flavobacterium* was enriched in tomato plants resistant to *Ralstonia solanacearum*. Interestingly, they found that plant associated *Flavobacterium* is enriched in several monosaccharide transporters and thus the authors hypothesized that competition for carbon sources could explain reduced pathogen fitness. This kind of protection resembles observations in other systems like the gut microbiome. Virulence gene expression of the pathogens *Citrobacter rodentium* and *Candida albicans* were downregulated by commensal bacteria in the intestine cells which compete for their same preferred nutrients (Kamada et al., 2012; Alonso-Roman et al., 2022). Similarly, the strain *E. coli* Nisle reduces *Salmonella* colonization by competing for iron (Deriu et al., 2013) and *Mucispirillum schaedleri* protects mice against colitis caused by *Salmonella enterica* by competing for anaerobic electron acceptors, thus, interfering with invasion gene expression (Herp et al., 2019). Since leaf pathogen virulence

can also be regulated by carbohydrates (Stauber et al., 2012), it is plausible that competition for these resources might lead to attenuated virulence.

The concept of exploitative competition has been generalized by the modern coexistence theory, which proposes that the closer the niche between a native community and an invader, the lower the success of the invasion. However, if the invader succeeds because it has a higher fitness than the native members, the impact on the community will be very significant (MacDougall et al., 2009). This theory was put to test in freshwater bacterial communities, and it turned out that for all three tested invaders, invasion success was best explained by niche differences, while invasion impact had a strong correlation with invader fitness (Li et al., 2019). Similarly in plants, Hu et al. (2016) showed that the establishment success of a probiotic consortia in the rhizosphere microbiome of tomato increased when the diversity and niche breadth in the consortia was higher. The community was also more successful in limiting a pathogen invasion when their niche overlap was higher. These interactions between native and invader organisms are likely also affected by the availability of resources.

In addition to competitive strategies, microorganisms also protect plants by enhancing the immune system of the plant (Fig. 1, lower right). One way this may occur is through the activation of an induced systemic response, where a microbe in the rhizosphere increases the above ground resistance to pathogens (van Loon et al., 1998). This may involve different host plant responses like cell wall fortification (De Vleeschauwer et al., 2008) and increased intensity of defense hormone responses against pathogens and herbivores (Hu et al., 2018). Strains in the rhizosphere may also cause plants to exude compounds that positively alter the surrounding microbiome. Such is the case of *P. simiae* strain WCS471-r, which induces MYB72 and BGLU42 in *Arabidopsis*. These proteins are involved in the secretion of scopoletin, a phenolic metabolite that can function as a siderophore to mobilize iron and which also has an inhibitory effect against soil pathogens (Stringlis et al., 2018).

Together, plants engage with diverse microorganisms that help them thrive in complex and challenging environments. Inter-microbial competition plays particularly important roles, whereby microbial adaptations serve to protect plant hosts by limiting potential pathogens. Plant immune systems regulate and shape colonization and help mediate inter-microbial interactions, ensuring that they stay beneficial. Whether the plant environment in turn modulates these adaptations is not clear.

Stabilizing microbe-microbe metabolic interactions may arise as a consequence of host nutrient constraints

Positive interactions between microbes, particularly cooperation based on metabolic exchanges are known to be widespread across environments and to provide fitness advantages under nutrient-limited conditions (Zelezniak et al., 2015). They may also stabilize microbiomes by limiting the availability of nutrients to

invaders (Herren, 2020). Given the importance of nutrition in leaves, such interactions in the leaf microbiome seem plausible (Hassani et al., 2018).

Nutrient limitation, as expected in the leaf apoplast, could increase the pressure for cooperative interactions. This has been reported in other systems (D'Souza et al., 2018; Zimmermann et al., 2020), regardless of the existence of a coevolution history of the strains (Hoek et al., 2016). For example, Hoek et al. (2016) showed that low amino acid concentration favored cooperation between two yeast strains, but very high concentrations led to competition and competitive exclusion. A similar nutrient effect was observed in the microbiome of *Caenorhabditis elegans*: both *in silico* predictions and *in vitro* experiments showed that mutualism between *Pseudomonas lurida* and *Ochrobactrum* dominated in minimal medium, while a richer substrate promoted competition between the strains (Zimmermann et al., 2020).

The effect of resources on microbe-microbe interactions and thus, on the overall diversity of the microbiome, is likely dynamic and subject to fluctuations in the environment and external factors. The presence of an invader, for example, could lead to increased competition for available resources or alternatively, to the emergence of new niches. Such changes could allow the arrival of certain groups of microbes (Fukami, 2015) or the appearance of new microbe-microbe interactions. Ultimately, these relations might affect how the microbiome interacts with the host.

Linking microbial community structure and function to plant leaf metabolites to discover key processes shaping host health

In the last decades, major efforts have been taken to characterize the diversity of the plant microbiome. Amplicon sequencing of universal genes like 16S rRNA gene for bacteria, the internal transcribed spacer (ITS, a conserved intergenic region) for fungi and 18S rRNA for other eukaryotes has provided insight into the taxonomic composition of microorganisms in the different sections of the plant (Agler et al., 2016; Coleman-Derr et al., 2016; Xiong et al., 2021). Correlating this information with metadata from the environment and host phenotypes has allowed a better understanding of the structuring of the microbiome (Wagner et al., 2016). However, this approach has strong limitations especially for leaf samples because these genes are also present in plastids of the plant, and therefore most of the sequences obtained correspond to host signals (Regalado et al., 2020). This is especially important when studying the community of the whole leaf (epiphytes and endophytes together) or when studying the endophytes only, because the leaf tissue must be broken, releasing the host's genetic material.

Exploring the plant microbiome beyond taxonomy is crucial to understand host-microbe interactions. Functional traits can give better clues on how endophytes adapt to the leaf environment. Nobori et al. (2018) recently designed a protocol to recover the transcriptome of a bacterial pathogen from *Arabidopsis* leaves, revealing important insights into bacterial activity in the apoplast and in response to plant immunity. However, the procedure requires large amounts of leaves to overcome host signals, making it impossible

to look at variation between individual leaves. In another study, Li et al. (2022) used metagenomics to identify enriched functions in endophytic communities of *Citrus unshiu* after a pathogen infection. Metagenomics wastes large amount of sequencing power on host DNA and require the removal of host reads *in silico*, which is only possible when a curated genome of the plant is available. As an alternative, a method that allows separation of the microbe and host signals, if it is sensitive to small samples, could enable resolution of microbial functions and activity at small scales.

Correlating functional data with a description of the leaf apoplast, where host-microbe interactions occur, has the potential of better explaining changes in the microbiome. For example, shifts in abundance of certain bacterial groups or expression of specific genes could respond to fluctuations in the metabolic environment. Untargeted metabolomics could be used to describe and compare the chemical landscape of the host at a given time. Characterizing the composition of the apoplast, rather than the whole leaf, is a more faithful representation of what the endophytic community is actually exposed to, which may help reveal why microorganisms respond the way they do. Furthermore, since the apoplast is crucial for the plant's physiology and its response to abiotic stimuli (Sattelmacher, 2001), focusing on this space could allow exploring host-microbiome-environment interactions.

In the first part of this work, a method to simultaneously obtain a snapshot of host physiology and microbial diversity in single samples was developed. The infiltration-centrifugation method (Lohaus et al., 2001) to recover apoplast fluid from several species was adapted and optimized by assessing different buffers and conditions. The efficiency of the method to recover microbial diversity in the apoplast fractions vs. in whole leaves was tested. In the second part of this work, the optimized method was used to compare the diurnal shifts in the metabolic profiles and microbial communities in the apoplast fluid of three *Flaveria* species with different photosynthesis strategies.

Given the importance of microbe-microbe interactions, increasing the understanding of how they arise and how diversity influences them may help develop better strategies to increase plant resilience. Therefore, in the third and fourth sections of this work, how the apoplast nutrient environment may impact the arisal of metabolic interactions among bacteria was investigated. To address this question, an *in vitro* community enrichment approach was used. Inter-bacterial interactions were dissected at the strain level using metabolomics, genomics and molecular tools. Finally, to understand whether these interactions likely arise in the context of a fluctuating environment like the leaf apoplast, the evolution of the metabolic interactions were assessed under contrasting nutrient conditions.

Chapter 2: Methods

Propagation of plant material

The plants used in this study were propagated under controlled lab conditions. Seeds (Supplementary Table 1) were grown in pots containing a mixture of soil (66%), perlite (33%) and Substral Osmocote® NPK (Mg) 17-9-11 (2) (4 g/L of soil mixture) and kept at an average day/night temperature of 25 °C/22 °C and a photoperiod of 16 hours. The germination rate was very low for some species and therefore, cuttings from adult plants (hereafter, mother plants) were used as propagating material (Supplementary Table 1). To propagate *Alternanthera* species, leaves and flowers were removed from the second or third nodes (top to bottom) of several branches. Next, the nodes were wrapped in aluminum foil and wetted soil to promote rooting. About four days later, segments of the branches including the rooted nodes were detached from the plant and placed on trays with soil. The trays were covered with a plastic lid to ensure humid conditions and promote further growth. One week later, the cuttings were transferred to individual pots. For cuttings of *Flaveria* species, lateral branches with two or three nodes were detached from the plant. To avoid dehydration, the largest leaves, and flowers (if present) were removed. The cuttings were placed in individual pots with soil, and these were placed in a tray and covered with a plastic lid to maintain humidity and promote further growth.

To allow natural colonization, *Flaveria* plants were grown outside in a common garden. For this, cuttings in pots were first acclimated for two weeks in a space protected from direct sunlight outside the lab and after this, they were transplanted to the garden.

Plant sampling

Optimization of infiltration-centrifugation method to obtain leaf apoplast fluid wash (AFW)

For optimizing the technique to obtain AFW for multiple species, leaves from *Zea mays* and different *Alternanthera* and *Flaveria* species were used. The plants were watered at least 2 hours before sampling to make sure the leaves were well hydrated. For all species, fully developed leaves at the second or third node from the top to the bottom were sampled. In the case of *Zea mays*, due to the large size of the leaves, middle sections (about 5 cm long) were used.

Growth of multiple Flaveria species for AFW collection

The AFW metabolite profile of *Flaveria robusta*, *F. linearis* and *F. trinervia* was compared in two different settings: under controlled lab conditions and in common gardens. The lab plants were sampled over the course of two years: *F. linearis* was sampled on two occasions four months apart (March 2020 and July 2020); each time five plants were grown, and ten samples were collected (2 samples per plant). *F. robusta* was sampled on three occasions: March 2020 (4 different plants, sampled twice each), June 2020 (two

different plants sampled twice) and April 2021 (2 different plants, sampled four and two times). *F. trinervia* was sampled in March 2020 along with the other two species (four plants, 2 samples per plant).

The plants grown in the common garden were sampled during summer 2020. Acclimatized cuttings were transplanted on the second week of July to a 3.0 m x 1.35 m common garden. To keep weeds from overgrowing the garden, a polypropylene fleece cover was used (Supplementary Fig. 1a). Two months after transplanting, the three species were sampled at three different points during the day (at 5:00 am, right before sunrise, at noon and at 8:00 pm, right after the sunset). Weather conditions for Jena on the sampling day were downloaded from the Saaleaue Weather Station (https://www.bgc-jena.mpg.de/wetter/weather_data.html). The same five plants of each species were sampled throughout the day. In addition, and to control for the possible effect of continuous sampling might have on the plant's metabolite distribution, two more plants from each species were sampled at 5:00 am (seven in total), and two more at 1:00 pm (nine in total) (Supplementary Fig. 1b). Each sample consisted of a pool of well-developed leaves without visible signs of disease or herbivory, whenever possible. For *F. linearis* and *F. robusta*, only leaves from the middle section of the plant were taken. For *F. trinervia*, which grows close to the ground, leaves that may have been in direct contact with the soil were avoided. The leaf pools had slightly different weights: 1.15 ± 0.25 g (*F. trinervia*), 1.73 ± 0.57 g (*F. linearis*) and 2.07 ± 0.27 g (*F. robusta*).

Comparison of whole leaf and AFW-based methods for characterization of bacterial diversity in Flaveria

To determine if AFW can be used to characterize endophytic bacterial diversity, samples from *F. robusta* and *F. linearis* were taken in October 2019. The plants were grown in a small garden plot at Neugasse 23, Jena (*F. robusta*) or at the Botanical Garden in Jena (*F. linearis*). After two months, 20 samples of fully developed leaves were collected from each species. In the case of *F. robusta*, each sample consisted of an individual leaf, while for *F. linearis* (where only small AFW volumes could be recovered per leaf), each sample consisted of three leaves. The samples were immediately transported in a cooler with ice to the lab and processed further. Under the sterile bench, the leaves were weighed, sterilized first with 2% bleach + 0.02% Triton, followed by 70% ethanol and finally washed three times with sterile water. Half of the samples of each species were destined for whole leaf-based microbiome analysis and were stored in 2 mL screw-cap tubes and frozen at -80 °C. The other half were destined for AFW-based microbiome analysis and were processed immediately (see AFW extraction below). The two species and locations together represent a treatment that is very likely to affect leaf bacterial communities, so this approach allows testing whether one approach is superior in ability to detect a treatment effect.

Enrichment of leaf microbiomes from Flaveria

For enrichment of bacterial communities *in vitro*, leaf extracts were prepared from *F. robusta* and *F. trinervia* plants grown outside in a small garden plot at Neugasse 23, Jena during the summer of 2018. Fully developed leaves from different plants were collected and pooled together into single samples. The leaves were weighed and washed three times in sterile water to remove dirt and insects. Extracts of leaf

microorganisms were prepared as detailed below. This sampling was part of the MSc. thesis of Hafiz Syed Muhammad Abdullah, student from the Plant Microbiosis Lab.

Leaf apoplast fluid wash (AFW) extraction

Recovery of AFW from leaves using the syringe technique

Apoplast fluid wash (AFW) was obtained from leaves by a modified version of the infiltration-centrifugation method (O'Leary et al., 2014; Gentzel et al., 2019). After removing the petiole, the leaves were weighed to record their pre-infiltration weight and were placed in a 60-cc syringe, which was filled with the infiltration solution (Fig. 2). After releasing air, the plunger was placed at the 50-cc mark, the tip was sealed with a gloved finger, and the plunger was pulled until the 55-cc mark and released; this was repeated several times until the leaves lost buoyancy. The leaves were then wiped dry and weighed to record the after-infiltration weight. To recover the AFW, the dried leaves were placed on a sheet of parafilm and gently rolled around a 15-mL tube, which was placed inside a 50 mL tube. Next, the tubes were centrifuged; the resulting liquid (AFW) was transferred to a 2 mL microcentrifuge tube and the leaves were weighed again. The AFW was centrifuged at high speed (15000 x g) for 5 minutes to remove cell particles and the supernatant was transferred to a clean tube and stored at -20°C until metabolomic analysis.

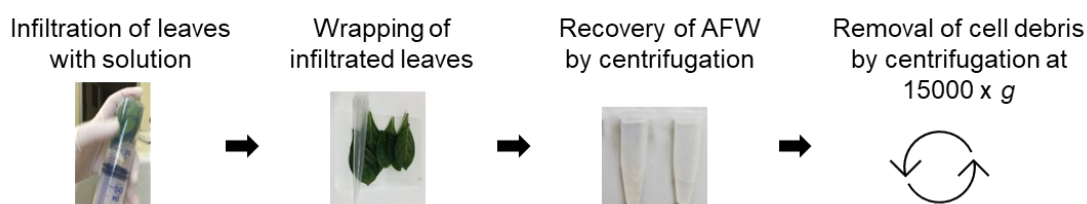


Figure 2. General procedure of Infiltration-centrifugation used to obtain the apoplast fluid wash (AFW) from leaves of different plants.

Since metabolite peak areas will depend on the amount of buffer infiltrated into the leaf, an infiltration ratio was calculated considering the amount of buffer that entered the leaf:

$$\text{Infiltration ratio (inf. ratio)} = \frac{(\text{Leaf weight after infiltration} - \text{Initial leaf weight})}{\text{Initial leaf weight}}$$

This value was used to normalize all samples to allow comparisons between them (See Metabolomics section).

Selection of infiltration solution

To find infiltration-centrifugation conditions that were robust for most plant species, and which allow recovery of clear AFW, the infiltration buffer was assessed. In a first attempt, five different species were infiltrated with either autoclaved sodium phosphate buffer (100 mM, pH 6.2) or 2-(N-morpholino)ethanesulfonic acid buffer (180 mM, pH 5.5) and centrifuged at 7500 x g (Experiment 1, Table

1). Each condition was tested in duplicates. In a separate test (Experiment 2, Table 1), *Flaveria robusta* leaves were infiltrated with either deionized water or sodium phosphate buffer and centrifuged at either 2500 x g or 7745 x g respectively. The resulting AFW after centrifugation was analyzed by untargeted metabolomics.

Table 1. Testing infiltration buffers to obtain a clear apoplast fluid wash from the leaves of different species.

Species	Infiltration solution*	Centrifugation time (min)	Centrifugation speed
Experiment 1			
<i>Zea mays</i>	MES	5	7500 x g
<i>Zea mays</i>	SP	5	7500 x g
<i>Alternanthera tenella</i>	MES	10	7500 x g
<i>Alternanthera tenella</i>	SP	10	7500 x g
<i>Flaveria trinervia</i>	MES	1	7500 x g
<i>Flaveria trinervia</i>	SP	1	7500 x g
<i>Flaveria linearis</i>	MES	1	7500 x g
<i>Flaveria linearis</i>	SP	1	7500 x g
<i>Flaveria robusta</i>	MES	1	7500 x g
<i>Flaveria robusta</i>	SP	1	7500 x g
Experiment 2			
<i>Flaveria robusta</i>	Deionized water	1	2500 x g
<i>Flaveria robusta</i>	SP	1	7745 x g

*MES: 2-(N-morpholino)ethanesulfonic acid buffer, SP: Sodium phosphate buffer.

Selection of centrifugation conditions

Infiltrated *F. linearis* and *F. robusta* leaves were centrifuged at 7745 x g or 2500 x g for different lengths of time (*F. linearis*: 1, 3, 5, 10, 15, 30 and 60 minutes and *F. robusta*: 1, 3, 5 and 6 minutes). After each centrifugation event, the apoplast fluid was collected and the leaves were weighed again to calculate the recovery efficiency (see below). Since a small fraction of the volume released from the leaves is lost during centrifugation and cannot be recovered as liquid, the recovery efficiency was theoretically calculated based on the masses of the leaf before and after centrifugation as follows:

$$\text{Recovery efficiency} = \frac{(\text{Weight after infiltration} - \text{weight after centrifugation})}{(\text{Weight after infiltration} - \text{initial weight})} \times 100$$

In the case of *F. linearis*, the AFW recovered after centrifugation was stored at -20 °C until it was analyzed by untargeted metabolomics.

Assessing the dilution effect on amino acid recovery from AFW

Since the AFW volume recovered from some samples was not enough to use directly for UHPLC-HRMS for metabolomic analysis, dilution was necessary. Therefore, the effect of diluting the AFW on the recovery

of compounds of interest was assessed by studying the effect on amino acids. For this, AFW from *F. linearis* was diluted with sodium phosphate buffer 0, 1.6, 2.5 and 5 times. After dilution, 40 μL of the AFW were mixed with 10 μL of a 100 $\mu\text{g}/\text{mL}$ mix of deuterated amino acids (U-D 98% Cell Free Amino Acid Mix from Cambridge Isotope Laboratories, Inc. Supplementary Table 2), for a final concentration of 20 $\mu\text{g}/\text{mL}$, to provide an internal standard so that each unlabeled amino acid in the samples could be quantified. The diluted samples were analyzed by untargeted metabolomics (See Metabolomics section).

Recovery of apoplast fluid wash from leaves using a vacuum pump

For the diurnal AFW experiment, to process more samples in parallel, a vacuum pump was used instead of a syringe to infiltrate the leaves. The leaves were placed in 50 mL conic tubes and submerged in the infiltration solution. The tubes were sealed with parafilm and arranged inside a modified pressure cooker with an attached vacuum pump and a manometer (Fig. 3). After closing the vessel, the pump was turned on and vacuum was generated until -0.8 Pa. The valve was then slightly opened allowing the infiltration of the leaves over the course of ~ 10 min until reaching normal pressure. After this time, the leaves were visually inspected and if they still had patches indicating no infiltration, the procedure was repeated until they were fully infiltrated. The centrifugation steps to recover the AFW and the calculation of the infiltration ratio were done as described earlier for the syringe method.



Figure 3. System to generate vacuum and infiltrate several leaves simultaneously.

AFW recovery for bacterial community analysis

To test whether it can be used to study diversity of apoplast bacterial communities, AFW was recovered either with the syringe or vacuum pump technique, but under sterile conditions. The leaves were weighed, surface sterilized with bleach and ethanol and washed twice with sterile water in a laminar flow hood. The leaves were infiltrated with sterile sodium phosphate buffer and using sterile instruments (syringes or 50 mL conic tubes depending on the technique). Drying of excess water from the infiltrated leaves was

conducted under the laminar flow hood using sterile cotton buds. The AFW was recovered using sterile tubes and parafilm to reduce contact with foreign DNA and it was stored at -80 °C until DNA extraction.

***In vitro* experiments**

Leaf extract preparation and enrichment of leaf microbiomes

Leaf extracts were prepared as indicated here (Hafiz, 2020) by macerating the washed leaves of *F. robusta* and *F. trinervia* with a sterile pestle, adding 1 mL of 1X PBS with 0.02% Silwet and vortexing for 10 seconds, followed by quick centrifugation to precipitate down the larger segments. The supernatant was transferred to a clean tube, mixed with glycerol to a final 20% v/v concentration and stored at -80 °C. The number of live bacterial cells in the glycerol stocks was estimated by plating serial dilutions on Reasoner's 2A agar (R2A) and calculating CFU/mL.

Leaf extracts were enriched *in vitro* over 12 passages on two different sucrose-based minimal media as indicated here (Hafiz, 2020). First, a preculture was generated by inoculating ~1000 cells in 15 mL of basic M9 minimal broth (See recipe in Supplementary Table 3) supplemented with 0.2% w/v casamino acids (Difco), 200 mM NH₄Cl and 200 µg/mL of cycloheximide to limit eukaryotic growth. The cultures were incubated at 26 °C and 220 rpm for 72 hours. One fraction of the culture was used to prepare 20% glycerol stocks. The remaining volume was centrifuged (5000 x g for 5 min), the pellet washed twice in 1X PBS and resuspended to a final OD_{600nm} of 0.3 in fresh 1x PBS. Five microliters of the suspension were inoculated in a 2 mL 96 well plate containing 1 mL M9 media supplemented with NH₄Cl (33 mM) and either no casamino acids (S-CA) or 0.2% m/v casamino acids (S+CA). Each enrichment condition was run in triplicates. The plate was incubated at 26 °C and 220 rpm. Every 48 hours, the cultures were homogenized by pipetting up and down and 5 µL of each well were transferred to a new plate with fresh media. The OD_{600nm} was measured at each passage and the procedure was repeated 12 times. In the first two passages, cycloheximide was added in the same concentration as before. At the last passage, 700 µL of each well were pelleted for DNA extraction by centrifuging at 20000 x g for 10 min. The remaining volume of the three replicates were combined to prepare 20% glycerol stocks, which were stored at -80 °C.

Isolation of bacteria and first assessment of carbon preference

Bacterial isolates were recovered from the initial leaf extract glycerol stocks and from the glycerol stocks from the twelfth enrichment passage of each condition as indicated here (Hafiz, 2020). Serial dilutions from the glycerol stocks were prepared in 1x PBS and spread in R2A plates. From each condition, 25 random isolates were selected and restreaked to recover pure cultures (150 isolates total– the number of strains were chosen based on feasibility to handle and characterize them). To minimize isolate adaptation to plating, no further cultivation was performed, and pure isolates were grown in R2A broth for 48 hours to prepare 20% glycerol stocks. Isolates enriched from *F. robusta* are named FrLE_X, Fr-CA_X or Fr+CA_X

(from leaf extracts, LE, or from enrichments without or with casamino acids, respectively) and similarly from *F. trinervia*, FtLE, Ft-CA_X or Ft+CA_X.

To quickly test whether the isolates could grow individually in either enrichment media (S-CA or S+CA), single colonies from all 150 isolates were inoculated in triplicates in a 96-well flat bottom plate, containing 200 μ L of media per well as indicated here (Hafiz, 2020). The plates were incubated at 26 °C and 220 rpm in a shaker for 24 h. The OD_{600nm} was measured right after inoculation and after 24 h. A second passage on the same media was done to verify the observed growth was not due to carryover of R2A nutrients. An isolate was categorized as growing if its final OD_{600nm} was ≥ 0.1 .

Preparation of isolates' precultures and testing their growth in different media

Pseudomonas sp. isolates (hereafter *Ps*) were precultured from the glycerol stocks into R2A broth overnight at 28 °C and 220 rpm. The cells were washed twice with 1X PBS, centrifuging each time at 5000 rpm for 5 min and discarding the supernatant in between. The washed cells were resuspended in 1X PBS and diluted to an OD_{600nm} of 0.2. To test their growth in a specific medium, 20 μ L of the bacterial suspension were mixed with 180 μ L of the media in 96-well plates. All strains were inoculated in triplicates or quadruplicates and uninoculated media was always used as negative controls. The plate was shaken at 28 °C and 220 rpm and the OD_{600nm} was measured in a VersaMax Tunable Microplate Reader (Molecular Devices, CA) after 48 hours as an indicator for growth. If a growth curve was required, the plate was incubated in the plate reader and OD_{600nm} measurements were taken every 10 min with mixing in between reads.

A similar procedure was followed to determine the growth rate on S-CA of several isolates that did not require casamino acid supplementation, except their precultures were prepared in S-CA medium to avoid nutrient carryover. To determine the growth rate of each curve, the first derivative was calculated in 30-min windows across the length of the curve, then the periods where the derivative was larger than the average variation were selected (i.e., the exponential phase) and the mean derivative for those time points was calculated.

Production of spent media and evaluation of cross-feeding interactions

To generate spent media, several *Pantoea* sp. (hereafter *Pa*) and FrLE and FtLE isolates were grown in 250 mL flasks with 80 mL of S-CA (sucrose-only) media at 26 °C and 220 rpm. After 48 hours, the cells were separated from the spent media by centrifuging at 5000 x g for 5 minutes and filter sterilizing twice through a 0.22 μ m PES filter. Spent media were subjected to untargeted metabolomics (see details in the Metabolomics section) or used further to assay growth of *Pseudomonas*. To measure taken up compounds after growth of *Pseudomonas* in the spent medium, the plate was centrifuged at 5000 x g for 5 min and the supernatants were transferred to a clean plate and stored at -20 °C until metabolomic analysis. To test for inhibitory effects, the spent media of *Pantoea* isolates *Pa* Ft-CA_14 and *Pa* Ft+CA_17 were diluted 1:2 with

spent media of *Pa* Fr-CA_6 and inoculated as before. Growth (OD_{600nm}) of the *Pseudomonas* strains was monitored over the course of 24 hours.

Tagging of Pseudomonas strains with fluorescent proteins

The *Pseudomonas* isolates Fr-CA_5 and Fr+CA_3 were tagged with the fluorophores mTagBFP2 and mOrange2 respectively, using the delivery plasmid systems pMRE-Tn7-140 and pMRE-Tn7-144 developed by Schlechter et al. (2018). In brief, *E. coli* ST18 containing each plasmid was grown overnight in LB broth + 5-aminolevulinic acid (ALA) (50 µg/mL) + Amp (100 µg/mL) + Gen (15 µg/mL) + Cam (15 µg/mL) at 30°C and 200 rpm. The *Pseudomonas* strains were grown in LB without ALA or antibiotics under the same conditions. On the following day, 25 mL of LB + ALA + antibiotics were inoculated with 500 µL of the overnight *E. coli* cultures and 25 mL of LB without antibiotics were inoculated with 1000 µL of the *Pseudomonas* culture and grown until an OD_{600nm} of about 0.7. The cells were harvested by mild centrifugation (2000 x g for 5 min) and after discarding the supernatant, they were resuspended in 1X PBS by gently inverting the tubes. The *Pseudomonas* and the corresponding *E. coli* cultures were mixed in three different ratios; 1:3, 1:1 and 3:1 based on the OD_{600nm} of the suspensions. The combined cells were harvested as before, resuspended in 100 µL of 1X PBS and drop spotted in an LB + ALA plate containing 100 µL 1 M CaCl₂ to increase conjugation efficiency. The plates were incubated at 30 °C overnight; the next day a colony was suspended in 1 mL 1X PBS and different volumes (100-300 µL) were spread in LB plates containing antibiotics and without ALA to counter-select against *E. coli*, and incubated at 30 °C for several days. Colonies that appeared in the plates were checked for fluorescence under UV light (330 nm) and in a Axiozoom Stereomicroscope (Zeiss). To rule out that the selected colonies were *E. coli*, a PCR using the specific primers FWD_uidA and REV_uidA (Supplementary Table 4) was performed. Additionally, the presence of the fluorescent protein coding sequence was checked with the primers FWD_Tn5/7_gt and REV_Tn5/7_g (Supplementary Table 4). In both cases, the Mastermix consisted of: 1x Buffer B (Biodeal), 0.4 mM dNTPs (Carl Roth), 0.2 µM of each primer, 2.5 mM MgCl₂ and 0.25 µL Taq DNA polymerase (Biodeal). For the *E. coli* specific primers, the cycling conditions were: 3 min at 95 °C, 30 cycles of 30 sec at 95°C, 30 sec at 57.5 °C and 1 min at 72 °C, followed by a 5 min final extension at 72 °C. For the set of primers targeting the transposon, the cycling was the same, except for the annealing temperature, which was set at 56.5 °C. Once the colonies were confirmed to be fluorescent *Pseudomonas*, they were grown in liquid LB + Gen + Cam at 37 °C and 200 rpm to cure them from the plasmid. The efficacy of this last step was confirmed via PCR with the primers FWD_Tn7_gt, Rev_Tn7_gt and Tn7_gt (Supplementary Table 4) targeting the plasmid backbone. The Mastermix was prepared as detailed before, except the concentration of each primer was 0.12 µM and the annealing temperature was set to 62 °C. To determine whether producing the fluorescent protein had a fitness cost on the strains, their growth on PaFr-CA_6 spent media was assessed. Precultures of the tagged and non-tagged strains were prepared in R2A broth as usual, inoculated to a final OD_{600nm} of 0.02 in 180 µL of spent media and incubated for 24 hours at 220 rpm and 28 °C. To check if either tagged strain had a fitness advantage over the other one, they were combined in

different ratios considering their OD_{600nm} (1:1, 3:1 and 1:3) and grown in R2A broth under the same conditions as before. CFU's were counted after 24 h under UV light to differentiate each strain.

Co-culture experiments using tagged Pseudomonas strains

Using the tagged *Pseudomonas* strains, four co-culture assays were established. As detailed before, 180 µL of the corresponding media were inoculated with 20 µL of bacterial suspension (see Supplementary Table 5 for details on the inoculation volumes of each strain in each experiment). 1) The tagged strains were combined individually with *Pantoea* Fr-CA_6 or together in a full mix (both *Pseudomonas* strains and *Pantoea*) and inoculated in black 96-well plates containing S-CA media. The plate was covered with an AeraSeal (Excel Scientific, Inc.) sealing film and incubated in a BioLector I (m2p-labs Beasweiler, Germany) at 500 rpm, 30 °C with humidity control. Every 24 hours, 5 µL of each well were transferred to 195 µL of fresh media. This was repeated for a total of eight passages. The full community was also subcultured every 48 hours for four passages. Each community had four replicates. After each round, the plate was opened under sterile conditions and 2 µL were sampled for serial dilutions and CFU counts, whereby *Pseudomonas* Fr-CA_5 and *Pseudomonas* Fr+CA_3 were distinguishable due to their fluorescence. During the runs, growth of *Pseudomonas* Fr+CA_3 was monitored by normalizing the red filter channel signal (gain=100) to the biomass signal (gain= 35) and this ratio was used to compare its growth in co-culture with *Pantoea* alone vs. together with *Pseudomonas* Fr-CA_5. 2) The tagged strains were passaged in S+CA in either monoculture or two co-culture ratios (1:1 and 1:3, being the latter higher on PsFr+CA_3:mOrange2) every 24 hours for a total of four passages. Each community had four replicates. 3) Both tagged strains and two additional *Pseudomonas* isolates: PsFr+CA_2 and PsFr+CA_18, were grown in either monoculture or in each possible pair on the spent media of PaFr-CA_6 for six 24-h passages. Samples for metabolite uptake were collected after the first passage following the same procedure described above. CFUs of each community were counted in LB plates after the last passage. Colonies were checked for fluorescence under UV light (330 nm) and in a Axiozoom Stereomicroscope (Zeiss). In the Ps Fr+CA_2 and Ps Fr+CA_18 combination, only total number of cells were recorded as it is not possible to distinguish their morphologies. 4) The two tagged *Pseudomonas* (Fr-CA_5:mBFP2 and Fr+CA_3:mOrange2) were evolved together with either *Pantoea* Fr-CA_6 or *Pantoea* Fr+CA_20 on the two different media: S-CA or S+CA. Each community had four replicates. The communities were passaged every 24 hours for 25 days and in every passage; at passage 22, the communities that had been growing in S-CA were transferred to S+CA and vice versa. The second, third and fourth experiments were carried out in 96-well plates (Greiner 655101), which were incubated at 28 °C and 220 rpm in an orbital shaker. For the passages, 5 µL of each well were transferred to a new plate containing 195 µL/well of the corresponding fresh media. The OD_{600nm} and the fluorescence of the mTagBFP2 and mOrange2 were read at the time of passaging in a Tecan Plate Reader (Infinite 200Pro), using the following parameters for each channel: (blue signal: Ex/Em WL: 402/457 nm, gain = 78, 3 flashes, Z-position: 18249 µm, 2x2 reads per well with a border of 1250 µm; red signal: Ex/Em WL: 550/580 nm, gain = 134, 10 flashes, Z-position: 19190

µm, 2x2 reads per well with a border of 1400 µm; biomass signal: absorbance at 600 nm, 10 flashes, 1 reading per well).

***In-planta* experiments**

In-planta testing of *Pseudomonas* colonization

The *Pseudomonas* strains Fr-CA_5 and Fr+CA_3 and the *Pantoea* strains Fr-CA_6 and Fr-CA_14 were grown in LB agar plates. The colonies were resuspended in 10 mM MgCl₂ with 0.02% silwet and diluted to an OD_{600nm} of 0.002. *F. robusta* and *F. trinervia* cuttings were prepared in the lab under controlled conditions as described earlier. Three-week old plants were used for testing bacterial inoculation. *F. robusta* cuttings were inoculated with each *Pseudomonas* alone or in combination with *Pantoea* Fr-CA_6. *F. trinervia* cuttings were inoculated in the same way, except using *Pantoea* Fr-CA_14 for the co-inoculations. In all inoculations, the bacterial suspensions contained each strain at an OD_{600nm} of 0.002.

Two fully independent experiments were performed for each species; for *F. robusta*, a total of 8 leaves were infiltrated and for *F. trinervia*, 12 leaves. The inoculation was carried out on the underside of the leaves, using a needles syringe. After inoculation, the plants were arranged in trays (each treatment in a separate tray to avoid cross-contamination), covered with a plastic lid to maintain humidity and placed inside a growth chamber with a photoperiod of 16 hours and a day/night temperature of 22/18 °C. After 10 days, the leaves which had been inoculated were harvested, surface sterilized first with 2% bleach + 0.02% Triton, followed by 70% ethanol and then washed three times with sterile water. Next, the leaves were thoroughly crushed and suspended in 10 mM MgCl₂. Serial dilutions were prepared and plated in LB agar. The plates were incubated for 48 h after which the colonies were counted.

DNA extraction and amplification

DNA extraction and amplification from whole leaves and AFW

Whole leaves (WL) were thoroughly crushed using a sterile pestle and transferred to a 1.5 mL screw cap tube containing 2 metal beads (3 mm diameter) and 0.2 g of 0.25-0.5 mm glass beads. The tubes were beat for two 30 sec cycles at 1400 rpm using a Mini-Beadbeater-96 (Biospec Products). The AFW samples were also bead-beat for the same time but using only glass beads. DNA was extracted by incubating the samples with 600 µL of CTAB buffer at 37 °C for 10 min, followed by 5 min centrifugation at 20000 x g. To precipitate the DNA, 200 µL of cold 100% ethanol were added, and the samples were centrifuged at 20000 x g for 30 min at 4 °C. The supernatant was discarded, and the pellet was washed twice with 70% ethanol, each time centrifuging for 5 min. Finally, the pellet was rehydrated in 100 µL of 1x TE buffer and purified with 1.5x volume of Sera-Mag™ magnetic carboxylate modified particles and eluted in 40 µL of 10 mM TrisHCl (pH 8.0). The purified DNA was stored at -20 °C until further amplification. The precipitation step was slightly modified for the AFW recovered from *Flaveria* plants sampled throughout the day: 3M sodium acetate was added to the 100% ethanol and the samples were left overnight at -20 °C. All the other steps were not changed.

Amplification of the 16S V3-V4 region was carried out in a two-step PCR with an enzymatic cleanup in between. In the first step, the AFW samples were amplified with the bacterial primers 341F/806R (Supplementary Table 4). The WL samples were also amplified with the same primer pair, but a pair of blocking oligos (Supplementary Table 4) was also added to limit the chloroplast amplification. On the second step, the PCR products were further amplified using concatenated primers designed to include an Illumina adapter P5 (forward) or P7 (reverse), an index sequence, a linker and the 341F/806R sequences. Full details on the reagents and PCR conditions are shown in Supplementary Table 6. All products of the second PCR were purified with 0.65 – 0.8x volume magnetic beads and eluted in 10 mM TrisHCl. The AFW amplicons were run in a 1.5% agarose gel; the brightness of each band in the gel was measured in ImageJ (version 1.52a) and used to adjust the volume of each sample in the final library. The WL samples were combined based on their fluorescence, which was measured with 200x diluted PicoGreen (Invitrogen™). After combining, the pools were purified once again with 0.8x volume Sera-Mag beads and quantified in a Qubit (Thermo Fisher Scientific, Inc). The libraries were denatured and then loaded onto a MiSeq lane spiked with 10% PhiX genomic DNA to ensure high sequence diversity and were sequenced in 500 cycles (2x250 bp).

The AFW samples from the *Flaveria* sampling throughout the day were amplified with a modified protocol; in the first step, the primers 341F-OH/799R-OH were used (Supplementary Table 4). In the second step, the concatenated primers included the Illumina adapter, the index sequence, and an overhang region (OH) which matched the overhang sequence in the bacterial primers. Full details on the PCR reactions for each step are shown in Supplementary Table 6. After the second step, all PCR products were purified with 0.65 – 0.8x volume magnetic beads and eluted in 10 mM TrisHCl. Measurement of fluorescence, pooling, purification, denaturing and loading into the MiSeq was done as described above.

DNA extraction and amplification of enriched bacterial communities

Cell pellets were resuspended in 600 μ L SDS extraction buffer (SDS extraction buffer, 10% filter sterilized SDS, 100 mM Tris pH 8.0, 200 mM NaCl, 2 mM EDTA) and transferred to 2 mL screw-cap tubes containing ~0.2 g of glass beads (0.25-0.50 mm, ROTH). The mixture was incubated at 37°C for 10 min, followed by bead beating for 30 sec at 1400 rpm in a BioSpec Mini-Beadbeater-96 and 4 minutes centrifugation at 13400 rpm. The supernatant was transferred into new 1.5 mL sterile Eppendorf tubes and tubes and 200 μ L of 5 M potassium acetate were added to precipitate the SDS. After 5 minutes of centrifugation at 13400 rpm, the supernatant was discarded, and the pellet was resuspended in 1x TE buffer. The DNA was purified using 1.5x volume of Sera-Mag™ magnetic carboxylate modified particles and eluted in 50 μ L of 10 mM TrisHCl (pH 8.0). The purified DNA was amplified using the same method described above for the AFW samples using the bacterial primers 341F/806R. More details are shown in Supplementary Table 6.

DNA extraction and amplification from bacterial isolates

All isolates were identified via DNA extraction and Sanger sequencing of the 16S rRNA gene. Liquid cultures in R2A broth of each isolate were pelleted by centrifugation at 20000 x g for 10 min and stored at -20 °C. DNA extraction and purification were done as described for the enriched bacterial communities. Whole 16S rRNA gene was amplified with the universal primers 8F and 1492R (Supplementary Table 4). Details of the PCR reaction are given in Supplementary Table 6.

DNA extraction and whole genome sequencing of bacterial isolates

Whole genome sequencing was performed on three *Pseudomonas* isolates (Fr-CA_5, Fr+CA_3 and Fr+CA_18) and four *Pantoea* isolates (Fr-CA_6, Fr+CA_20, Ft-CA_14 and Ft+CA_17). For this, a modified DNA extraction protocol was used. The isolates were grown overnight in 2 mL of R2A broth at 28 °C and 220 rpm. The cultures were pelleted by centrifugation (20000 x g for 5 min), resuspended in 600 µL SDS extraction buffer and transferred to 2 mL screw-cap tubes containing ~0.2 g of glass beads (0.25-0.50 mm, ROTH). The mixture was incubated at 37 °C for 10 min, followed by bead beating for 30 sec at 1400 rpm in a BioSpec Mini-Beadbeater-96 and five minutes centrifugation at 20000 x g. The supernatant was transferred into new 1.5 mL sterile Eppendorf tubes to which 100 µg/mL of Proteinase K (Sigma) were added. The mixture was incubated for one hour at 37 °C, followed by 10 min at 80 °C to deactivate the enzyme. Once the tube had cooled, 10 µg/mL RNase A was added and they were incubated at 37 °C for 30 min. To remove proteins, an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) were added to each sample, followed by 10 min centrifugation at 4 °C and 20000 x g. The top layer was transferred to a clean tube where an equal volume of chloroform / isoamyl alcohol (24:1) was added. The centrifugation step was repeated and again the top layer was transferred to a new tube. One-tenth volume of 3M sodium acetate and 2.5 volumes of 100% ethanol were added, and the tubes were left overnight at -20 °C. To pellet the DNA, the tubes were spun for 40 minutes at 20000 x g and 4 °C; the supernatant was removed, and the pellet was washed with 200 µL of 70% ethanol and left to air dry. Finally, the DNA was resuspended in 100 µL of 10 mM TrisHCl. The samples were sent to Microbial Genome Sequencing Center (Pittsburgh, USA) for sequencing on the NextSeq 2000 (Illumina) platform at a depth of 300 MBp (~50x for *Pseudomonas* strains and ~60x for *Pantoea* strains).

Metabolomics

Ultra-high performance liquid chromatography - high resolution mass spectrometry

Ultra-high performance liquid chromatography coupled with high resolution mass spectrometry (UHPLC-HRMS) was used to analyze the metabolic profile of the AFW from *Flaveria* plants, as well as the samples from the cross-feeding and co-cultures assays. Samples that were directly compared were analyzed in the same HPLC-HRMS run (Supplementary Table 7), and all runs were with the same conditions at the Mass Spectrometry Platform of the Friedrich Schiller University Jena. As internal standard, samples were spiked with a 100 µg/mL mix of deuterated amino acids (U-D 98% Cell Free Amino Acid Mix from Cambridge

Isotope Laboratories, Inc., Supplementary Table 2) before injection, to a final concentration of 20 µg/mL. UHPLC-HRMS was carried out using a THERMO (Bremen, Germany) UltiMate HPG-3400 RS binary pump, WPS-3000 auto sampler which was set to 10 °C and which was equipped with a 25 µL injection syringe and a 100 µL sample loop. The column was kept at 25 °C within the column compartment TCC-3200. The chromatography column used was a THERMO Accucore® C-18 RP (100 × 2.1 mm; 2.6 µm) with the following gradient: Eluent A (water with 2% acetonitrile and 0.1% formic acid) and Eluent B (pure acetonitrile), respectively at a constant flow rate of 0.4 mL/min: 0 min (100% and 0%), 0.2 min (100% and 0%), 8 min (0% and 100%), 11 min (0% and 100%), 11.1 min (100% and 0%), 12 min (100% and 0%). The injection volume was 5 µL.

Mass spectrums were recorded with THERMO Q Exactive plus orbitrap mass spectrometer coupled to a heated electrospray source (HESI). Column flow was switched at 0.5 min from waste to the MS and at 11.5 min again back to the waste, to prevent source contamination. For monitoring two full scan modes were selected with the following parameters. Polarity: positive; scan range: 80 to 1200 m/z; resolution: 70,000; AGC target: 3×10^6 ; maximum IT: 200 ms. General settings: sheath gas flow rate: 60; auxiliary gas flow rate 20; sweep gas flow rate: 5; spray voltage: 3.0 kV; capillary temperature: 360 °C; S-lens RF level: 50; auxiliary gas heater temperature: 400 °C; acquisition time frame: 0.5 - 11.5 min. For negative mode, all values were kept the same except for the spray voltage which was set to 3.3 kV.

Metabolomics data analysis

Raw data files were converted into mzML format in MSConvert (ProteoWizard, Version 3.0.19246-075ea16f5). To reduce the size of the files, they were filtered by peak picking with the Vendor algorithm, followed by threshold peak filtering using absolute intensity at $1.0E2$. Next, they were imported into MzMine (version 2.40.1) for processing. The spectrums were first separated into positive and negative ionization modes. The baseline was corrected at 1.0 m/z bin width and using the asymmetric correction method. The peaks were then selected by centroid detection and by setting the noise level at $1.0E2$. The chromatograms were built with the ADAP Chromatogram builder. The settings to build the chromatograms were slightly modified for each run by verifying the right integration of the deuterated amino acids peaks. The minimum group size in a number of scans was set to 2-3, the group intensity threshold was either $1.0E4$ or $6.0E4$, the minimum highest intensity was set at the same value as the group intensity and the m/z tolerance was set at 0.001 m/z. The generated chromatograms were deconvoluted with the Wavelets (ADAP) algorithm, setting a signal to noise threshold of 10.0, a minimum feature height of $1.0E4$, a coefficient/area threshold of 50, peak duration range 0-1 and a RT wavelet range of 0-0.1. Furthermore, the peaks were aligned together using the Join Aligner with m/z tolerance of 0.001, retention time tolerance of 0.1 min and setting the weight for m/z and RT at 75 ppm and 25 ppm respectively. Finally, the matrix of negative and positive mode peaks, with their corresponding ID, m/z, retention time and area were exported into R (version 3.6.0).

Using an in-house script, uncommon (in less than 50% of the samples in the run) and small peaks (area <1.0E4) were filtered out.

In the case of the AFW samples, the areas were corrected in the following way:

$$\text{Corrected area} = \text{Area under the curve} \times \text{Inf. Ratio} \times \frac{1}{\text{Dil. factor}} \times \text{Dil. in vial} \times \frac{1}{\text{Inj. Vol}} \times 1000$$

Where:

Inf. Ratio = infiltration ratio of the leaf (See “Apoplast fluid wash recovery” section)

Dil. factor = if the AFW was diluted before injection, the dilution factor was calculated as follows:

$$\text{Dilution factor (Dil. factor)} = \frac{\text{Volume of AFW}}{(\text{Volume of AFW} + \text{volume of Millipore water})}$$

Dil. in vial = if a mix of deuterated amino acids (10 μL) was added to the sample (40 μL) before injection.

This value was always 1.25.

Inj. Vol = volume of sample injected in the UHPLC-HRMS. This value was always 5 μL .

To get a first overview of the distribution of the peaks across plant species, a principal coordinate analysis, constrained by plant species using Euclidean distance on the corrected peak matrix was conducted with the R (R Core Team 2020, version 4.0.4) packages *phyloseq* (McMurdie and Holmes 2013, version 1.34.0) and *vegan* (Oksanen et al. 2020, version 2.5-7). Before plotting the data, the areas under the curve of each peak were transformed by $\text{glog2}(x) = \log_2\left\{\frac{x + \sqrt{x^2 + a^2}}{2}\right\}$, where “a” is a constant with a default value of 1. This transformation was used as it has been shown to emphasize biological variation over technical variation in metabolomic data (Parsons et al., 2007). To compare the amino acids areas, the peak matrix was first annotated against a custom library of 570 compounds including all proteinogenic amino acids (MSMLS™, Sigma Aldrich) which had been developed in the same equipment and with the same UHPLC-MS method. This annotation was carried out in R with an in-house script (See data availability) by setting the m/z tolerance at 0.002 and the RT tolerance at 0.2 min. The area of each amino acid was corrected towards its deuterated internal standard as follows:

$$\text{Corrected area} = \text{Area of amino acid in sample} * \frac{\bar{x} \text{ area of deuterated amino acid in all samples in the run}}{\text{area of deuterated amino acid in sample}}$$

Next, the concentration of each amino acid was calculated based on the known concentrations of the standard (Supplementary Table 2).

To compare the profile of *Flaveria* apoplast leaves at different sampling times, the peak matrices were: Filtered to exclude small peaks (<1.0E4) and uncommon peaks (present in less than half of the samples), log transformed and subjected to differential analysis using the R package *DESeq2* (Love, Huber, and Anders 2014, version 1.30.1). The function estimates size factors, the dispersion, performs a negative binomial generalized linear model fitting using Wald tests. This results in a data frame with log2 fold change (log2FC), *p* value and the adjusted *p* value for each peak, which was used as input for the volcano plots. The plots were generated with the package *EnhancedVolcano* (Blighe, Rana, and Lewis 2020, version 1.8.0), setting the cutoff values for log2FC and *p* value at 2 and 0.01, respectively. The significantly different peaks that surpassed the log2FC cutoff in each case were used to annotate for possibly enriched metabolic pathways. This analysis was carried out in Metaboanalyst (Xia et al. 2009, version 5.0) using the workflow “Functional Analysis”. The input data included the m/z and retention time of each peak. The Mummichog algorithm, version 2.0 was used, setting the *p* value at 0.05. The pathway library of *Arabidopsis thaliana* was used to annotate the peaks and calculate enriched pathways. The heatmaps were generated directly in the program, clustering by Ward’s method. The compound hits were annotated against the KEGG database in R.

To find out which compounds had been taken up via cross-feeding, the spent media was analyzed before and after growth of the consuming isolates. To focus exclusively on uptake of large peaks, only peaks present in high amounts in the spent media (before growth of the consumers) were considered. This cutoff was defined for each dataset and ranged between 1.0E4 and 1.0E5. The log2 fold change, *p* value and FDR-corrected *p* value were calculated for each peak. Metabolites with a fold change lower than -2 (after growth / before growth) and an FDR *p*<0.05 were defined as significantly taken up. Venn Diagrams were created using the online tool Venny (version 2.1.0) and heatmaps were built in R with the package *ComplexHeatmap* (Gu, Eils, and Schlesner 2016, version 2.6.2). The resulting peak tables were annotated against the custom library as described above.

Bioinformatics

Processing of amplicon sequencing data

Amplicon sequencing data was split on the indices and the adapters were trimmed from the read ends using Cutadapt. The data was clustered into amplicon sequencing variants “ASVs” using the R package dada2 (Callahan et al. 2016, version 1.18.0), setting the quality filter at (truncLen=c(200,200)) for the ASW vs WL samples, and at (truncLen=278) for the rest. Next, the sequences were dereplicated to remove redundant reads, followed by denoising using the error rate and calling of ASVs in the forward and reverse reads. When merging the forward and reverse reads, only overlapping regions were kept. Chimeric sequences were removed and obtained a sequence table with the merged data. For taxonomy assignment, the Silva database (v132) was used. Diversity analyses were carried out in R with the packages *phyloseq* and *vegan* and *DeSeq2*.

Processing of Sanger sequencing data

The sequences were trimmed in R using the package *sangeranalyseR* (Chao et al. 2021, version 1.26.0) with default parameters. The resulting fasta file was “BLASTed” against the NCBI 16S rRNA gene database. The first ten hits were used to identify taxonomy at the genus level using a least common ancestor algorithm in MEGAN6 with percent to cover set to 70 and the minimum percent identity to 97 (Huson et al., 2016).

Whole genome assembly, annotation, ANI calculation and single nucleotide polymorphism detection

For each isolate, the raw forward and reverse reads were trimmed using Trimmomatic (version 0.39), which removed the Illumina adapters and trimmed at quality below 15 in a 4-base wide sliding window. The assembly of the genomes was done using SPADes (3.14.1) with the default parameters in “isolate” mode. Average nucleotide identity between the different isolates of each genera was calculated in Kbase (Arkin et al., 2018). First, the final scaffolds from the SPADes assembly were imported into Kbase the average nucleotide identity was calculated using the FastANI app. Sequence variants (SNPs and indels) compared to the *Pseudomonas siliginis* D26 reference genome assembly accession GCF_001605965.1 (*Pseudomonas* Genome Database, Winsor et al. 2016) were then called. To do so, SNIPPY (version 4.6.0) was used with default parameters in “Multi” mode to map the raw sequencing reads from each of the three *Pseudomonas* strains. This resulted in one variant table for each strain including the chromosome location, the gene it was in (if any) and the predicted effect. An R script was then used to parse these tables to determine which variants were shared or unique for each *P. siliginis* isolate. Unique variants were extracted for each isolate and were manually filtered for potentially disruptive mutations (disruptive in frame insertions/deletions, frameshift variants, missense variants, stop gains and initiator codon variants). These lists were used to make word clouds (<http://www.wordclouds.com>) where the word size represents the relative frequency of disruptive variants in a particular gene product.

Data availability and scripts to generate figures from raw data

Supplementary figures and tables are provided in the Supplementary Information section. Supplementary Tables 8, 9, 14, 16-18, 20-22 were too large to include here and are available in the electronic folder Supplementary Material, along with scripts and data used to recreate all figures and tables. mzML files for Chapters 3.1 and 3.2 are also provided in the electronic folder. Raw data from amplicon and whole genome sequencing of Chapters 3.3 and 3.4 have been made publicly available in the NCBI BioProject PRJNA778092 and raw metabolomic data for Chapters 3.3 and 3.4 are being made publicly available in MetaboLights under the study MTBLS3719.

Chapter 3.1 Parallel characterization of apoplast chemical composition and microbial diversity

A critical unanswered question facing researchers in understanding plant health is related to how microorganisms affect plants in an ecological context (Harris et al., 2020). A major challenge in understanding the ecological mechanisms that underlie plant colonization by microorganisms is the lack of simultaneous insight both microbial diversity and plant physiology (the environment microorganisms face). Such insights could help enable predictions about how and why microorganisms interact to shape plant health. Characterizing the nutrient environment in leaves that endophytic commensal bacteria face requires a metabolomics approach but accessing the intercellular environment these microbes have access to is challenging. In combination with microbial diversity data, this technique could allow better understanding the dynamics of plant-microbe interactions under natural conditions. Here, the infiltration-centrifugation method (Lohaus et al. 2001) was adapted to simultaneously obtain a snapshot of host physiology and microbial diversity in single samples. In short, optimized infiltration buffers and centrifugation conditions were used to characterize both microbial diversity and the metabolome of the apoplastic space of multiple species within the *Flaveria* genus.

Sodium phosphate buffer allows UHPLC-HRMS analysis of apoplast fluid wash (AFW) from several plant species

Infiltration-centrifugation is a two-step method that allows a quick recovery of the apoplast fluid from leaves (Lohaus et al., 2001). In the first step, leaves are infiltrated with a solution to dissolve compounds in the apoplastic space, and in the second step, the solution is recovered by centrifugation. The recovered solution can then be analyzed, for example for metabolome characterization (O'Leary et al., 2014). Here, the focus was on amino acids and other small, polar compounds that have been reported to be important in bacterial nutrition in the apoplast (Rico and Preston, 2008). To get an accurate picture of the apoplast composition, the leaf cell membranes should remain intact. How to achieve this may vary between plant species and conditions of the sample and thus, adjustments to the method may be required (O'Leary et al., 2014).

It was previously reported that deionized water infiltration and gentle centrifugation did not rupture maize cells (Gentzel et al., 2019). However, tests performed on *Flaveria robusta* leaves with deionized water and slow centrifugation (2500 x g) resulted in a green AFW (Fig. 4a, left panel). This was likely due to the low osmolarity of the water, which could drive compounds across the cell membrane, even if cells were not ruptured. Thus, sodium phosphate buffer (SP, 100 mM, pH 6.2) with a slightly acidic pH similar to reported apoplast pH levels and therefore, more compatible with the leaf apoplast (Geilfus, 2017), was tested. This trial resulted in a clear AFW, even when fast centrifugation (7500 x g) was used. Untargeted metabolomics revealed that overall, the metabolic profile was different than that obtained with water (PERMANOVA p

value = 0.02857, Fig. 4b, Supp Fig. 2), but none of the peaks with significantly different areas ($\log_2FC > 2$, FDR < 0.05) could be identified (Supplementary Table 8).

The same infiltration solution was then tested in four more species with different leaf characteristics: *Alternanthera tenella*, *F. linearis* and *F. trinervia*, which are all dicotyledons with C₃, C₃-C₄ and C₄ photosynthesis respectively, and *Zea mays*, a monocotyledon. This time, an alternative buffer 2-(N-morpholino)ethanesulfonic acid buffer (MES, 180 mM, pH 5.5), which has been used in different studies (Ruhe et al., 2016; Geilfus et al., 2021) was also tested. Although both buffers resulted in clear AFW, MES was not suitable because its elution time and ionization patterns were similar to compounds of interest, including all proteinogenic amino acids (Fig. 4c). Correspondingly, the metabolic profile obtained with these two buffers was very different for most species tested (PERMANOVA p value=0.0071, Fig. 4d). Therefore, although deionized water was previously shown to be a good apoplast extraction medium, this is not true for all plant species and for small, polar compound characterization. As a result, SP was selected for subsequent tests.

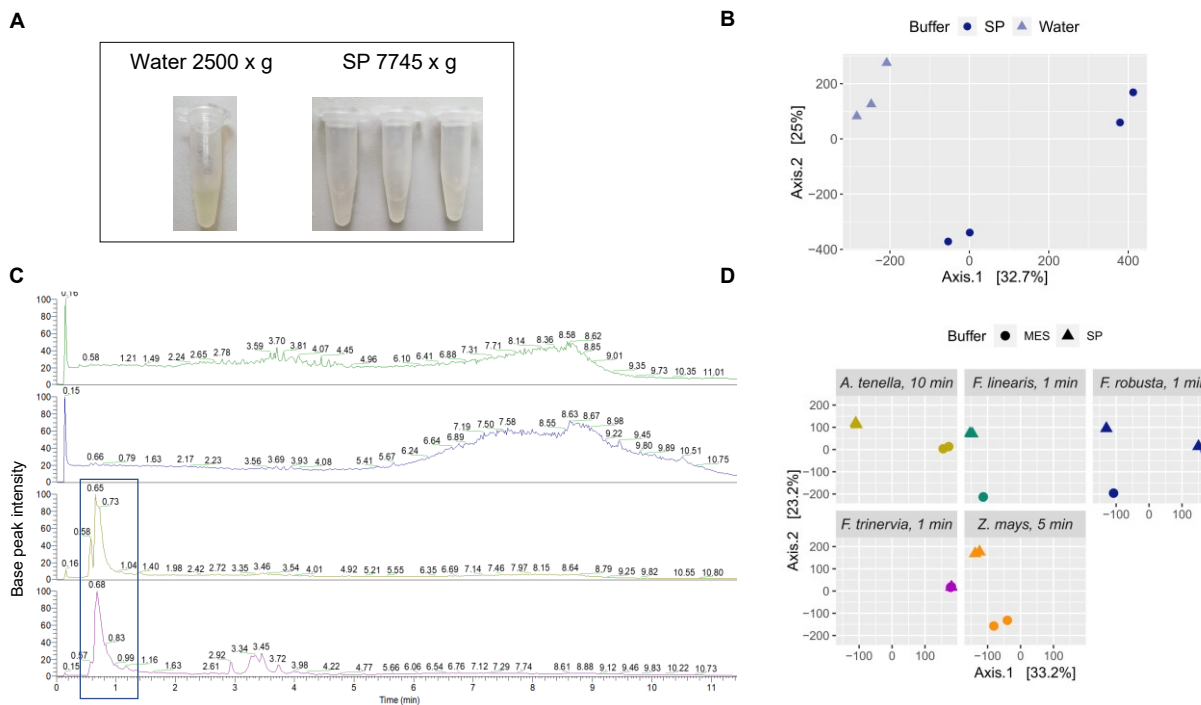


Figure 4. Sodium phosphate (SP) buffer is suitable to obtain apoplast fluid wash (AFW) from plant species with different leaf characteristics. A) Green vs. clear AFW obtained from *Flaveria robusta* leaves infiltrated either with MilliQ water or SP buffer, respectively, and centrifuged for 1 minute at 2500 x g and 7745 x g correspondingly. B) MDS based on Euclidean distance of all UHPLC-HRMS peaks (area >1.0E4) detected by untargeted metabolomics of the extracted AFW of *F. robusta* with either MilliQ water or SP buffer under the same conditions as in panel A. N=3 for MilliQ water infiltrations, N=4 for SP buffer. C) Comparison of ionization profiles of the AFW of *Flaveria linearis*. The first two chromatograms show positive and negative ionization of AFW obtained with SP buffer, and the following two with 2-(N-morpholino)ethanesulfonic acid (MES) buffer. The blue box marks the MES interference with the ionization. D) MDS metabolic profiles on the apoplast fluid wash (AFW) from leaves of different species infiltrated with MES or SP buffer. The centrifugation conditions are detailed for each species. N=2 for each buffer. A. = *Alternanthera*, F. = *Flaveria*, Z. = *zea*. In panels B and D, a generalized logarithm transformation (Parsons et al, 2007) was applied to the data.

Gentle centrifugation allows recovery of clean apoplast fluid wash from infiltrated leaves

The apoplast space includes an aqueous layer that is used for transport processes (Sattelmacher, 2001). The water content is variable due to several factors, including the plant species, the physiological state of the plant and microbial activity (López-Millán et al., 2000; Gentzel et al., 2019; Nakano et al., 2020). When the leaves are infiltrated, the infiltrate mixes with the fluid already present in the apoplast, diluting the compounds. Different methods have been tested to account for this dilution, but they are generally laborious (Lohaus et al., 2001; O’Leary et al., 2014). Recently, the calculation of the apoplast hydration index was proposed as an alternative. This method is based on weight differences of the leaf before and after infiltration and centrifugation and it requires that the leaves are centrifuged until a constant weight is reached that is lower than the pre-infiltration weight. This indicates that all infiltrated liquid plus the liquid present in the apoplast before infiltration is fully recovered (Gentzel et al. 2019). For some species this point

might be achieved quickly, but for others, centrifugation might need to be extended or carried out at higher speeds, which could affect the integrity of the leaf and thus the composition of the AFW. To determine optimum centrifugation conditions for the species used in this study, and to determine whether the apoplast hydration index could be used to calculate the dilution effect of infiltration, *F. robusta* and *F. linearis* leaves were infiltrated with SP buffer and centrifuged at two different speeds for increasing times. After each centrifugation event the leaves were weighed to calculate the recovery efficiency (see Methods) and the resulting liquid was collected for metabolomic analyses.

In both species, high speed centrifugation allowed a full recovery of the infiltrated volume after just 3 minutes (red arrows in Fig. 5a and Supplementary Fig. 3). However, to reach a constant weight after spinning as suggested by Gentzel et al. (2019), the leaves had to be centrifuged further (Fig. 5a and Supplementary Fig. 3), which led to visible damage of the leaf tissue. Under lower speed, *F. linearis* leaves stayed intact but the infiltrated volume was harder to recover. Indeed, no further liquid could be collected after 5 minutes, but at this point the recovery efficiency was only 77% (Fig 5a). Based on weight differences, full recovery was attained at 30 minutes (red arrows in Fig. 5a), but there was no liquid to analyze. In *F. robusta*, the recovery efficiency plateaued at 3 minutes and no further liquid could be collected afterwards (Supplementary Fig. 3). Beyond this point, the leaves began to appear ruptured, thus the centrifugation was terminated, and full recovery of the infiltrated volume was not reached.

To establish whether the differences in AFW recovery from *F. linearis* under the two centrifugation conditions were due to contamination with cytoplasm, the extracts were characterized with untargeted metabolomics. The overall composition of the liquid recovered under high speed changed significantly over time (Fig. 5b, PERMANOVA p value= 0.0261), especially after 3 minutes of centrifugation. Furthermore, a comparison of the first two samplings (1 and 3 minutes) against the last two (30 and 60 minutes) showed that almost all enriched compounds had a larger peak area (FDR<0.05, log₂FC> 2.0) in the later time points (Supplementary Fig. 4). Only a few of these compounds could be annotated (Supplementary Table 9) and one of them, 3-Hydroxy-3-Methylglutarate, involved in the degradation of lysine to acetyl-CoA is not expected in the apoplast (Hildebrandt et al., 2015). Together, these results suggest that the apoplast hydration index cannot be calculated for *Flaveria* as indicated by Gentzel et al. (2019) since high-speed centrifugation damages the leaves, possibly contaminating extracts with cytoplasm metabolites. Therefore, the dilution effect of the infiltration was calculated using the infiltration ratio instead (see Methods). Since liquid collected at 1, 3 and 5 minutes under low speed did not significantly differ in composition (Fig. 5c, PERMANOVA p value = 0.06667), 2500 x g for 3 minutes was chosen as an optimum centrifugation condition. Although only ~70% of the infiltrated volume is recovered with this procedure (Fig. 5a and Supplementary Fig. 3), the integrity of leaves and of the metabolic landscape of the apoplast is assured.

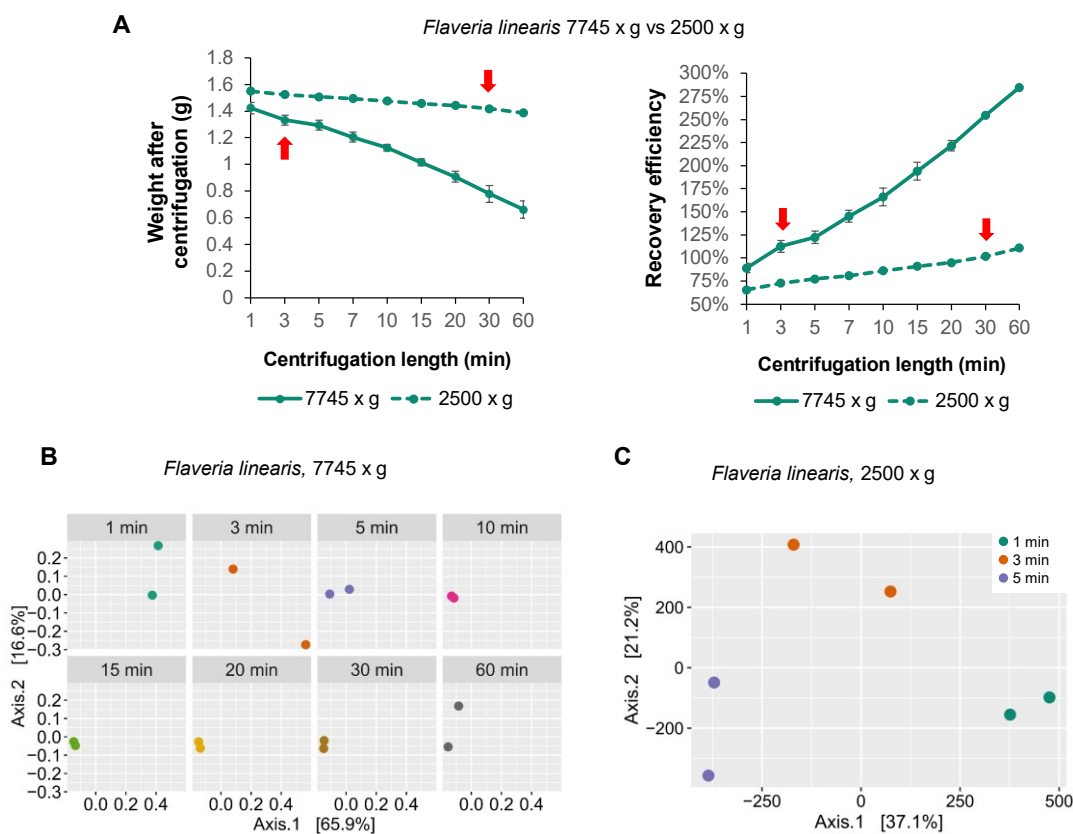


Figure 5. Centrifugation speed and duration affects recovery efficiency and composition of apoplast fluid wash (AFW) A) Recovery of AFW from infiltrated *Flaveria linearis* leaves at increasing centrifugation times under either high (7745 x g) or low (2500 x g) speed. The left panel shows the weight of the leaves after each centrifugation event. The right panel shows the recovery efficiency (see Methods for the calculation details). The red arrows indicate the centrifugation time in which the weight after centrifugation equals the weight before infiltration, meaning full recovery of the infiltrated volume or 100% recovery efficiency. Each point shows the average of two replicates and the standard error. B) MDS plot based on all UHPLC-HRMS peaks (area >1.0E4) in untargeted metabolomics of the recovered AFW of *F. linearis* leaves after centrifugation for the indicated times at high speed. N=2. C) MDS based on all UHPLC-HRMS peaks (area >1.0E4) in untargeted metabolomics of the recovered AFW of *F. linearis* leaves after each centrifugation event at low speed. N=2. A generalized logarithm transformation (Parsons et al, 2007) was applied to the data in panels B and C.

Dilution of samples does not hinder detection of compounds of interest

Ideally, the apoplast metabolic landscape could be measured with the resolution of single leaves. However, the apoplastic volume of individual leaves varies widely across species (McKown and Dengler, 2010), thus the AFW obtained after centrifugation may be very limited in some cases and would require dilution into a larger volume before analysis. To test whether dilution affects the detection of compounds of interest, AFW recovered from *F. linearis* leaves under the selected conditions (SP buffer, followed by 3 min centrifugation at 2500 x g) was diluted several times and amino acids were quantified in the apoplast fluid. After correcting the peak areas with the corresponding dilution factor, most amino acid concentrations were not significantly affected up to a dilution of 5x. Only L-Val and L-Glu were very slightly overestimated at intermediate

dilutions (Fig. 6). In addition, an untargeted analysis of all detected metabolite peaks in the same run showed that differences between the undiluted sample and the dilutions were not significant (Supplementary Fig. 5, PERMANOVA p value = 0.07959). Thus, metabolite profiles in single leaf resolution are possible even when samples are small and require dilution.

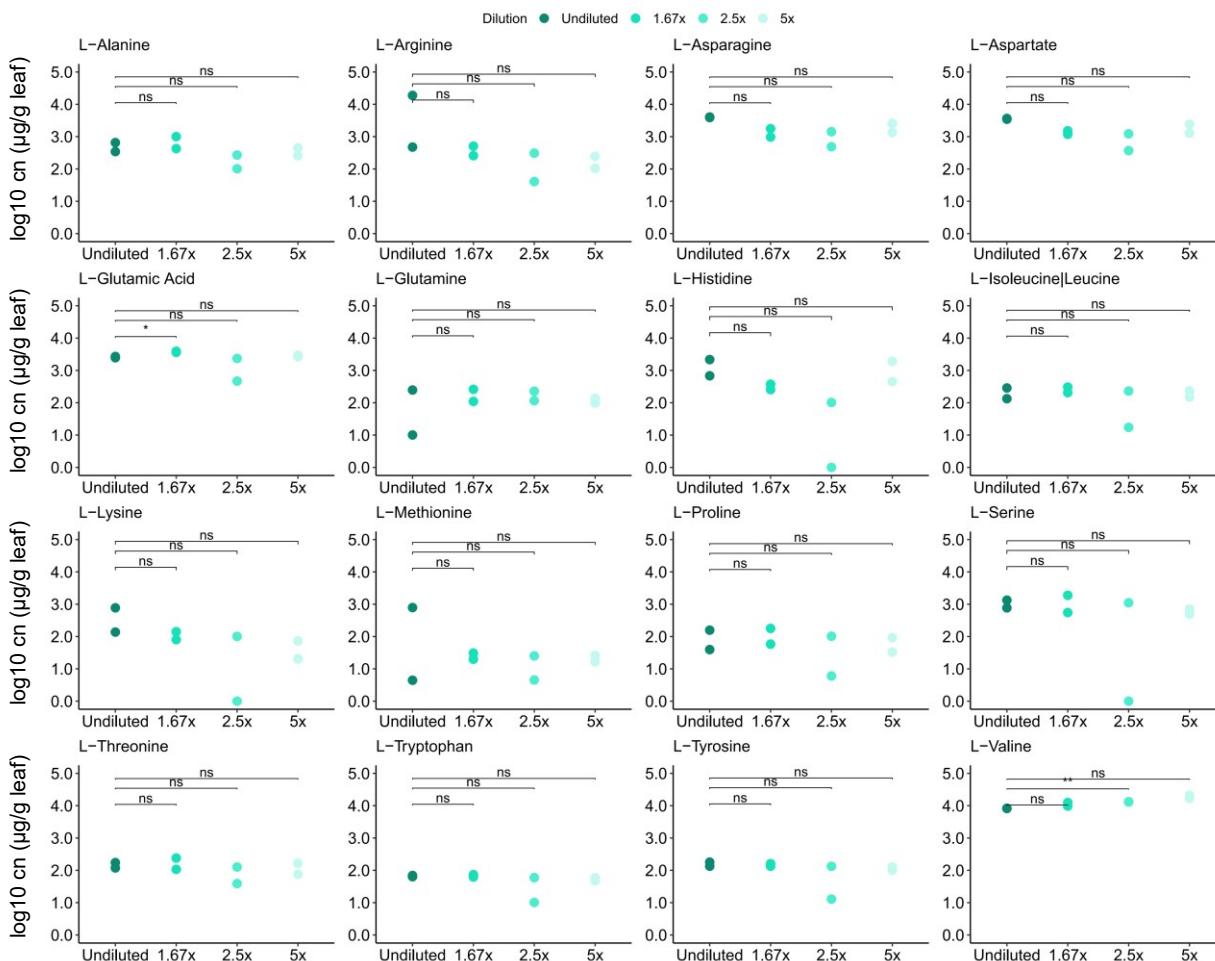


Figure 6. Correction by dilution factor does not overestimate amino acid concentrations in apoplast fluid wash (AFW). Amino acid concentrations in the apoplast fluid wash (AFW) of *F. linearis* diluted up to 5x in sodium phosphate (SP) buffer are shown. Concentrations were calculated based on an internal standard. Significance values are based on t-tests comparing concentrations between dilutions (ns: p value > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001). N=2 for each dilution.

Potential resources for microorganisms in the apoplast differ between *Flaveria* species

The optimized method to recover apoplast fluid wash was used to characterize the apoplastic space of three species of *Flaveria*, each of which utilizes a different photosynthesis pathway. *F. robusta* utilizes C₃ photosynthesis, *F. trinervia* utilizes C₄ photosynthesis, and *F. linearis* uses an “intermediate” C₃-C₄ photosynthesis type. The plants were grown under controlled laboratory conditions and were sampled at the same time of the day to avoid differences due to diurnal fluctuations. Samples from each species were taken from several plants over the course of a year to avoid nesting of the results by sampling event. Compounds were identified as distinct peaks in the untargeted UHPLC-HRMS profiles. Constrained analysis of principal coordinates based on all detected compounds showed an overall clear separation of the plant species (Fig. 7a, PERMANOVA p value= 9.999e-05). Using a targeted approach in parallel, amino acids were quantified in the extracted apoplastic fluid. Among all detected amino acids, there were significant differences between plant species (Kruskal Wallis p value<0.05) in all except L-Glu and L-Trp. Among seven amino acids whose production has highest metabolic costs (Akashi and Gojobori, 2002) and which therefore maybe especially interesting for colonizing microorganisms, the aromatics (L-Tyr, L-Trp and L-Phe) were significantly more abundant in *F. trinervia* than in *F. linearis* and tended to be higher in *F. robusta* (Fig. 7b). Three of the other four could also be detected (L-Met, L-Ile, L-Lys, but not L-His) and all were significantly higher abundance in *F. trinervia* than the other two species (Fig. 7b). These results show that the nutrient landscape for potential leaf apoplast colonizers differs between plant species, with some like *F. trinervia* potentially offering richer resources.

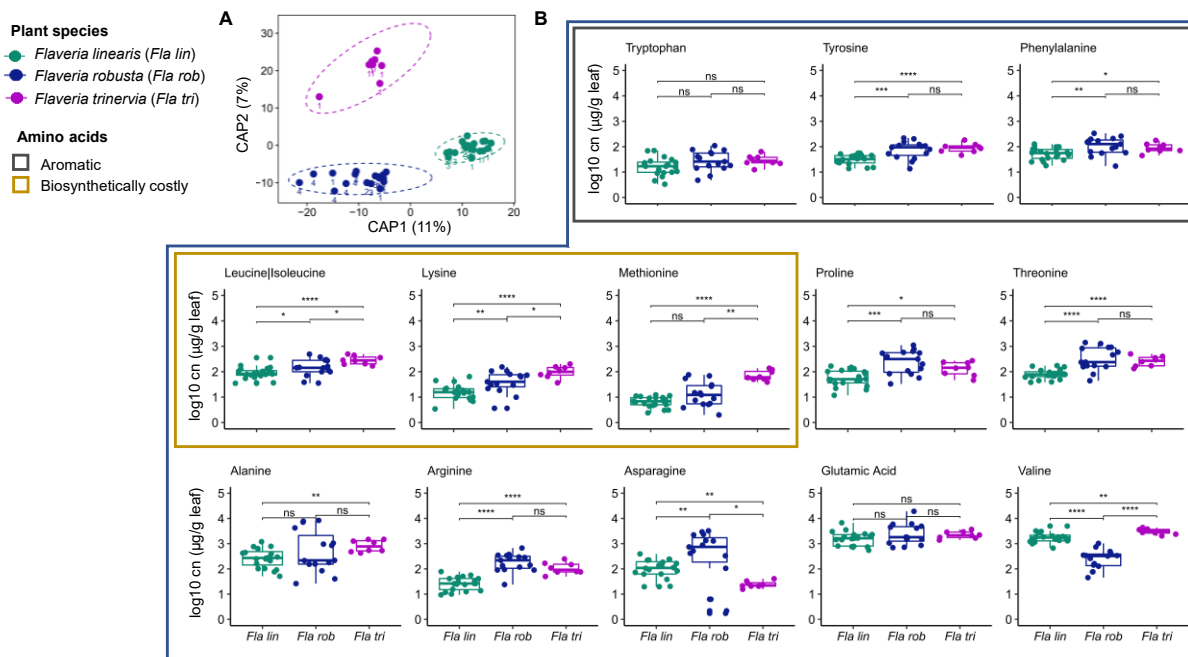


Figure 7. *Flaveria* species have distinct metabolic composition of the leaf apoplast fluid wash. A) Principle coordinate analysis on Euclidean distances, constrained by plant species and based on all UHPLC-HRMS peaks (area > 1.0E4) in untargeted metabolomics of the extracted leaf apoplast. Points represent replicate sampling at a given time point, ellipses show 95% confidence intervals, numbers indicate the sampling event: N = 20 *F. linearis*, N = 18 *F. robusta*, N = 8 *F. trinervia*. A generalized logarithm transformation (Parsons et al. 2007) was applied to the data. B) Amino acids concentrations in the apoplast of *F. trinervia*, *F. robusta* and *F. linearis*. Concentrations were calculated based on an internal standard. Significance values are based on Kruskal Wallis test comparing concentrations between species (ns: p value > 0.05, *p < 0.05, **p < 0.01, ***p < 0.001).

Higher bacterial diversity can be recovered from AFW than from whole leaves.

Conventional methods to study endophytic bacterial community structures in leaves consist of surface sterilization followed by crushing to recover the microbial DNA embedded in the tissue, followed by sequencing 16S rRNA gene amplicons to characterize microbiota. This results, however, in a large amount of plant reads derived from chloroplast and mitochondrial 16S rRNA, which can obscure bacterial diversity information. Here, the possibility to extract bacterial DNA and characterize its diversity directly from the AFW of naturally colonized *F. linearis* and *F. robusta* plants was tested by comparing this approach to the conventional one. Using the conventional approach, 98.0% and 99.6% of the reads from *F. robusta* and *F. linearis* respectively, corresponded to plant mitochondria and chloroplast 16S. This was despite using blocking oligos (Mayer et al., 2021) to reduce the host signal and reflects the very low bacterial loads typical in the apoplast (Regalado et al., 2020). In AFW samples, these values decreased to 66.5% and 60.8%, respectively (Supplementary Table 10). The reason for some plant contamination in AFW samples is unclear, but it is not surprising given natural damage that occurs in the garden, including full exposure to insects. At any rate, although total read depths were similar, usually only <100 bacterial reads per sample were available for WL (many were 0), while many more were recovered from AFW. This difference, unsurprisingly, enabled better resolution of diversity patterns. Specifically, the estimated alpha diversity in

the AFW was significantly higher for both species than in the crushed leaves (Fig. 8a). This can be explained because more bacterial genera could be identified in the AFW samples than in the WL ones in both *Flaveria* species (Supplementary Table 11, Supplementary Fig. 6). A constrained analysis of beta diversity showed that using the AFW method also allowed better separation of the two species (PERMANOVA p value = 0.002 for AFW and 0.4989 for the WL method, Fig. 8b). In addition, the diversity recovered with each method was clearly different for both species (PERMANOVA p value = 0.003 for *F. robusta* and 0.006 for *F. linearis* Supplementary Fig. 7). It is possible that with equal sampling depth the methods would have performed similarly, but for the whole leaf samples this would have required extremely deep sequencing to overcome the host signal.

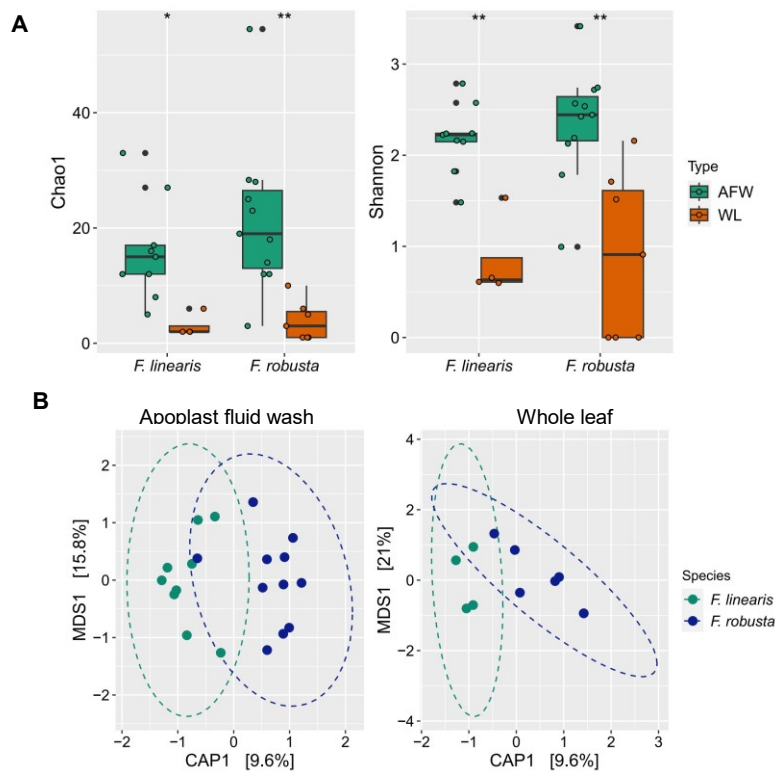


Figure 8. Bacterial diversity trends were clearer using the apoplast fluid wash (AFW) extraction method than the conventional whole crushed leaves (WL) method. A) Alpha diversity estimates of total amplicon sequencing variants (ASV) richness (Chao1) and richness and evenness (Shannon). Significance values are based on t-tests (ns: p value > 0.05, *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001). B) Principle coordinate analysis of Bray-Curtis distances, constrained by species. The left panel shows samples obtained by the apoplast fluid wash (AFW) and the right panel, whole leaf samples. Ellipses indicate 95% confidence intervals. In both panels, A and B, only samples with bacterial reads are shown; N=7 *F. robusta* WL, N=11 *F. robusta* AFW, N=4 *F. linearis* WL, N=9 *F. linearis* AFW.

Chapter 3.2: The leaf microbiome and chemical landscape simultaneously cycle diurnally

Availability of nutrients in leaves are likely to be transient and variable, partly because of diurnal fluctuations (Lohaus et al., 1995). Whether this variability can also shape the microbial community on similar timescales is not known. However, studies in roots for example, have suggested that oscillations throughout the day of the microbiome may be linked to changes in the composition of root exudates (Staley et al., 2017; Lu et al., 2021). To begin to address this question, the methods developed in Chapter 3.1 were used to ask whether plant leaf apoplast metabolomes and microbiomes fluctuate on diurnal time scales.

Leaf metabolic profiles fluctuate throughout the day in a species-dependent manner

Flaveria robusta, *F. linearis* and *F. trinervia* were grown together in a common garden in summer 2020 and sampled across one day at 5:00 am, 1:00 pm and 8:00 pm. Apoplast fluid wash collected from the leaves was analyzed via UHPLC-HRMS and 16S rRNA amplicon sequencing to study fluctuations in the chemical landscape and the bacterial community. 5-9 samples were taken per timepoint from specific plants so that differences between individual plants could be detected. Discolored AFW (likely due to cell rupture) was discarded. During the early morning, the temperature was near 10 °C and the relative humidity was at 100%. As the morning progressed, the temperature increased reaching its maximum (30 °C) at 3:00 pm, while the relative humidity decreased reaching the lowest point (10%) also around the same time. Towards sunset (8:00 pm), the temperature decreased to 15 °C and the relative humidity reached again its maximum at 100% (Supplementary Fig. 8). Temperature and relative humidity have a direct impact on many plant processes, including the opening and closing of stomata. Since these changes are nested with the diurnal cycle, they may influence the results observed here.

Although only healthy leaves were sampled, some were damaged during the processing, resulting in a yellowish AFW, indicative of cell membrane rupture (indicated as empty spaces in Table 2). These samples were considered in neither the metabolomic analysis, nor in the amplicon sequencing runs.

Table 2. Samples that were subjected to metabolomic analysis (UHPLC-HRMS) from each species at each time point. Each number represents an individual plant that was sampled at one or more time points (color coded accordingly). Missing x's indicate samples discarded due to cell rupture (see methods).

Plant ID	5:00 AM							1:00 PM									8:00 PM					
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8	9	1	2	3	4	5	
<i>F. linearis</i>	x	x	x	x	x	x	x	x	x			x	x	x	x	x	x	x				x
<i>F. robusta</i>	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
<i>F. trinervia</i>	x	x	x	x	x	x	x	x							x	x	x					

An ordination analysis of all the peak data together showed that the three *Flaveria* species clearly differed and that the samples grouped by time of day (Fig. 9a). PERMANOVA analysis showed that species identity had the largest effect (15% of variation, $p = 0.001$, Fig. 9b). Only 2.3% of variation ($p = 0.001$) on the other hand, correlated to sampling time (Fig. 9c). This was likely because the effect of sampling time differed by species ($p < 0.001$ for the interaction of the variables): Whereas in *F. linearis* 8:00 pm was most different from 5:00 am, in *F. robusta* and *F. trinervia*, 1:00 pm and 5:00 am had the largest differentiation (Fig. 9a).

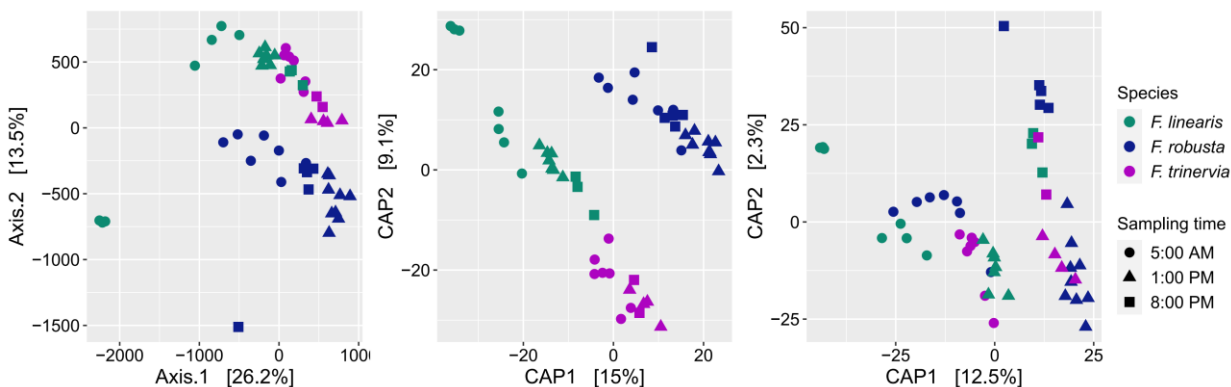


Figure 9. Metabolic composition of apoplast fluid wash differed between *Flaveria* species and grouped by sampling time. A) MDS based on all UHPLC-HRMS peaks (area $> 1.0E4$) in untargeted metabolomics of the recovered AFW from leaves of three *Flaveria* species grown in a common garden and sampled at different time points during the day. B) Principle coordinate analysis of Euclidean distances, constrained by species. C) Principle coordinate analysis of Euclidean distances, constrained by sampling event. A generalized logarithm transformation (Parsons et al, 2007) was applied to the data before analysis.

Indeed, within each species apoplast metabolites varied strongly with the sampling time (explaining 23.6% of the variation, $p < 0.001$ for *F. robusta*, 33.1%, $p = 0.0022$ for *F. linearis* and 32.6%, $p < 0.001$ for *F. trinervia*, Fig. 10). Additionally, all individual plants within a species exhibited similar diurnal patterns: The differences between individuals were generally not significant. The slight difference observed between plants in *F. robusta* was possibly due to plant #1 sampled at 8:00 pm, which separated from the rest of the night samples (Supplementary Table 12). Even so, the effect of sampling time was much stronger. In addition, the patterns were robust to multiple sampling events since plants that were sampled only once (Table 2) had the same patterns as plants that were sampled throughout the day.

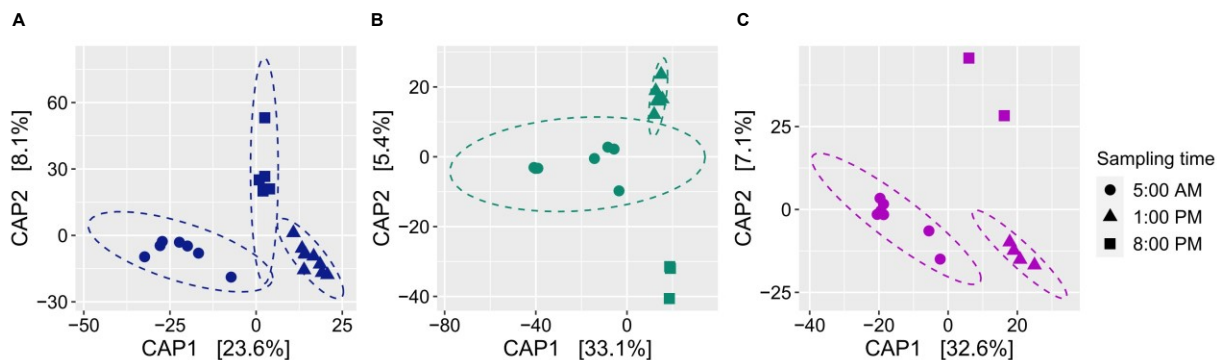


Figure 10. Apoplast fluid wash of *Flaveria* species showed different metabolic composition throughout the day. Principle coordinate analysis based on all UHPLC-HRMS peaks (area >1.0E4) in untargeted metabolomics of the recovered AFW from leaves of three *Flaveria* species grown in a common garden and sampled at different time points during the day. The data was constrained in each case for the sampling time and a generalized logarithm transformation (Parsons et al, 2007) was applied to the data before analysis. A) *Flaveria robusta* B) *F. linearis* C) *F. trinervia*.

To determine which peaks or compounds were responsible for the observed differences over time, metabolite peak areas were compared in a differential analysis (Fig. 11). In *F. robusta* and *F. linearis*, the 1:00 pm sampling had more peaks with significantly larger areas than the 5:00 am (p value < 0.01 and $\log_2FC < -2$) or 8:00 pm sampling (p value < 0.01 and $\log_2FC > 2$), while the 8:00 pm sampling had more peaks with larger areas than the 5:00 am (p value < 0.01 and $\log_2FC < -2$). In *F. trinervia*, the calculations were hindered by the low number of replicates in the latest time point, but overall, the trend was opposite: the 5:00 am and 8:00 pm samplings had more significantly larger peaks than 1:00 pm (Supplementary Table 13). Thus, metabolites that could be detected across multiple time points tended to maximize in abundance at 1:00 pm in *F. robusta* and *F. linearis* but to minimize at the same time point in *F. trinervia*.

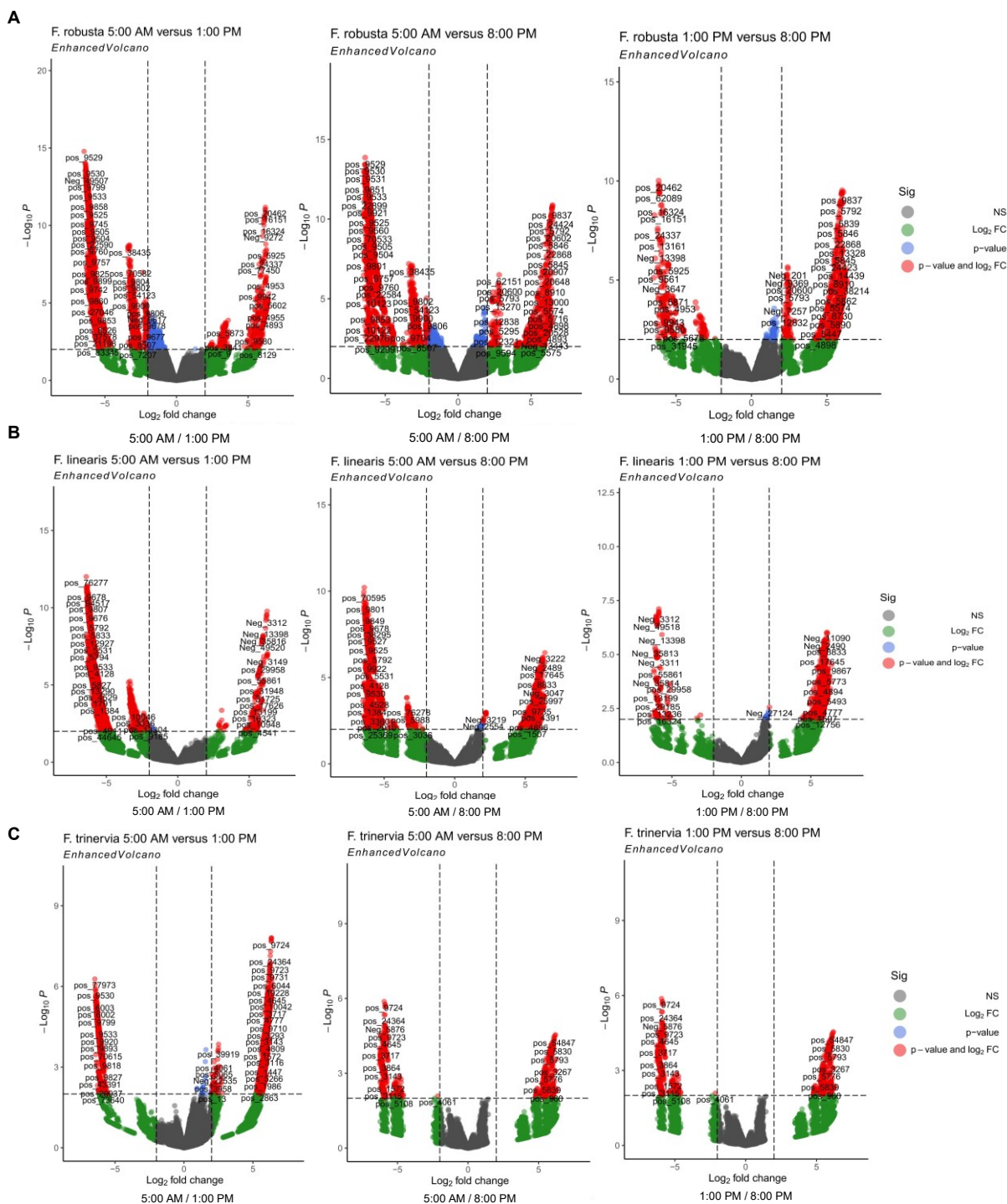


Figure 11. Abundance of peaks in *F. robusta* and *F. linearis* show similar trends, but different from *F. trinervia*. Comparison of peak areas between sampling points. Vertical dashed lines indicate the cutoff of the log₂FC |2| and the horizontal line indicates the cutoff for the statistical significance (*p* value <0.05). Gray dots indicate peaks not significantly different between the two time points, blue dots show peaks with *p* value significance, green dots show peaks above the fold change threshold and red dots show peaks with significance differences and above the fold change threshold. A generalized logarithm transformation (Parsons et al, 2007) was applied to the data before analysis. The *p* values are based on Wald significance test. A) *Flaveria robusta* B) *Flaveria linearis* C) *Flaveria trinervia*.

The lists of peaks that were significantly different between sampling points were filtered further to reduce the number of false positives ($FDR < 0.01$) (Supplementary Table 14). The m/z and retention time of these peaks were annotated against a reference metabolome from *A. thaliana* to search for enriched pathways.

In the 1:00 pm samples of *F. robusta*, compounds involved in the biosynthesis of certain amino acids, zeatin and terpenoids as well as pyrimidine and alpha-linolenic acid metabolism were enriched in comparison to the earlier samples (Fig. 12a). At 8:00 pm pyrimidine metabolism remained active, but porphyrin and chlorophyll metabolism was less active than at 5:00 am (Fig. 12b). Although the metabolite abundances peaked at the same time as in *F. robusta*, the enriched pathways were different in *F. linearis*. Specifically, at 1:00 pm compounds related to amino acid were enriched similar to in *F. robusta*, but also sugar metabolism pathways, as well as purine and glutathione metabolism were more enriched than at 5:00 am (Fig. 12c). For the 8:00 pm sampling, none of the annotated compounds suggested enrichment of specific pathways. In the case of *F. trinervia*, only the comparison between the earliest time point and 1:00 pm had significantly different peaks ($FDR < 0.01$), however none of them got annotated and therefore no pathway enrichment could be determined.

To compare whether *F. robusta* and *F. trinervia*, which have the most contrasting photosynthesis strategies, differed in the enriched pathways at each sampling point, the same procedure was followed. In the early morning, pathways involving the synthesis of terpenoids and aromatic amino acids were enriched in *F. robusta*, while purine, starch and sugar metabolism were enriched in *F. trinervia*. At 1:00 pm, diverse pathways were highly enriched in *F. robusta* involving metabolism of sugars, purines, pyrimidines, amino acids and a variety of other molecules. (Supplementary Fig. 9 and 10).

Overall, the metabolomic analyses show that the composition of the leaf apoplast fluid changes throughout the day in species-specific ways. *F. robusta* and *F. linearis* have a more active metabolism towards noon, but the specifically enriched pathways differ. The maximum activity of *F. trinervia*, on the other hand, seems to rather be in the early morning or night.

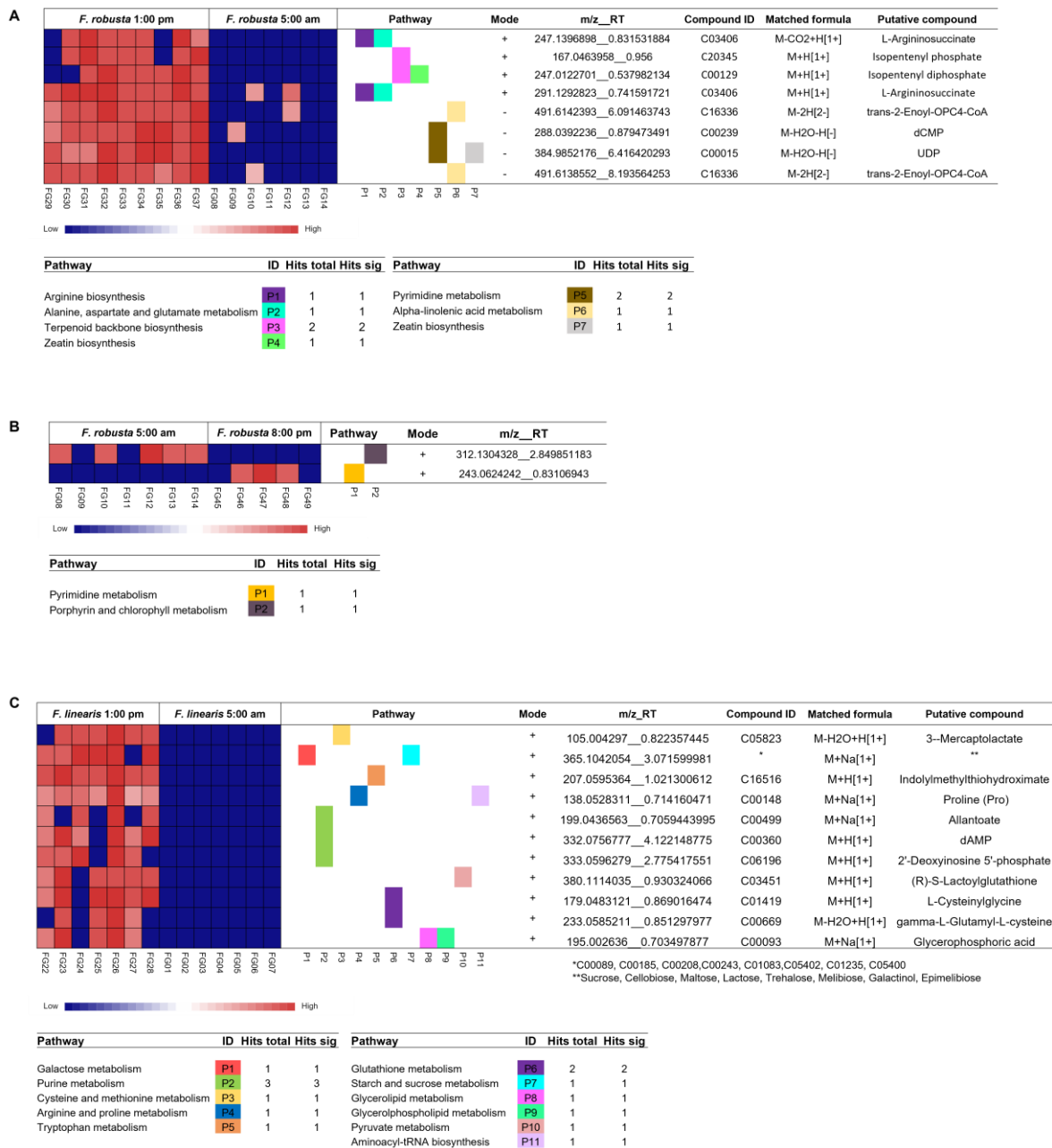


Figure 12. Different metabolic pathways were enriched in different species at different sampling times. Metabolic pathways enriched at timepoints in either *Flaveria robusta* samples (Panels A and B) or *F. linearis* samples (panel C). The heatmaps show the abundance of each compound (rows) in each sample (columns) based on Ward distance. The ionization mode, m/z value, retention time in minutes and the closest annotation is shown for each compound to the right side of the heatmap. The compounds are colored according to the pathway (P) they were related to. The table beneath each heatmap shows the pathways that were enriched with the corresponding number of hits.

Bacterial communities in the leaf apoplast may vary throughout the day

The bacterial community was next characterized with 16S rRNA gene amplicon sequencing in the same samples that were analyzed with untargeted metabolomics. On average, 73% of the reads corresponded to bacteria, but the total number of sequencing reads per sample was highly variable, with several samples having less than 100 reads. The low number of reads in these samples could have been due to low bacterial load. Unfortunately, however, most of these low-read samples were amplified on a plate together, including all morning samples from the three species and 1:00 pm samples from *F. linearis* (Supplementary Table 15). The negative controls (no template or DNA extraction buffer) had < 300 bacterial reads on average, while positive controls had about 40,000 reads on average (Supplementary Fig 11), however these were not amplified in the plate with the morning samples. Thus, we cannot exclude a batch amplification problem and to remain conservative pending further experiments, the diversity analyses were carried out after excluding these samples. This means that only comparisons between *F. robusta* and *F. trinervia* at 1:00 pm and 8:00 pm are possible.

Overall, the sampling time exerted a weakly detectable influence on the alpha diversity. *F. robusta* communities at 8:00 pm appeared richer (Chao1) but less even (Shannon index) than 1:00 pm samples, while in *F. trinervia*, 1:00 pm communities appeared richer and more even. None of these differences however, were significant with a conservative Wilcox test. For beta diversity analyses, samples were rarefied at 1200 reads. Sampling time explained about 50% of the bacterial community variation in *F. robusta* and about 40% in *F. trinervia*, but the differences between times were not significant (Fig. 13b and 13c, PERMANOVA p value = 0.2381 and p value = 0.67, respectively), probably due to the low number of samples at each sampling point. The two species shared most of the high-abundant genera, with a few exceptions (Supplementary Fig. 12 and 13). A differential analysis showed that in *F. robusta*, the abundance of several genera like *Cellvibrio*, *Plantibacter* and *Tychonema*, a Cyanobacteria, decreased towards the night, while Bacillales tended to increase, but were not significantly different (Supplementary Fig. 12). In *F. trinervia*, the genus *Aquabacterium* was more abundant at night, while *Paeniglutamicibacter* was more abundant at 1:00 pm (Supplementary Fig. 13). Although it was not possible to compare these results against the morning samples, it seems like the enrichment of taxa in each *Flaveria* species corresponded with increasing metabolite complexity of the AFW. In *F. robusta* the metabolic diversity peaked at 1:00 pm and the alpha diversity seemed to do so at 8:00 pm. While in *F. trinervia*, the microbial diversity seemed to peak at 1:00 pm, following the high metabolic diversity from the morning. This may suggest a possible correlation between microbial diversity and the chemical environment of the leaf but will require further testing.

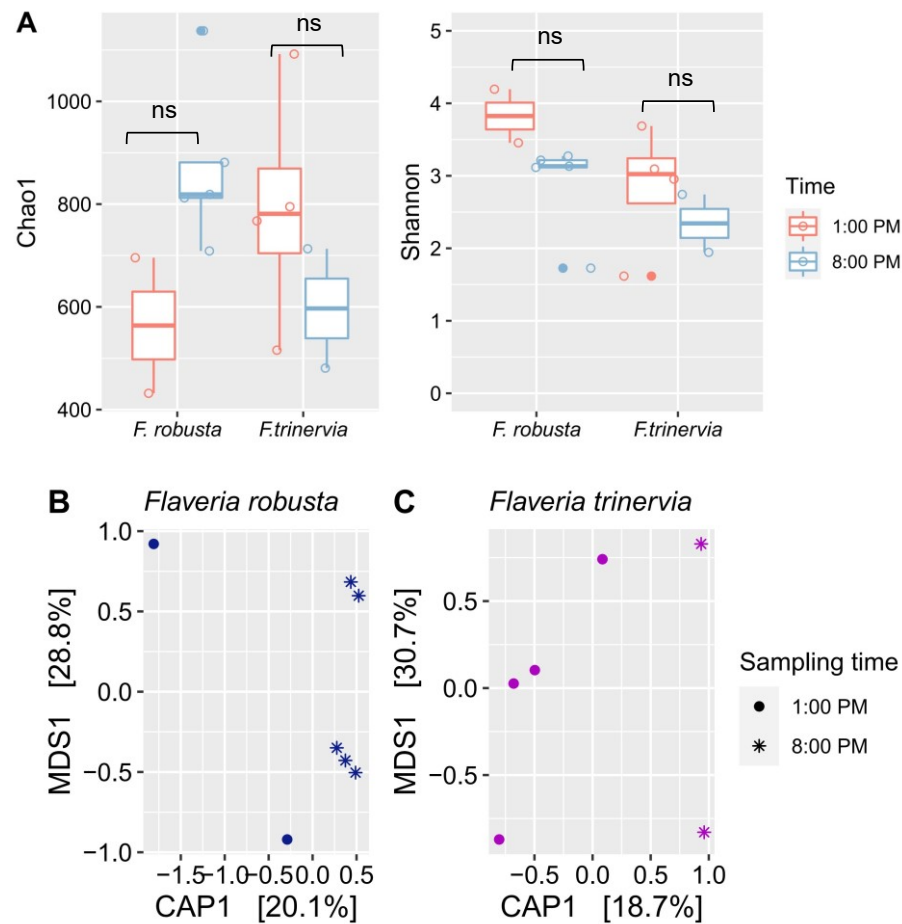


Figure 13. Bacterial diversity in the apoplast fluid wash of *Flaveria* leaves may have changed throughout the day. A) Alpha diversity indices Chao1 and Shannon of leaf apoplast samples of two *Flaveria* species sampled at different times of the day. Significance values are based on Wilcox test (ns: p value > 0.05, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$). B and C) Principle coordinate analysis of Bray-Curtis distances, constrained by sampling time. The left panel shows *Flaveria robusta* samples and the right panel, *F. trinervia* samples.

Chapter 3.3: Leaf bacterial strain-level diversity is supported by niche separation during cross-feeding

Leaf commensals are found in very low cell numbers in the apoplast (Chapter 3.2 and Regalado et al. 2020). Since they lack many of the strategies found in pathogens to hijack nutrients from the leaf, they probably persist by scavenging what is available in the colonizable space, i.e., the apoplast. Nutrient availability in the apoplast seems to vary drastically between hosts and over time (Chapters 3.1 and 3.2); thus, microbes in the apoplast could increase their chances for survival by depending on other microbes, at least occasionally. Here, this hypothesis was addressed by enriching leaf bacterial communities from two *Flaveria* species with contrasting apoplast metabolomic profiles (Chapter 3.1) in two nutrient regimes and studying the interdependences that emerged. These nutrient regimes not only reflect possible scenarios in the leaf, but also are carbon sources that are known to be involved in bacterial nutrition and virulence (Rico and Preston, 2008; Yu et al., 2013; O'Leary et al., 2016).

Nutrient and taxonomic diversity shape microbiome function in *Flaveria* leaf enrichments

Bacteria derived from leaves of *F. trinervia* and *F. robusta* were enriched in a simple minimal media with sucrose as the base carbon source and either no amino acids (S-CA) or addition of casamino acids (S+CA) at low levels similar to those previously reported in the apoplast (Lohaus et al., 2001) (Fig. 14a). The final OD was used as a measure of productivity (function) of the enrichments (Fig. 14b). In the S-CA condition where sucrose was the only carbon source, the *F. robusta* enrichment was slightly more productive than the *F. trinervia* enrichment. Additionally, enrichments from *F. robusta* tended to increase their productivity when the potential carbon and nitrogen sources were richer (S+CA), but those from *F. trinervia* did not.

To have a broader picture of their taxonomic composition, amplicon sequencing was conducted on the enriched communities (Fig. 14d). Despite diverse bacteria characterized in the original leaves using a culture-dependent approach (Fig. 14b), enrichments selected very few taxa, which was not affected by additional resource availability in the form of amino acids. *F. trinervia* enrichments were fully dominated by *Pantoea* sp. (Fig. 14d). *F. robusta* enrichments also contained a high prevalence of *Pantoea*, but in combination with *Pseudomonas* (Fig. 14d). To investigate why the productivity differed between the enrichments, 25 random bacterial isolates were collected from the *F. robusta* and *F. trinervia* S-CA and S+CA communities (isolates designated as Fr-CA_X, Fr+CA_X, Ft-CA_X or Ft+CA_X, where X is a unique isolate number), identified taxonomically (Fig. 14e) and tested for growth of each individual in S-CA and S+CA medium (Supplementary Table 16). The isolates included multiple individuals of both *Pseudomonas* and *Pantoea*, the only genera that were found using 16S amplicon data. In the S-CA community, the fraction of *Pantoea* was higher than in S+CA (67% vs. 21%, respectively, Fig. 14e). All tested *Pantoea* isolates could grow on sucrose or amino acids as a sole carbon source (Supplementary Table 16). The fact that productivity did not increase in *F. trinervia* enrichments dominated by *Pantoea* when casamino acids were added suggests that another resource must have limited additional growth. All tested *Pseudomonas*

isolates could use amino acids as a sole carbon source (Supplementary Table 16). Thus, the increased productivity in the *F. robusta* S+CA enrichment suggests that they could make use of the additional resources. *Pseudomonas* isolates did not use sucrose as a sole carbon source even though they were present in the S-CA (sucrose-only) enrichment. Therefore, *Pseudomonas* in this enrichment may have been somehow dependent on *Pantoea*, which could explain the higher productivity in *F. robusta* enrichments compared to *F. trinervia* when sucrose was the only carbon source.

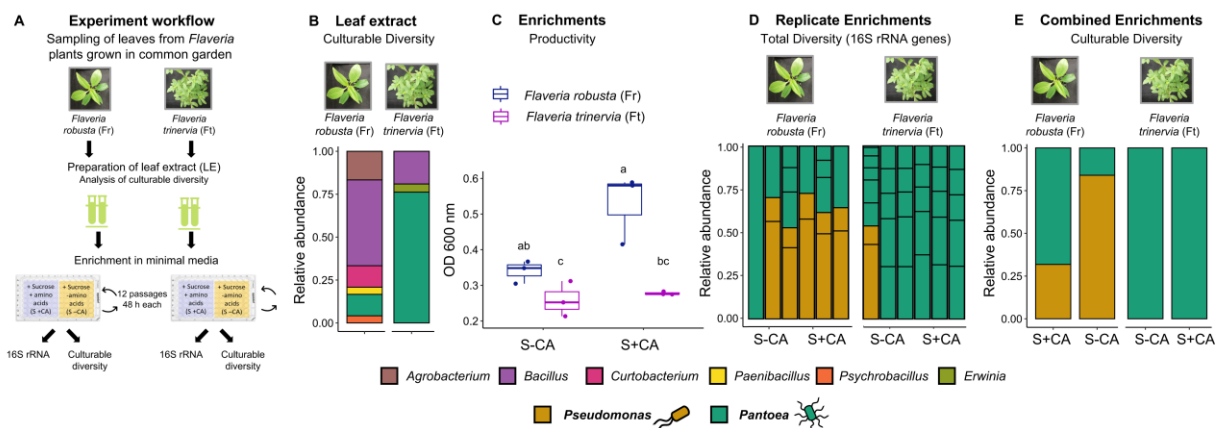


Figure 14. Bacterial community enrichments from *F. robusta* leaves show higher productivity and more functional responsiveness to nutrient diversity than *F. trinervia*. A) Enrichment of leaf bacterial communities in two different minimal media. B) Culturable diversity in the leaf extracts used for enrichment from *Flaveria robusta* and *F. trinervia*, based on 16S rRNA Sanger sequencing; n=25 isolates. C) Productivity (final optical density at OD_{600nm}) of the enrichments from each species. Different letters indicate significantly different values (Kruskal-Wallis test with FDR adjustment <0.05, N=3) D) Diversity obtained in each enrichment replicate after 12 passages based on amplicon sequencing of 16S rRNA gene. E) Culturable diversity obtained in each enrichment after 12 passages based on Sanger sequencing of 16S rRNA gene (Three replicates were combined for isolation; n=25 isolates). The taxonomy color legend corresponds to panels B, D and E.

Cross-feeding on diverse metabolites sustains *Pseudomonas* in the absence of a primary carbon source.

One possibility how *Pseudomonas* isolates survived in the S-CA enrichment is that they were auxotrophs and could only grow on sucrose when amino acids were available in the environment (either in S+CA or provided from *Pantoea* in S-CA). However, all tested *Pseudomonas* isolates could grow on glucose without amino acid supplementation (Supplementary Table 16), suggesting that they were not strict auxotrophs. Therefore, it was tested whether *Pantoea* produced other metabolic by-products that *Pseudomonas* could grow on (Fig. 15a). Indeed, all *Pseudomonas* isolates from the sucrose-only enrichment (*Pseudomonas* Fr-CA) could grow on spent media of a *Pantoea* isolated from the same enrichment (*Pantoea* Fr-CA_6, Fig. 15b). To determine what *Pseudomonas* consumed, untargeted metabolomics was performed on the spent media before and after growth of three *Pseudomonas* isolates. Metabolite peaks were considered consumed if their area was strongly reduced ($\log_2FC < -2$, FDR < 0.05). With this strict cutoff, uptake of between 23-25 metabolites by each strain was detected, including 2-hydroxybutyric acid, hypoxanthine,

spermidine and, in one of the isolates, alanine (Supplementary Table 17). This indicates that *Pseudomonas* is capable of a complex metabolic dependency on *Pantoea* in the sucrose-only enrichments. Trace levels of glucose and/or fructose were detected in the *Pantoea* spent medium (hexoses are indistinguishable with this method), but they were not significantly taken up ($\log_2FC = -0.85$, p value = 0.36 on average, not shown).

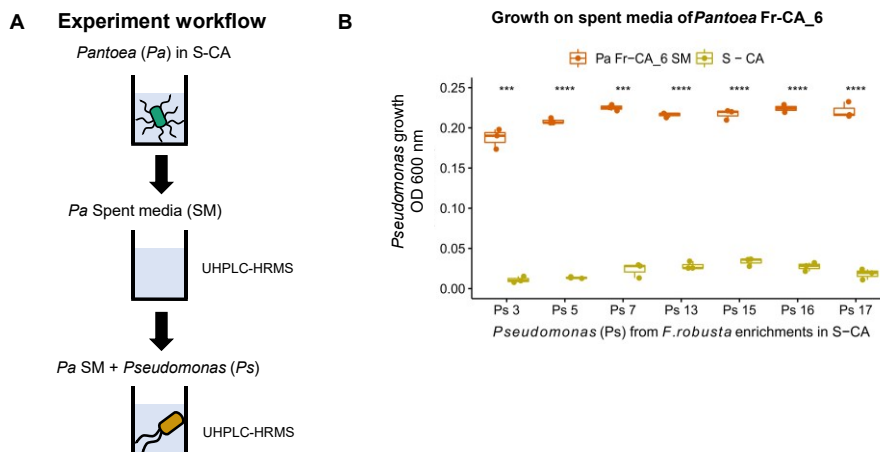
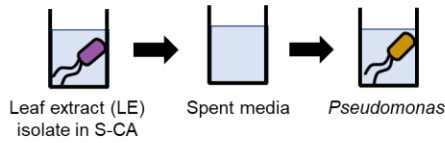


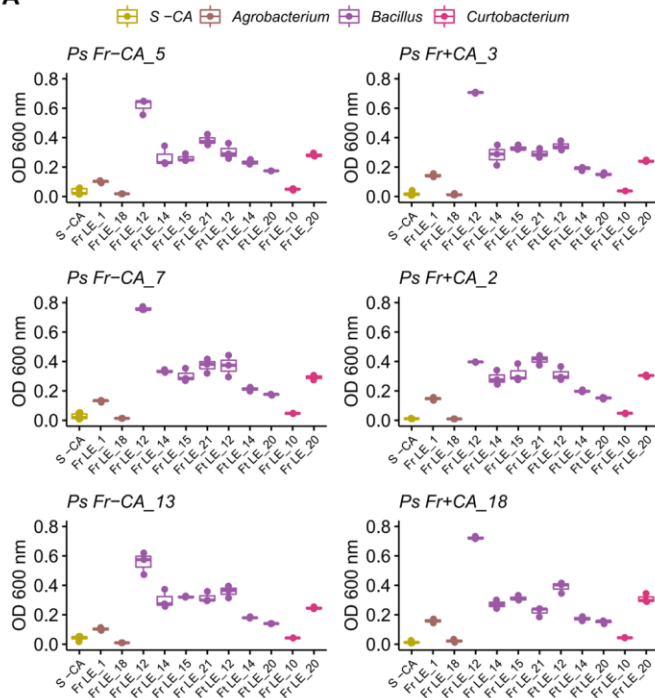
Figure 15. *Pseudomonas* isolates could persist in the *F. robusta* S-CA enrichment by cross-feeding metabolites from *Pantoea* isolates. A) Procedure to obtain *Pantoea* (*Pa*) spent media and evaluate its potential to support *Pseudomonas* (*Ps*) growth. B) Growth of *Ps* isolates from the Fr-CA enrichment on *Pa* Fr-CA_6 spent medium. Significance values are based on a t-test comparing growth in S-CA and in *Pa* Fr-CA_6 spent medium (ns: p value > 0.05, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, $N=3$).

Next, it was tested whether the cross-feeding was specific to *Pantoea* or if other isolates from the original leaf extracts (FrLE and FtLE) could feed *Pseudomonas* too. Indeed, various strains could support *Pseudomonas* growth, but interestingly, it was not genus dependent. In *Agrobacterium* and *Curtobacterium*, only one out of two tested strains induced growth. While in *Bacillus*, all strains could cross-feed, but to different degrees. The growth on *Bacillus* FrLE_12 spent medium, for example, was especially high, even compared to growth on *Pantoea* spent medium (Fig. 16a). Metabolomic analyses suggested that *Pseudomonas* may utilize a diverse array of metabolites from this isolate ($\log_2FC < -2$, $FDR < 0.05$) with little or no overlap with compounds utilized from *Pantoea* spent media (Supplementary Table 18). None of the taxa with potential to cross-feed *Pseudomonas* were found in the enrichments, probably due to their lower growth rate in sucrose, when compared to *Pantoea* (Fig. 16b and 16c). Overall, these results suggest that when competition for sucrose dominates dynamics, *Pseudomonas* will be limited in cross-feeding partners but otherwise may be able to persist by feeding on exudates of diverse leaf bacteria.

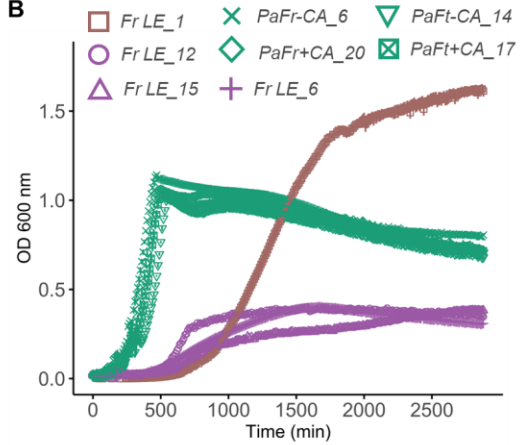
Experiment workflow



A



B



C

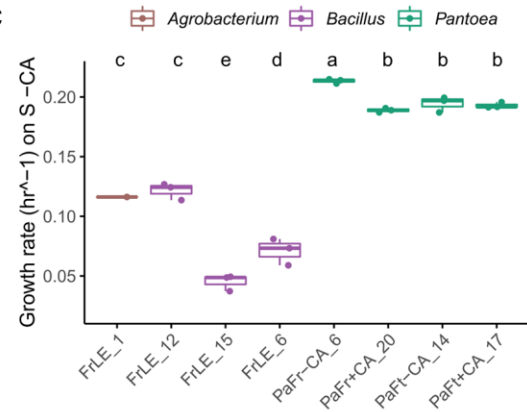


Figure 16. Diverse isolates could cross-feed *Pseudomonas*, but they were probably outcompeted on sucrose by *Pantoea* (Pa). A) Growth of several *Pseudomonas* (Ps) isolates on spent media from different *Agrobacterium*, *Bacillus* and *Curtobacterium* isolates obtained from the leaf extracts (LE) of *F. robusta* and *F. trinervia*. B) Growth in S-CA medium of different *Agrobacterium* and *Bacillus* isolates versus the *Pantoea* isolates. C) Growth rates in S-CA medium of the same isolates as in panel B. Different letters indicate statistical differences after a Tukey test (p value <0.05).

Diverse plant-derived *Pseudomonas* strains cross-feed in parallel from *Pantoea*

In the S-CA enrichment, *Pseudomonas* survived exclusively by cross-feeding on diverse resources from *Pantoea* exudates, but in the S+CA enrichment, it would have been able to either cross-feed or utilize amino acids or both. Therefore, it was hypothesized that multiple *Pseudomonas* strains may exist to optimally utilize this niche diversity. To investigate this, four *Pseudomonas* isolates (one from the S-CA enrichment and three from the S+CA enrichment) were grown in spent media of *Pantoea* Fr-CA_6. Interestingly, two distinct growth phenotypes were observed across the isolates. *Pseudomonas* Fr-CA_5 and Fr+CA_18 grew earlier and reached maximum OD_{600nm} faster compared to *Pseudomonas* Fr+CA_3 and Fr+CA_2 (Fig. 17a). A correlated phenotype was observed in R2A medium, where only the slow cross-feeders switched to more rapid growth when supplemented with vitamins that were present in the enrichment medium (Fig. 17b). Thus, it was predicted that in the S-CA enrichment, where cross-feeding was required, faster cross-feeders would outcompete slower ones to dominate the mix, while the more diverse nutrient conditions in the S+CA enrichment would result in a more balanced mix. However, when checking the phenotype in R2A or R2A supplemented with vitamins across all *Pseudomonas* isolates, it was observed that the strains were in similar ratios in both enrichments (Fig. 17c, $\chi^2 p=1$).

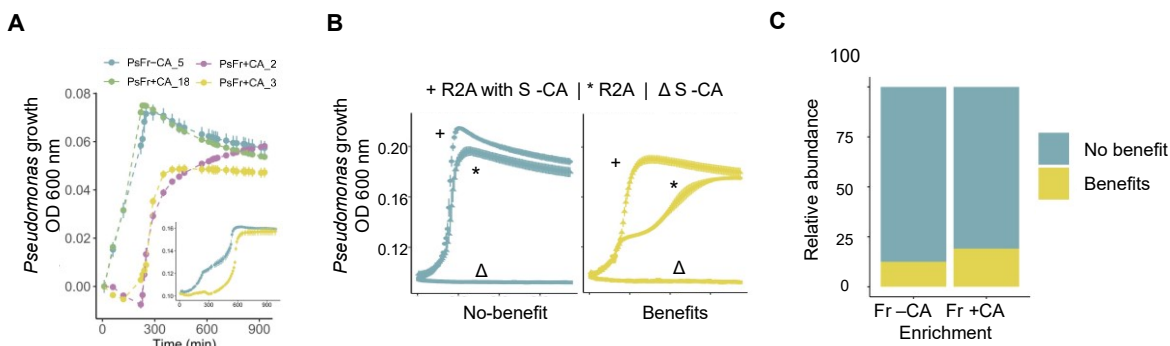


Figure 17. Both enrichments harbored metabolically distinct *Pseudomonas* isolates. A) Growth curve of two fast (*Ps* Fr-CA_5 and *Ps* Fr+CA_18) and two slow (*Ps* Fr+CA_2 and *Ps* Fr+CA_3) growing *Pseudomonas* isolates on *Pa* Fr-CA_6 spent medium. The insert shows the growth of *Ps* Fr-CA_5 and *Ps* Fr+CA_3 in a separate experiment with more data points. B) Two different phenotypes observed across the *Pseudomonas* isolates when nutrients from the S-CA medium were added to their growth on R2A broth. Here *Ps* Fr-CA_5 and *Ps* Fr+CA_3 are shown as an example. C) Distribution of the two growth phenotypes across the *Pseudomonas* isolates of each enrichment. In panels A and B each data point shows the average of three replicates and its corresponding standard error.

According to classical competition theory, if two strains compete for the same resources with different growth rates, one should completely outcompete the other (Hardin, 1960). Thus, in this case the *Pseudomonas* strains should be occupying different niches in parallel. To test whether faster and slower-growing strains could simultaneously cross-feed from *Pantoea*, the slow and fast cross-feeder (*Pseudomonas* Fr+CA_3 and *Pseudomonas* Fr-CA_5) were labeled with fluorescent tags, which did not alter their growth patterns on *Pantoea* spent media (Supplementary Fig 14).

The tagged strains were combined and grown directly with *Pantoea* Fr-CA_6 in sucrose-only media (Fig. 18a). Although the fast cross-feeder was inoculated with only half the cells as the slow cross-feeder, they

were approximately equivalent within 48 hours (Fig. 18c) and neither strain overtook the other after four 48h or eight 24h passages (Fig. 18c, 18d, raw signals provided in Supplementary Fig. 15). In the 24h passages, the slow grower did reach higher levels with *Pantoea* alone than it did together with the fast grower (Fig. 18b), consistent with a hypothesis of partially overlapping niches. Total biomass was higher when *Pantoea* was with *Pseudomonas* than alone (Supplementary Fig. 15), similar to productivity observations in the original enrichments (Fig. 14c). While cross feeding from *Pantoea* allowed survival of both strains when sucrose was the only carbon source, it was not necessarily required in the S+CA media, where both strains could persist without *Pantoea* (Supplementary Fig. 16).

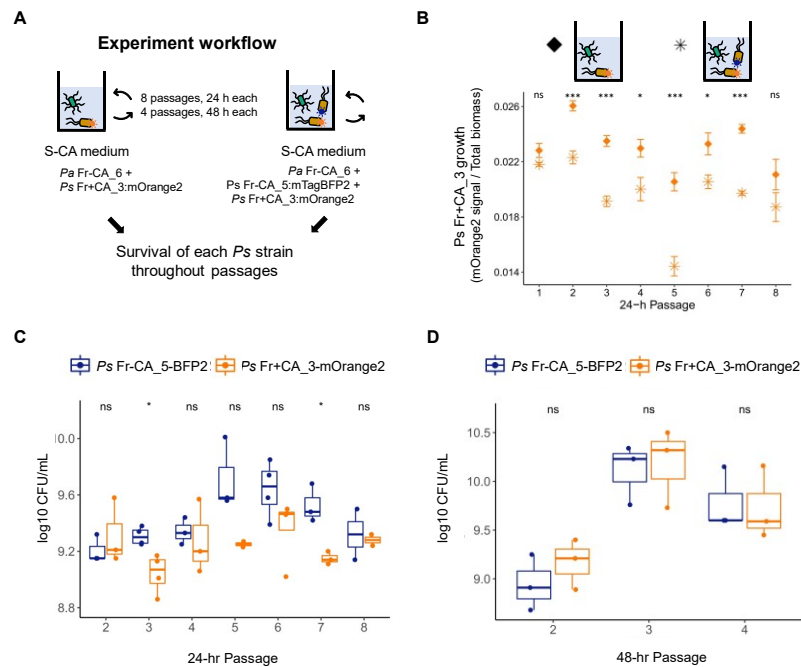


Figure 18. Different *Pseudomonas* (*Ps*) isolates can persist together in S-CA with *Pantoea* (*Pa*) Fr-CA_6. A) Passaging of tagged *Ps* strains along with *Pa* Fr-CA_6 on S-CA medium. B) Signal of *Ps* Fr+CA_3:mOrange2 fluorescence normalized by the biomass signal in either co-culture with *Pa* Fr-CA_6 or in the full community with *Pa* Fr-CA_6 and *Ps* Fr-CA_5:mTagBFP2. C) CFU counts of each tagged *Pseudomonas* in the full community passaged every 24 hours D) CFU counts of each tagged *Pseudomonas* in the full community passaged every 48 hours. N=4 in panels B and C and D.

Genetic analysis also supported partially differential niches between *Pseudomonas* strains. Draft genomes of two fast-growing (Fr-CA_5 and Fr+CA_18) and one slow-growing isolate (Fr+CA_3) were generated. All three strains were identified as *P. siliginis* and shared 99.98 to 99.99% average nucleotide identity between them (Supplementary Table 19). Alignment to the closest *Pseudomonas* reference genome showed that all strains shared about 97,000 sequence variants (98.9% of total variants) and that only a few hundred were unique to each genome (Fig. 19). Among the distinctive variants, several with disruptive potential were located in genes related to nutrient acquisition, including porins, carboxylate, sugar and peptide transport proteins. The resulting gene lists are provided in Supplementary Table 20. This could suggest a scenario

of niche differentiation whereby the fast-growing strains would dominate some resources, limiting growth of the slow-growers when growing together, but still enabling slow-growers to persist on unique resources.

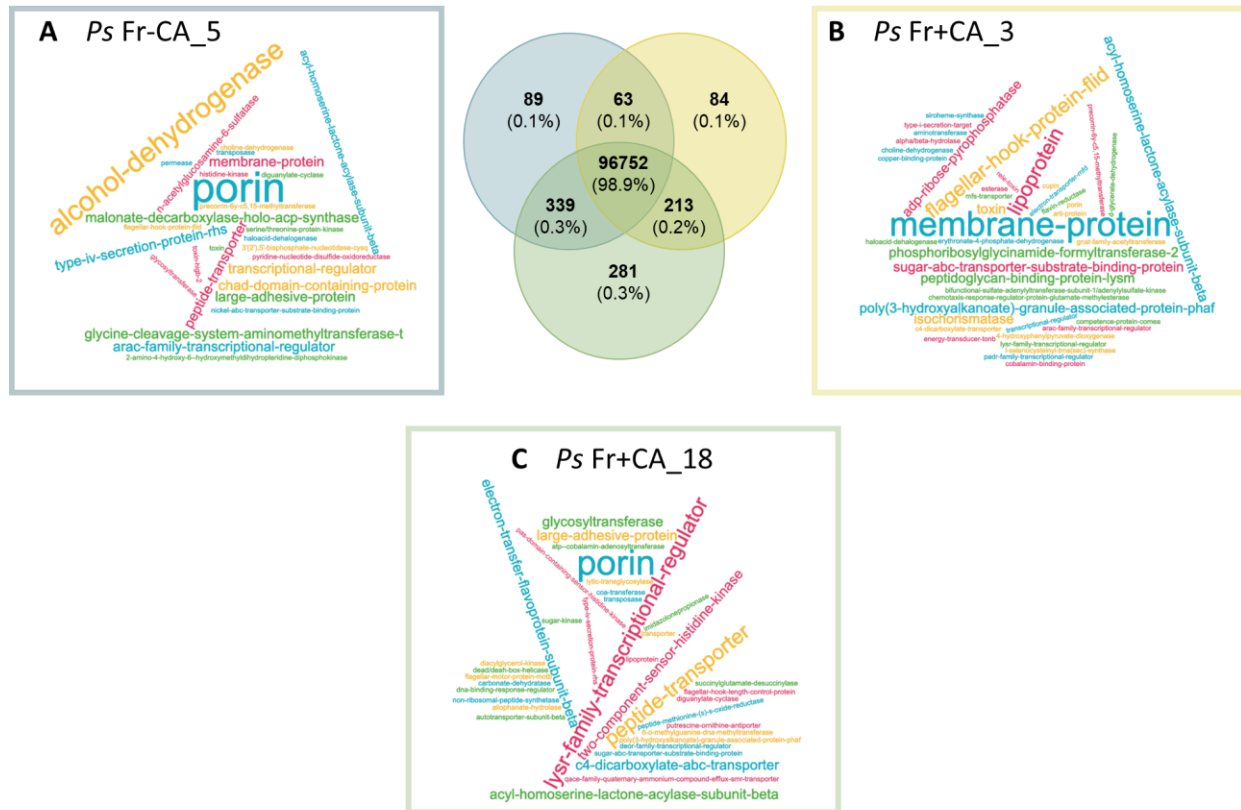


Figure 19. *Pseudomonas* (*Ps*) isolates with distinct metabolic phenotypes have extremely similar genomes. The Venn Diagram in the center shows the number of single nucleotide variants across three *Pseudomonas* isolates after annotation with the reference genome of *P. siliginis* D26. Unique SNPs for each isolate with potentially disruptive mutations are shown for each *Pseudomonas* strain. The word size represents the relative frequency of disruptive SNPs in a particular gene product.

Niche differentiation among distinct *Pseudomonas* strains maintains diversity during cross-feeding

To test whether and how *Pseudomonas* strains exploit complementary niches, the two labeled *Pseudomonas* (Fr-CA_5 and Fr+CA_3) and two additional fast and slow growing isolates (*Pseudomonas* Fr+CA_18 and *Pseudomonas* Fr+CA_2, respectively) were cultured individually and in all possible pairs in *Pantoea* Fr-CA_6 spent media (Fig. 20a) and evaluated for metabolite uptake (niches) and growth patterns. The pattern of metabolite uptake by the strains growing alone was consistent with partially overlapping niches. Several significantly taken up metabolite peaks were shared by all four isolates (common niches)

and each also significantly depleted unique sets of metabolite peaks (unique niches) (Fig. 20b, 20c and Supplementary Table 21).

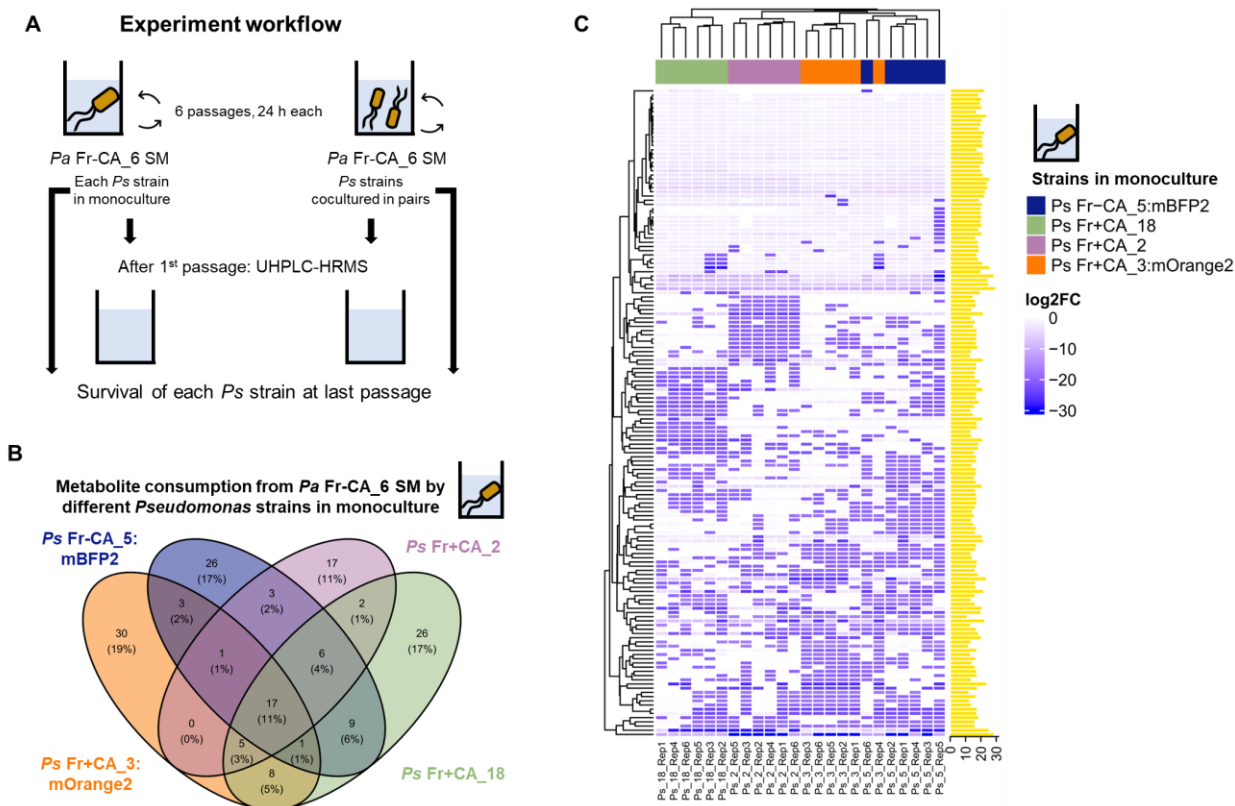


Figure 20. Different *Pseudomonas* (*Ps*) isolates had specialized niches when feeding from *Pantoea* (*Pa*) Fr-CA_6. A) Each *Pseudomonas* strain was passaged either alone or in pairs with each other on *Pa* Fr-CA_6 spent medium. B and C) Taken up metabolites from *Pa* Fr-CA_6 spent media by each *Ps* strain when grown in monoculture. Only metabolites with significant uptake are shown ($\log_2 \text{FC} < -2$, $\text{FDR} < 0.05$). In the heatmap, each column represents a different repetition ($N=6$ for each strain). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (\log_{10}) of the peak in the spent media of *Pantoea* Fr-CA_6 before growth of the *Ps*. The information of each peak ionization mode, retention time, m/z value and putative annotation) is given in Supplementary Table 21. The order of the peaks shown in the heatmap matches exactly the order of the peaks in the Table.

When grown together (Supplementary Fig. 17-22), most niches that the two strains shared were still utilized. Around half of the niches that were unique to each strain individually were still utilized when they were grown together, further supporting partially overlapping niches. However, new compounds were also taken up, suggesting some cooperative niche exploration. Similar to previous results (Supplementary Table 17), some peaks that could be annotated matched to purines and pyrimidines and to amino acids Trp and Ile/Leu in some combinations (Supplementary Table 21). Fluorescence signals and CFU counts showed that after 6 24-hr growth cycles, growth of tagged strains was reduced in co-cultures compared to monocultures (Fig. 21 and Supplementary Fig. 23) but all strains successfully persisted in co-cultures, further supporting a lack of competitive exclusion.

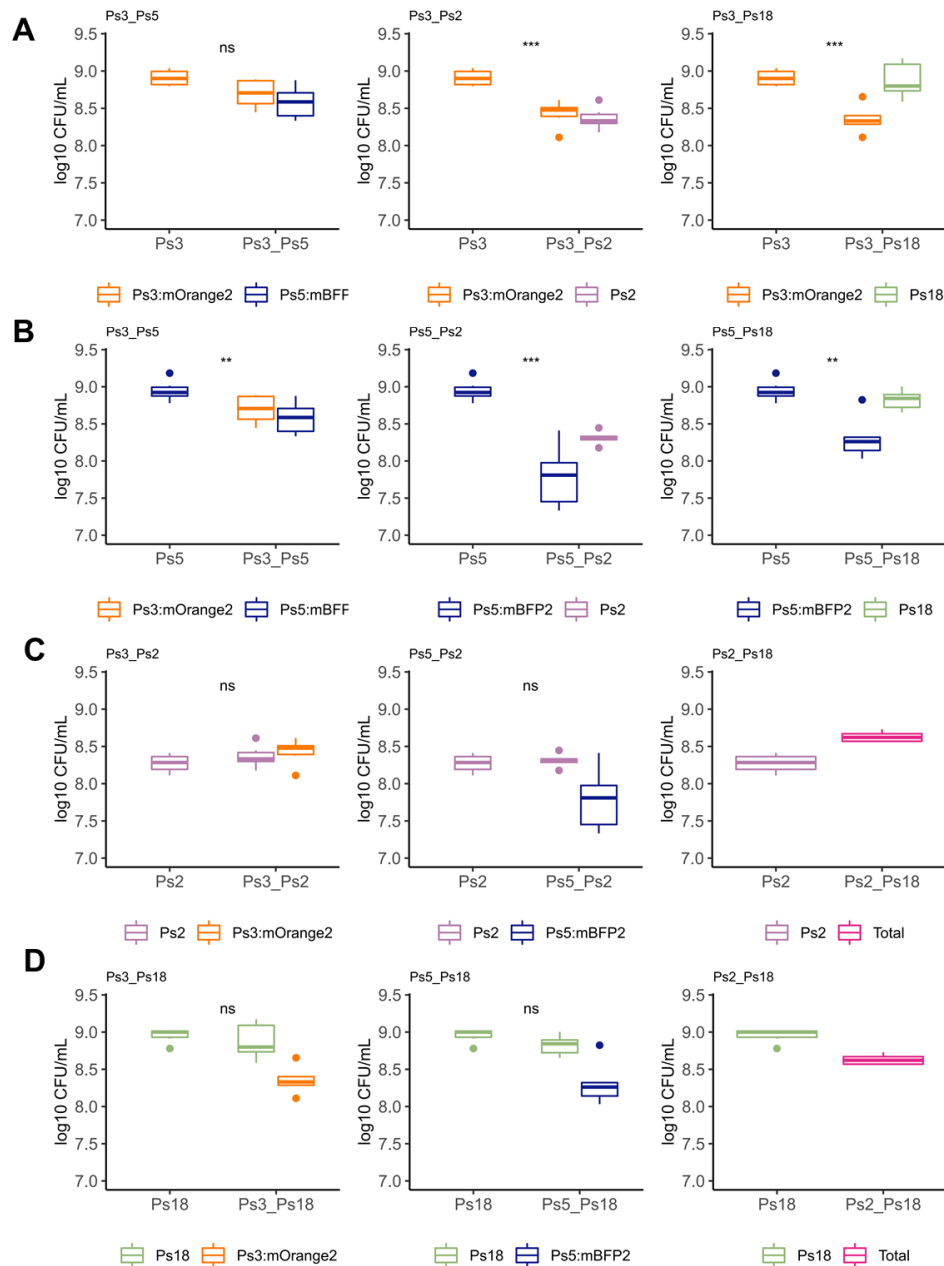


Figure 21. *Pseudomonas* strains with different growth patterns can coexist while cross-feeding from *Pantoea* Fr-CA_6 spent media. Colony counts of four *Pseudomonas* strains growing either in monocultures or in co-culture with each of the other strains. Counts were taken after six passages on *Pantoea* Fr-CA_6 spent media. A) *P. siliginis* Fr+CA_3:mOrange2 counts. B) *P. siliginis* Fr+CA_5:mBFP2 counts. C) *Pseudomonas* Fr+CA_2 counts. D) *P. siliginis* Fr+CA_18 counts. In each case a t-test was performed between the counts in the monoculture and the counts in the co-culture (N=6, ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$). *Pseudomonas* Fr+CA_2 and *P. siliginis* Fr+CA_18 colonies are not distinguishable, therefore only the total counts are reported and no significance test is shown.

Notably, *Pseudomonas* Fr+CA_2 alone or in co-culture had lower total growth (CFUs) than other combinations (Fig. 21). This was however, not supported by a decrease in overall biomass (Supplementary Fig. 23c) which suggests this strain may form cell aggregates that would affect OD_{600nm} readings. To check whether the lower cell counts were due to inhibition, spent media of Fr+CA_2 was collected after growth in the spent media of *Pantoea* Fr-CA_6, supplemented with R2A broth and used as a culture media for the other *Pseudomonas* isolates (Fig. 22a). In all strains, growth was promoted compared to growth in R2A broth alone (Fig. 22b), indicating Fr+CA_2 was not inhibitory and can even provide substrates to other strains. A likely explanation is that Fr+CA_2 inefficiently used some limiting resource in the spent media thereby reducing growth of itself and co-colonizers. Regardless, this further highlights the extent of *Pseudomonas* commensal strain diversity.

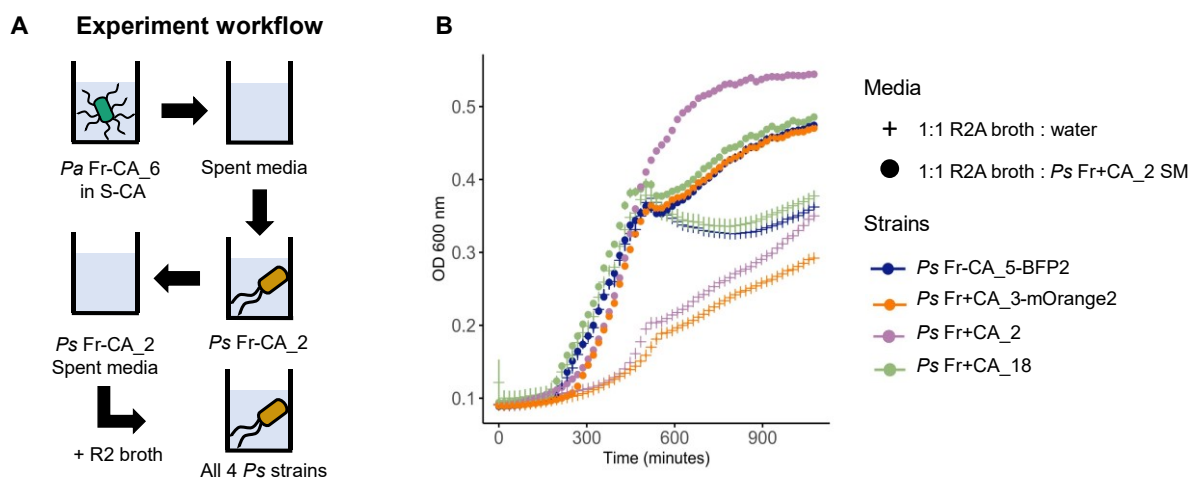


Figure 22. Growth of *Pseudomonas* (*Ps*) strains is not inhibited by *Ps* Fr+CA_2. A) Production of spent media from *Ps* Fr+CA_2 after growth on the spent media of *Pa* Fr-CA_6 and testing its inhibitory effect on *Ps* growth. B) Growth of *Pseudomonas* (*Ps*) strains on the spent media of *Ps* Fr+CA_2 or in diluted R2A broth as a control. Each data point shows the average of three replicates and its corresponding standard error.

Taken together, this chapter showed that *Pseudomonas* can survive even when they could not consume the main carbon source by cross-feeding from different possible partners. It also shows that their diversity is relevant for accessing multiple niches while cross-feeding.

Chapter 3.4 Cross-feeding is selected for in a leaf nutrient-dependent manner

Previous studies have shown that nutrient limitation may lead to the appearance of metabolic interactions between microbes (Chapter 3.3 and Hoek et al., 2016; D'Souza et al., 2018; Zimmermann et al., 2020). In the apoplast of leaves, where the nutrient availability is constantly oscillating (Chapter 3.2 and Lohaus et al., 2001), this might be especially relevant for the assembly of the microbial community. *Flaveria robusta*-associated bacterial communities are well-suited to cross-feed, which can enable survival in a nutrient-limited environment (Chapter 3.3). In opposition to competition theory, in which the fastest grower outcompetes the slower (Hardin, 1960), niche diversification allowed the co-existence of several *Pseudomonas* strains which were genetically very similar but had different growth patterns. Here, the question of whether this diversification is likely to arise in response to different nutrient and cross-feeding contexts was addressed. First, the cross-feeding potential of different *Pantoea* strains was compared. Next, the dependence between two *Pseudomonas* and two *Pantoea* strains for colonizing plants with contrasting amino acid availability was tested. Finally, the question of whether the nutrient and strain-pair context shapes adaptation of cross-feeding interactions was addressed.

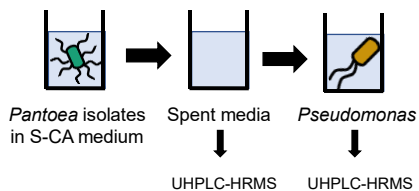
Cross-feeding potential emerged as a convergent trait in bacterial isolates from *F. robusta*

Enriching *F. robusta* and *F. trinervia* leaf bacteria over several generations on base media with or without amino acids led to communities dominated by *Pantoea* and *Pseudomonas* in *F. robusta* and only *Pantoea* in *F. trinervia*. *Pantoea* isolated from *F. robusta* was shown to cross-feed *Pseudomonas* with different metabolites (Chapter 3.3). Since no *Pseudomonas* were enriched in *F. trinervia*, it was hypothesized that *Pantoea* isolates from *F. trinervia* were not as efficient cross-feeders. To test this, spent media from four *Pantoea* isolates (one from each enrichment) was used to grow several *Pseudomonas* strains (Fig. 23a). The spent media was subjected to metabolomic analysis before and after growth of *Pseudomonas*. Interestingly, *Pantoea* from *F. robusta* fed *Pseudomonas* best (Fr-CA_6 better than Fr+CA_20) while both *Pantoea* from *F. trinervia* hardly supported *Pseudomonas* or not at all (Fig. 23b).

To elucidate whether the difference in growth was due to inhibitory compounds from *F. trinervia* isolates, the spent media of *Pantoea* Ft-CA_14 and Ft+CA_17 were used to culture several *Pseudomonas* strains either in full strength or combined with spent medium of *Pantoea* Fr-CA_6 (Fig. 23c). None of the *Pantoea* inhibited *Pseudomonas*; on the contrary, the addition of the *F. trinervia* isolates spent media to that of the best cross-feeder (*Pantoea* Fr-CA_6) increased the growth of all tested *Pseudomonas* (Fig. 23d).

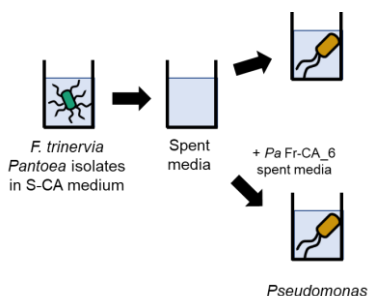
A Experiment workflow

Production of spent media from different *Pantoea* isolates

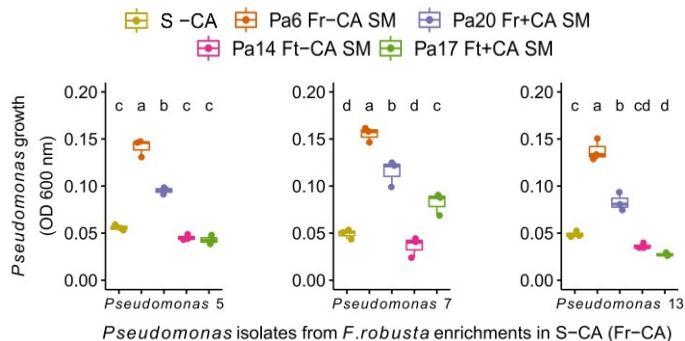


C Experiment workflow

Testing inhibitory potential of *Pantoea* isolates



B Spent media of *Pantoea* isolates from different enrichments



D Growth of *Pseudomonas* isolates on *Pantoea* spent media

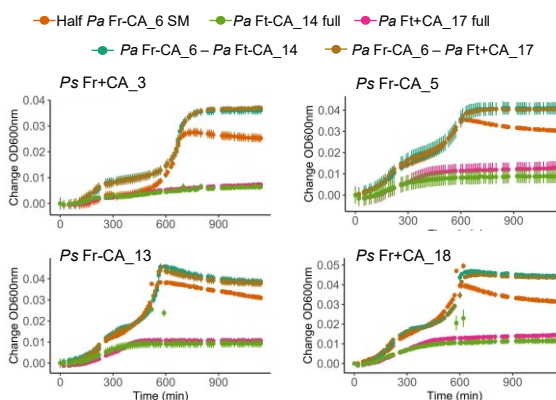


Figure 23. *Pantoea* isolates differ in their ability to feed *Pseudomonas* but none is inhibitory A) Production of spent media from four different *Pantoea* (Pa) isolates. Samples were taken to analyze metabolite consumption. B) Potential of *Pantoea* isolates to support growth of three *Pseudomonas* isolates. Growth was recorded after 48 hours. Different letters indicate significant differences after a Tukey test (N=3). C) Production of Pa spent media and testing its inhibitory effect on Ps growth. D) Growth of several Ps isolates on *Pantoea* Ft-CA_14 and Ft+CA_17 spent media, either at full concentration, or in 1:2 combination with spent media of *Pantoea* Fr-CA_6. The curves are the average of three independent samples. The error bars at each time point show the standard error.

Using untargeted metabolomics, the composition of the spent medium from each of the four *Pantoea* strains was analyzed. Overall, their composition was not significantly different (Fig. 24a, Constrained analysis by strain, p value=0.7713) and they all contained several compounds previously annotated and shown to feed *Pseudomonas* like hypoxanthine and spermidine (Supplementary Table 22). To gain more insight of the feeding behavior of *Pseudomonas* in each spent medium, the consumption of metabolites by the strain Fr-CA_5 from each *Pantoea* was compared (Fig. 24b). Coinciding with the growth patterns, the number of significantly consumed compounds ($\log_2FC > 2$, $FDR < 0.01$) from *F. robusta* *Pantoea* isolates was higher than from the *F. trinervia* isolates. Even though common metabolites had been detected in all four media, there was no overlap between consumption from *F. robusta* and *F. trinervia* isolates (Fig. 24b). One possible explanation is that *F. trinervia* isolates lack some important compounds that would be required in the early stages of growth, hindering the later consumption of the common compounds. Indeed, there were

a few compounds taken up uniquely from either of the *F. robusta Pantoea* spent media that were significantly more abundant there than in the spent media from the other *Pantoea* strains (Fig. 24c and Supplementary Table 22). These are good metabolite candidates for what drove *Pseudomonas* growth but could not yet be annotated.

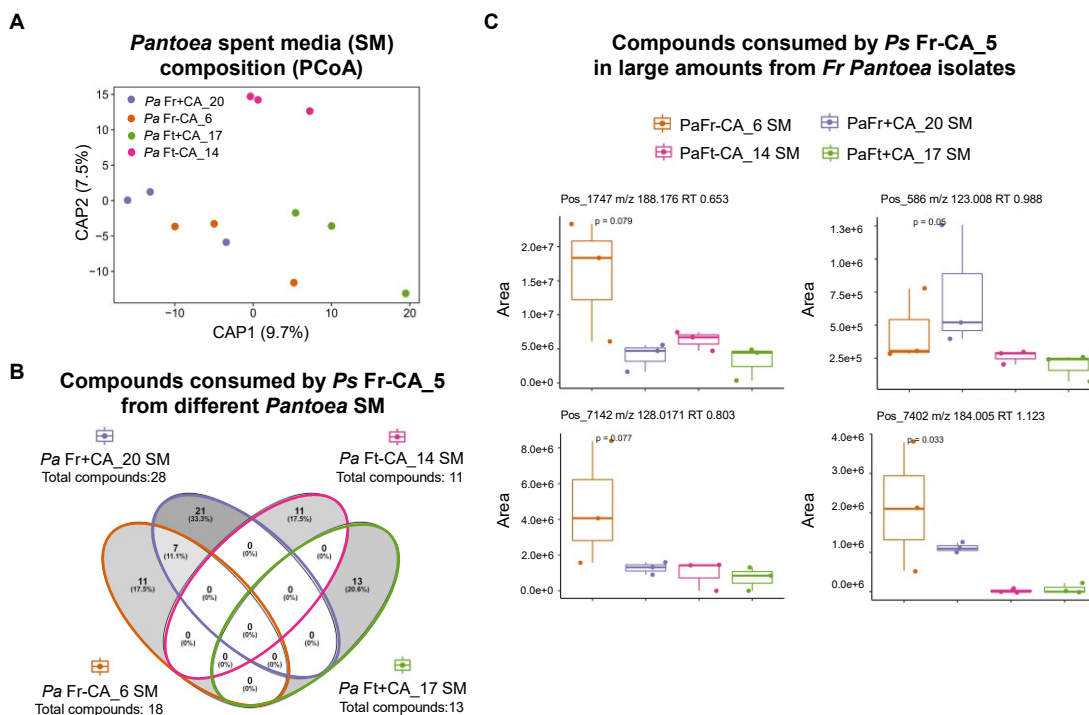


Figure 24. Low growth of *Pseudomonas* (*Ps*) isolates on spent media of *Pantoea* from *Flaveria trinervia* may be due to small differences in metabolite composition. A) Constrained analysis of principal components of all peaks (area > 1.0E4) detected in the spent media of the different *Pantoea* isolates. A generalized logarithm transformation (Parsons et al., 2007) was applied to the data and the analysis was based on Euclidean distance matrix. (N=3). B) Number of compounds significantly taken up (FDR < 0.05, log₂FC < -2) by *Pseudomonas* Fr-CA_5 from the different *Pantoea* spent media. C) Areas of compounds commonly taken up uniquely from the spent media of the two *F. robusta Pantoea* isolates. For each peak, the ionization mode in which it was detected (Pos=positive, Neg=negative), the m/z value and the retention time are shown. Each media had three replicates; p values are based on Kruskal-Wallis test.

Interestingly, the cross-feeding differences between the *Pantoea* strains also did not correlate to their genetic relatedness: the two *F. robusta Pantoea* strains that best supported *Pseudomonas* growth shared only ~81% ANI (consistent with genus-level similarity), but *Pantoea* Fr-CA_6 and the poor cross-feeding *Pantoea* Ft isolates shared ~98.6% ANI (species-level similarity Supplementary Table 23). Together, these results are consistent with good cross-feeding emerging as a convergent trait of *Pantoea* in *F. robusta*.

The leaf nutrient landscape might actively select on cross-feeding interactions

The relevance of potential inter-microbial metabolic interactions for colonization might depend on the chemical environment of the leaf. Since *F. robusta* and *F. trinervia* have contrasting apoplast chemical profiles, with several high-cost amino acids being more abundant in the latter (Chapter 3.2), it was

hypothesized that cross-feeding would not be as necessary for *F. trinervia* colonization as for *F. robusta*. This was tested by inoculating either plant with *Pseudomonas* Fr+CA_3 (slow-grower on *Pantoea* spent media) or Fr-CA_5 (fast-grower) with or without a *Pantoea* partner strain. As partner strain, a *Pantoea* isolate from the matching enrichment was used; *Pantoea* Fr-CA_6 and *Pantoea* Ft-CA_14 were used for *F. robusta* and *F. trinervia* inoculations, respectively (Fig. 25a).

When inoculated alone, each *Pseudomonas* strain only occasionally colonized *F. robusta* leaves (1/8 replicates for Fr-CA_5 and 4/8 replicates for Fr+CA_3), but colonized *F. trinervia* at high levels and much more consistently (11/12 replicates for both strains; 2-sided t-test, p value < 0.0001 and 0.087, respectively). When inoculated together with *Pantoea*, the colonization did not increase for either strain in either plant, however this is possibly due to high variability among replicates. (Fig. 25b). Taken together, these results suggest that *F. trinervia* may better support *Pseudomonas* alone, reducing the need for cross-feeding partners. It was not possible however, to conclusively answer whether colonization of *F. robusta* by either *Pseudomonas* benefited from *Pantoea* colonization.

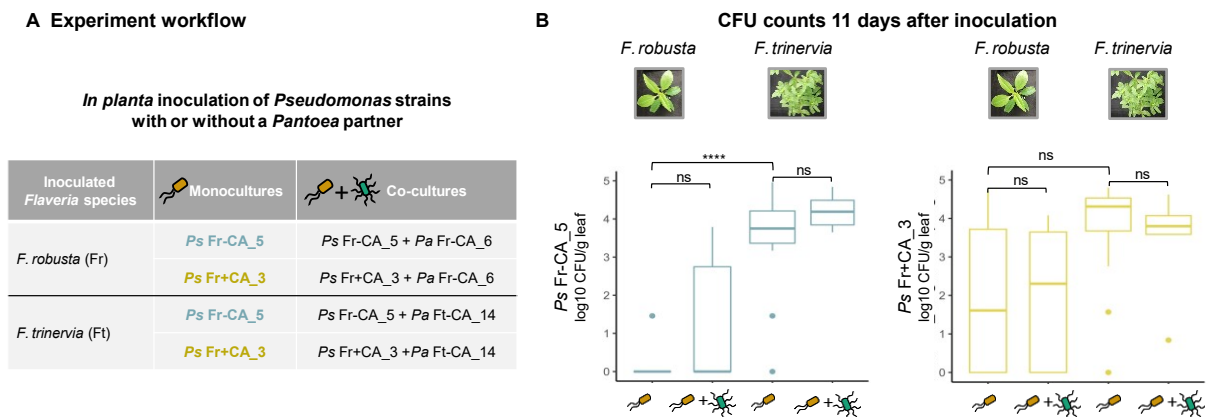


Figure 25. Plant nutrient landscape could select on cross-feeding interactions. A) *In-planta* assay to test the ability of two *Pseudomonas* (Ps) strains to colonize leaves of *Flaveria robusta* and *F. trinervia* either when inoculated alone, or together with a *Pantoea* (Pa) isolate. B) Recovery of *Pseudomonas* isolates (Ps Fr-CA_5 or Ps Fr+CA_3) from *F. robusta* and *F. trinervia* leaves 11 days after inoculation. The experiments were performed twice fully independently (N= 8 leaves for *F. robusta*, N=12 leaves for *F. trinervia*). Significance values are based on 2-sided t-test comparing CFU counts (ns: p value > 0.05. * p <= 0.05. ** p <= 0.01. *** p <= 0.001).

Factors driving selection for cross-feeding depend on the partner strains

The differences in cross-feeding potential among *Pantoea* strains suggested that cross-feeding emerged as a convergent trait in *F. robusta* and that this could have been due to the relatively poor nutrient environment compared to *F. trinervia*. Therefore, the hypothesis that the cross-feeding interaction can be selected on by the nutrient environment was tested using experimental evolution in two contrasting nutrient environments (S-CA where cross feeding is required for *Pseudomonas* or S+CA where it is not necessarily needed, Fig. 26a and 26d). Additionally, two different “feeding” *Pantoea* strains were used to test whether the outcome depends on the interaction partners. Thus, two different communities were experimentally

evolved across 25 24-hr passages: the tagged *P. siliginis* Fr-CA:5mBFP2 and *P. siliginis* Fr+CA_3:mOrange2 which can feed in parallel together with either *Pantoea* Fr-CA_6, the best feeder, or *Pantoea* Fr+CA_20, an intermediate cross-feeder (Fig. 26). Using the tagged *Pseudomonas* strains enabled tracking of the interaction fitness of both partners even though they were cultured together with *Pantoea*. At the 22nd passage, the evolved communities were switched to the opposite environment for 4 cycles. Thus, two outputs were obtained: 1) How growth of each *Pseudomonas* strain with *Pantoea* changed (based on fluorescence measurements over the 25 cycles) and 2) How changes affected growth in the other environment (based on fluorescence after switching). Since *Pseudomonas* growth could change due to adaptation of either *Pantoea* or *Pseudomonas* or both, here only interaction fitness is referred to. To address the first question, the fluorescence of each *Pseudomonas* was compared at passage 1 (naïve strain) versus passage 22 (evolved strain). For the second question, the fluorescence of the evolved strain after the switch to the new environment was compared against the fluorescence of the naïve strain in this environment.

Although each *Pantoea* always evolved together with two *Pseudomonas* strains, each interaction clearly followed distinct paths. The *Pseudomonas* Fr-CA:5mBFP2 interaction generally fit the hypothesis for how environment should select on cross-feeding. When evolved in the S-CA environment with *Pantoea* Fr-CA_6, its growth oscillated throughout the passages but the starting and ending point were very similar (Supplementary Fig. 24a), indicating that the cross-feeding interaction did not exhibit major changes (Fig. 26b, p value = 0.0554). With *Pantoea* Fr+CA_20, the interaction increased towards the end of the experiment (Fig. 26c and Supplementary Fig. 24b, p value = 0.0002). The switch to +CA led to either no change with *Pantoea* Fr-CA_6 (p value = 0.0877, Fig. 26b) or to a decrease in fitness with *Pantoea* Fr+CA_20 (p value = 0.0007 Fig. 26c). When evolved in the richer +CA environment, fitness over 25 cycles decreased with *Pantoea* Fr-CA_6 (Fig. 26e, p value = 0.0037) and did not change with *Pantoea* Fr+CA_20 (Fig. 26f, p value = 0.6956). In both cases, the cross-feeding interaction significantly degraded, evidenced by lower growth of the evolved strain after the switch compared to the naïve strain (Fig. 26e p value = 0.0044 and Fig. 26f p value <0.0001, respectively).

The *Pseudomonas* Fr+CA_3:mOrange2 interaction, on the other hand, evolved in a way that was very dependent on the *Pantoea* isolate. With *Pantoea* Fr-CA_6, fitness increased over the experiment in both environments (3/4 replicates in -CA, Fig. 26b p value = 0.1301 and 4/4 replicates in +CA, Fig. 26e). This was apparently because fitness was coupled across the environments (evolved replicates that grew better in one environment also grew better in the other environment after switching). With *Pantoea* Fr+CA_20, fitness decreased over the course of the experiment (3/4 replicates in -CA, Fig. 26c p value = 0.3442 and 4/4 replicates in +CA, Fig. 26f p value <0.0001) with evolved fitness in each environment again strongly correlated to fitness after the switch between environments. Together, these results demonstrate a complex but consistent story of selection: Selection for cross-feeding depended on amino acid richness for some strains, while for others it is independent of richness but highly dependent on the cross-feeding partner.

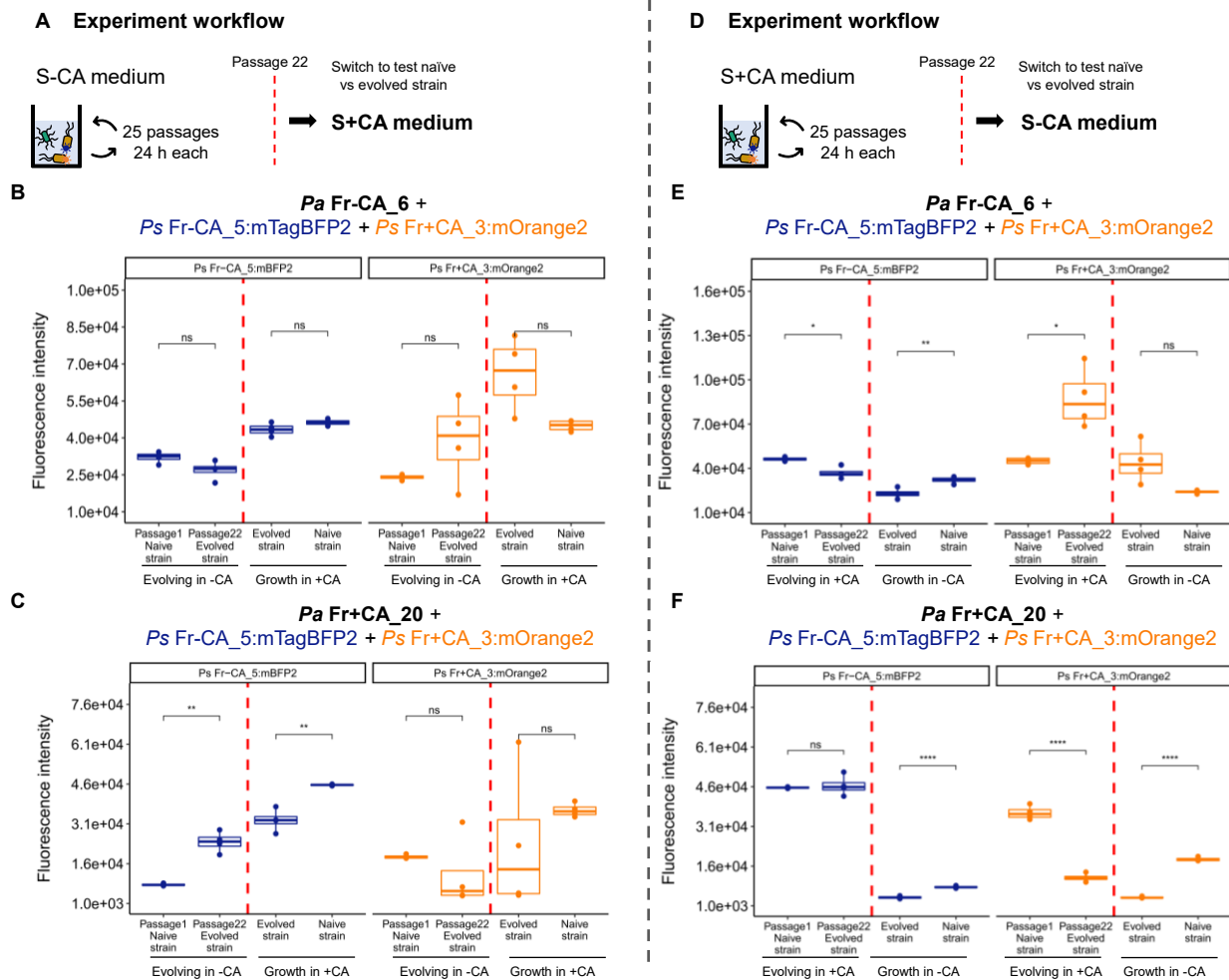


Figure 26. Cross-feeding interactions between *Pantoea* and *Pseudomonas* followed distinct paths during evolution in two nutrient environments. Passaging of the two tagged *Pseudomonas* (*Ps*) strains with either *Pantoea* (*Pa*) Fr-CA_6 or *Pa* Fr+CA_20 on S-CA (A-C) or S + CA medium (D-F). The red vertical line marks passage 22, when the communities were switched to the opposite media to evaluate the effects of their adaptation. B C, E and F) Fluorescence signals of *Ps*Fr-CA_5:mTagBFP2 and *Ps* Fr+CA_3:mOrange2 detected in the communities. Passage 1 naïve strain: fluorescence intensity of the strain at the beginning of the experimental adaptation, Passage 22 evolved strain: fluorescence intensity after 22 passages in the experimental adaptation right before switching media. Evolved strain after the switch of media: fluorescence intensity after the evolved community was switched to the other nutrient condition and allowed to equilibrate for 4 passages. Naïve strain: fluorescence intensity of the naïve strain in the media that the evolved strain was switched to. Significance values are based on t-tests comparing the fluorescence intensity between Passage 1 and Passage 22 (evolving media) and the evolved and naïve strains in the switch media. ns: p value > 0.05, *: p <= 0.05, **: p <= 0.01, ****: p <= 0.001, N=4.

Chapter 4: Discussion

Metabolites play significant roles in the communication between plants and microbes and thus in shaping the plant microbiome (Cadot et al., 2021). Through root exudates for example, plants can recruit beneficial organisms (Carvalhais et al., 2013) or exclude potential pathogens (Hu et al., 2018). Plant metabolites can also work as signaling cues in bacteria to initiate multiple cellular processes and can be used as main nutrient sources in the rhizosphere (Moormann et al., 2022).

The effect of plant metabolites on the assembly of the leaf microbial community is less clear. It is known that microbes that establish in the leaf are exposed to complex responses from the plant and that these are partly mediated by secondary metabolites (Tierens et al., 2001). In addition, these microbes are thought to face nutrient limitations in the leaf because the plant exerts a strict control over cellular transporters to avoid uncontrolled leakage of metabolites. While pathogens have evolved different strategies to manipulate the plant to obtain nutrients, non-pathogenic strains or commensals may rely more on scavenging transient nutrients or on metabolic interactions with other microbes (Paulsen et al., 2005).

Commensal bacteria can play important roles in protecting plants from pathogens and abiotic stress (Trivedi et al., 2020), therefore it is crucial to understand the processes governing their assembly. However, characterizing how plant physiology and plant-microbe interactions together influence the leaf environment and the resources available for colonizers has remained a difficult and complex task. Thus, techniques to simultaneously dissect plant physiology and microbial diversity are needed. In this study, this was addressed in two ways: first, by developing a method to characterize the leaf environment and the corresponding microbial community and second, through an *in vitro* approach comparing the effect of two contrasting nutrient conditions, that reflect possible scenarios in the leaf, on the emergence of bacterial interactions and enhancement of diversity.

4.1 Apoplast fluid washes offer the potential to simultaneously study microbial diversity and the apoplast environment

The leaf apoplast is crucial for plant physiology; it plays important roles in different processes, for example the synthesis of the cell wall (Liu et al., 2018). Additionally, transport of water, sucrose and amino acids utilize several steps involving the apoplast (Sattelmacher, 2001; Zhang and Turgeon, 2018). While the apoplast metabolic composition is not well understood, its composition changes according to environmental conditions such as light and biotic or abiotic stresses (López-Millán et al., 2000). Furthermore, the apoplast is an important arena for plant-microbe communication (Godson and van der Hoorn, 2021). Thus, better characterizing this complex environment could provide valuable information to understand pathogen growth, or promotion and establishment of potentially beneficial microbes.

High-throughput characterization of metabolite landscapes (i.e., metabolomics) largely relies on mass spectrometry or the measurement of mass to charge ratio of compounds. Metabolomic methods can be coarsely divided into imaging techniques for metabolite localization and bulk sample analysis, usually in combination with chromatography separation. Current mass spectrometry imaging techniques like matrix assisted laser desorption (MALDI) or laser activated electron tunneling (LAET) have been used to study metabolites in leaves (Huang et al., 2016; Dreisbach et al., 2021) and to register changes that occur in the leaf during the interaction with microbes (Maia et al., 2022). However, their spatial resolution (~20 μm) is not enough to characterize compact spaces like the apoplast and they generally require even surfaces. Other techniques, like the use of FRET-based nanosensors can allow tracking temporal and spatial metabolite changes at the subcellular level, but they are not suitable for getting overall metabolic profiles because each metabolite requires a localized sensor protein to be detected (Kueger et al., 2012). To study the bulk metabolite composition of the apoplast, previous studies have largely relied on extraction of fluid from the tissue. The resulting liquid can be subjected to different analytical techniques, like NMR and chromatography (O'Leary et al., 2014). In this study, the liquid chromatography technique that was used was optimized to separate low molecular weight metabolites as they are known to be especially relevant for nutrition of plant-associated microbes (Gu et al., 2020).

Metabolites in the leaf apoplast were recovered by the infiltration-centrifugation technique (Lohaus et al., 1995; O'Leary et al., 2016; Gentzel et al., 2019). This method requires optimization for each plant species to avoid cell membrane rupture and ensure that the apoplast fluid wash (AFW) is not contaminated with cytoplasmic metabolites. One critical point is the infiltration solution itself because major differences in osmotic pressure and pH can lead to membrane rupture and leakage of cytoplasmic compounds (Witzel et al., 2011). Out of the three solutions tested in this work, sodium phosphate best preserved the integrity of the cell membrane and was most compatible with the UHPLC-HRMS method. Double distilled water has been successfully used in different species (O'Leary et al., 2014), but in this study infiltration with Millipore water led to leakage in some of the species. This is likely due to a low ionic strength and/or to a higher sensitivity to osmotic differences in the species used. It is important to note that the infiltration solution possibly influences which metabolites are recovered, partly because compounds are not equally soluble in different solutions (Witzel et al. 2011). For example, to recover more non-polar or hydrophobic metabolites, organic solvents like chloroform or alcohols such as ethanol and methanol are frequently added to water. This has proven to be an effective extraction method for flavonoids, alkaloids and different plant hormones (De Vos et al., 2007; Zhang et al., 2008).

The infiltration process dilutes the apoplast concentration of metabolites (Lohaus et al., 2001; O'Leary et al., 2014). This dilution must be considered to correct the mass spectrometry peak areas. One way to calculate a dilution factor is to infiltrate leaves with indigo carmine and spectrophotometrically measure the dilution of the dye after centrifugation (Solomon and Oliver, 2001; Rico and Preston, 2008). While this is theoretically an accurate method, it is in effect only an estimation. Since the dye is not compatible with downstream steps, infiltration is done on different leaves than the actual sample. A recent study offered a

simple solution that calculates the percentage of apoplastic space corresponding to liquid before infiltration (i.e., apoplast hydration). For this, the leaf must be centrifuged until it reaches a constant weight which should be lower than the pre-infiltration weight. This indicates that all infiltrated liquid plus the liquid present in the apoplast before infiltration is fully recovered (Gentzel et al. 2019). Since this can be calculated for the actual samples, it was reasoned that this would be more accurate.

In *F. linearis*, the centrifugation had to be extended for 30 min under low speed just to recover all the infiltrated volume. To calculate the hydration index thus, longer centrifugation would have been required. This is very time-consuming and not feasible when working with many samples. At high centrifugation speed 1 min was enough to recover all infiltrated fluid, but constant weight was only achieved after 30 min. Metabolomic data already showed contamination from the cytoplasm at this point. These differences between studies are not surprising since Gentzel et al. (2019) used leaf tips of young maize plants, a monocotyledon species which has a very different vascular structure and apoplast air space than the dicotyledon *F. linearis* (Taiz et al., 2015). At any rate, these results suggest the hydration method may not be applicable for all species. Instead, incomplete AFW recovery was selected (mild spin at 2500 x g for 3 min) which did not fully empty the leaves but guaranteed the integrity of the cytoplasmic membrane and overall allowed enough volume recovery for downstream processing.

Because it was not possible to calculate a hydration value, in this study the infiltrated volume relative to the leaf mass was used to correct the MS peak areas (infiltration ratio). This value works as expected, in the sense that if two leaves have the same initial weight but one is infiltrated with more liquid, the apoplast dilution will be more and the corresponding correction factor will be larger. To further test this infiltration ratio, it would be useful to inject a known concentration of a compound that would not be expected in the plant (a deuterated standard, for example) to the leaf, quantify its recovery after centrifugation and correct the value by the corresponding dilution factor. Theoretically, no matter how much volume was infiltrated, the final concentration should be similar.

Many plants have very small leaves and/or highly packed tissues, with little apoplastic space. In such cases, the recovered liquid is usually lower than the minimum required for injection in an HPLC and thus, must be diluted into a larger volume. The resulting peak areas must then be corrected by this dilution factor. Amino acids were selected as a proxy to evaluate the effect of dilution as they play important roles for bacterial nutrition in the apoplast (Rico and Preston, 2008). Using a deuterated amino acid standard in a real apoplast sample, all 16 amino acids were detected in all dilutions with similar concentrations. Thus, the correction did not overestimate their areas in the higher dilutions. For the samples analyzed here, it seems that diluting up to 5x is possible without losing the overall profile of the AFW. Thus, the method is suitable for species with small leaves and/or with small apoplastic space. However, it is possible that individual compounds are more affected than the overall composition. This would be especially problematic for compounds that are in very low concentrations in the apoplast or whose detection limits are higher. This is likely to also depend

on the UHPLC equipment setup, so it is recommended to define the maximum dilution ratio for compounds of interest in different setups.

The apoplast fluid wash can also be used to explore structures of endophytic bacterial communities in naturally colonized plants. These communities are usually studied by crushing surface sterilized leaves, extracting DNA from the crushed tissue and amplifying 16S rRNA genes. This leads to a disproportionate signal from the host chloroplast and very limited signals from leaf bacteria (Regalado et al., 2020). In addition, it is possible that some bacteria remain protected by leaf fragments, preventing the extraction buffer from reaching them. Amplification of the host can be blocked via different strategies, including the use of blocking oligos which target regions of the 16S gene that are specific of the plant (Mayer et al., 2021). Even though this works well for different plant species, it seems that for *Flaveria* it would require further modifications because the host signal still represented over 98% of the reads in the samples from crushed leaves. Instead, using the AFW considerably reduced the signal from the host (~65% of reads), allowed recovery of more taxa and led to a better separation of the *Flaveria* species. Still, host signals were detected in the AFW (~35% of reads), possibly indicating that there was cell membrane rupture to some extent at least. One possible explanation is that these plants were grown outside a controlled environment and were more exposed to insect and pathogen damage.

Overall, the method could separate *Flaveria* species based on both their metabolic profiles and the bacterial communities of the leaf apoplasts. In plants grown under controlled conditions, the metabolic differences between species were consistent across multiple plants grown independently at different times and were not dependent on which plant was sampled. While it is not yet clear why some amino acids and other metabolites differ strongly between *Flaveria* species, one possibility is their different photosynthesis mechanisms. Evolution of C4 photosynthesis in *Flaveria* has had diverse effects. For example, higher glutathione turnover in sulfate assimilation (Gerlich et al., 2018) has led to higher cysteine levels in *F. trinervia* leaves. While it was not possible to reliably quantify cysteine with this method, it is the direct precursor to methionine, which was significantly elevated in *F. trinervia* apoplast. Thus, links between photosynthesis and apoplast metabolites are plausible. At any rate, the differences observed in compounds that are known to be important for bacterial nutrition in the apoplast, like amino acids, open an exciting avenue to simultaneously study the chemical and nutrient landscape that microbes are exposed to (plant physiology) and the composition of the endophytic bacterial community. Combined with other techniques like metagenomics or RNAseq, this method could provide more information on the metabolic capabilities of commensals. For example, it would be interesting to know whether bacteria established in the apoplast are auxotrophs for specific compounds. The method could also be used to study gene expression to understand adaptations that leaf commensals undergo in response to nutrient fluctuations, such as the assembly of secretion systems that would allow them to engage in microbe-microbe interactions.

4.2 Metabolite and microbial dynamics follow species-specific and distinct diurnal patterns

Leaves are dynamic organs subject to complex diurnal plant cycles. The onset of daylight activates the photosynthetic machinery of plants; carbon fixed throughout the day via the Calvin-Cycle can be stored as starch or can be used to synthesize sucrose, which is allocated to multiple biosynthetic routes. Sugars are translocated throughout the plant to reach non-photosynthetically active tissues, and this occurs partly through the apoplast (Lemoine et al., 2013). During the night, starch can be used as an energy reserve to fulfill the plant's demands (Taiz et al., 2015). Correspondingly, metabolic profiles in leaves generally show increasing sucrose levels from dawn to sunset and decreasing starch throughout the night (Sulpice et al., 2014). In addition, plants experience metabolic shifts throughout the day which correspond to the plant's inner clock (Venkat and Muneer, 2022).

In the previous section, a “baseline” profile of the chemical landscape of three *Flaveria* species was generated by sampling lab-grown plants at one point in the daytime. In addition, it was shown that the apoplast fluid can be used to study diversity of bacterial communities. However, to ultimately understand how microbes thrive in the leaf, potential fluctuations of the communities and host metabolites across diurnal cycles should be considered. A lot of processes in the plant, including photosynthesis, are regulated diurnally (Dodd et al., 2005; Cervela-Cardona et al., 2021) and so it is possible that commensal endophytes adapt their metabolic activity or even change in abundance throughout the day in response to metabolic fluctuations in the plant. Previous studies suggest that such adaptations occur in rhizosphere communities (Staley et al., 2017), but there is practically no corresponding studies in leaves, where diurnal resource shifts are likely to be equally drastic. Here, the metabolic profile and the community composition of *Flaveria* leaves from plants grown outdoors were simultaneously studied in three critical moments of the day: before sunrise, in the middle of the photoperiod and at the end of the day.

All three *Flaveria* species that were studied showed a fluctuation of the overall metabolic profiles throughout the day. Overall, the midday samples were especially different from the early and late samples. Interestingly, these patterns were not homogenous in the three species; while in *F. robusta*, a plant with C3 photosynthesis, and *F. linearis* (C3/C4 intermediate) more compounds were highly present at 1:00 pm, in *F. trinervia* (C4) more compounds were high at 5:00 am and 8:00 pm (Fig. 11). It has been previously reported that C3 plants have slower export rates of photosynthates than C4 plants (Leonardos and Grodzinski, 2000). Lower export rates could cause a backup of metabolites in C3 plants, which could explain enrichment of many metabolites in *F. robusta* relative to *F. trinervia* at 1:00 pm.

Metabolic pathway enrichment analyses identified several diurnal patterns that agree with previous reports, suggesting their possible utility for building hypotheses on nutrition in the apoplast. In *F. robusta*, compounds related to chlorophyll/porphyrin were enriched at dawn and depleted at the end of the day. Chlorophyll biosynthesis typically begins at dawn, peaks at midday and decreases afterwards (García-Plazaola et al., 2017). On the other hand, amino acids levels in leaves seem to steadily increase from dawn

and peak towards the end of the day and this is thought to be due to circadian regulation of protein synthesis for growth (Espinoza et al., 2010; Sulpice et al., 2014). In both *F. robusta* and *F. linearis*, compounds related to amino acid biosynthesis were indeed higher at midday compared to dawn but did not show any particular trend at the latest time point. Compounds related to terpene and zeatin biosynthesis were also more abundant in the 1:00 pm samples of *F. robusta*, possibly indicating a response to heat stress (Singsaas, 2000; Prerostova et al., 2020). Similarly, metabolism of glutathione, an antioxidant compound (Hasanuzzaman et al., 2017), was higher at midday in the *F. linearis* samples. In *Arabidopsis*, it was shown that expression of several genes related to generation of reactive oxygen species (ROS) peak towards the late morning (Jiménez et al., 2021). Since glutathione and other compounds related to ROS are known to play important roles in plant immunity (Hiruma et al., 2013), it would be interesting to evaluate whether their presence impacts the microbial community assembly. Observations on stress pathways could be complemented with physiological measurements of the plant, like chlorophyll fluorescence, which gives an idea of how stressed the plant is (Guidi et al., 2019). Unfortunately, some of the comparisons between species were hindered by the small number of samples available at specific time points due to rupture of the cell membrane. This is a limitation when sampling outdoor plants, which are exposed to insect and/or pathogen attack and can only be circumvented by increasing the sampling effort.

F. robusta and *F. trinervia*, which have completely different photosynthesis strategies (C_3 and C_4 , respectively), showed some clearly distinct patterns, which could be related to their fixation strategies. At dawn, compounds related to starch and sugar metabolism were more enriched in *F. trinervia*. This early carbon fixation could be explained by the high efficiency of C_4 plants in concentrating CO_2 (Sage et al., 2012). Alternatively, it could reflect the differences between the two species in their strategies to store and allocate starch. C_4 plants for example are known to store more starch in the bundle sheath cells than in the mesophyll, contrary to C_3 plants (Weise et al., 2011). Furthermore, higher levels of starch at dawn could suggest that transitory starch was not fully depleted during the night. In addition, compounds related to purine metabolism were also enriched in *F. trinervia*. Purine catabolism was shown to be enhanced during darkness in *A. thaliana* (Baccolini and Witte, 2019), possibly as an alternative N source. At midday, all annotated biosynthetic and catabolic pathways were enriched in *F. robusta*. One possible explanation for this is, as mentioned above, the slower export rates in C_3 plants of photosynthetic products, possibly leading to their accumulation towards midday (Leonardos and Grodzinski, 2000). Previous studies have already shown that C_3 and C_4 *Flaveria* species can be separated based on the leaf content of primary metabolites. However, they used whole grinded leaves instead of apoplast fluid (Gowik et al., 2011; Borghi et al., 2022). Determining whether photosynthesis type has a generalized effect on the metabolite landscape of the leaf apoplast would require including more examples of closely related species with different photosynthesis strategies. Ideally, this group of plants would also include monocotyledon species, which have different venation patterns and different transpiration and photosynthesis efficiencies (Ehleringer et al., 1997). Nevertheless, this preliminary comparison already shows some interesting differences.

These results from the metabolic pathway analyses are useful to get a first idea of the differences between sampling points and plant species, but they have several limitations. First, the analysis uses as reference a metabolomic library from *Arabidopsis* plants which was not generated using the exact same UHPLC-HRMS method used in this study, which could result in variations in the retention time of the compounds (Wieder et al., 2021). In addition, the library includes compounds that are not necessarily expected in *Flaveria* plants, like glucosinolates and most probably misses others specific to *Flaveria*. Thus, to confidently guarantee that certain pathways were enriched at specific time points, it would be necessary to identify these compounds. For this, a subsequent ionization of the compounds of interest would be necessary (MS/MS) to better characterize the compound. In addition, potentially important compounds could be separated to test their effects in-vitro and use for example, nuclear magnetic resonance to definitively identify them.

Oscillations throughout the day of plant-associated microbial communities have been previously reported, especially in rhizosphere communities (Staley et al., 2017; Hubbard et al., 2018; Lu et al., 2021; Newman et al., 2022). Given that stomata aperture and plant defense mechanisms are in part diurnally regulated (Zhang et al., 2019) as well as apoplast metabolites (this study), it is conceivable that leaf bacteria also oscillate throughout the day. It has been observed, for example, that entry to the leaf tissue of sprayed *P. syringae* is hindered at night by stomata closure (Zhang et al., 2013). Furthermore, changes in apoplastic pH and redox state throughout the day (Geilfus, 2017; Karpinska et al., 2018) are known to impact the establishment of microbes in the leaf (Alkan et al., 2013). Thus, the same samples used for metabolic analyses were used to assess bacterial diversity. The bacterial reads were overall very low. This is not surprising, because previous studies have already shown low bacterial loads in endophytic communities (Regalado et al., 2020). The morning samples were especially limited in reads, but it was not possible to determine if this corresponded to lower bacterial loads. If this was indeed true, a possible explanation is that there is extensive bacterial turnover across the day. Turnover rates of bacteria in leaves are limited in literature, but reports of root-associated bacterial communities have shown they can be lower than 24 hours (Olsson et al., 1987; Söderberg and Bååth, 1998). The reasons for turnover could be many but could include digestion by plants as nutrient sources. *Arabidopsis* and tomato roots, for example, were reported to take up and digest *E. coli* and *Saccharomyces cerevisiae* to release nitrogen (Paungfoo-Lonhienne et al., 2010). Understanding possible roles of turnover will require future studies using internal standards that give more certainty about absolute abundances at different timepoints. Considering these uncertainties, a conservative approach was preferred and samples with too few reads were not included in the diversity analyses. The analyses were then limited to comparing *F. robusta* and *F. trinervia* in the afternoon and night.

In *F. trinervia*, the alpha diversity peak seemed to occur at 1:00 pm and was marked by a slight increase in Actinomycetes. In *F. robusta*, 1:00 pm samples were also enriched in Actinomycetes, while Bacilli tended to increase towards the end of the day. Both of these taxa were previously shown to have circadian

oscillations in the rhizosphere (Newman et al., 2022). Diversity enrichments towards the dark period were also observed by Staley et al. (2017) in *Arabidopsis* rhizosphere communities. By doing functional predictions of the community, the authors found that bacterial genes related to carbohydrate catabolism were less abundant during the dark period. Since sugar catabolism is on the contrary more active during the dark period in plants, the authors suggested that rhizosphere communities could be providing complementary functions to the plant throughout the day. Although this is an interesting hypothesis, no further evidence was provided to support it. What other studies have certainly shown is that disruptions in the circadian rhythm of the plant can lead to different compositions of the rhizosphere bacterial community (Hubbard et al., 2018). Zhao et al. (2021) showed, for example, that interrupting the cyclic exposure of rice plants to light decreased the oxygen secretion by roots. This directly affected the metabolic niches in the soil, resulting in higher bacterial alpha diversity.

In *F. trinervia* and *F. robusta*, the apparent increment of alpha diversity happened right after the peaks of metabolic complexity (5:00 am and 1:00 pm, respectively), which could suggest a link between nutrient availability and complexity and microbial community structures. However, given the limited data at earlier timepoints, an experiment with internal standards as suggested above is also needed to provide more concrete conclusions. Interestingly, however, examples outside plants have suggested such links between the nutrient environment of the host and fluctuations of the microbiome. In the human and mouse gut, for example, changes in the composition of the microbiome throughout the day are controlled by the rhythmicity of food ingestion and nutrient availability (Thaiss et al., 2014; Zarrinpar et al., 2014).

If leaf commensals are indeed turned over during the day as suggested above, most bacterial reads probably correspond to live cells. However, since DNA is stable over time it is not possible to rule out that some reads could be from dead cells. A possible solution to this could be to account for live cells only (dormant and metabolically active) by sequencing RNA instead of DNA. Alternatively, dead cells could be identified and removed by using DNA-intercalating agents, like propidium monoazide which only penetrate membrane-compromised cells (Li et al., 2017). Both alternatives have their technical complications but may lead to more robust conclusions. In any case, to strengthen the results shown here, it would be necessary to include an internal standard that would allow accurate quantification of the bacterial load (Guo et al., 2020).

Most of the information on metabolic diurnal shifts so far has been generated in model plants like *A. thaliana* and *Nicotiana tabacum* using whole leaf tissue (including cytoplasm contents). Although direct comparisons between whole leaves and apoplast fluid are apparently non-existent in literature, it seems reasonable to expect differences in their metabolic profile. Many transporters in the plant cell are regulated by the circadian rhythm of the plant and thus could lead to differences in solute availability in the apoplast (Haydon et al., 2011). In addition, the apoplastic pH changes with light exposure and water availability, which in turn oscillate throughout the day (Mühling and Läuchli, 2000; Geilfus, 2017). These changes are likely to affect

nutrient availability in the apoplast, so it is hard to know how to translate results for whole leaves to the apoplast.

Regardless of the technical limitations, this study shows that the leaf apoplast metabolic composition varies throughout the day, with possible implications for microbial community assembly. If this is the case, such oscillations should be considered when sampling leaves to assess microbial diversity. Furthermore, understanding diurnal patterns in the microbial community of plants could lead to the identification of biomarkers for dysbiosis. In the human gut microbiome, for example, rhythmicity rupture of 13 bacterial taxa was associated to type 2 diabetes patients (Reitmeier et al., 2020). Such biomarkers could in turn provide information on the plant's health or about its susceptibility to stress and this could help direct early responses.

4.3 Leaf bacteria can employ metabolic cross-feeding to survive in nutrient-limited environments

It is increasingly clear that interactions between microbial colonizers play major roles in community structuring in plants, whereby the presence of specific microorganisms can drastically alter leaf communities (Aglar et al., 2016). However, the complications of studying microbe-microbe interactions *in vivo* in leaves severely limits the understanding of the types of interactions that are important. To overcome this, leaf bacteria were enriched *in vitro* under defined nutrient regimes. This approach demonstrated that leaf bacteria interact with one another via metabolic cross-feeding. This supports previous work that showed that pervasive cross-feeding is possible between leaf-derived bacteria, enabling a surprising diversity can subsist on single carbon sources (Goldford et al., 2018). It was also shown that cross-feeding can support bacterial taxa who have no direct utilizable carbon source and that in this case, they can survive by utilizing only secreted or leaked metabolites from diverse bacteria.

The cross-feeders in the enrichments consistently were only *Pantoea* and *Pseudomonas*. This was surprising, since in previous studies, the diversity of cross-feeding taxa enriched on single carbon sources was high due to promiscuous cross-feeding interactions (Goldford et al., 2018; Dal Bello et al., 2021) and because of the results showing that multiple other bacterial taxa isolated from the leaf extracts could have fed *Pseudomonas*. The most likely explanation is simple resource competition. *Pantoea* grew much faster than other leaf isolates on sucrose, allowing it to become the sole feeder of *Pseudomonas*. *Pseudomonas* also grew rapidly on *Pantoea* spent medium and could have thereby outcompeted others for key resources. These results do not necessarily mean that cross-feeding partners would be this restricted in leaves. In contrast to the well-mixed and homogeneous *in vitro* enrichment environment where competition would dominate dynamics, the leaf apoplast is highly compartmentalized and heterogeneous (Sattelmacher, 2001). Additionally, most endophytic commensal bacteria in leaves reach only low colonization density (Regalado et al., 2020). These factors would decrease the importance of resource competition and thereby diverse taxa would be able to establish a niche consuming primary plant-derived resources and *Pseudomonas* or others could establish to take up leaked metabolites. Therefore, cross-feeding

interactions in leaves can potentially arise between taxonomically much more diverse members than what was observed in the enrichments.

It is important to note that the findings that bacterial diversity is relevant for metabolic interactions among very common leaf bacteria is based on a culture-dependent approach and a limited set of isolates. While such culture-dependent work is critical, especially given technological barriers that limit the ability to directly observe metabolic interactions in leaves, it also means that it is not possible yet to fully weigh the relative importance of metabolic or other (e.g., competitive) interactions among the isolates or especially uncultivated microbes. On the other hand, culture independent techniques and modeling can help provide some insight into the relevance of such interactions. Generally, inference of inter-bacterial interactions in leaves based on culture-independent taxa correlations has suggested extensive negative interactions, but recent improvements that utilize abundance information suggests that positive interactions have probably been underestimated (Regalado et al., 2020). Similarly, experimental results suggest frequent positive interactions among co-colonizing leaf bacteria (Remus-Emsermann et al., 2014; Morella et al., 2020). This might seem unlikely, since ecological models have predicted that cooperative interactions can destabilize microbial communities and that competition should therefore benefit hosts (Coyte et al., 2015). However, the instability in these models is caused by strong species dependencies that can easily be interrupted. In contrast, cross-feeding among leaf bacteria seems to involve weak coupling and high promiscuity.

This type of cross-feeding involving promiscuous interactions could even benefit host plants by stabilizing microbiomes to invasion if cross-feeding networks are redundant enough to leave few resources for invaders (Herren, 2020). Additionally, apoplast nutrients are important regulators of pathogen virulence, so full occupation of resources by cross-feeding could directly limit damage caused by them even if they successfully invade (Rahme et al., 1992). This could explain, for example, how *Pantoea* protects crops from pathogenic *P. syringae* by decreasing its virulence without eliminating it from the leaf microbiome (Morella et al., 2019). Additionally, cross-feeding would be beneficial if it allows bacteria to establish who can, upon invasion, switch to competitive behaviors that protect plants, such as antibiotic production (Cha et al., 2016). Therefore, more thorough investigations into the role of cross-feeding in interactions between leaf bacteria and pathogens and its general role in shaping and stabilizing co-colonizing leaf microbiota are needed, including dissection of how metabolic networks arise from host resources through the microbiome.

4.4 Strain-level leaf bacterial diversity is linked to cross-feeding niches

Leaf microbiomes are known to be diverse. A previous study showed that a single OTU (99% 16S rRNA gene sequence similarity) of the leaf-associated *Pseudomonas viridiflava* harbored at least 82 distinct strains (99.9% genome identity), with clearly different host interaction phenotypes (Karasov et al., 2018). Commensal *Sphingomonas* bacteria were also shown to exhibit extensive diversity in the presence of secretion systems likely relevant for microbe-microbe interactions (Lundberg et al., 2021). The three *Pseudomonas siliginis* isolates that were sequenced in this study were genetically similar with ANI of ~99.98% to 99.99% and ~97,000 shared SNPs against the most similar *P. siliginis* reference genome. While

more information is needed to understand this diversity, it appears to be functionally relevant. The sequenced strains had distinct cross-feeding niches with only partially overlapping metabolite uptake profiles, making it possible to cross-feed in parallel despite different growth rates. In addition, two of the strains each showed clearly distinct patterns of adaptation when evolved with *Pantoea* in two nutrient environments. This diversity must have originated in the original leaf communities since the phenotypes appeared both in sucrose-only enrichments where cross-feeding was required and in enrichments with amino acids as additional resources. Given that diverse leaf bacteria from multiple plant species have been shown to readily engage in cross-feeding (Goldford et al., 2018), these results strongly suggest that leaf bacterial genetic diversity has important functional relevance for inter-bacterial metabolic interactions in leaves and probably wherever they persist. Since nutrition plays key roles in virulence (Wei et al., 1992; Ikawa and Tsuge, 2016), it also raises the intriguing question of whether interactions like cross-feeding may ultimately influence interactions with hosts.

4.5 Complex and variable leaf nutrient environments can affect selection on metabolic interactions

Pantoea isolates also exhibited interesting and surprising functional diversity. Those from *F. robusta* enrichments, where cross-feeding occurred, fed *Pseudomonas* better than those from *F. trinervia* enrichments, probably by producing more of key metabolites. However, two “good” cross-feeders enriched from *F. robusta* were only distantly related (81% ANI), consistent with inter-species differences (Jain et al., 2018) compared to 98-99% ANI between good and bad cross-feeders. This result is consistent with the *F. robusta* environment selecting for *Pantoea* with increased metabolite secretion, benefitting cross-feeders. This agrees with *in-planta* data, where individual bacteria persisted in *F. trinervia* leaves alone at higher levels than in *F. robusta*, possibly due to relatively lower levels of key nutrients like amino acids methionine, isoleucine, lysine and valine. Except for the aromatic amino acids, methionine, isoleucine and lysine are three of the four biosynthetically highest-cost amino acids (together with histidine, (Akashi and Gojobori, 2002)). The availability of these amino acids could also directly affect exchange of the diverse metabolites observed here. For example, cross-fed purines guanine and hypoxanthine are connected to methionine via the THF cycle. Thus, differences in the apoplast metabolite landscape could alter selection for traits underlying cross-feeding. Indeed, experimental evolution of cross-feeding interactions in contrasting nutrient environments confirmed that the nutrient environment and the specific cross-feeding interaction determine the adaptive path followed by the strains.

In extreme cases, increased metabolite secretion could result from evolution of reciprocal cross-feeding between *Pantoea* and *Pseudomonas* specifically, which can in turn lead to strong co-adaptation (D'Souza et al., 2018). However, this seems unlikely in leaves since conditions including compartmentalization, low bacterial density and high bacterial diversity are likely to make interactions between specific taxa transient, decreasing the likelihood that reciprocal cross-feeding will arise (Oña and Kost, 2022). Thus, if more metabolite secretion is a beneficial adaptation in the *F. robusta* apoplast environment, it is more likely because *Pantoea* derives benefits from cross-feeding with diverse taxa rather than *Pseudomonas* alone.

It is important to note that these results were based on a single set of enrichments from *Flaveria* plants and therefore represent only a sliver of the interactions that probably happen in a more diverse leaf microbiome and across more diverse host plants. However, given the clear evidence that microorganisms influence one another's colonization patterns in leaves (Agler et al., 2016; Durán et al., 2018) and the promiscuity of cross-feeding among leaf bacteria, it seems highly unlikely that the vast genetic diversity and stable phenotypes known to exist among leaf bacteria (Karasov et al., 2018) only developed in response to interactions with hosts. Whether the most important interactions shaping this diversity are cooperative metabolic interactions or some other interactions type remain to be seen and this will be an important future avenue of research. However, the results that nutrient differences that are relevant in host plants can drive divergent adaptive cross-feeding responses in closely related bacterial strains add weight to the argument that metabolic interactions may play important roles. This is reminiscent of roots, where the exudate nutrient landscape is key to shaping microbial communities (Voges et al., 2019; Cadot et al., 2021) and differs between different plant species (Aulakh et al., 2001; Dietz et al., 2020). The results shown here strongly suggest that apoplast exudates influence the arisal of microbial interactions, which in turn help shape leaf microbiomes (Agler et al., 2016). If so, this is an exciting prospect because it could offer targets for manipulation by plant breeders.

Conclusions and outlook

- Current methods to study bacterial diversity and nutrition in leaves have various economic and technical limitations. This study provided a solution to simultaneously study the leaf apoplast environment and the bacterial communities that establish there. With this method it is possible to identify components in the apoplast that may be critical for bacterial nutrition, even in small samples. This information can be further used, for instance, to design growth media that better represents the leaf environment. Further improvements of the method should consider adding controls to better assess the bacterial load in the apoplastic space. Ongoing work suggests that the apoplast fluid wash recovered from the leaves can be used in other workflows (i.e., metagenomics) to assess functional diversity of the microbial community.
- In nature, plants are exposed to mechanical damage caused by different external factors (wind, insects, crop management, etc.). This is usually overlooked in lab studies, although it may have important implications in the colonization processes. With the method established in this work this question could be addressed by comparing the metabolite profile and the microbial community of plants grown in controlled conditions versus plants grown outside.
- Since the apoplast reliably reflects the physiological state of the plant, it would also be possible to use this method to correlate metabolic responses to external stresses (drought, salinity, light fluctuation, temperature changes, pathogen attack, etc.) with changes in the microbiome. This information can ultimately be used to design more stable microbial communities that show resilience to disturbance or that express traits valuable to plants under stress.
- The diurnal sampling of *Flaveria* leaves showed that the metabolic profile of the apoplast clearly changes throughout the day, and that these changes are, to some extent, species dependent. This also confirms that analyzing the apoplast composition is a promising strategy to capture physiological changes in the plant that may affect the microbial community.
- The matching patterns of metabolic complexity and alpha diversity suggest an interesting link between resources in the leaf and the microbial community. However, proving this would require further testing. First it would be important to sample for at least two day/night cycles to check whether the patterns are repeated in the next cycle. Another possibility would be to disrupt the circadian rhythm of the plant. In *Arabidopsis* this is possible by using circadian clock mutants (Lu et al., 2021); in *Flaveria* it would probably be easier to test whether metabolic and microbial fluctuations are still present under an elongated light period, as done by Newman et al. (2022). Gaining more insight into whether leaf commensals are diurnally regulated and if it corresponds to availability of certain metabolites in the plant could have important implications for correctly timing foliar applications of phytoprotectants and probiotics.

- According to current knowledge, one important driver of diversity in leaf microbiomes is the interaction with the host (Karasov et al., 2018). This work shows that metabolic interactions are common in the leaf microbiome and possibly also contribute to diversity.
- Plant breeding for microbiomes has proven to be a challenging task. Linking metabolites to certain kinds of interactions and thus, stable microbiomes, would open the door to breeding for these metabolites, bypassing the need to breed for enrichment of specific taxa.
- The *in-planta* experiments suggest that cross-feeding may be important for the colonization of *Pseudomonas* bacteria and that its relevance may be environment dependent. Specifically, the higher availability of certain amino acids in the *F. trinervia* leaves could lower the pressure for cross-feeding. A possible experiment to further address this hypothesis would be to add these amino acids to *F. robusta* plants and evaluate colonization of individual strains. Alternatively, the experiments could be set up with *A. thaliana* amino acid mutants (for example, At3g17390 for overexpression of methionine).
- The outcome of the microbe-microbe interactions seem to be shaped partly by the environment and partly by who takes part in the interaction. An open question is whether diversity that emerges from cross-feeding interactions in turn shape plant-microbe interactions in a resource-dependent manner. One possibility to explore this would be by bringing pathogens into the system and asking whether cross-feeding between a commensal and a pathogen would lead to attenuation of pathogenicity. This seems likely especially if the commensal, like *Pantoea* in this study, can quickly consume apoplast metabolites that can function as regulators of pathogenicity (i.e., sucrose). Understanding these mechanisms could have significant applications for biocontrol strategies.

Supplementary Information

Methods Supp Figures

Supplementary Figure 1

A



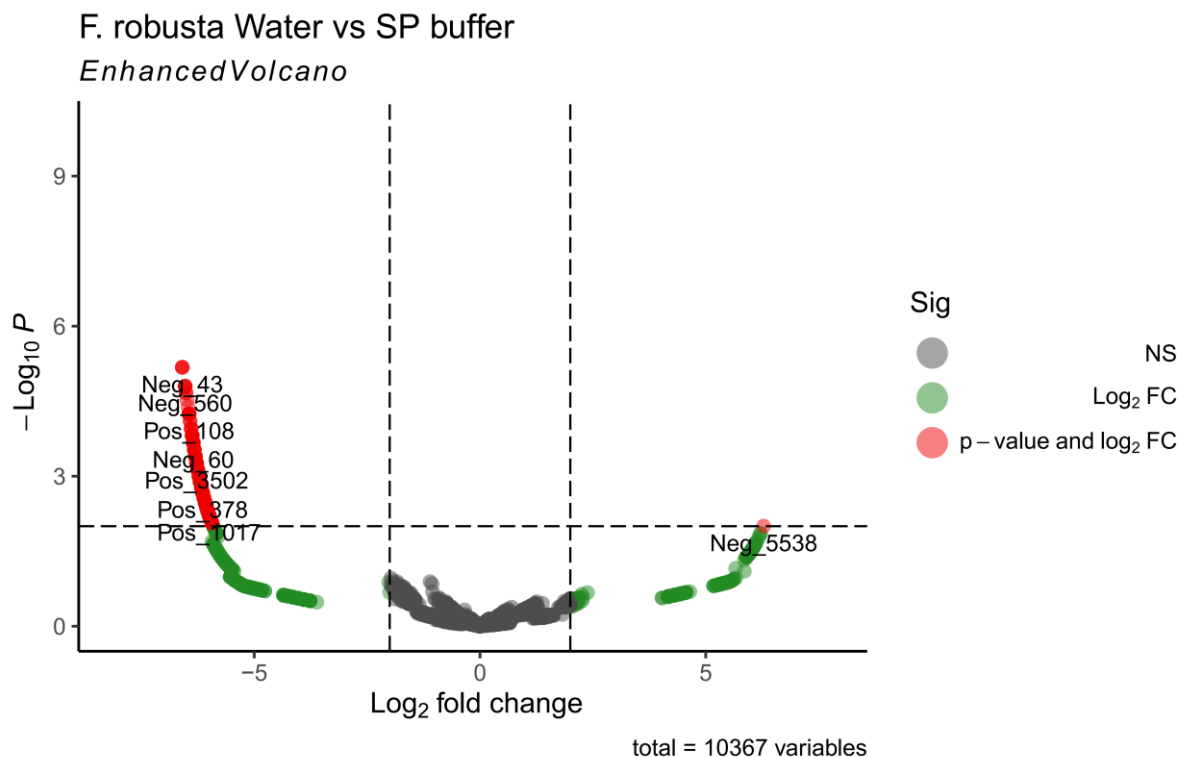
B

Plant ID	5:00 AM							1:00 PM									8:00 PM				
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	8	9	1	2	3	4	5
<i>F. linearis</i>	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
<i>F. robusta</i>	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
<i>F. trinervia</i>	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Supplementary Figure 1. A) *Flaveria* common garden established in summer 2020 to compare microbial communities and metabolic profile in leaves throughout the day. B) Planned sampling schedule for each species at each time point. Black numbers indicate plants 1-5, sampled at every time point. The colored numbers indicate plants that were not sampled at other time points.

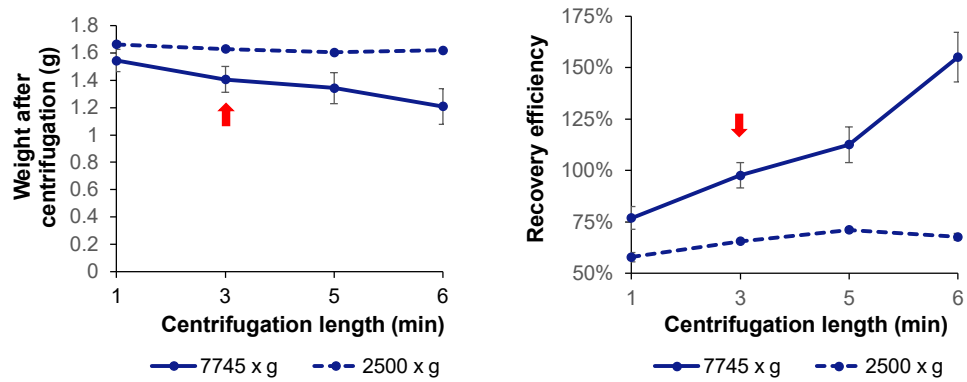
Chapter 3.1 Supp. Figures

Supplementary Figure 2



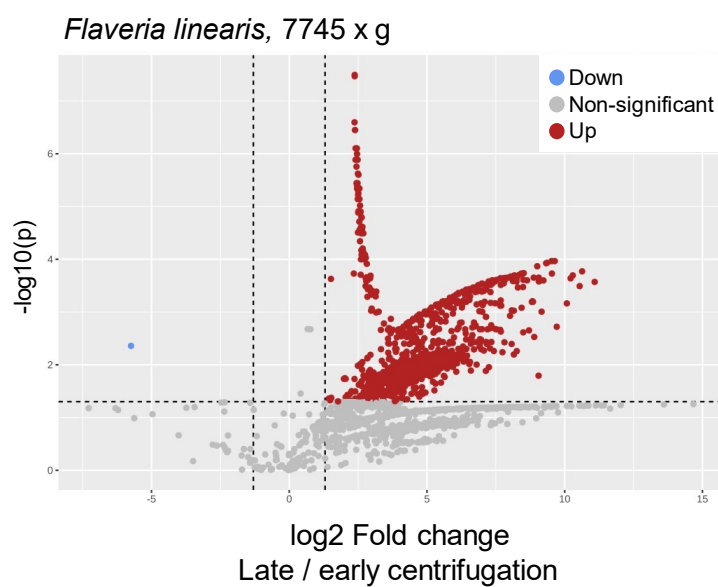
Supplementary Figure 2. Untargeted metabolomics of *Flaveria robusta* apoplast fluid wash obtained by infiltrating the leaves with either water or sodium phosphate buffer. The peak areas were compared between treatments. Vertical dashed lines indicate the cutoff of the \log_2FC $|2|$ and the horizontal line indicates the cutoff for the statistical significance (p value <0.05). Gray dots indicate peaks not significantly different between the two time points, blue dots show peaks with p value significance, green dots show peaks above the fold change threshold and red dots show peaks with significance differences and above the fold change threshold. A generalized logarithm transformation (Parsons et al, 2007) was applied to the data before analysis. The p values are based on Wald significance test.

Supplementary Figure 3



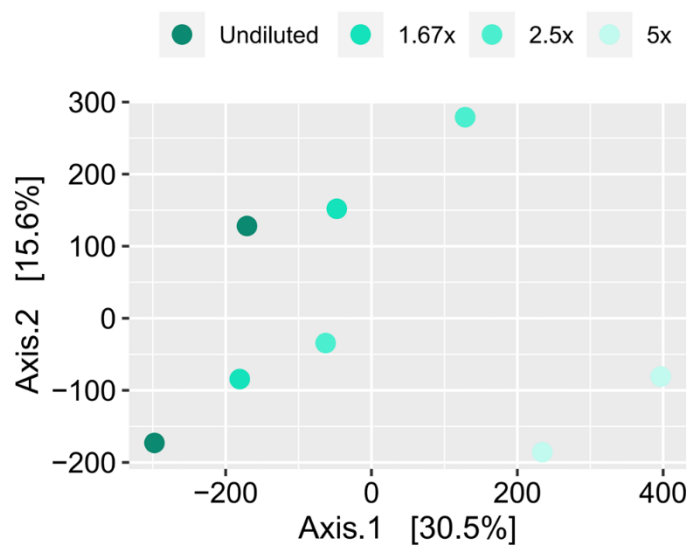
Supplementary Figure 3. Recovery of apoplast fluid wash from infiltrated *Flaveria robusta* leaves at increasing centrifugation times under either high (7745 x g) or low (2500 x g) speed. The left panel shows the weight of the leaves after each centrifugation event. The right panel shows the recovery efficiency (see Methods for the calculation details). The red arrows indicate the centrifugation event in which the weight after centrifugation equals the weight after infiltration, meaning full recovery of the infiltrated volume or 100% recovery efficiency. Each point shows the average of two replicates and the standard error.

Supplementary Figure 4



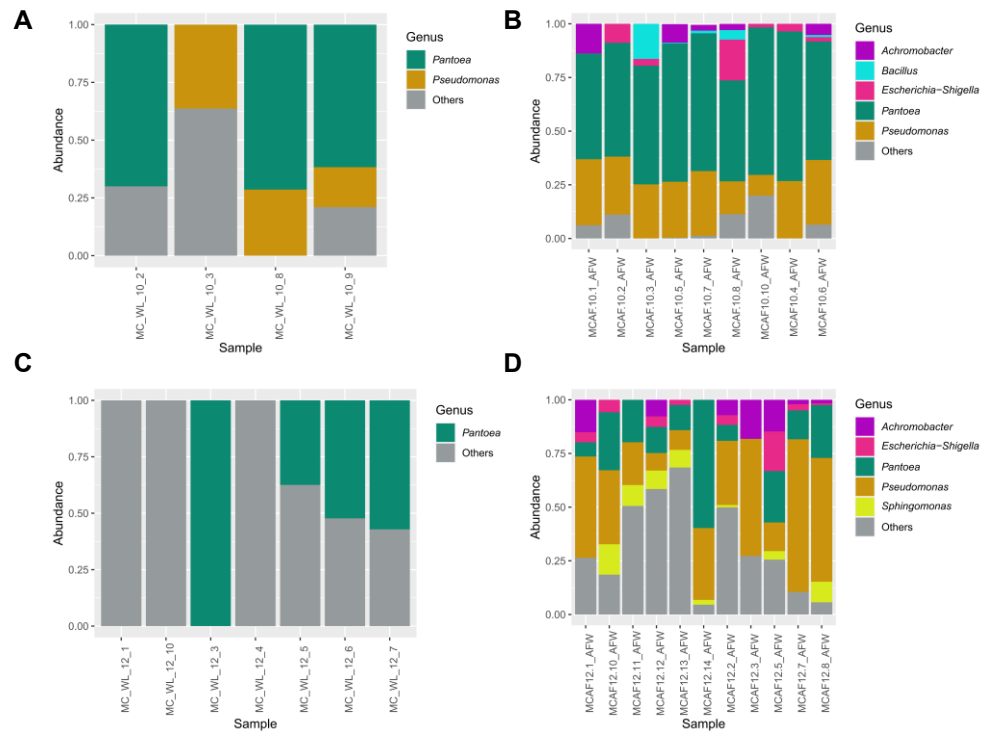
Supplementary Figure 4. Metabolic analysis of *Flaveria linearis* apoplast fluid wash obtained by centrifuging the infiltrated leaves at a high speed (7745 x g). The volcano plot shows the top 2500 peaks differentiating the groups of early (1 and 3 min) and late (30 and 60 min) time points. Blue points indicate peaks that were higher in the early group and the red points, peaks that were higher in the late group. Grey dots were not significantly different. The vertical dotted lines indicate the cutoff for \log_2FC (± 2.0) and the horizontal line indicates the cutoff for the p adjusted value (1.3).

Supplementary Figure 5



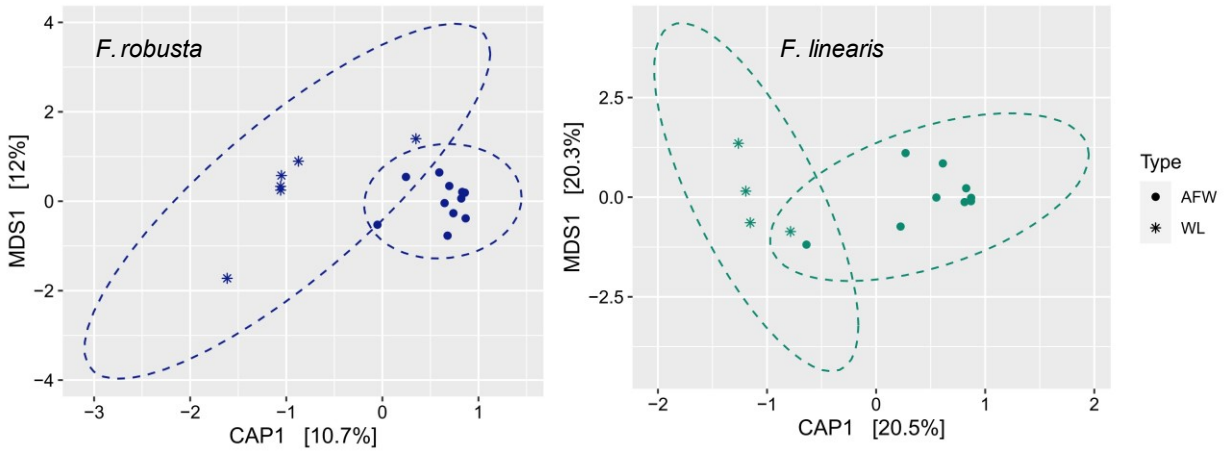
Supplementary Figure 5. MDS based on all UHPLC-HRMS peaks (area >1.0E4) in untargeted metabolomics of the recovered AFW of *F. linearis* leaves and diluted several times in sodium phosphate (SP) buffer: 1.67x= 30 μ L AFW:20 μ L SP, 2.5x= 20 μ L AFW:30 μ L SP, 5x= 10 μ L AFW:50 μ L SP. N=2 for each dilution. A generalized logarithm transformation (Parsons et al, 2007) was applied to the data.

Supplementary Figure 6



Supplementary Figure 6. Relative abundance of 16S taxa in either whole leaf (WL) or apoplast fluid wash (AFW) samples from *Flaveria linearis* (A and B) and *F. robusta* (C and D). Genera with a median abundance ≤ 0.001 were grouped under the category “Others”.

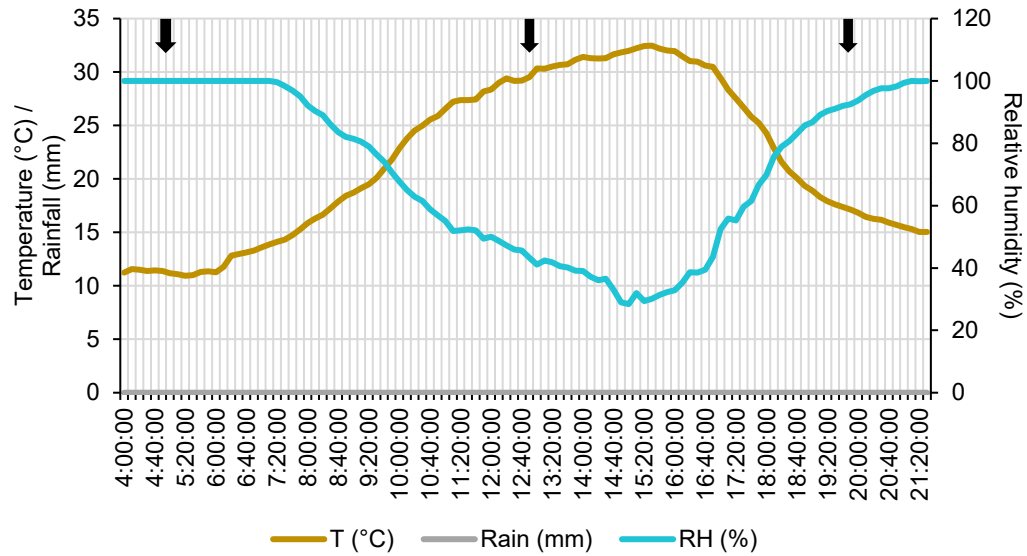
Supplementary Figure 7



Supplementary Figure 7. Principle coordinate analysis of Bray-Curtis distances, constrained by method. The left panel shows *Flaveria robusta* samples and the right panel, *F. linearis* samples. Ellipses indicate 95% confidence intervals. In both panels, A and B, only samples with bacterial reads are shown; N=11 *F. robusta* AFW, N=9 *F. linearis* AFW, N=7 *F. robusta* WL, N=4 *F. linearis* WL.

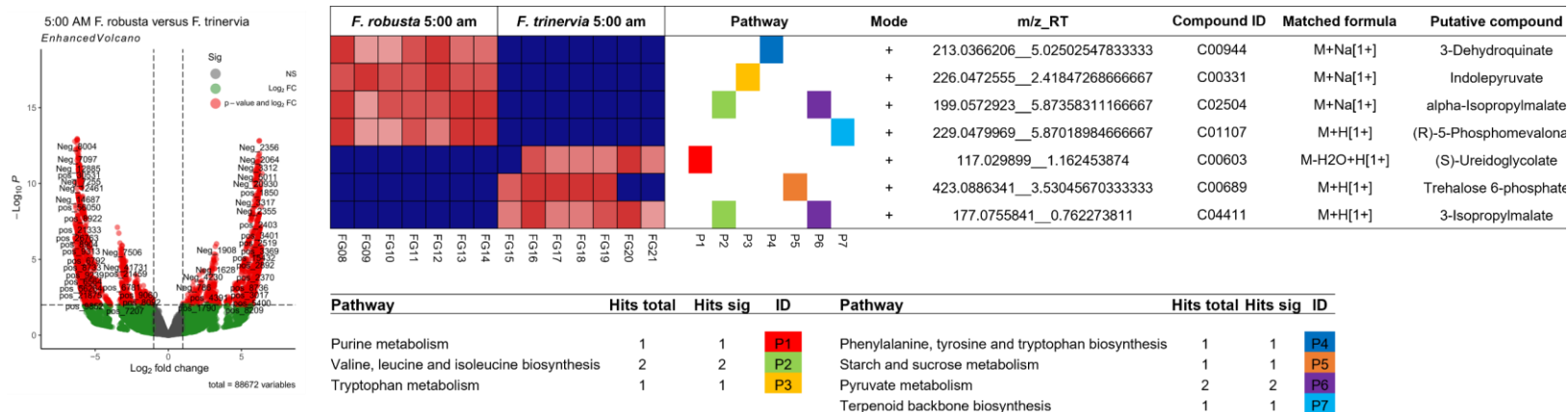
Chapter 3.2 Supp. Figures

Supplementary Figure 8



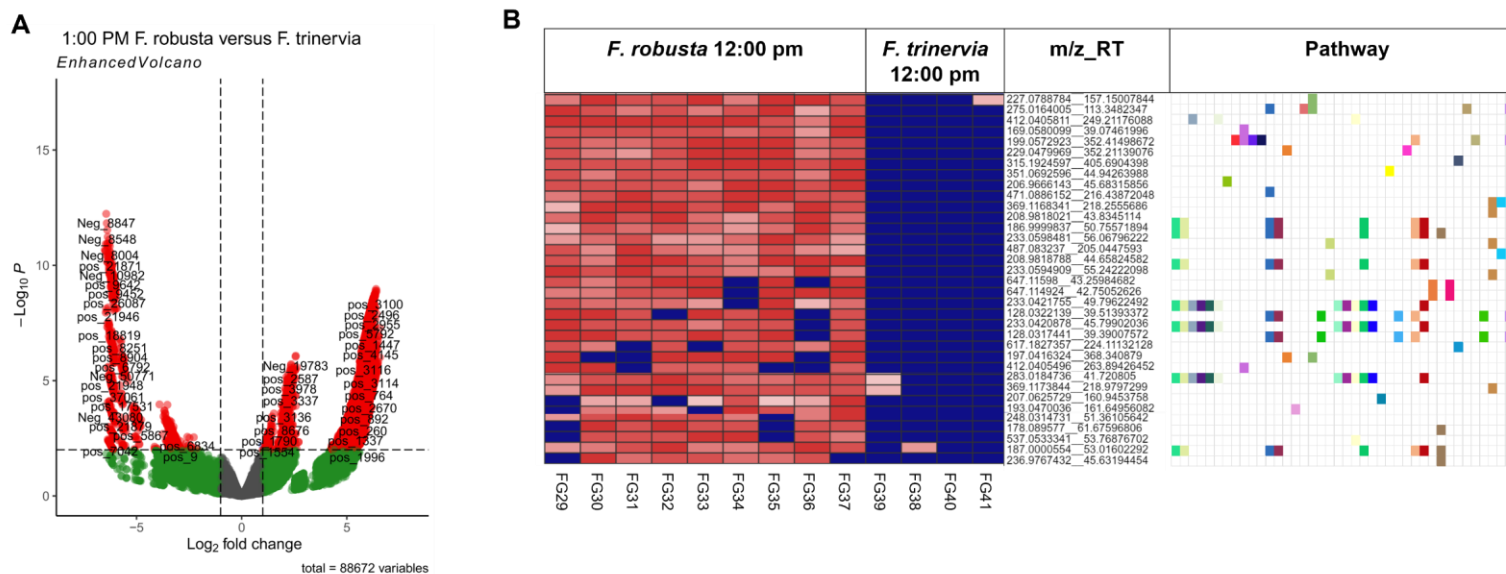
Supplementary Figure 8. Temperature (T), relative humidity (RH) and rainfall throughout the sampling day. The black arrows indicate the three sampling points at 5:00 am, 1:00 pm and 8:00 pm. Data taken from weather station Saaleaue, Max-Planck-Institut für Biogeochemie, Jena.

Supplementary Figure 9



Supplementary Figure 9. A) Differential analysis between *F. robusta* and *F. trinevia* morning samples. B) Metabolic pathways enriched in either *F. robusta* or *F. trinevia* 5:00 am samples. The heatmap shows the abundance of each compound (rows) in each sample (columns) based on Ward distance. The ionization mode, m/z value, retention time in minutes and the closest annotation is shown for each compound to the right side of the heatmap. The compounds are colored according to the pathway (P) they were related to. The table beneath the heatmap shows the pathways that were enriched with the corresponding number of hits.

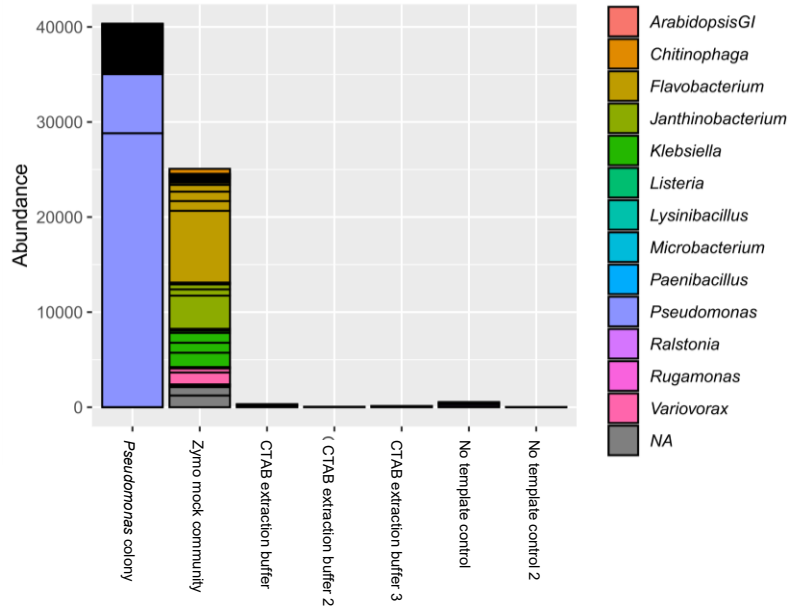
Supplementary Figure 10



Pathway	ID	Pathway	ID	Pathway	ID	Pathway	ID
Glycolysis or Gluconeogenesis	P1	Glycine, serine and threonine metabolism	P12	Amino sugar and nucleotide sugar metabolism	P23	Nicotinate and nicotinamide metabolism	P34
Pentose phosphate pathway	P2	Cysteine and methionine metabolism	P13	Glycerolipid metabolism	P24	Porphyrin and chlorophyll metabolism	P35
Pentose and glucuronate interconversions	P3	Valine, leucine and isoleucine biosynthesis	P14	Inositol phosphate metabolism	P25	Terpenoid backbone biosynthesis	P36
Fructose and mannose metabolism	P4	Tyrosine metabolism	P15	Glycerophospholipid metabolism	P26	Indole alkaloid biosynthesis	P37
Galactose metabolism	P5	Tryptophan metabolism	P16	alpha-Linolenic acid metabolism	P27	Nitrogen metabolism	P38
Ascorbate and aldarate metabolism	P6	Phenylalanine, tyrosine and tryptophan biosynthesis	P17	Sphingolipid metabolism	P28	Sulfur metabolism	P39
Ubiquinone and other terpenoid-quinone biosynthesis	P7	Selenocompound metabolism	P18	Pyruvate metabolism	P29	Phenylpropanoid biosynthesis	P40
Arginine biosynthesis	P8	Cyanoamino acid metabolism	P19	Glyoxylate and dicarboxylate metabolism	P30	Flavone and flavonol biosynthesis	P41
Purine metabolism	P9	Glutathione metabolism	P20	Carbon fixation in photosynthetic organisms	P31	Aminoacyl-tRNA biosynthesis	P43
Pyrimidine metabolism	P10	Starch and sucrose metabolism	P21	Thiamine metabolism	P32	Phosphatidylinositol signaling system	P44
Alanine, aspartate and glutamate metabolism	P11	N-Glycan biosynthesis	P22	Vitamin B6 metabolism	P33		

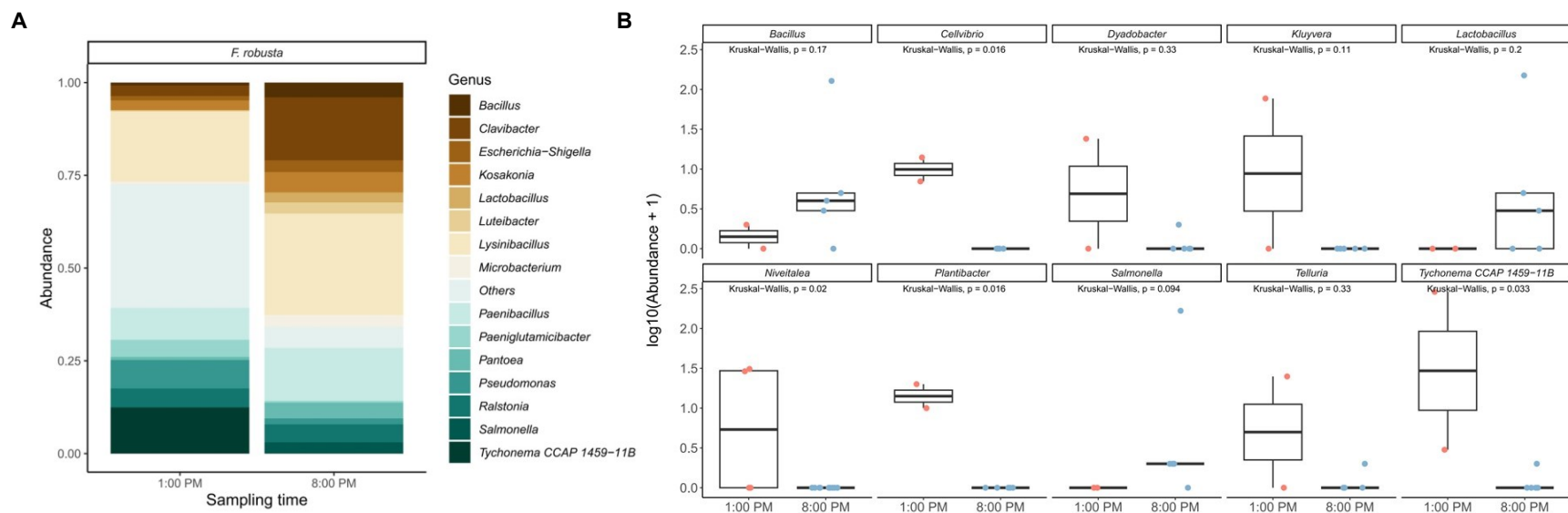
Supplementary Figure 10. A) Differential analysis between *F. robusta* and *F. trinervia* 12:00 pm samples. B) Metabolic pathways enriched in either *F. robusta* or *F. trinervia* 12:00 pm samples. The heatmap shows the abundance of each compound (rows) in each sample (columns) based on Ward distance. The m/z value and retention time in minutes is shown for each compound to the right side of the heatmap. The compounds are colored according to the pathway (P) they were related to. The table beneath the heatmap shows the pathways that were enriched.

Supplementary Figure 11



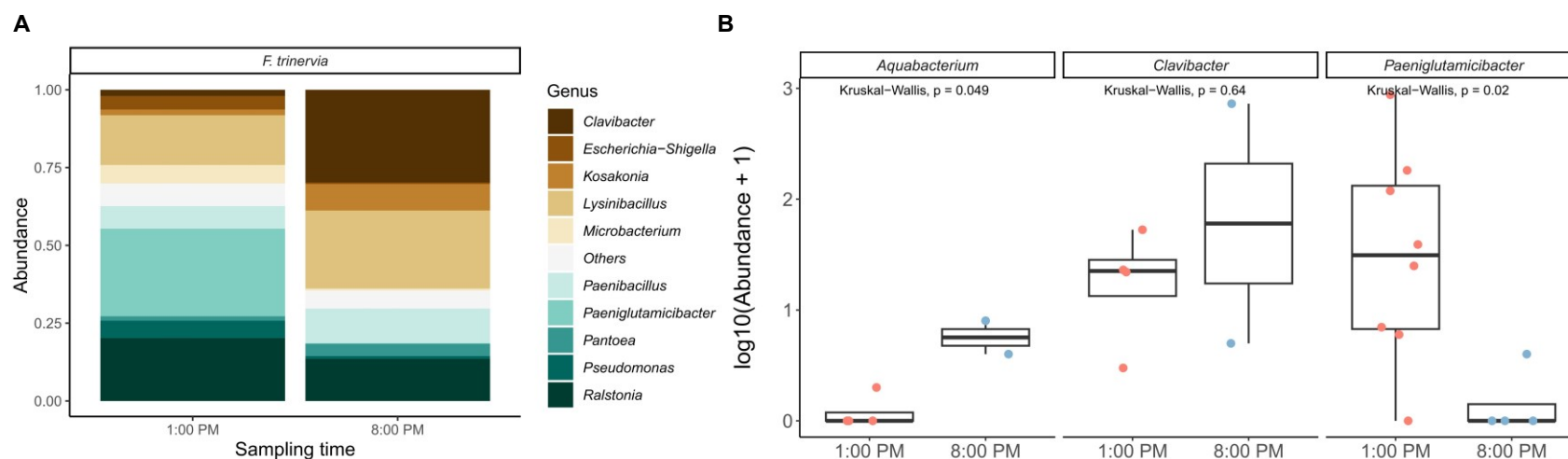
Supplementary Figure 11. Abundance of 16S taxa in the positive and negative controls of the amplicon sequencing run of the *Flaveria* diurnal samples.

Supplementary Figure 12



Supplementary Figure 12. A) Relative abundance of main bacterial genera present in *F. robusta* samples at each sampling point. The mean abundance across samples is shown. Genera which had a mean abundance of <0.1% throughout samples were grouped under “Others” category. B) Abundance of genera that had the most significant changes between sampling times (p value <0.06), according to a differential analysis (Wald-test). The points indicate the different amplicon sequence variants (ASV). The mean was compared using a Kruskal-Wallis test.

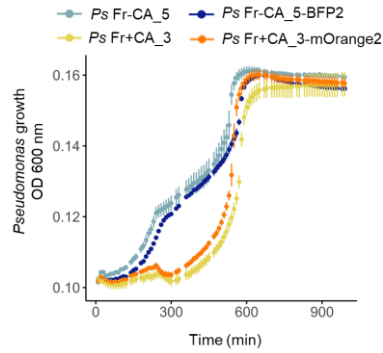
Supplementary Figure 13



Supplementary Figure 13. A) Relative abundance of main bacterial genera present in *F. trinervia* samples at each sampling point. The mean abundance across samples is shown. Genera which had a mean abundance of <0.1% throughout samples were grouped under “Others” category. B) Abundance of genera that had the most significant changes between sampling times (p value <0.06), according to a differential analysis (Wald-test). The points indicate the different amplicon sequence variants (ASV). The mean was compared using a Kruskal-Wallis test.

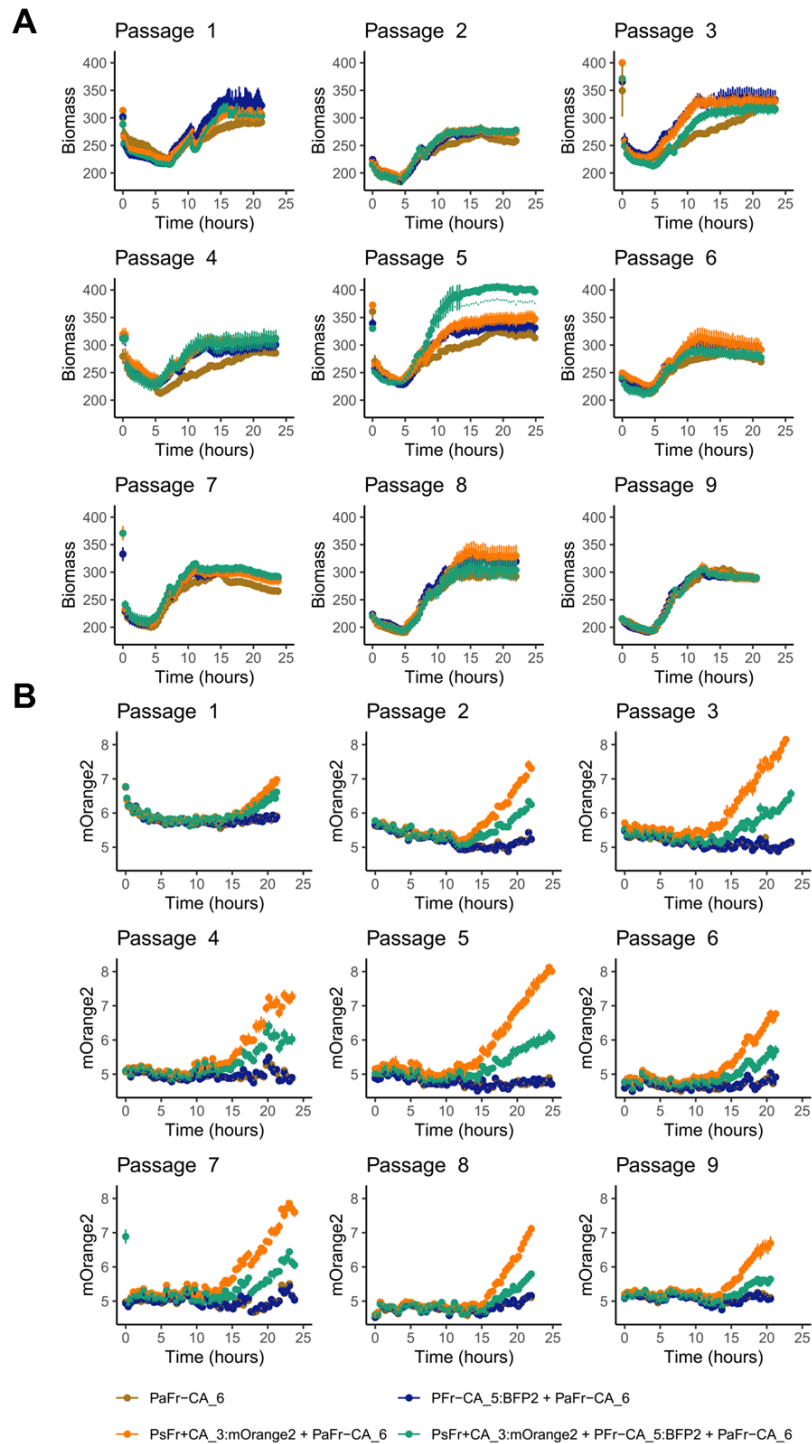
Chapter 3.3 Supp. Figures

Supplementary Figure 14



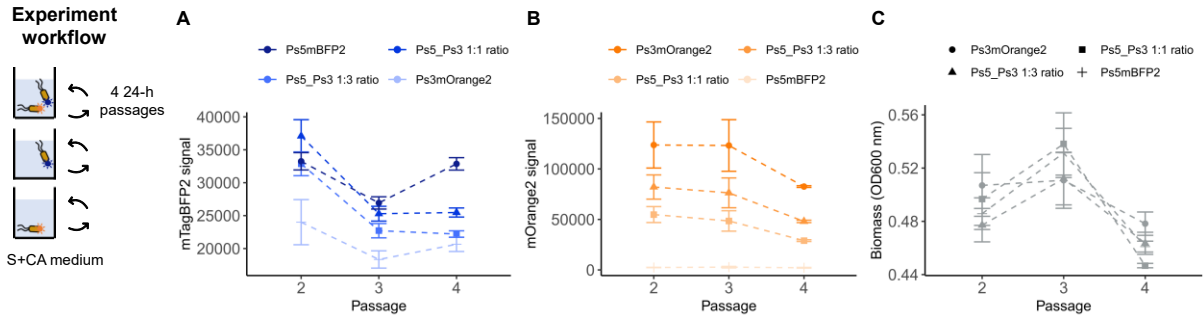
Supplementary Figure 14. Tagging did not affect the strains phenotypes. Growth of tagged versus non-tagged *Pseudomonas* (Ps) isolates on *Pantoea* (Pa) Fr-CA_6 spent media. Each data point is the average of three replicates and the corresponding standard error.

Supplementary Figure 15



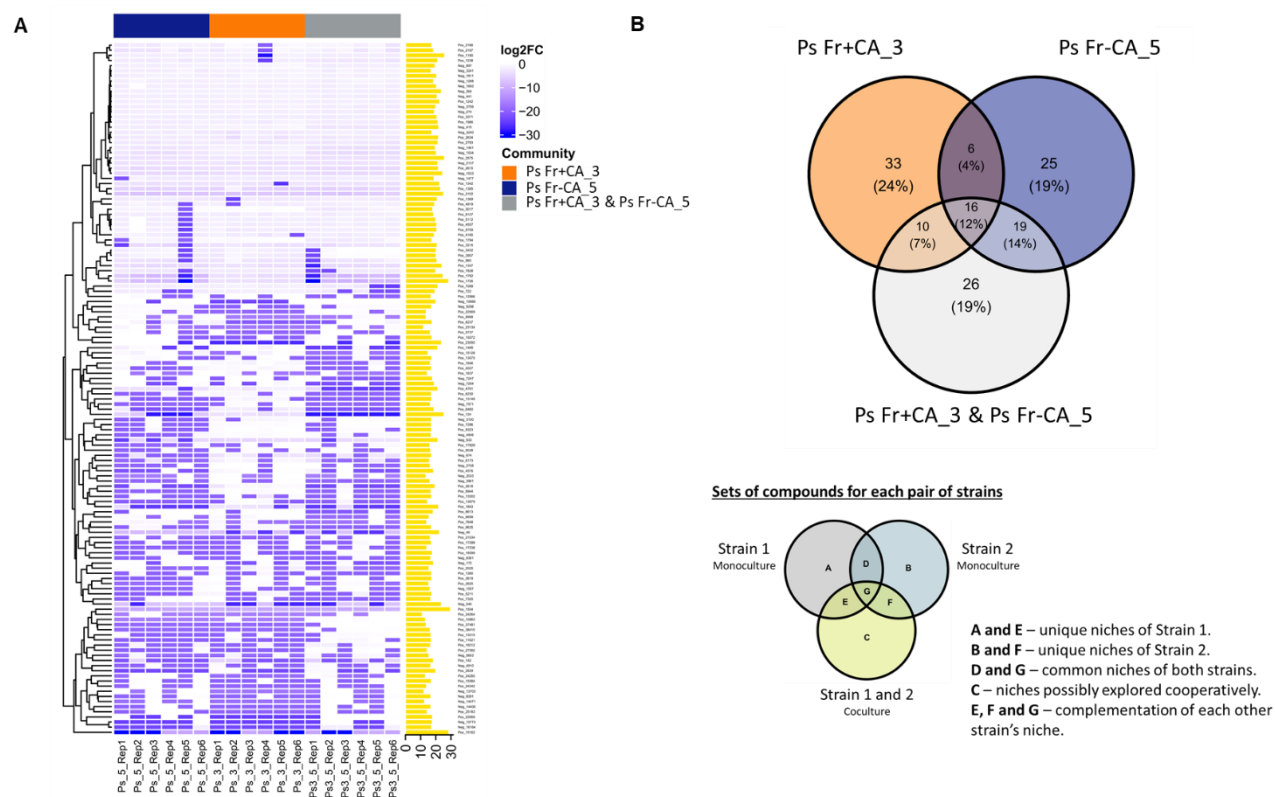
Supplementary Figure 15. A) Biomass signal in the co-cultures (each tagged *Pseudomonas* individually with *Pantoea*) and full community (both tagged *Pseudomonas* strains and *Pantoea* Fr-CA_6) throughout the passages. B) mOrange2 signal in the co-cultures and full community throughout the passages. Each data point is the average of four replicates with standard error.

Supplementary Figure 16



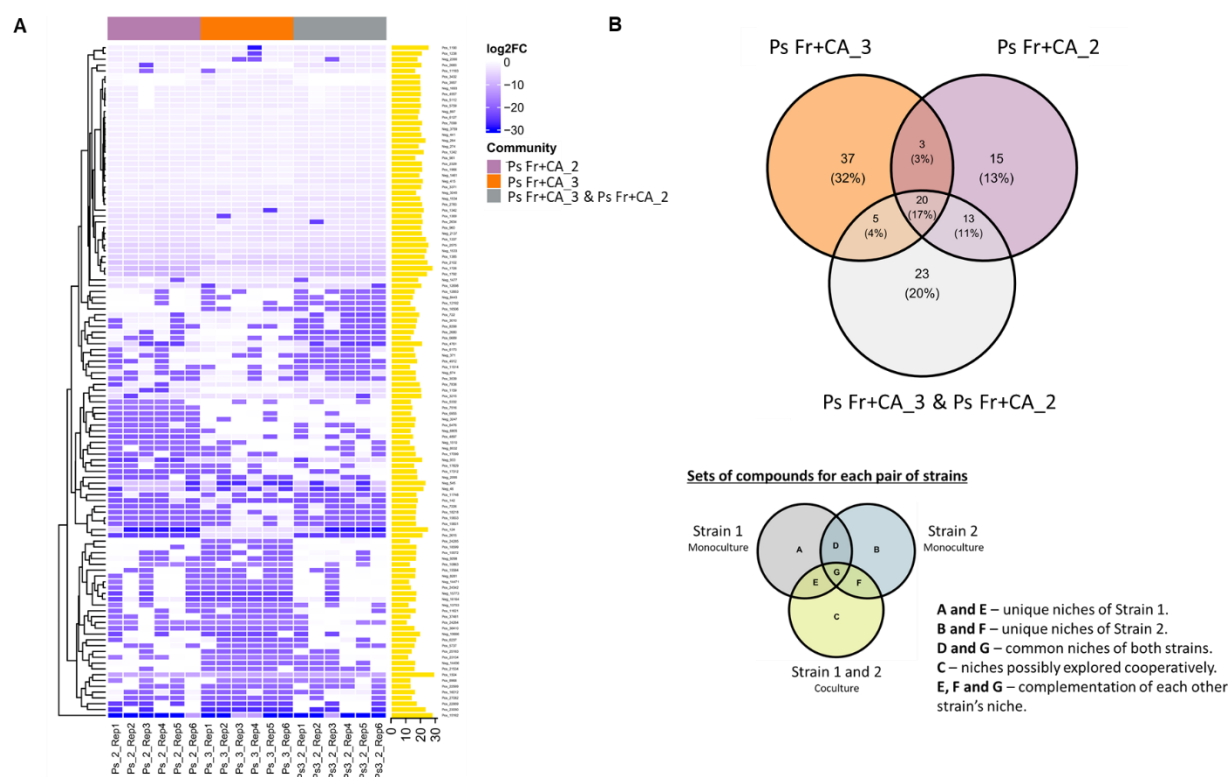
Supplementary Figure 16. Tagged *Pseudomonas* strains can coexist in S+CA medium. Fluorescence (A-B) and biomass (C) signals over four 24-h passages of the two tagged *Pseudomonas* growing on S+CA medium either in monoculture or together in two different ratios (1:1 or 1:3). In each plot data points show the average of three replicates and the corresponding standard error.

Supplementary Figure 17



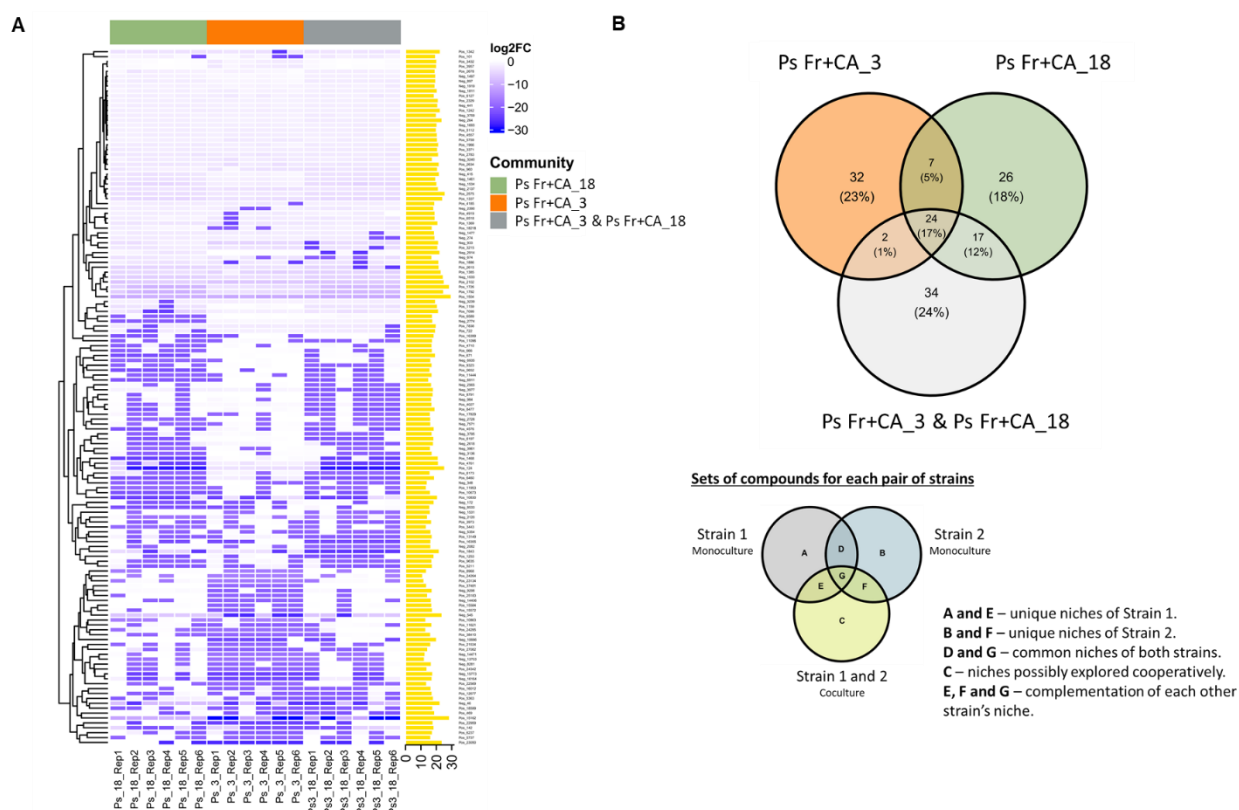
Supplementary Figure 17. Compounds taken up from the spent media of *Pantoea* Fr-CA_6 by either *P. siliginis* Fr+CA_3:mOrange2 and *Pseudomonas* Fr-CA_5:mTagBFP2 in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (\log_{10}) of the peak in the spent media of *Pantoea* Fr-CA_6 before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. B) Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown (\log_2 FC < -2, FDR < 0.05).

Supplementary Figure 18



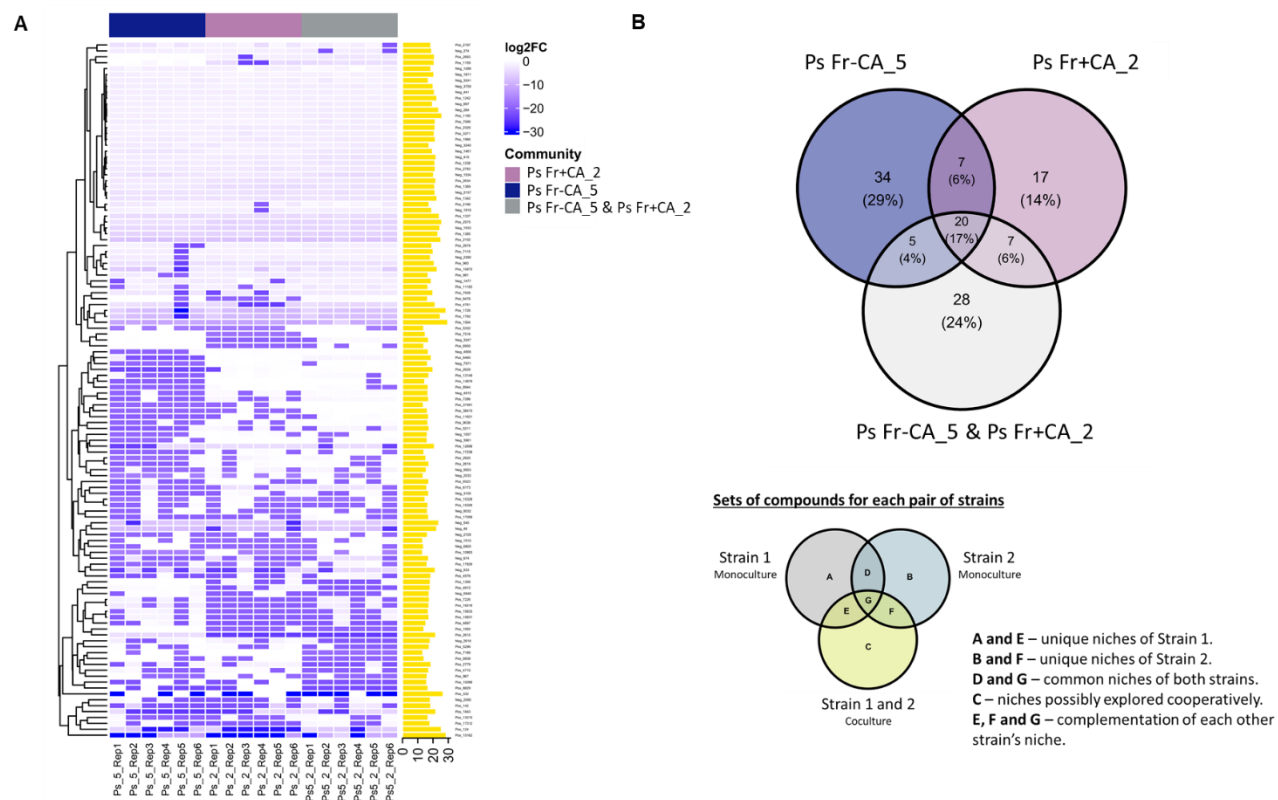
Supplementary Fig 18. Compounds taken up from the spent media of *Pantoea* Fr-CA₆ by either *P. siliginis* Fr+CA₃:mOrange2 and *Pseudomonas* Fr +CA₂ in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (log₁₀) of the peak in the spent media of *Pantoea* Fr-CA₆ before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. **B)** Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown (log₂ FC <-2, FDR<0.05).

Supplementary Figure 19



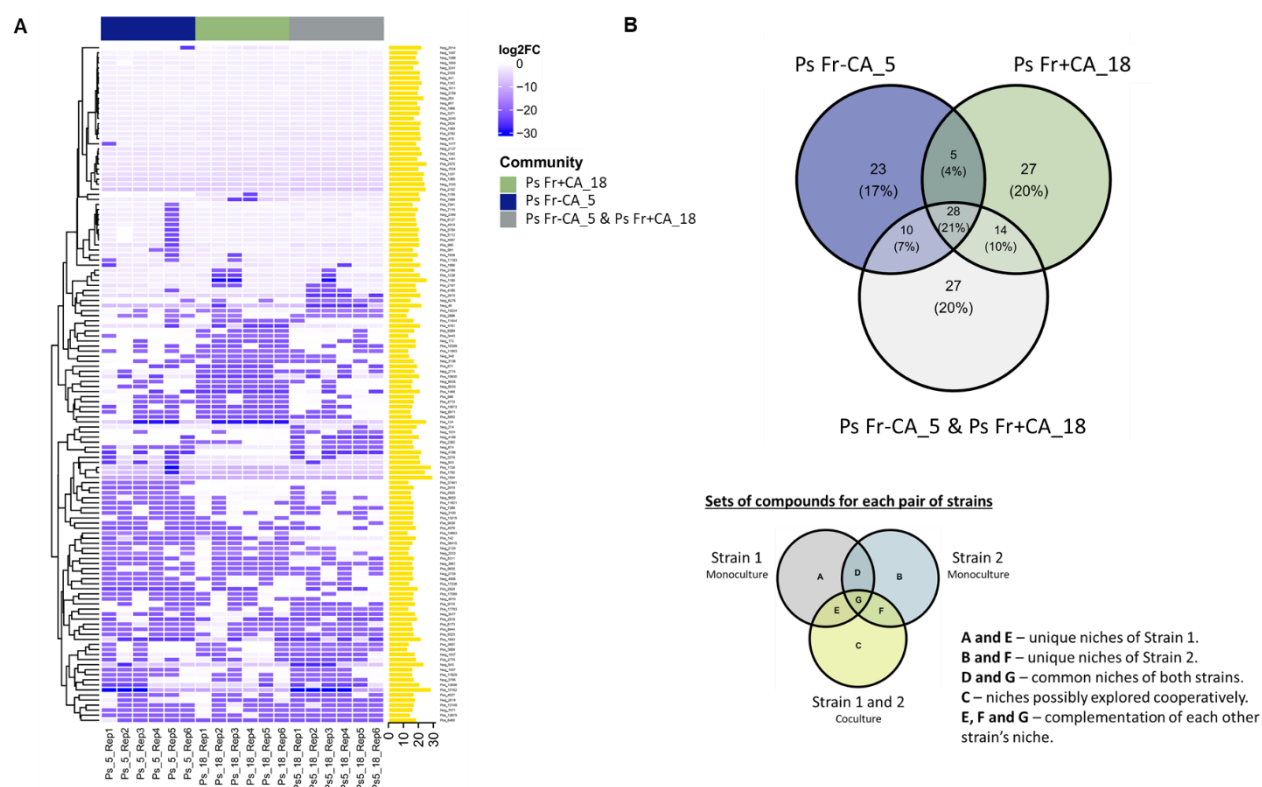
Supplementary Figure 19. Compounds taken up from the spent media of *Pantoea* Fr-CA_6 by either *P. siliginis* Fr+CA_3:mOrange2 and *P. siliginis* Fr +CA_18 in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (\log_{10}) of the peak in the spent media of *Pantoea* Fr-CA_6 before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. B) Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown ($\log_2 FC < -2$, $FDR < 0.05$).

Supplementary Figure 20



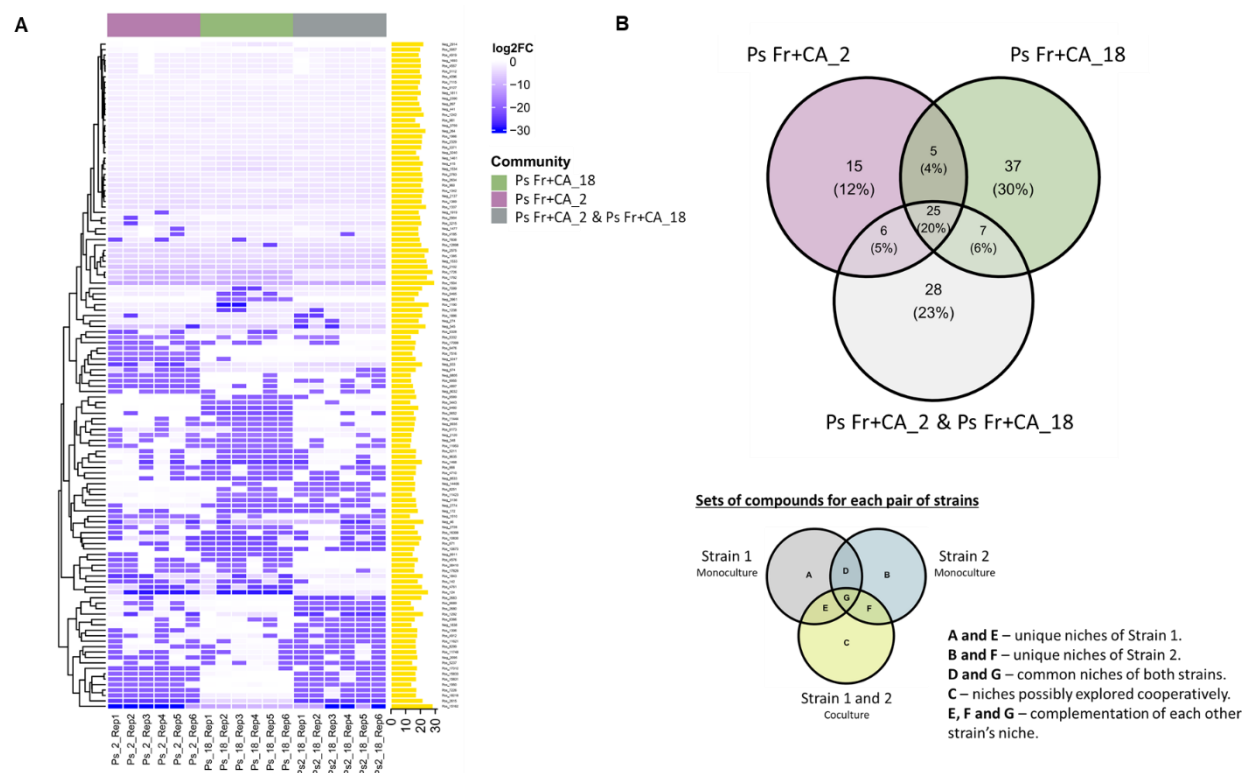
Supplementary Fig 20. Compounds taken up from the spent media of *Pantoea* Fr-CA_6 by either *P. siliginis* Fr-CA_5:mBFP2 and *P. siliginis* Fr +CA_2 in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (\log_{10}) of the peak in the spent media of *Pantoea* Fr-CA_6 before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. B) Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown (\log_2 FC < -2 , FDR < 0.05).

Supplementary Figure 21



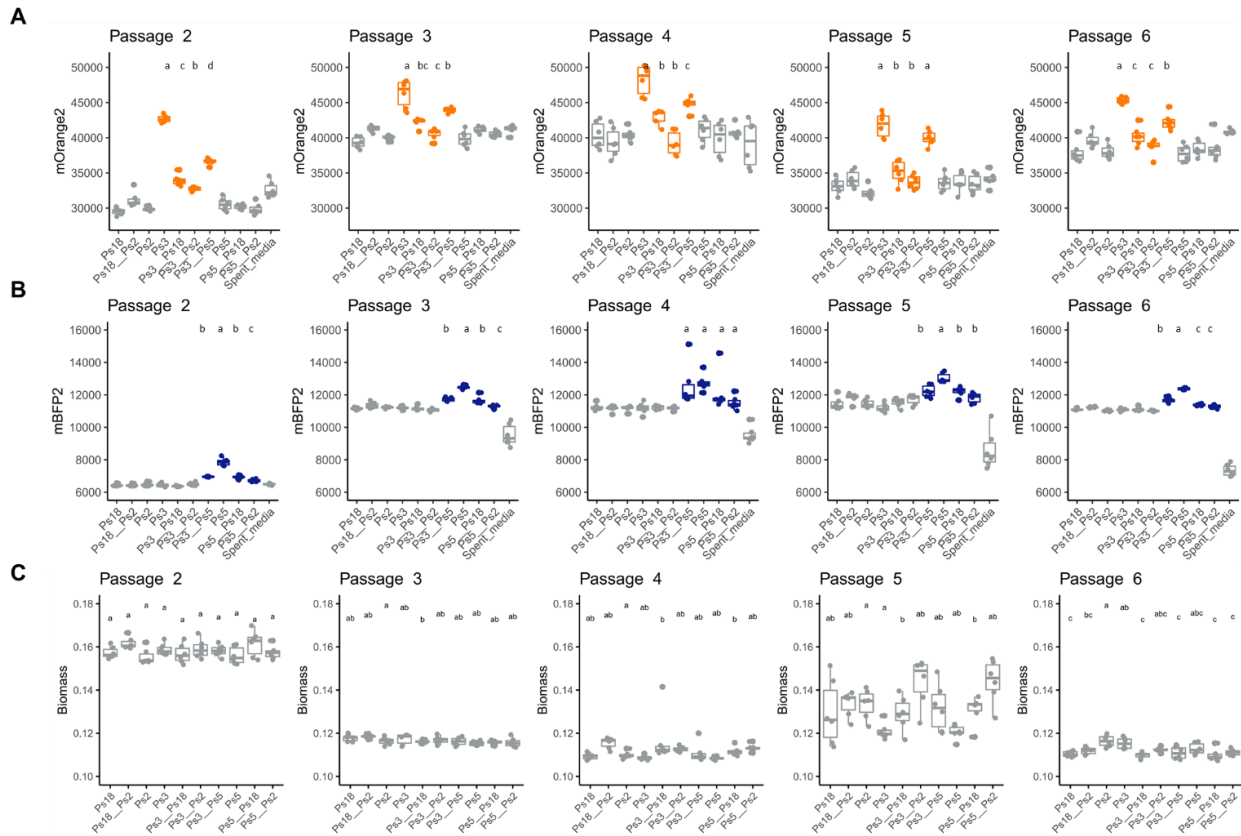
Supplementary Figure 21. Compounds taken up from the spent media of *Pantoea* Fr-CA_6 by either *P. siliginis* Fr-CA_5:mTagBFP2 and *P. siliginis* Fr +CA_18 in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (\log_{10}) of the peak in the spent media of *Pantoea* Fr-CA_6 before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. B) Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown (\log_2 FC < -2, FDR < 0.05).

Supplementary Figure 22



Supplementary Fig 22. Compounds taken up from the spent media of *Pantoea* Fr-CA₆ by either *P. siliginis* Fr+CA₂ and *P. siliginis* Fr +CA₁₈ in monoculture, or both strains growing in co-culture. A) Heatmap of taken up compounds. Each column represents a different repetition (N=6). The strains and peaks have been clustered by Euclidean distance. The yellow bars on the right indicate the area (log₁₀) of the peak in the spent media of *Pantoea* Fr-CA₆ before growth of the *Ps*. Each peak is identified with a unique number; for information of each one (ionization mode, retention time, m/z value and putative annotation) see Supplementary Table 21. B) Venn Diagram showing the sets of compounds taken up by each strain alone and when co-cultured. An explanation for easier interpretation of the different areas of the diagram is provided. In both panels A and B only metabolites with significant uptake are shown (log₂ FC < -2, FDR < 0.05).

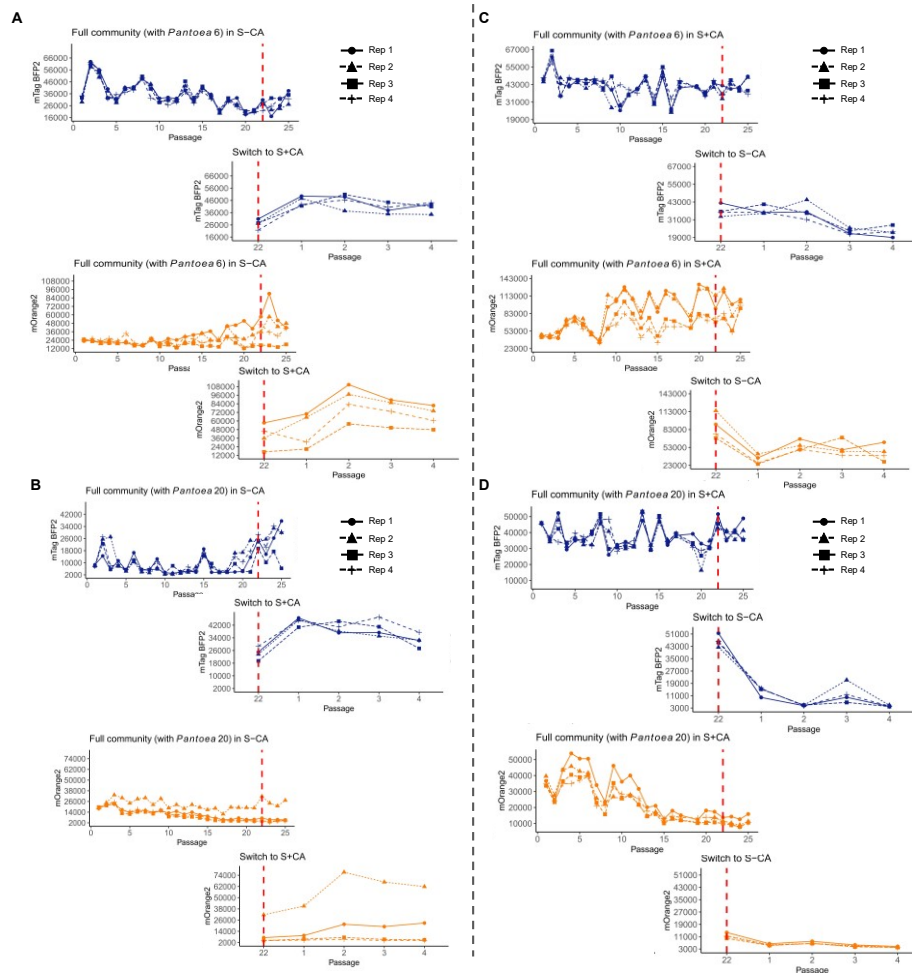
Supplementary Figure 23



Supplementary Figure 23. Growth of tagged *Pseudomonas* (*Ps*) strains was reduced when co-cultured with the other strains, but overall biomass remained constant. Fluorescence and biomass measurements throughout the passages for each community of *Pseudomonas* (*Ps*). A) mOrange2 signal; only the communities including *Ps* Fr+CA_3:mOrange2 are colored in orange. B) mBFP2 signal; only the communities including *Ps* Fr-CA_5:mTagBFP2 are shown in blue. C) Biomass signal (OD 600 nm) for each community. In each plot different letters indicate significant differences after a Tukey test, N=6.

Chapter 3.4 Supp. Figures

Supplementary Figure 24



Supplementary Figure 24. Interactions between *Pantoea* and *Pseudomonas* followed distinct paths in each environment. Passaging of the two tagged *Pseudomonas* (*Ps*) strains with either *Pantoea* (*Pa*) Fr-CA_6 or *Pa* Fr+CA_20 on S-CA or S+CA medium. Fluorescence signals (mTagBFP2 and mOrange2) in each community are shown. Panels A and B show the communities passaged from S-CA to S+CA with either with *Pantoea* Fr-CA_6 or *Pantoea* Fr+CA_20 respectively. Panels C and D show the communities passaged from S+CA to S-CA with either *Pantoea* Fr-CA_6 or *Pantoea* Fr+CA_20, respectively. The red vertical line marks passage 22, when the communities were switched to the opposite media to evaluate the effects of their adaptation. After the switch, the fluorescence was monitored for four more passages. Each condition had four replicates; all the raw values are shown in the figures.

Methods Supp.Tables

Supplementary Table 1. Details of the plant species used in this study.

Species	Family	Native to*	Photosynthesis type	Source	Collection location	Collected by	Propagation method
<i>Alternanthera caracasana</i>	Amaranthaceae	Mexico to Tropical America	C4	Rowan Sage	Alpine, Texas	AM Powell	Cuttings
<i>Alternanthera sessilis</i>	Amaranthaceae	Tropical and Subtropical Asia to N. & E. Australia, S. Mexico to Tropical	C3	Rowan Sage		A. Ragavendra	Seeds/cuttings
<i>Alternanthera ficoidea</i> syn:	Amaranthaceae	Tropical America	C3/C4	Rowan Sage		A. Ragavendra	Cuttings
<i>Flaveria linearis</i>	Asteraceae	Florida, Bahamas, Cuba, Mexico to Belize	C3/C4	Rowan Sage	Estero Blvd, FL	Rowan Sage	Cuttings
<i>Flaveria trinervia</i>	Asteraceae	Tropical & Subtropical America	C4	Rowan Sage	Plantation Key, FL	Rowan Sage	Seeds/cuttings
<i>Flaveria robusta</i>	Asteraceae	Mexico	C3	Rowan Sage	Summit of Colima to	Rowan Sage	Cuttings
<i>Zea mays</i>	Poaceae	Guatemala, Mexico	C4	Quedlinburger Saatgut	Commercial seeds		Seeds

*KEW Royal Botanic Gardens: <https://powo.science.kew.org/>

Supplementary Table 2. Concentration of deuterated amino acids in CELL FREE AMINO ACID MIX (U-D, 98%), from Cambridge Isotope Laboratories, Inc. (MA, USA). Lot# PR-29964

Deuterated amino acid	Weight percentage (m/m)
Alanine_D	6.80%
Arginine_D	3.10%
Asparagine_D	4.40%
Aspartate_D	4.90%
Glutamine_D	4.50%
Glutamic Acid_D	10.30%
Histidine_D	1.20%
Lysine_D	8.70%
Leucine_D Isoleucine_D	18.90%
Methionine_D	1.90%
Phenylalanine_D	6.30%
Proline_D	2.60%
Serine_D	3.80%
Threonine_D	5.10%
Tryptophan_D	2.60%
Tyrosine_D	3.00%
Valine_D	6.50%

Supplementary Table 3. Recipes of growth media used in this work.

M9 Minimal media					R2A broth		R2A broth + S-CA				LB broth	
Component	Enrichment preculture	S-CA	S+CA	+ Glucose	Component	g/L	Component	1:0:0.8	1:0.2:0.6	0:0.2:1.6	Component	g/L
M9 Salts*	1 X	1 X	1 X	1 X	Yeast extract	0.5	S-CA	56%	56%	0%	Peptone	10
NH ₄ Cl	200 mM	33 mM	33 mM	33 mM	Peptone	0.5	R2A broth	0%	11%	11%	Yeast extract	5
Sucrose	11 mM	11 mM	11 mM	0 mM	Casamino acids	0.5	Water	44%	33%	89%	NaCl	5
Glucose	0 mM	0 mM	0 mM	22 mM	Glucose	0.5						
MgSO ₄	2 mM	2 mM	2 mM	2 mM	Starch	0.5						
CaCl ₂	0.2mM	0.2mM	0.2mM	0.2mM	Na-pyruvate	0.3						
Biotin	1 µg/mL	1 µg/mL	1 µg/mL	1 µg/mL	K ₂ HPO ₄	0.3						
Thiamine	1 µg/mL	1 µg/mL	1 µg/mL	1 µg/mL	MgSO ₄ ·7H ₂ O	0.05						
Trace elements**	1 X	1 X	1 X	1 X	pH adjusted to 7.2							
Casamino acids	0.20%	0%	0.20%	0%								
Cycloheximide	200 µg/mL	0 µg/mL	0 µg/mL	0 µg/mL								

*M9 salts - for 10X stock: 75.2 g/L Na₂HPO₄ · 2H₂O, 30 g/L KH₂PO₄, 5 g/L NaCl

**Trace elements - for 100x stock: 13.4 mM EDTA, 3.1 mM FeCl₃·6H₂O, 0.62 mM ZnCl₂, 76 µM CuCl₂·2H₂O, 42 µM CoCl₂·2H₂O, 162 µM H₃BO₃, 8.1 µM MnCl₂·4H₂O

Supplementary Table 4. Oligonucleotides used in this study.

Name	Target	Sequence 5' to 3'	References
B341F	16S rRNA V3/V4	CCTACGGGNGGCWGCAG	(Muyzer et al., 1993)/(Caporaso et al., 2010)
B806R	16S rRNA V3/V4	GGACTACHVGGGTWTCTAAT	
B341F-OH*	16S rRNA V3/V4	CCTACGGGNGGCWGCAGACACTCTTTCCCTACACGACGCTCTTC	
B799R-OH*	16S rRNA V3/V4	AACMGGATTAGATACCCGGGTGACTGGAGTTCAGACGTGTGCTC	(799R (Bulgarelli et al., 2012))
BLc_16S_F5	Blocks chloroplast in plants	AACTTCTTTTCYHGGAGAAGAARCAAT	(Mayer et al., 2021)
BLc_16S_R1	Blocks chloroplast in plants	GCTTTCGCCRTTGGTGTTCYTTCCGATCTC	
8F	16S rRNA	AGAGTTTGATCCTGGCTCAG	(Turner et al., 1999)
1492R	16S rRNA	GGTTACCTTGTTACGACTT	
FWD Tn5/7_gt	Verifying the Tn7 insertion	ATGGTGAGCAAGGGCGAG	(Schlechter et al., 2018)
REV_Tn5/7_gt	Verifying the Tn7 insertion	CAACAGGAGTCCAAGCTCAG	
Tn7_gt1	Genotyping the Tn7 backbone	GAATTACAACAGTACTGCGATGAG	
Tn7_gt2	Genotyping the Tn7 backbone	GATCAACTCTATTTCTCGCGGG	
Tn7_gt3	Genotyping the Tn7 backbone	TACATAACGGACTAAGAAAAACTACAC	
FWD_uidA	<i>uidA</i> gene (<i>E. coli</i>)	AACAGGTGGTTGCAACTGGA	
REV_uidA	<i>uidA</i> gene (<i>E. coli</i>)	TTGCTGAGTTTCCCCGTTGA	

*Sequence shows the Universal 16S Primer and the overhang (OH) region.

Supplementary Table 5. Volume (μL) of bacterial suspension of *Pantoea* (Pa) and *Pseudomonas* (Ps) added as inoculum per well in each of the co-culture assays

Experiment	Community	Pa Fr-CA_6	Pa Ft-CA_14	Ps FrCA_5:mBFP2	Ps Fr+CA_3:mOrange2	Ps Fr+CA_2	Ps Fr+CA_18	1x PBS
Growth of tagged <i>Ps</i> and <i>Pa</i> Fr-CA_6 in S-CA media over 8 and 4 passages (24 and 48h, respectively).	Pa Fr-CA_6 + Ps Fr-CA_5:BFP2	10		3.3	---	---	---	6.6
	Pa Fr-CA_6 + Ps Fr+CA_3:mOrange2	10		---	6.6	---	---	3.3
	Pa Fr-CA_6 + Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2	10		3.3	6.6	---	---	---
	Pa Fr-CA_6	10		---	---	---	---	9.9
Growth of tagged <i>Ps</i> in either monoculture or co-culture in S+CA over 4 24h-passages	Ps Fr-CA_5:BFP2 monoculture	---		10	0	---	---	10
	Ps Fr+CA_3:mOrange2 monoculture	---		0	10	---	---	10
	Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2 (1:1 ratio)	---		10	10	---	---	---
	Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2 (1:3 ratio)	---		5	15	---	---	---
Growth of <i>Ps</i> in monoculture or in pairs on PaFr-CA_6 spent media over 6 24-h passages	Ps Fr+CA_3:mOrange2 monoculture	---		---	10	---	---	10
	Ps Fr-CA_5:BFP2 monoculture	---		10	---	---	---	10
	Ps Fr+CA_2 monoculture	---		---	---	10	---	10
	Ps Fr+CA_18 monoculture	---		---	---	---	10	10
	Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2	---		10	10	---	---	---
	Ps Fr+CA_2 + Ps Fr+CA_3:mOrange3	---		---	10	10	---	---
	Ps Fr+CA_18 + Ps Fr+CA_3:mOrange3	---		---	10	---	10	---
	Ps Fr+CA_2 + Ps Fr-CA_5:BFP2	---		10	---	10	---	---
	Ps Fr+CA_18 + Ps Fr-CA_5:BFP2	---		10	---	---	10	---
Ps Fr+CA_2 + Ps Fr+CA_18	---		---	---	10	10	---	
Tagged <i>Ps</i> with either <i>Pa</i> Fr-CA_6 or <i>Pa</i> Ft-CA_14 in either S+CA or S-CA for 25 24-h passages	Pa Fr-CA_6 + Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2	10	---	5	5	---	---	---
	Pa Ft-CA_14 + Ps Fr-CA_5:BFP2 + Ps Fr+CA_3:mOrange2	---	10	5	5	---	---	---

Supplementary Table 6. Two-step PCR conditions to amplify bacterial communities from different sets of samples.

Sample set	1. AFW vs WL	2. AFW from <i>Flaveria diurnal</i> sampling	3. Enrichments in minimal media (Passage 12)	4. Bacterial isolates from Enrichments in minimal media
1st step PCR				
Bacterial primers	B341/B806	B341-Overhang/B806-Overhang	B341F/B806R	8F/1492R
Reagents concentrations	40 µL reaction: 1x KAPA GC Buffer, 0.3 mM KAPA dNTP, 0.075 µM primers, 0.02 U KAPA HiFi polymerase, 1 µL template.	10 µL reaction: 1.25x KAPA Buffer GC, 0.375 mM dNTP, 0.1 µM primers, 0.025 U KAPA HiFi polymerase, 2 µL template	25 µL reaction: 1.0 KAPA Buffer GC, 0.3 mM dNTP, 0.075 µM primers, 0.02 U KAPA HiFi polymerase, 1 µL template	37.2 µL reaction: 1.08 µL KAPA Buffer GC, 0.32 mM dNTPs, 0.27 µM primers, 0.024 U KAPA HiFi polymerase, 1 µL template
PCR conditions	3 min denaturing at 95 °C. 15 cycles of denaturation at 98 °C 20 sec, annealing at 55 °C 30 sec, elongation at 72 °C for 40 sec. Final extension at 72 °C for 2 min.	3 min denaturing at 95 °C. 30 cycles of: denaturation at 98 °C 20 sec, annealing at 55 °C 1 min, elongation at 72 °C for 2 min. Final extension at 72 °C for 2 min.	3 min denaturing at 95 °C. 15 cycles of denaturation at 98 °C 20 sec, annealing at 55 °C 30 sec, elongation at 72 °C for 40 sec. Final extension at 72 °C for 2 min.	2 min denaturing at 95 °C. 30 cycles denaturation at 95 °C 30 sec, annealing at 50 °C for 30 sec, elongation at 72 °C for 1:30 min. Final extension at 72 °C for 5 min.
Enzymatic cleanup				
Reagents concentrations	0.5 µL Exonuclease, 0.5 µL Antartic phosphatase, 1.22 µL Antartic phosphatase buffer, 10 µL template	0.5 µL Exonuclease, 0.5 µL Antartic phosphatase, 1.22 µL Antartic phosphatase buffer, 10 µL template	0.5 µL Exonuclease, 0.5 µL Antartic phosphatase, 1.22 µL Antartic phosphatase buffer, 10 µL template	
Incubation conditions	37 °C for 30 min, 80 °C for 15 min.	37 °C for 30 min, 80 °C for 15 min.	37 °C for 30 min, 80 °C for 15 min.	
2nd step PCR				
Concatenated primers	p5 or p7+ Illumina+ Index+ Heterogeneity spacer + 16S Standard primer	p5 or p7+ Index+ Overhang	p5 or p7+ Illumina+ Index+ Heterogeneity spacer + 16S Standard primer	
Reagents concentrations	15 µL reaction: 1x KAPA GC Buffer, 0.3 mM KAPA dNTP, 0.3 µM primers, 0.02 U KAPA HiFi polymerase, 2 µL template	In 20 µL reactions: 1.0x KAPA Buffer GC, 0.3 mM dNTP, 0.3 µM primers, 0.02 U KAPA HiFi polymerase, 10 µL template	15 µL reaction: 1.0 KAPA Buffer GC, 0.3 mM dNTP, 0.3 µM primers, 0.02 U KAPA HiFi polymerase, 1 µL template	
PCR conditions	3 min denaturing at 95 °C. 30 cycles of denaturation at 98 °C 20 sec, annealing at 55 °C 30 sec, elongation at 72 °C for 40 sec. Final extension at 72 °C for 2 min.	3 min denaturing at 95 °C. 5 cycles of denaturation at 98 °C 20 sec, annealing at 60 °C 1 min, elongation at 72 °C for 1 min. Final extension at 72 °C for 2 min.	3 min denaturing at 95 °C. 25 cycles of: denaturation at 98 °C 20 sec, annealing at 55 °C 30 sec, elongation at 72 °C for 40 sec. Final extension at 72 °C for 2 min.	

Chapter 3.1 Supp. Tables

Supplementary Table 7. Distribution of apoplast fluid wash and cross-feeding samples in UHPLC-HRMS runs.

Set of samples	Date of HPLC run
<u>Apoplast fluid wash (AFW) samples</u>	
Different plant species. Infiltrated with either MES or sodium phosphate buffer.	23.05.2019
<i>F. linearis</i> infiltrated with sodium phosphate buffer. Centrifuged at 7500 x g for 1, 3, 5, 10, 15, 20, 30 and 60 min.	05.07.2019
<i>F. robusta</i> . Water vs sodium phosphate infiltration.	31.01.2020
<i>F. linearis</i> infiltrated with sodium phosphate buffer. Centrifuged at 2500 x g for 1, 3 and 5 min.	31.01.2020
<i>F. linearis</i> infiltrated with sodium phosphate buffer. Undiluted vs diluted samples.	31.01.2020
<i>F. robusta</i> , <i>F. trinervia</i> and <i>F. linearis</i> plants grown in common garden and sampled throughout the day. Optimized protocol.	11.06.2020
<i>F. robusta</i> , <i>F. trinervia</i> and <i>F. linearis</i> plants grown in lab. Optimized protocol.	20.07.2021
<u>Cross-feeding samples</u>	
Different <i>Pseudomonas</i> isolates in <i>Pantoea</i> Fr-CA_6 spent media	18.12.2019
<i>Pseudomonas</i> Fr-CA_5 in spent media of different <i>Pantoea</i> isolates	11.06.2020
<i>Pseudomonas</i> Fr-CA_5 in spent media of leaf extract isolates	11.06.2020
<i>Pseudomonas</i> isolates grown in monoculture or co-culture in <i>Pantoea</i> Fr-CA_6 spent media.	25.02.2022

Supplementary Table 10. Number of reads annotated to chloroplast, mitochondria or bacteria in either whole leaves or apoplast fluid wash samples from *Flaveria robusta* and *F. linearis*.

Sample	Species	Type of sample	Total reads	Chloroplast reads	Mitochondria reads	Bacteria reads
MC_WL_10_1	<i>Flaveria linearis</i>	Whole leaf	185	185	0	0
MC_WL_10_10	<i>Flaveria linearis</i>	Whole leaf	1840	1840	0	0
MC_WL_10_2	<i>Flaveria linearis</i>	Whole leaf	13647	13573	50	10
MC_WL_10_3	<i>Flaveria linearis</i>	Whole leaf	2767	2750	3	11
MC_WL_10_4	<i>Flaveria linearis</i>	Whole leaf	610	608	2	0
MC_WL_10_5	<i>Flaveria linearis</i>	Whole leaf	1503	1503	0	0
MC_WL_10_6	<i>Flaveria linearis</i>	Whole leaf	2621	2612	5	0
MC_WL_10_7	<i>Flaveria linearis</i>	Whole leaf	4631	4621	2	0
MC_WL_10_8	<i>Flaveria linearis</i>	Whole leaf	4469	4444	9	14
MC_WL_10_9	<i>Flaveria linearis</i>	Whole leaf	6338	6236	12	81
MC_WL_12_1	<i>Flaveria robusta</i>	Whole leaf	476	471	3	2
MC_WL_12_10	<i>Flaveria robusta</i>	Whole leaf	725	721	0	4
MC_WL_12_2	<i>Flaveria robusta</i>	Whole leaf	44	44	0	0
MC_WL_12_3	<i>Flaveria robusta</i>	Whole leaf	1761	1735	24	2
MC_WL_12_4	<i>Flaveria robusta</i>	Whole leaf	4073	3999	33	40
MC_WL_12_5	<i>Flaveria robusta</i>	Whole leaf	10255	10217	22	16
MC_WL_12_6	<i>Flaveria robusta</i>	Whole leaf	1613	1415	24	174
MC_WL_12_7	<i>Flaveria robusta</i>	Whole leaf	3474	3451	2	21
MC_WL_12_9	<i>Flaveria robusta</i>	Whole leaf	493	493	0	0
MCAF.10.1_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	311	246	0	65
MCAF.10.2_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	2163	1947	35	181
MCAF.10.3_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	1135	743	7	385
MCAF.10.5_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	4011	283	10	3718
MCAF.10.7_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	788	380	0	408
MCAF.10.8_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	5414	3755	23	1636
MCAF10.10_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	1536	852	13	611
MCAF10.4_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	1400	947	31	422
MCAF10.6_AFW	<i>Flaveria linearis</i>	Apoplast fluid wash	1177	767	8	402
MCAF12.1_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	803	675	22	106
MCAF12.10_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	2431	503	4	1924
MCAF12.11_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	5276	3978	813	481
MCAF12.12_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	2467	956	2	1509
MCAF12.13_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	3206	2114	11	1081
MCAF12.14_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	1008	653	0	355
MCAF12.2_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	953	573	13	351
MCAF12.3_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	24	13	0	11
MCAF12.4_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	3	3	0	0
MCAF12.5_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	831	120	6	705
MCAF12.6_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	12	12	0	0
MCAF12.7_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	4504	4241	58	201
MCAF12.8_AFW	<i>Flaveria robusta</i>	Apoplast fluid wash	8646	7945	135	565

Supplementary Table 11. 16S taxa present in at least one sample of either whole leaf or apoplast fluid wash samples from *Flaveria robusta* and *F. linearis* leaves.

	<i>Flaveria robusta</i>		<i>Flaveria linearis</i>	
	Whole leaf	AFW	Whole leaf	AFW
<i>Achromobacter</i>	<i>Achromobacter</i>		<i>Candidatus Tremblaya</i>	<i>Achromobacter</i>
<i>Bacillus</i>	<i>Acidibacter</i>		<i>Lactobacillus</i>	<i>Anoxybacillus</i>
<i>Enterococcus</i>	<i>Acinetobacter</i>		<i>Massilia</i>	<i>Bacillus</i>
<i>Kluyvera</i>	<i>Afipia</i>		<i>Pantoea</i>	<i>Blastococcus</i>
<i>Lactobacillus</i>	<i>Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium</i>		<i>Pseudomonas</i>	<i>Carnobacterium</i>
<i>Listeria</i>	<i>Aminobacter</i>		<i>Sphingomonas</i>	<i>Cloacibacterium</i>
<i>Pantoea</i>	<i>Anaerocolumna</i>			<i>Curtobacterium</i>
<i>Pseudomonas</i>	<i>Anaerostignum</i>			<i>Cutibacterium</i>
<i>Roseovarius</i>	<i>Anaerotruncus</i>			<i>Desulfovermiculus</i>
<i>Samsonia</i>	<i>Angustibacter</i>			<i>Enhydrobacter</i>
<i>Staphylococcus</i>	<i>Bacillus</i>			<i>Escherichia-Shigella</i>
	<i>Blastococcus</i>			<i>Gemella</i>
	<i>Bosea</i>			<i>Haliangium</i>
	<i>Bradyrhizobium</i>			<i>Hymenobacter</i>
	<i>Brevibacterium</i>			<i>Lysobacter</i>
	<i>Brevundimonas</i>			<i>Mesorhizobium</i>
	<i>Candidatus Nucleicultrix</i>			<i>Methylobacterium-Methylorubrum</i>
	<i>Candidatus Obscuribacter</i>			<i>Micrococcus</i>
	<i>Candidatus Ovatusbacter</i>			<i>Okibacterium</i>
	<i>Candidatus Riegeria</i>			<i>Pantoea</i>
	<i>Caulobacter</i>			<i>Pseudomonas</i>
	<i>Chryseobacterium</i>			<i>Pygmaibacter</i>
	<i>Clavibacter</i>			<i>Rathayibacter</i>
	<i>Cloacibacterium</i>			<i>Rhodococcus</i>
	<i>Conexibacter</i>			<i>Samsonia</i>
	<i>Corynebacterium</i>			<i>SN8</i>
	<i>Curtobacterium</i>			<i>Sphingomonas</i>
	<i>Cutibacterium</i>			<i>Staphylococcus</i>
	<i>Deinococcus</i>			<i>Sulfurospirillum</i>
	<i>Delftia</i>			<i>Variovorax</i>
	<i>Desulfovibrio</i>			<i>Xanthomonas</i>
	<i>Devosia</i>			
	<i>Duganella</i>			
	<i>Enhydrobacter</i>			
	<i>Enterobacillus</i>			
	<i>Escherichia-Shigella</i>			
	<i>Exiguobacterium</i>			
	<i>Gemmatimonas</i>			
	<i>Glutamicibacter</i>			
	<i>Granulicatella</i>			
	<i>Hymenobacter</i>			
	<i>Leuconostoc</i>			
	<i>Marmoricola</i>			
	<i>Mesorhizobium</i>			
	<i>Methylobacterium-Methylorubrum</i>			
	<i>Neochlamydia</i>			
	<i>Nitrospira</i>			
	<i>Noviherbaspirillum</i>			
	<i>Ochrobactrum</i>			
	<i>Pantoea</i>			
	<i>Paracoccus</i>			
	<i>Pedobacter</i>			
	<i>Pedomicrobium</i>			
	<i>Pseudomonas</i>			
	<i>Pseudonocardia</i>			
	<i>Quadrisphaera</i>			
	<i>Ralstonia</i>			
	<i>Reyranella</i>			
	<i>Rhizobacter</i>			
	<i>Rubritepida</i>			
	<i>Samsonia</i>			
	<i>Sediminibacterium</i>			
	<i>SM1A02</i>			
	<i>Sphingobium</i>			
	<i>Sphingomonas</i>			
	<i>Streptococcus</i>			
	<i>Sulfurospirillum</i>			
	<i>Tepidimonas</i>			
	<i>TM7a</i>			
	<i>Tychonema CCAP 1459-11B</i>			
	<i>UCG-008</i>			
	<i>Vibrionimonas</i>			

Chapter 3.2 Supp. Tables

Supplementary Table 12. Permutational multivariate analysis of variance of metabolomic data (UHPLC-HRMS) from *Flaveria* species grown in a common garden, sampled throughout the day. Permutations set to 10000. Significance p values for each factor are indicated (p value 0: *** 0.001: ** 0.01: * 0.05: . 0.1 ' '1).

All data						
	Df	SumofSqs	R2	F	Pr(>F)	
Species	2	24118698	0.2406	10.3725	0.001	***
Sampling time	2	15181080	0.15144	6.5288	0.001	***
Species*Sampling time	4	12113201	0.12084	2.6047	0.001	***
Residual	42					
Total	50					

<i>F. robusta</i>						
	Df	SumofSqs	R2	F	Pr(>F)	
Plant	1	2117391	0.06687	1.974	0.0283	*
Sampling time	2	9422775	0.2976	4.3924	0.0001	***
Plant*Sampling time	2	4032385	0.12736	1.8797	0.0109	*
Residual	15	16089501	0.50816			
Total	20					

<i>F. linearis</i>						
	Df	SumofSqs	R2	F	Pr(>F)	
Plant	1	1385863	0.04319	0.9174	0.401	
Sampling time	2	12060628	0.37588	3.992	0.0016	**
Plant*Sampling time	2	2023327	0.06306	0.6697	0.7122	
Residual	11	16616460	0.51787			
Total	16	32086278	1			

<i>F. trinervia</i>						
	Df	SumofSqs	R2	F	Pr(>F)	
Plant	1	946445	0.07647	1.2504	0.2143	
Sampling time	2	4670283	0.37735	3.085	0.001	***
Plant*Sampling time	2	1461202	0.11806	0.9652	0.4788	
Residual	7	5298577	0.42812			
Total	12	12376506	1			

Supplementary Table 13. Differential analysis of metabolic data (UHPLC-HRMS peaks) detected in the leaf apoplast of three *Flaveria* species grown in a common garden and sampled throughout the day. The number of peaks with significantly different areas (p value <0.01 , $\log_2FC \geq 2$ or $\log_2FC \leq -2$) between sampling points is shown. The number of peaks considered in the metabolic pathways (FDR <0.01) are shown in parenthesis.

	5:00 am vs 1:00 pm		5:00 am vs 8:00 pm		1:00 pm vs 8:00 pm	
	Larger area at 5:00 am	Larger area at 1:00 pm	Larger area at 5:00 am	Larger area at 8:00 pm	Larger area at 1:00 pm	Larger area at 8:00 pm
<i>F. robusta</i>	774 (495)	4017 (3408)	1086 (364)	3477 (2863)	922 (251)	267 (58)
<i>F. linearis</i>	215 (132)	8747 (6940)	811 (284)	7626 (5934)	1055 (48)	146 (41)
<i>F. trinervia</i>	2937 (1354)	1807 (1330)	1512 (*)	1570 (*)	350 (*)	644 (*)

*Not enough replicates

Supplementary Table 15. Sequencing reads per sample in the leaf apoplast of samples from the *Flaveria* common garden

Sample ID	Species	Plant ID	Sampling time	Amplification plate	Total reads	Chloroplast reads	Mitochondria reads	Bacteria reads
FG2128	<i>F. robusta</i>	8	1:00 PM	Plate 1	4052	492	723	2478
FG2131	<i>F. trinervia</i>	1	1:00 PM	Plate 1	35566	4930	234	26720
FG2137	<i>F. trinervia</i>	7	1:00 PM	Plate 1	23137	8601	449	5240
FG2138	<i>F. trinervia</i>	8	1:00 PM	Plate 1	13840	2396	12	10326
FG2139	<i>F. trinervia</i>	9	1:00 PM	Plate 1	1845	108	15	1496
FG3101	<i>F. linearis</i>	1	8:00 PM	Plate 1	6271	2870	9	2963
FG3102	<i>F. linearis</i>	2	8:00 PM	Plate 1	2503	109	22	2061
FG3105	<i>F. linearis</i>	5	8:00 PM	Plate 1	2901	78	6	2150
FG3121	<i>F. robusta</i>	1	8:00 PM	Plate 1	11314	1146	29	9661
FG3122	<i>F. robusta</i>	2	8:00 PM	Plate 1	45689	4327	152	38255
FG3123	<i>F. robusta</i>	3	8:00 PM	Plate 1	43372	10538	514	23444
FG3124	<i>F. robusta</i>	4	8:00 PM	Plate 1	16432	1455	3	14377
FG3125	<i>F. robusta</i>	5	8:00 PM	Plate 1	14107	405	4	12775
FG3131	<i>F. trinervia</i>	1	8:00 PM	Plate 1	15838	2546	1	12668
FG3132	<i>F. trinervia</i>	2	8:00 PM	Plate 1	6102	56	0	5811
FG1101	<i>F. linearis</i>	1	5:00 AM	Plate 2	24	0	0	22
FG1102	<i>F. linearis</i>	2	5:00 AM	Plate 2	2	0	0	2
FG1103	<i>F. linearis</i>	3	5:00 AM	Plate 2	732	0	0	726
FG1104	<i>F. linearis</i>	4	5:00 AM	Plate 2	1127	0	0	1121
FG1106	<i>F. linearis</i>	6	5:00 AM	Plate 2	9	0	0	8
FG1122	<i>F. robusta</i>	2	5:00 AM	Plate 2	3	0	0	2
FG1123	<i>F. robusta</i>	3	5:00 AM	Plate 2	66	26	0	39
FG1124	<i>F. robusta</i>	4	5:00 AM	Plate 2	89	0	0	87
FG1125	<i>F. robusta</i>	5	5:00 AM	Plate 2	5	0	0	5
FG1126	<i>F. robusta</i>	6	5:00 AM	Plate 2	262	0	0	262
FG1127	<i>F. robusta</i>	7	5:00 AM	Plate 2	7	0	0	7
FG1131	<i>F. trinervia</i>	1	5:00 AM	Plate 2	3	0	0	3
FG1132	<i>F. trinervia</i>	2	5:00 AM	Plate 2	18	0	0	18
FG1133	<i>F. trinervia</i>	3	5:00 AM	Plate 2	2	0	0	2
FG1134	<i>F. trinervia</i>	4	5:00 AM	Plate 2	2	0	0	2
FG1135	<i>F. trinervia</i>	5	5:00 AM	Plate 2	71	1	0	70
FG1136	<i>F. trinervia</i>	6	5:00 AM	Plate 2	2	0	0	2
FG1137	<i>F. trinervia</i>	7	5:00 AM	Plate 2	172	0	10	156
FG2101	<i>F. linearis</i>	1	1:00 PM	Plate 2	22	2	0	20
FG2102	<i>F. linearis</i>	2	1:00 PM	Plate 2	16	2	0	11
FG2105	<i>F. linearis</i>	5	1:00 PM	Plate 2	1	1	0	0
FG2106	<i>F. linearis</i>	6	1:00 PM	Plate 2	54	5	1	35
FG2107	<i>F. linearis</i>	7	1:00 PM	Plate 2	8	3	0	2
FG2109	<i>F. linearis</i>	9	1:00 PM	Plate 2	4	1	0	2
FG2121	<i>F. robusta</i>	1	1:00 PM	Plate 2	3431	4	2	3320
FG2122	<i>F. robusta</i>	2	1:00 PM	Plate 2	49	11	0	35
FG2123	<i>F. robusta</i>	3	1:00 PM	Plate 2	18	4	0	5
FG2124	<i>F. robusta</i>	4	1:00 PM	Plate 2	22	7	0	8
FG2125	<i>F. robusta</i>	5	1:00 PM	Plate 2	10	7	0	1
FG2126	<i>F. robusta</i>	6	1:00 PM	Plate 2	1	0	0	1
FG2127	<i>F. robusta</i>	7	1:00 PM	Plate 2	12	5	1	6
FG1105	<i>F. linearis</i>	5	5:00 AM	Plate 4	590	357	0	158
FG2108	<i>F. linearis</i>	8	1:00 PM	Plate 4	429	74	0	333
FG2129	<i>F. robusta</i>	9	1:00 PM	Plate 4	420716	194084	12457	3813

Chapter 3.3 Supp. Tables

Supplementary Table 19. Average nucleotide identity (%) between the genomes of several *Pseudomonas* (*Ps*) isolates.

	Ps Fr-CA_5	Ps Fr+CA_3	Ps Fr+CA_18
Ps Fr-CA_5		99.9909	99.9906
Ps Fr+CA_3	99.9909		99.9841
Ps Fr+CA_18	99.9906	99.9841	

Chapter 3.4 Supp. Tables

Supplementary Table 23. Average nucleotide identity (%) between the genomes of several *Pantoea* (*Pa*) isolates.

	Pa Ft-CA_14	Pa Fr+CA_20	Pa Ft+CA_17	Pa Fr-CA_6
Pa Ft-CA_14		81.160	99.997	98.685
Pa Fr+CA_20	81.160		81.292	81.207
Pa Ft+CA_17	99.997	81.292		98.678
Pa Fr-CA_6	98.685	81.207	98.678	

References

- Agler, M. T., J. Ruhe, S. Kroll, C. Morhenn, S.-T. Kim, D. Weigel, and E. M. Kemen. 2016. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS biology* 14: e1002352.
- Akashi, H., and T. Gojobori. 2002. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* 99: 3695–3700.
- Alkan, N., E. A. Espeso, and D. Prusky. 2013. Virulence regulation of phytopathogenic fungi by pH. *Antioxidants & Redox Signaling* 19: 1012–1025.
- Alonso-Roman, R., A. Last, M. H. Mirhakkak, J. L. Sprague, L. Möller, P. Großmann, K. Graf, et al. 2022. *Lactobacillus rhamnosus* colonisation antagonizes *Candida albicans* by forcing metabolic adaptations that compromise pathogenicity. *Nature Communications* 13: 3192.
- Arkin, A. P., R. W. Cottingham, C. S. Henry, N. L. Harris, R. L. Stevens, S. Maslov, P. Dehal, et al. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology* 36: 566–569.
- Aulakh, M. S., R. Wassmann, C. Bueno, J. Kreuzwieser, and H. Rennenberg. 2001. Characterization of root exudates at different growth stages of ten rice (*Oryza sativa* L.) cultivars. *Plant Biology* 3: 139–148.
- Ayre, B. G. 2011. Membrane-transport systems for sucrose in relation to whole-plant carbon partitioning. *Molecular Plant* 4: 377–394.
- Baccolini, C., and C.-P. Witte. 2019. AMP and GMP catabolism in *Arabidopsis* converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. *The Plant Cell* 31: 734–751.
- Bar-On, Y. M., R. Phillips, and R. Milo. 2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences* 115: 6506–6511.
- Beattie, G. A. 2011. Water Relations in the Interaction of Foliar Bacterial Pathogens with Plants. *Annual Review of Phytopathology* 49: 533–555.
- Berg, G., and J. M. Raaijmakers. 2018. Saving seed microbiomes. *The ISME Journal* 12: 1167–1170.
- Bittar, T. B., P. Pound, A. Whitetree, L. D. Moore, and J. T. Van Stan II. 2018. Estimation of throughfall and stemflow bacterial flux in a subtropical oak-cedar forest. *Geophysical Research Letters* 45: 1410–1418.
- Blighe, K., S. Rana, and M. Lewis. 2020. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package, version 1.8.0. <https://github.com/kevinblighe/EnhancedVolcano>.

- Bodenhausen, N., M. Bortfeld-Miller, M. Ackermann, and J. A. Vorholt. 2014. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLOS Genetics* 10: e1004283.
- Borghgi, G. L., S. Arrivault, M. Günther, D. Barbosa Medeiros, E. Dell'Aversana, G. M. Fusco, P. Carillo, et al. 2022. Metabolic profiles in C3, C3-C4 intermediate, C4-like, and C4 species in the genus *Flaveria*. *Journal of Experimental Botany* 73: 1581–1601.
- Brodmann, D., A. Schuller, J. Ludwig-Müller, R. A. Aeschbacher, A. Wiemken, T. Boller, and A. Wingler. 2002. Induction of trehalase in *Arabidopsis* plants infected with the trehalose-producing pathogen *Plasmodiophora brassicae*. *Molecular Plant-Microbe Interactions* 15: 693–700.
- Brown, S. P., M. A. Grillo, J. C. Podowski, and K. D. Heath. 2020. Soil origin and plant genotype structure distinct microbiome compartments in the model legume *Medicago truncatula*. *Microbiome* 8: 139.
- Brühl, C. A., and J. G. Zaller. 2019. Biodiversity decline as a consequence of an inappropriate Environmental Risk Assessment of pesticides. *Frontiers in Environmental Science* 7.
- Bulgarelli, D., M. Rott, K. Schlaeppli, E. Ver Loren van Themaat, N. Ahmadinejad, F. Assenza, P. Rauf, et al. 2012. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488: 91–95.
- Bziuk, N., L. Maccario, B. Straube, G. Wehner, S. J. Sørensen, A. Schikora, and K. Smalla. 2021. The treasure inside barley seeds: microbial diversity and plant beneficial bacteria. *Environmental Microbiome* 16: 20.
- Cadot, S., H. Guan, M. Bigalke, J.-C. Walser, G. Jander, M. Erb, M. G. A. van der Heijden, and K. Schlaeppli. 2021. Specific and conserved patterns of microbiota-structuring by maize benzoxazinoids in the field. *Microbiome* 9: 103.
- Callahan, B., P. J. McMurdie, M. Rosen, A. Han, A. Johnson, and S. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *13*: 581–583.
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335–336.
- Caradonia, F., D. Ronga, M. Catellani, C. V. Giaretta Azevedo, R. A. Terrazas, S. Robertson-Albertyn, E. Francia, and D. Bulgarelli. 2019. Nitrogen fertilizers shape the composition and predicted functions of the microbiota of field-grown tomato plants. *Phytobiomes Journal* 3: 315–325.
- Carlström, C. I., C. M. Field, M. Bortfeld-Miller, B. Müller, S. Sunagawa, and J. A. Vorholt. 2019. Synthetic microbiota reveal priority effects and keystone strains in the *Arabidopsis* phyllosphere. *Nature Ecology & Evolution* 3: 1445–1454.
- Carrión, V. J., J. Perez-Jaramillo, V. Cordovez, V. Tracanna, M. de Hollander, D. Ruiz-Buck, L. W. Mendes, et al. 2019. Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science (New York, N.Y.)* 366: 606–612.

- Carvalho, L. C., P. G. Dennis, B. Fan, D. Fedoseyenko, K. Kierul, A. Becker, N. von Wieren, and R. Borriss. 2013. Linking plant nutritional status to plant-microbe interactions. *PLOS ONE* 8: e68555.
- Carvalho, C. R., A. C. Dias, S. K. Homma, and E. J. Cardoso. 2020. Phyllosphere bacterial assembly in citrus crop under conventional and ecological management. *PeerJ* 8: e9152.
- Cervela-Cardona, L., B. Alary, and P. Mas. 2021. The *Arabidopsis* circadian clock and metabolic energy: a question of time. *Frontiers in Plant Science* 12.
- Cha, J.-Y., S. Han, H.-J. Hong, H. Cho, D. Kim, Y. Kwon, S.-K. Kwon, et al. 2016. Microbial and biochemical basis of a *Fusarium* wilt-suppressive soil. *The ISME Journal* 10: 119–129.
- Chao, K.-H., K. Barton, S. Palmer, and R. Lanfear. 2021. sangeranalyseR: Simple and interactive processing of sanger sequencing data in R. *Genome Biology and Evolution* 13: evab028.
- Chaparro, J. M., D. V. Badri, M. G. Bakker, A. Sugiyama, D. K. Manter, and J. M. Vivanco. 2013. Root exudation of phytochemicals in *Arabidopsis* follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PloS One* 8: e55731.
- Chaparro, J. M., D. V. Badri, and J. M. Vivanco. 2014. Rhizosphere microbiome assemblage is affected by plant development. *The ISME Journal* 8: 790–803.
- Chase, A. B., C. Weihe, and J. B. H. Martiny. 2021. Adaptive differentiation and rapid evolution of a soil bacterium along a climate gradient. *Proceedings of the National Academy of Sciences* 118: e2101254118.
- Chaudhry, V., P. Runge, P. Sengupta, G. Doehlemann, J. E. Parker, and E. Kemen. 2021. Shaping the leaf microbiota: plant-microbe-microbe interactions. *Journal of Experimental Botany* 72: 36–56.
- Chave, J. 2004. Neutral theory and community ecology. *Ecology Letters* 7: 241–253.
- Chen, L.-Q., B.-H. Hou, S. Lalonde, H. Takanaga, M. L. Hartung, X.-Q. Qu, W.-J. Guo, et al. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468: 527–532.
- Chen, T., K. Nomura, X. Wang, R. Sohrabi, J. Xu, L. Yao, B. C. Paasch, et al. 2020. A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature* 580: 653–657.
- Chen, Y., J. Wang, N. Yang, Z. Wen, X. Sun, Y. Chai, and Z. Ma. 2018. Wheat microbiome bacteria can reduce virulence of a plant pathogenic fungus by altering histone acetylation. *Nature Communications* 9: 3429.
- Colaiani, N. R., K. Parys, H.-S. Lee, J. M. Conway, N. H. Kim, N. Edelbacher, T. S. Mucyn, et al. 2021. A complex immune response to flagellin epitope variation in commensal communities. *Cell Host & Microbe* 29: 635-649.e9.

- Cole, B. J., M. E. Feltcher, R. J. Waters, K. M. Wetmore, T. S. Mucyn, E. M. Ryan, G. Wang, et al. 2017. Genome-wide identification of bacterial plant colonization genes. *PLoS Biology* 15: e2002860.
- Coleman-Derr, D., D. Desgarennnes, C. Fonseca-Garcia, S. Gross, S. Clingenpeel, T. Woyke, G. North, et al. 2016. Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytologist* 209: 798–811.
- Corkley, I., B. Fraaije, and N. Hawkins. 2022. Fungicide resistance management: Maximizing the effective life of plant protection products. *Plant Pathology* 71: 150–169.
- Coyte, K. Z., J. Schluter, and K. R. Foster. 2015. The ecology of the microbiome: Networks, competition, and stability. *Science (New York, N.Y.)* 350: 663–666.
- Dal Bello, M., H. Lee, A. Goyal, and J. Gore. 2021. Resource-diversity relationships in bacterial communities reflect the network structure of microbial metabolism. *Nature Ecology & Evolution* 5: 1424–1434.
- De Vleeschauwer, D., M. Djavaheri, P. A. H. M. Bakker, and M. Höfte. 2008. *Pseudomonas fluorescens* WCS374r-induced systemic resistance in rice against *Magnaporthe oryzae* is based on pseudobactin-mediated priming for a salicylic acid-repressible multifaceted defense response. *Plant Physiology* 148: 1996–2012.
- De Vos, R. C., S. Moco, A. Lommen, J. J. Keurentjes, R. J. Bino, and R. D. Hall. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols* 2: 778–791.
- Delmotte, N., C. Knief, S. Chaffron, G. Innerebner, B. Roschitzki, R. Schlapbach, C. von Mering, and J. A. Vorholt. 2009. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 106: 16428–16433.
- Deriu, E., J. Z. Liu, M. Pezeshki, R. A. Edwards, R. J. Ochoa, H. Contreras, S. J. Libby, et al. 2013. Probiotic bacteria reduce *Salmonella typhimurium* intestinal colonization by competing for iron. *Cell Host & Microbe* 14: 26–37.
- Dietz, S., K. Herz, K. Gorzalka, U. Jandt, H. Bruelheide, and D. Scheel. 2020. Root exudate composition of grass and forb species in natural grasslands. *Scientific Reports* 10: 10691.
- Dodd, A. N., N. Salathia, A. Hall, E. Kévei, R. Tóth, F. Nagy, J. M. Hibberd, et al. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* 309: 630–633.
- Dreisbach, D., G. Petschenka, B. Spengler, and D. R. Bhandari. 2021. 3D-surface MALDI mass spectrometry imaging for visualising plant defensive cardiac glycosides in *Asclepias curassavica*. *Analytical and Bioanalytical Chemistry* 413: 2125–2134.
- D'Souza, G., S. Shitut, D. Preussger, G. Yousif, S. Waschina, and C. Kost. 2018. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports* 35: 455–488.

- Duong, D. A., R. V. Jensen, and A. M. Stevens. 2018. Discovery of *Pantoea stewartii* ssp. *stewartii* genes important for survival in corn xylem through a Tn-Seq analysis. *Molecular Plant Pathology* 19: 1929–1941.
- Durán, P., T. Thiergart, R. Garrido-Oter, M. Agler, E. Kemen, P. Schulze-Lefert, and S. Hacquard. 2018. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell* 175: 973-983.e14.
- Dutta, B., A. K. Barman, R. Srinivasan, U. Avci, D. E. Ullman, D. B. Langston, and R. D. Gitaitis. 2014. Transmission of *Pantoea ananatis* and *P. agglomerans*, causal agents of center rot of onion (*Allium cepa*), by onion thrips (*Thrips tabaci*) through feces. *Phytopathology* 104: 812–819.
- Dutta, B., Y. Ha, J. T. Lessl, U. Avci, A. C. Sparks, K. L. Johnson, and R. R. Walcott. 2015. Pathways of bacterial invasion and watermelon seed infection by *Acidovorax citrulli*. *Plant Pathology* 64: 537–544.
- Ehleringer, J. R., T. E. Cerling, and B. R. Helliker. 1997. C4 photosynthesis, atmospheric CO₂, and climate. *Oecologia* 112: 285–299.
- Espinoza, C., T. Degenkolbe, C. Caldana, E. Zuther, A. Leisse, L. Willmitzer, D. K. Hinch, and M. A. Hannah. 2010. Interaction with diurnal and circadian regulation results in dynamic metabolic and transcriptional changes during cold acclimation in *Arabidopsis*. *PLOS ONE* 5: e14101.
- FAO. 2022. Information Note - The importance of Ukraine and the Russian Federation for global agricultural markets and the risks associated with the war in Ukraine. <https://www.fao.org/3/cb9013en/cb9013en.pdf>.
- FAO. 2021. Pesticides use. Global, regional and country trends, 1990–2018. <https://www.fao.org/3/cb3411en/cb3411en.pdf>.
- Fitzpatrick, C. R., I. Salas-González, J. M. Conway, O. M. Finkel, S. Gilbert, D. Russ, P. J. P. L. Teixeira, and J. L. Dangl. 2020. The plant microbiome: from ecology to reductionism and beyond. *Annual Review of Microbiology* 74: 81–100.
- Foster, A. J., J. M. Jenkinson, and N. J. Talbot. 2003. Trehalose synthesis and metabolism are required at different stages of plant infection by *Magnaporthe grisea*. *The EMBO Journal* 22: 225–235.
- Freeman, B. C., C. Chen, and G. A. Beattie. 2010. Identification of the trehalose biosynthetic loci of *Pseudomonas syringae* and their contribution to fitness in the phyllosphere. *Environmental Microbiology* 12: 1486–1497.
- Fukami, T. 2015. Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics* 46: 1–23.
- Gabriel, R., and J. Kesselmeier. 1999. Apoplastic solute concentrations of organic acids and mineral nutrients in the leaves of several Fagaceae. *Plant and Cell Physiology* 40: 604–612.

- Gao, M., C. Xiong, C. Gao, C. K. M. Tsui, M.-M. Wang, X. Zhou, A.-M. Zhang, and L. Cai. 2021. Disease-induced changes in plant microbiome assembly and functional adaptation. *Microbiome* 9: 187.
- García-Plazaola, J. I., B. Fernández-Marín, J. P. Ferrio, J. G. Alday, G. Hoch, D. Landais, A. Milcu, et al. 2017. Endogenous circadian rhythms in pigment composition induce changes in photochemical efficiency in plant canopies. *Plant, Cell & Environment* 40: 1153–1162.
- Geilfus, C.-M. 2017. The pH of the apoplast: Dynamic factor with functional impact under stress. *Molecular Plant* 10: 1371–1386.
- Geilfus, C.-M., L. Wang, J. Wu, and C. Xue. 2020. The pH of the leaf apoplast is critical for the formation of *Pseudomonas syringae*-induced lesions on leaves of the common bean (*Phaseolus vulgaris*). *Plant Science* 290: 110328.
- Geilfus, C.-M., X. Zhang, A. Mithöfer, L. Burgel, G. Bárdos, and C. Zörb. 2021. Leaf apoplastic alkalization promotes transcription of the ABA-synthesizing enzyme Vp14 and stomatal closure in *Zea mays*. *Journal of Experimental Botany* 72: 2686–2695.
- Gentzel, I., L. Giese, W. Zhao, A. P. Alonso, and D. Mackey. 2019. A simple method for measuring apoplast hydration and collecting apoplast contents. *Plant Physiology* 179: 1265–1272.
- Gerlich, S. C., B. J. Walker, S. Krueger, and S. Kopriva. 2018. Sulfate metabolism in C4 *Flaveria* species is controlled by the root and connected to serine biosynthesis. *Plant Physiology* 178: 565–582.
- Godson, A., and R. A. L. van der Hoorn. 2021. The front line of defence: a meta-analysis of apoplastic proteases in plant immunity. *Journal of Experimental Botany* 72: 3381–3394.
- Goldford, J. E., N. Lu, D. Bajić, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segrè, et al. 2018. Emergent simplicity in microbial community assembly. *Science* 361: 469–474.
- Gowik, U., A. Bräutigam, K. L. Weber, A. P. M. Weber, and P. Westhoff. 2011. Evolution of C4 photosynthesis in the genus *Flaveria*: How many and which genes does it take to make C4? *The Plant Cell* 23: 2087–2105.
- Gu, Y., X. Wang, T. Yang, V.-P. Friman, S. Geisen, Z. Wei, Y. Xu, et al. 2020. Chemical structure predicts the effect of plant-derived low-molecular weight compounds on soil microbiome structure and pathogen suppression. *Functional Ecology* 34: 2158–2169.
- Gu, Z., R. Eils, and M. Schlesner. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Genome Biology* 17: 2847–9.
- Guidi, L., E. Lo Piccolo, and M. Landi. 2019. Chlorophyll fluorescence, photoinhibition and abiotic stress: Does it make any difference the fact to be a C3 or C4 species? *Frontiers in Plant Science* 10.

- Guo, X., X. Zhang, Y. Qin, Y.-X. Liu, J. Zhang, N. Zhang, K. Wu, et al. 2020. Host-associated quantitative abundance profiling reveals the microbial load variation of root microbiome. *Plant Communications* 1: 100003.
- Hafiz, S. M. A. 2020. Microbial interactions and the arisal of amino acid auxotrophy in phyllosphere bacteria of C3 and C4 plants. [Master of Sciences, Friedrich-Schiller-Universität].
- Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. H. Martiny. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology* 10: 497–506.
- Hardin, G. 1960. The Competitive Exclusion Principle. *Science* 131: 1292–1297.
- Harris, J. M., P. Balint-Kurti, J. C. Bede, B. Day, S. Gold, E. M. Goss, L. J. Grenville-Briggs, et al. 2020. What are the top 10 unanswered questions in molecular plant-microbe interactions? *Molecular Plant-Microbe Interactions* 33: 1354–1365.
- Hasanuzzaman, M., K. Nahar, T. I. Anee, and M. Fujita. 2017. Glutathione in plants: biosynthesis and physiological role in environmental stress tolerance. *Physiology and Molecular Biology of Plants* 23: 249–268.
- Hassani, M. A., P. Durán, and S. Hacquard. 2018. Microbial interactions within the plant holobiont. *Microbiome* 6: 1–17.
- Hatch, M. D., T. Kagawa, and S. Craig. 1975. Subdivision of C4-pathway species based on differing C4 acid decarboxylating systems and ultrastructural features. *Functional Plant Biology* 2: 111–128.
- Haydon, M. J., L. J. Bell, and A. A. R. Webb. 2011. Interactions between plant circadian clocks and solute transport. *Journal of Experimental Botany* 62: 2333–2348.
- Herp, S., S. Brugiroux, D. Garzetti, D. Ring, L. M. Jochum, M. Beutler, C. Eberl, et al. 2019. *Mucispirillum schaedleri* antagonizes *Salmonella* virulence to protect mice against colitis. *Cell Host & Microbe* 25: 681-694.e8.
- Herren, C. M. 2020. Disruption of cross-feeding interactions by invading taxa can cause invasional meltdown in microbial communities. *Proceedings of the Royal Society B: Biological Sciences* 287: 20192945.
- Hildebrandt, T. M., A. Nunes Nesi, W. L. Araújo, and H.-P. Braun. 2015. Amino acid catabolism in plants. *Molecular Plant* 8: 1563–1579.
- Hiruma, K., S. Fukunaga, P. Bednarek, M. Piślewska-Bednarek, S. Watanabe, Y. Narusaka, K. Shirasu, and Y. Takano. 2013. Glutathione and tryptophan metabolism are required for *Arabidopsis* immunity during the hypersensitive response to hemibiotrophs. *Proceedings of the National Academy of Sciences* 110: 9589–9594.
- Hoek, T. A., K. Axelrod, T. Biancalani, E. A. Yurtsev, J. Liu, and J. Gore. 2016. Resource availability modulates the cooperative and competitive nature of a microbial cross-feeding mutualism. *PLOS Biology* 14: e1002540.

- Hu, J., M. Liu, A. Zhang, Y. Dai, W. Chen, F. Chen, W. Wang, et al. 2022. Co-evolved plant and blast fungus ascorbate oxidases orchestrate the redox state of host apoplast to modulate rice immunity. *Molecular Plant* 15: 1347–1366.
- Hu, J., Z. Wei, V.-P. Friman, S. Gu, X. Wang, N. Eisenhauer, T. Yang, et al. 2016. Probiotic diversity enhances rhizosphere microbiome function and plant disease suppression. *mBio* 7: e01790-16.
- Hu, L., C. A. M. Robert, S. Cadot, X. Zhang, M. Ye, B. Li, D. Manzo, et al. 2018. Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature Communications* 9: 2738.
- Huang, L., X. Tang, W. Zhang, R. Jiang, D. Chen, J. Zhang, and H. Zhong. 2016. Imaging of endogenous metabolites of plant leaves by mass spectrometry based on laser activated electron tunneling. *Scientific Reports* 6: 24164.
- Hubbard, C. J., M. T. Brock, L. T. van Diepen, L. Maignien, B. E. Ewers, and C. Weinig. 2018. The plant circadian clock influences rhizosphere community structure and function. *The ISME Journal* 12: 400–410.
- Ikawa, Y., and S. Tsuge. 2016. The quantitative regulation of the *hrp* regulator HrpX is involved in sugar-source-dependent *hrp* gene expression in *Xanthomonas oryzae* pv. *oryzae*. *FEMS Microbiology Letters* 363.
- Jain, C., L. M. Rodriguez-R, A. M. Phillippy, K. T. Konstantinidis, and S. Aluru. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* 9: 5114.
- Jiménez, A., F. Sevilla, and M. C. Martí. 2021. Reactive oxygen species homeostasis and circadian rhythms in plants. *Journal of Experimental Botany* 72: 5825–5840.
- Jones, J. D. G., and J. L. Dangl. 2006. The plant immune system. *Nature* 444: 323–329.
- Kamada, N., Y.-G. Kim, H. P. Sham, B. A. Vallance, J. L. Puente, E. C. Martens, and G. Núñez. 2012. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* 336: 1325–1329.
- Kamo, T., S. Hiradate, K. Suzuki, I. Fujita, S. Yamaki, T. Yoneda, M. Koitabashi, and S. Yoshida. 2018. Methylobamine, a UVA-absorbing compound from the plant-associated bacteria *Methylobacterium* sp. *Natural Product Communications* 13: 1934578X1801300208.
- Karasov, T. L., J. Almario, C. Friedemann, W. Ding, M. Giolai, D. Heavens, S. Kersten, et al. 2018. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. *Cell Host & Microbe* 24: 168-179.e4.
- Karpinska, B., K. Zhang, B. Rasool, D. Pastok, J. Morris, S. R. Verrall, P. E. Hedley, et al. 2018. The redox state of the apoplast influences the acclimation of photosynthesis and leaf metabolism to changing irradiance. *Plant, Cell & Environment* 41: 1083–1097.

- Kembel, S. W., T. K. O'Connor, H. K. Arnold, S. P. Hubbell, S. J. Wright, and J. L. Green. 2014. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences* 111: 13715–13720.
- Koskella, B. 2020. The phyllosphere. *Current Biology* 30: R1143–R1146.
- Kueger, S., D. Steinhauser, L. Willmitzer, and P. Giavalisco. 2012. High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *The Plant Journal: For Cell and Molecular Biology* 70: 39–50.
- Kwak, M.-J., H. G. Kong, K. Choi, S.-K. Kwon, J. Y. Song, J. Lee, P. A. Lee, et al. 2018. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nature Biotechnology* 36: 1100–1109.
- Lemoine, R., S. La Camera, R. Atanassova, F. Dédaldéchamp, T. Allario, N. Pourtau, J.-L. Bonnemain, et al. 2013. Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science* 4.
- Leonardos, E. D., and B. Grodzinski. 2000. Photosynthesis, immediate export and carbon partitioning in source leaves of C3, C3-C4 intermediate, and C4 *Panicum* and *Flaveria* species at ambient and elevated CO₂ levels. *Plant, Cell & Environment* 23: 839–851.
- Leveau, J. H. J., and S. E. Lindow. 2001. Appetite of an epiphyte: Quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proceedings of the National Academy of Sciences* 98: 3446–3453.
- Levy, A., I. Salas Gonzalez, M. Mittelviefhaus, S. Clingenpeel, S. Herrera Paredes, J. Miao, K. Wang, et al. 2018. Genomic features of bacterial adaptation to plants. *Nature Genetics* 50: 138–150.
- Li, P.-D., Z.-R. Zhu, Y. Zhang, J. Xu, H. Wang, Z. Wang, and H. Li. 2022. The phyllosphere microbiome shifts toward combating melanose pathogen. *Microbiome* 10: 56.
- Li, R., H. M. Tun, M. Jahan, Z. Zhang, A. Kumar, W. G. Dilantha Fernando, A. Farenhorst, and E. Khafipour. 2017. Comparison of DNA-, PMA-, and RNA-based 16S rRNA Illumina sequencing for detection of live bacteria in water. *Scientific Reports* 7: 5752.
- Li, S., J. Tan, X. Yang, C. Ma, and L. Jiang. 2019. Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *The ISME Journal* 13: 402–412.
- Li, Y., X. Wu, T. Chen, W. Wang, G. Liu, W. Zhang, S. Li, et al. 2018. Plant phenotypic traits eventually shape its microbiota: a common garden test. *Frontiers in Microbiology* 9: 2479.
- Liu, H.-Q., X.-B. Lu, Z.-H. Li, C.-Y. Tian, and J. Song. 2021. The role of root-associated microbes in growth stimulation of plants under saline conditions. *Land Degradation & Development* 32: 3471–3486.

- Liu, Q., L. Luo, and L. Zheng. 2018. Lignins: biosynthesis and biological functions in plants. *International Journal of Molecular Sciences* 19: E335.
- Lohaus, G., K. Pennewiss, B. Sattelmacher, M. Hussmann, and K. Hermann Muehling. 2001. Is the infiltration-centrifugation technique appropriate for the isolation of apoplastic fluid? A critical evaluation with different plant species. *Physiologia Plantarum* 111: 457–465.
- Lohaus, G., H. Winter, B. Riens, and H. W. Heldt. 1995. Further studies of the phloem loading process in leaves of barley and spinach. The comparison of metabolite concentrations in the apoplastic compartment with those in the cytosolic compartment and in the sieve tubes. *Botanica Acta* 108: 270–275.
- van Loon, L. C., P. A. H. M. Bakker, and C. M. J. Pieterse. 1998. Systemic resistance induced by rhizosphere bacteria. *Annual Review of Phytopathology* 36: 453–483.
- López-Millán, A. F., F. Morales, A. Abadía, and J. Abadía. 2000. Effects of iron deficiency on the composition of the leaf apoplastic fluid and xylem sap in sugar beet. Implications for iron and carbon transport. *Plant Physiology* 124: 873–884.
- Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *15*: 550.
- Lu, T., Z. Zhang, Y. Li, Q. Zhang, H. Cui, L. Sun, W. J. G. M. Peijnenburg, et al. 2021. Does biological rhythm transmit from plants to rhizosphere microbes? *Environmental Microbiology* 23: 6895–6906.
- Lundberg, D. S., R. de P. Jové, P. P. N. Ayutthaya, T. L. Karasov, O. Shalev, K. Poersch, W. Ding, et al. 2022. Contrasting patterns of microbial dominance in the *Arabidopsis thaliana* phyllosphere. 2021.04.06.438366.
- Lundberg, D. S., R. de P. Jové, P. P. N. Ayutthaya, T. L. Karasov, O. Shalev, K. Poersch, W. Ding, et al. 2021. Contrasting patterns of microbial dominance in the *Arabidopsis thaliana* phyllosphere. *bioRxiv*: 2021.04.06.438366.
- Ma, S., Y. Li, X. Li, X. Sui, and Z. Zhang. 2019. Phloem unloading strategies and mechanisms in crop fruits. *Journal of Plant Growth Regulation* 38: 494–500.
- MacDougall, A. S., B. Gilbert, and J. M. Levine. 2009. Plant invasions and the niche. *Journal of Ecology* 97: 609–615.
- Maia, M., A. McCann, C. Malherbe, J. Far, J. Cunha, J. Eiras-Dias, C. Cordeiro, et al. 2022. Grapevine leaf MALDI-MS imaging reveals the localisation of a putatively identified sucrose metabolite associated to *Plasmopara viticola* development. *Frontiers in Plant Science* 13.
- Maier, B. A., P. Kiefer, C. M. Field, L. Hemmerle, M. Bortfeld-Miller, B. Emmenegger, M. Schäfer, et al. 2021. A general non-self response as part of plant immunity. *Nature Plants* 7: 696–705.

- Maignien, L., E. A. DeForce, M. E. Chafee, A. M. Eren, and S. L. Simmons. 2014. Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio* 5: e00682-13.
- Malcolm, G. M., G. A. Kuldau, B. K. Gugino, and M. del M. Jiménez-Gasco. 2013. Hidden host plant associations of soilborne fungal pathogens: an ecological perspective. *Phytopathology* 103: 538–544.
- Martins, S. J., F. H. V. Medeiros, V. Lakshmanan, and H. P. Bais. 2018. Impact of seed exudates on growth and biofilm formation of *Bacillus amyloliquefaciens* ALB629 in common bean. *Frontiers in Microbiology* 8: 2631.
- Matt, P., M. Geiger, P. Walch-Liu, C. Engels, A. Krapp, and M. Stitt. 2001. The immediate cause of the diurnal changes of nitrogen metabolism in leaves of nitrate-replete tobacco: a major imbalance between the rate of nitrate reduction and the rates of nitrate uptake and ammonium metabolism during the first part of the light period. *Plant, Cell & Environment* 24: 177–190.
- Mayer, T., A. Mari, J. Almario, M. Murillo-Roos, H. Syed M Abdullah, N. Dombrowski, S. Hacquard, et al. 2021. Obtaining deeper insights into microbiome diversity using a simple method to block host and nontargets in amplicon sequencing. *Molecular Ecology Resources* 21: 1952–1965.
- Mayer, T., M. Reichelt, J. Gershenzon, and M. Agler. 2022. Leaf bacterial community structure and variation in wild ruderal plants are shaped by the interaction of host species and defense chemistry with environment. 2022.03.16.484556.
- McGarvey, J. a., T. Tran, R. Han, R. Hnasko, and P. Brown. 2019. Bacterial population dynamics after foliar fertilization of almond leaves. *Journal of Applied Microbiology* 126: 945–953.
- McKown, A. D., and N. G. Dengler. 2010. Vein patterning and evolution in C4 plants. *Botany* 88: 775–786.
- McMurdie, P. J., and S. Holmes. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 8: e61217.
- Melotto, M., W. Underwood, J. Koczan, K. Nomura, and S. Y. He. 2006. Plant stomata function in innate immunity against bacterial invasion. *Cell* 126: 969–980.
- Mendes, R., M. Kruijt, I. de Bruijn, E. Dekkers, M. van der Voort, J. H. M. Schneider, Y. M. Piceno, et al. 2011. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 332: 1097–1100.
- Mercier, J., and S. E. Lindow. 2000. Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Applied and Environmental Microbiology* 66: 369–374.
- Meyer, S., C. Lauterbach, M. Niedermeier, I. Barth, R. D. Sjolund, and N. Sauer. 2004. Wounding enhances expression of AtSUC3, a sucrose transporter from *Arabidopsis* sieve elements and sink tissues. *Plant Physiology* 134: 684–693.

- Mitter, B., N. Pfaffenbichler, R. Flavell, S. Compant, L. Antonielli, A. Petric, T. Berninger, et al. 2017. A new approach to modify plant microbiomes and traits by introducing beneficial bacteria at flowering into progeny seeds. *Frontiers in Microbiology* 8: 11.
- Monier, J.-M., and S. E. Lindow. 2003. Differential survival of solitary and aggregated bacterial cells promotes aggregate formation on leaf surfaces. *Proceedings of the National Academy of Sciences* 100: 15977–15982.
- Moormann, J., B. Heinemann, and T. M. Hildebrandt. 2022. News about amino acid metabolism in plant–microbe interactions. *Trends in Biochemical Sciences* 47: 839–850.
- Morella, N. M., F. C.-H. Weng, P. M. Joubert, C. J. E. Metcalf, S. Lindow, and B. Koskella. 2020. Successive passaging of a plant-associated microbiome reveals robust habitat and host genotype-dependent selection. *Proceedings of the National Academy of Sciences of the United States of America* 117: 1148–1159.
- Morella, N. M., X. Zhang, and B. Koskella. 2019. Tomato seed-associated bacteria confer protection of seedlings against foliar disease caused by *Pseudomonas syringae*. *Phytobiomes Journal* 3: 177–190.
- Morris, C. E., and B. Moury. 2019. Revisiting the concept of host range of plant pathogens. *Annual Review of Phytopathology* 57: 63–90.
- Mortensen, M. S., M. A. Rasmussen, J. Stokholm, A. D. Brejnrod, C. Balle, J. Thorsen, K. A. Kroghfelt, et al. 2021. Modeling transfer of vaginal microbiota from mother to infant in early life. *eLife* 10: e57051.
- Mühling, K. H., and A. Läuchli. 2000. Light-induced pH and K⁺ changes in the apoplast of intact leaves. *Planta* 212: 9–15.
- Muyzer, G., E. C. de Waal, and A. G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* 59: 695–700.
- Nakano, R. T., N. Ishihama, Y. Wang, J. Takagi, T. Uemura, P. Schulze-Lefert, and H. Nakagami. 2020. Apoplastic fluid preparation from *Arabidopsis thaliana* leaves upon interaction with a nonadapted powdery mildew pathogen. *Methods in Molecular Biology* 2139: 79–88.
- Naseem, M., M. Kunz, and T. Dandekar. 2017. Plant–pathogen maneuvering over apoplastic sugars. *Trends in Plant Science* 22: 740–743.
- Nerva, L., M. Sandrini, L. Moffa, R. Velasco, R. Balestrini, and W. Chitarra. 2022. Breeding toward improved ecological plant–microbiome interactions. *Trends in Plant Science* 27: 1134–1143.
- Newman, A., E. Picot, S. Davies, S. Hilton, I. A. Carré, and G. D. Bending. 2022. Circadian rhythms in the plant host influence rhythmicity of rhizosphere microbiota. *BMC Biology* 20: 235.

- Nobori, T., Y. Cao, F. Entila, E. Dahms, Y. Tsuda, R. Garrido-Oter, and K. Tsuda. 2022. Dissecting the co-transcriptome landscape of plants and microbiota members. 2021.04.25.440543.
- Nobori, T., A. C. Velásquez, J. Wu, B. H. Kvitko, J. M. Kremer, Y. Wang, S. Y. He, and K. Tsuda. 2018. Transcriptome landscape of a bacterial pathogen under plant immunity. *Proceedings of the National Academy of Sciences* 115: E3055–E3064.
- Nordgaard, M., C. Blake, G. Maróti, G. Hu, Y. Wang, M. L. Strube, and Á. T. Kovács. 2022. Experimental evolution of *Bacillus subtilis* on *Arabidopsis thaliana* roots reveals fast adaptation and improved root colonization. *iScience* 25: 104406.
- Oksanen, J., G. F. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minching, et al. 2020. vegan: Community Ecology Package. R package, version 2.5-7. <https://CRAN.R-project.org/package=vegan>.
- O’Leary, B. M., C. P. Lee, O. K. Atkin, R. Cheng, T. B. Brown, and A. H. Millar. 2017. Variation in leaf respiration rates at night correlates with carbohydrate and amino acid supply. *Plant Physiology* 174: 2261–2273.
- O’Leary, B. M., H. C. Neale, C.-M. Geilfus, R. W. Jackson, D. L. Arnold, and G. M. Preston. 2016. Early changes in apoplast composition associated with defence and disease in interactions between *Phaseolus vulgaris* and the halo blight pathogen *Pseudomonas syringae* pv. phaseolicola. *Plant, Cell & Environment* 39: 2172–2184.
- O’Leary, B. M., A. Rico, S. McCraw, H. N. Fones, and G. M. Preston. 2014. The infiltration-centrifugation technique for extraction of apoplastic fluid from plant leaves using *Phaseolus vulgaris* as an example. *Journal of Visualized Experiments: JoVE*: e52113.
- Olsson, S., E. Bååth, and B. Söderström. 1987. Growth of *Verticillium dahlia* Kleb. hyphae and of bacteria along the roots of rape (*Brassica napus* L.) seedlings. *Canadian Journal of Microbiology* 33: 916–919.
- Oña, L., and C. Kost. 2022. Cooperation increases robustness to ecological disturbance in microbial cross-feeding networks. *Ecology Letters* PMID: 35384221.
- Panchal, S., R. Chitrakar, B. K. Thompson, N. Obulareddy, D. Roy, W. S. Hambright, and M. Melotto. 2016. Regulation of stomatal defense by air relative humidity. *Plant Physiology* 172: 2021–2032.
- Paulsen, I. T., C. M. Press, J. Ravel, D. Y. Kobayashi, G. S. A. Myers, D. V. Mavrodi, R. T. DeBoy, et al. 2005. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nature Biotechnology* 23: 873–878.
- Paungfoo-Lonhienne, C., D. Rentsch, S. Robatzek, R. I. Webb, E. Sagulenko, T. Näsholm, S. Schmidt, and T. G. A. Lonhienne. 2010. Turning the table: plants consume microbes as a source of nutrients. *PLOS ONE* 5: e11915.
- Penet, L., S. Guyader, D. Péto, M. Salles, and F. Bussi ere. 2014. Direct splash dispersal prevails over indirect and subsequent spread during rains in *Colletotrichum gloeosporioides* infecting yams. *PLOS ONE* 9: e115757.

- Pieterse, C. M. J., A. Leon-Reyes, S. Van der Ent, and S. C. M. Van Wees. 2009. Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology* 5: 308–316.
- Pracharoenwattana, I., W. Zhou, O. Keech, P. B. Francisco, T. Udomchalothorn, H. Tschoep, M. Stitt, et al. 2010. *Arabidopsis* has a cytosolic fumarase required for the massive allocation of photosynthate into fumaric acid and for rapid plant growth on high nitrogen. *The Plant Journal* 62: 785–795.
- Prerostova, S., P. I. Dobrev, B. Kramna, A. Gaudinova, V. Knirsch, L. Spichal, M. Zatloukal, and R. Vankova. 2020. Heat acclimation and inhibition of cytokinin degradation positively affect heat stress tolerance of *Arabidopsis*. *Frontiers in Plant Science* 11: 87.
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Raaijmakers, J. M., and M. Mazzola. 2012. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annual Review of Phytopathology* 50: 403–424.
- Rahme, L. G., M. N. Mindrinos, and N. J. Panopoulos. 1992. Plant and environmental sensory signals control the expression of *hrp* genes in *Pseudomonas syringae* pv. phaseolicola. *Journal of Bacteriology* 174: 3499–3507.
- Ramírez-Sánchez, D., C. Gibelin-Viala, B. Mayjonade, R. Duflos, E. Belmonte, V. Pailler, C. Bartoli, et al. 2022. Investigating genetic diversity within the most abundant and prevalent non-pathogenic leaf-associated bacteria interacting with *Arabidopsis thaliana* in natural habitats. *Frontiers in Microbiology* 13.
- Rashed, A., T. D. Nash, L. Paetzold, F. Workneh, and C. M. Rush. 2012. Transmission efficiency of *Candidatus Liberibacter solanacearum* and potato zebra chip disease progress in relation to pathogen titer, vector numbers, and feeding sites. *Phytopathology* 102: 1079–1085.
- Redford, A. J., R. M. Bowers, R. Knight, Y. Linhart, and N. Fierer. 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental microbiology* 12: 2885–2893.
- Regalado, J., D. S. Lundberg, O. Deusch, S. Kersten, T. Karasov, K. Poersch, G. Shirsekar, and D. Weigel. 2020. Combining whole-genome shotgun sequencing and rRNA gene amplicon analyses to improve detection of microbe–microbe interaction networks in plant leaves. *The ISME Journal* 14: 2116–2130.
- Reitmeier, S., S. Kiessling, T. Clavel, M. List, E. L. Almeida, T. S. Ghosh, K. Neuhaus, et al. 2020. Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. *Cell Host & Microbe* 28: 258-272.e6.
- Remus-Emsermann, M. N. P., S. Lücker, D. B. Müller, E. Potthoff, H. Daims, and J. A. Vorholt. 2014. Spatial distribution analyses of natural phyllosphere-colonizing bacteria on *Arabidopsis thaliana* revealed by fluorescence in situ hybridization. *Environmental Microbiology* 16: 2329–2340.

- Remus-Emsermann, M. N. P., R. Tecon, G. A. Kowalchuk, and J. H. J. Leveau. 2012. Variation in local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere. *The ISME journal* 6: 756–765.
- Rico, A., and G. M. Preston. 2008. *Pseudomonas syringae* pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. *Molecular plant-microbe interactions: MPMI* 21: 269–282.
- Rocheftort, A., M. Simonin, C. Marais, A.-Y. Guillermin-Erckelboudt, M. Barret, and A. Sarniguet. 2021. Transmission of seed and soil microbiota to seedling. *mSystems* 6: e0044621.
- Ruan, Y.-L., Y. Jin, Y.-J. Yang, G.-J. Li, and J. S. Boyer. 2010. Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Molecular Plant* 3: 942–955.
- Ruhe, J., M. T. Agler, A. Placzek, K. Kramer, I. Finkemeier, and E. M. Kemen. 2016. Obligate biotroph pathogens of the genus *Albugo* are better adapted to active host defense compared to niche competitors. *Frontiers in Plant Science* 7.
- Runge, P., F. Ventura, E. Kemen, and R. Stam. 2022. Distinct phyllosphere microbiome of wild tomato species in central Peru upon dysbiosis. *Microbial Ecology*: 1–16.
- Ryffel, F., E. J. Helfrich, P. Kiefer, L. Peyriga, J.-C. Portais, J. Piel, and J. A. Vorholt. 2016. Metabolic footprint of epiphytic bacteria on *Arabidopsis thaliana* leaves. *The ISME Journal* 10: 632–643.
- Sage, R. F., T. L. Sage, and F. Kocacinar. 2012. Photorespiration and the evolution of C4 photosynthesis. *Annual Review of Plant Biology* 63: 19–47.
- Sattelmacher, B. 2001. The apoplast and its significance for plant mineral nutrition. *New Phytologist* 149: 167–192.
- Schäfer, M., C. M. Vogel, M. Bortfeld-Miller, M. Mittelviefhaus, and J. A. Vorholt. 2022. Mapping phyllosphere microbiota interactions in planta to establish genotype–phenotype relationships. *Nature Microbiology* 7: 856–867.
- Schiro, G., G. Verch, V. Grimm, and M. E. H. Müller. 2018. *Alternaria* and *Fusarium* fungi: differences in distribution and spore deposition in a topographically heterogeneous wheat field. *Journal of Fungi* 4: 63.
- Schlechter, R. O., H. Jun, M. Bernach, S. Oso, E. Boyd, D. A. Muñoz-Lintz, R. C. J. Dobson, et al. 2018. Chromatic bacteria – A broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Frontiers in Microbiology* 9: 3052.
- Segarra, G., E. Casanova, M. Avilés, and I. Trillas. 2010. *Trichoderma asperellum* strain T34 controls *Fusarium* wilt disease in tomato plants in soilless culture through competition for iron. *Microbial Ecology* 59: 141–149.
- Singsaas, E. L. 2000. Terpenes and the thermotolerance of photosynthesis. *The New Phytologist* 146: 1–4.

- Söderberg, K. H., and E. Bååth. 1998. Bacterial activity along a young barley root measured by the thymidine and leucine incorporation techniques. *Soil Biology and Biochemistry* 30: 1259–1268.
- Solomon, P. S., and R. P. Oliver. 2001. The nitrogen content of the tomato leaf apoplast increases during infection by *Cladosporium fulvum*. *Planta* 213: 241–249.
- Staley, C., A. P. Ferrieri, M. M. Tfaily, Y. Cui, R. K. Chu, P. Wang, J. B. Shaw, et al. 2017. Diurnal cycling of rhizosphere bacterial communities is associated with shifts in carbon metabolism. 5: 1–13.
- Stauber, J. L., E. Loginicheva, and L. M. Schechter. 2012. Carbon source and cell density-dependent regulation of type III secretion system gene expression in *Pseudomonas syringae* pathovar tomato DC3000. *Research in Microbiology* 163: 531–539.
- Streubel, J., C. Pesce, M. Hutin, R. Koebnik, J. Boch, and B. Szurek. 2013. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *The New Phytologist* 200: 808–819.
- Stringlis, I. A., K. Yu, K. Feussner, R. de Jonge, S. Van Bentum, M. C. Van Verk, R. L. Berendsen, et al. 2018. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proceedings of the National Academy of Sciences* 115: E5213–E5222.
- Sulpice, R., A. Flis, A. A. Ivakov, F. Apelt, N. Krohn, B. Encke, C. Abel, et al. 2014. *Arabidopsis* coordinates the diurnal regulation of carbon allocation and growth across a wide range of photoperiods. *Molecular Plant* 7: 137–155.
- Taiz, L., E. Zeiger, I. M. Møller, and A. Murphy. 2015. Plant physiology and development. 6th ed. Sinauer Associates Incorporated.
- Tegeder, M., and U. Z. Hammes. 2018. The way out and in: phloem loading and unloading of amino acids. *Current Opinion in Plant Biology* 43: 16–21.
- Thaiss, C. A., D. Zeevi, M. Levy, G. Zilberman-Schapira, J. Suez, A. C. Tengeler, L. Abramson, et al. 2014. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell* 159: 514–529.
- Thiergart, T., P. Durán, T. Ellis, N. Vannier, R. Garrido-Oter, E. Kemen, F. Roux, et al. 2020. Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations. *Nature Ecology & Evolution* 4: 122–131.
- Tierens, K. F. M.-J., B. P. H. J. Thomma, M. Brouwer, J. Schmidt, K. Kistner, A. Porzel, B. Mauch-Mani, et al. 2001. Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of *Arabidopsis* to microbial pathogens. *Plant Physiology* 125: 1688–1699.
- Ting, A. S. Y., and J. Y. Chai. 2015. Chitinase and β -1,3-glucanase activities of *Trichoderma harzianum* in response towards pathogenic and non-pathogenic isolates: Early indications of compatibility in consortium. *Biocatalysis and Agricultural Biotechnology* 4: 109–113.

- Trivedi, P., J. E. Leach, S. G. Tringe, T. Sa, and B. K. Singh. 2020. Plant–microbiome interactions: from community assembly to plant health. *Nature Reviews Microbiology* 18: 607–621.
- Truchado, P., M. I. Gil, M. Moreno-Candel, and A. Allende. 2019. Impact of weather conditions, leaf age and irrigation water disinfection on the major epiphytic bacterial genera of baby spinach grown in an open field. *Food Microbiology* 78: 46–52.
- Tukey, H. B. 1966. Leaching of metabolites from above-ground plant parts and its implications. *Bulletin of the Torrey Botanical Club* 93: 385–401.
- Turner, S., K. M. Pryer, V. P. Miao, and J. D. Palmer. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *The Journal of Eukaryotic Microbiology* 46: 327–338.
- Vannier, N., M. Agler, and S. Hacquard. 2019. Microbiota-mediated disease resistance in plants. *PLOS Pathogens* 15: e1007740.
- Vannier, N., C. Mony, A.-K. Bittebiere, S. Michon-Coudouel, M. Biget, and P. Vandenkoornhuysen. 2018. A microorganisms' journey between plant generations. *Microbiome* 6: 79.
- Veillet, F., C. Gaillard, P. Coutos-Thévenot, and S. La Camera. 2016. Targeting the AtCWIN1 gene to explore the role of invertases in sucrose transport in roots and during *Botrytis cinerea* infection. *Frontiers in Plant Science* 7: 1899.
- Velásquez, A. C., J. C. Huguet-Tapia, and S. Y. He. 2022. Shared in planta population and transcriptomic features of nonpathogenic members of endophytic phyllosphere microbiota. *Proceedings of the National Academy of Sciences* 119: e2114460119.
- Venkat, A., and S. Muneer. 2022. Role of circadian rhythms in major plant metabolic and signaling pathways. *Frontiers in Plant Science* 13.
- Vogel, C. M., D. B. Potthoff, M. Schäfer, N. Barandun, and J. A. Vorholt. 2021. Protective role of the *Arabidopsis* leaf microbiota against a bacterial pathogen. *Nature Microbiology* 6: 1537–1548.
- Voges, M. J. E. E., Y. Bai, P. Schulze-Lefert, and E. S. Sattely. 2019. Plant-derived coumarins shape the composition of an *Arabidopsis* synthetic root microbiome. *Proceedings of the National Academy of Sciences of the United States of America* 116: 12558–12565.
- Vorholt, J. A. 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology* 10: 828–840.
- Wagner, M. R., D. S. Lundberg, T. G. Del Rio, S. G. Tringe, J. L. Dangl, and T. Mitchell-Olds. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature Communications* 7: 12151.

- Wahl, R., K. Wippel, S. Goos, J. Kämper, and N. Sauer. 2010. A novel high-affinity sucrose transporter is required for virulence of the plant pathogen *Ustilago maydis*. *PLoS Biology* 8: e1000303.
- Wang, C., X. Wang, Z. Jin, C. Müller, T. A. M. Pugh, A. Chen, T. Wang, et al. 2022. Occurrence of crop pests and diseases has largely increased in China since 1970. *Nature Food* 3: 57–65.
- Wang, L., A. Czedik-Eysenberg, R. A. Mertz, Y. Si, T. Tohge, A. Nunes-Nesi, S. Arrivault, et al. 2014. Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. *Nature Biotechnology* 32: 1158–1165.
- Wei, Z. M., B. J. Sneath, and S. V. Beer. 1992. Expression of *Erwinia amylovora* *hrp* genes in response to environmental stimuli. *Journal of Bacteriology* 174: 1875–1882.
- Weise, S. E., K. J. van Wijk, and T. D. Sharkey. 2011. The role of transitory starch in C3, CAM, and C4 metabolism and opportunities for engineering leaf starch accumulation. *Journal of Experimental Botany* 62: 3109–3118.
- Wieder, C., C. Frainay, N. Poupin, P. Rodríguez-Mier, F. Vinson, J. Cooke, R. P. Lai, et al. 2021. Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. *PLoS Computational Biology* 17: e1009105.
- Winsor, G. L., E. J. Griffiths, R. Lo, B. K. Dhillon, J. A. Shay, and F. S. L. Brinkman. 2016. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Research* 44: D646-653.
- Witzel, K., M. Shahzad, A. Matros, H.-P. Mock, and K. H. Mühling. 2011. Comparative evaluation of extraction methods for apoplastic proteins from maize leaves. *Plant Methods* 7: 1–11.
- Wright, C. A., and G. A. Beattie. 2004. *Pseudomonas syringae* pv. tomato cells encounter inhibitory levels of water stress during the hypersensitive response of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* 101: 3269–3274.
- Xia, J., N. Psychogios, N. Young, and D. S. Wishart. 2009. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research* 37: W652-660.
- Xin, X.-F., K. Nomura, K. Aung, A. C. Velásquez, J. Yao, F. Boutrot, J. H. Chang, et al. 2016. Bacteria establish an aqueous living space in plants crucial for virulence. *Nature* 539: 524–529.
- Xiong, C., Y.-G. Zhu, J.-T. Wang, B. Singh, L.-L. Han, J.-P. Shen, P.-P. Li, et al. 2021. Host selection shapes crop microbiome assembly and network complexity. *New Phytologist* 229: 1091–1104.
- Yamada, K., Y. Saijo, H. Nakagami, and Y. Takano. 2016. Regulation of sugar transporter activity for antibacterial defense in *Arabidopsis*. *Science* 354: 1427–1430.

- Yu, P., X. He, M. Baer, S. Beirinckx, T. Tian, Y. A. T. Moya, X. Zhang, et al. 2021. Plant flavones enrich rhizosphere Oxalobacteraceae to improve maize performance under nitrogen deprivation. *Nature Plants* 7: 481–499.
- Yu, X., S. P. Lund, R. A. Scott, J. W. Greenwald, A. H. Records, D. Nettleton, S. E. Lindow, et al. 2013. Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. *Proceedings of the National Academy of Sciences of the United States of America* 110: E425.
- Zarrinpar, A., A. Chaix, S. Yooseph, and S. Panda. 2014. Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. *Cell Metabolism* 20: 1006–1017.
- Zelezniak, A., S. Andrejev, O. Ponomarova, D. R. Mende, P. Bork, and K. R. Patil. 2015. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences of the United States of America* 112: 6449–6454.
- Zhang, C., M. Gao, N. C. Seitz, W. Angel, A. Hallworth, L. Wiratan, O. Darwish, et al. 2019. LUX ARRHYTHMO mediates crosstalk between the circadian clock and defense in *Arabidopsis*. *Nature Communications* 10: 2543.
- Zhang, C., J. He, H. Dai, G. Wang, X. Zhang, C. Wang, J. Shi, et al. 2021. Discriminating symbiosis and immunity signals by receptor competition in rice. *Proceedings of the National Academy of Sciences* 118: e2023738118.
- Zhang, C., and R. Turgeon. 2018. Mechanisms of phloem loading. *Current Opinion in Plant Biology* 43: 71–75.
- Zhang, C., Q. Xie, R. G. Anderson, G. Ng, N. C. Seitz, T. Peterson, C. R. McClung, et al. 2013. Crosstalk between the circadian clock and innate immunity in *Arabidopsis*. *PLOS Pathogens* 9: e1003370.
- Zhang, F. J., Y. J. Jin, X. Y. Xu, R. C. Lu, and H. J. Chen. 2008. Study on the extraction, purification and quantification of jasmonic acid, abscisic acid and indole-3-acetic acid in plants. *Phytochemical analysis: PCA* 19: 560–567.
- Zhang, L., M. Zhang, S. Huang, L. Li, Q. Gao, Y. Wang, S. Zhang, et al. 2022. A highly conserved core bacterial microbiota with nitrogen-fixation capacity inhabits the xylem sap in maize plants. *Nature Communications* 13: 3361.
- Zhao, K., B. Ma, Y. Xu, E. Stirling, and J. Xu. 2021. Light exposure mediates circadian rhythms of rhizosphere microbial communities. *The ISME Journal* 15: 2655–2664.
- Zhou, F., A. Emonet, V. Déneraud Tendon, P. Marhavy, D. Wu, T. Lahaye, and N. Geldner. 2020. Co-incidence of damage and microbial patterns controls localized immune responses in roots. *Cell* 180: 440–453.e18.
- Zhou, J., and D. Ning. 2017. Stochastic community assembly: does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews* 81: e00002-17.

- Zhu, H., S. Li, Z. Hu, and G. Liu. 2018. Molecular characterization of eukaryotic algal communities in the tropical phyllosphere based on real-time sequencing of the 18S rDNA gene. *BMC Plant Biology* 18: 365.
- Zimmermann, J., N. Obeng, W. Yang, B. Pees, C. Petersen, S. Waschina, K. A. Kissoyan, et al. 2020. The functional repertoire contained within the native microbiota of the model nematode *Caenorhabditis elegans*. *The ISME Journal* 14: 26–38.

Acknowledgments

These five years have been quite a journey.

I am first of all thankful with God for the opportunity given and for constantly renewing my hopes.

I also thank myself for taking the challenge and committing to it till the end.

This experience would not have been possible without the support from the amazing people who have been around:

My supervisor Matt Agler. He made the lab environment welcoming from the first day and I soon discovered how dedicated and committed to our projects he was. I am thankful for his support at many levels, for respecting the pace of my development as a scientist and for taking the time to really listen to us. It's been a challenging process, but Matt helped me keep my motivation high.

Prof. Erika Kothe who since the beginning of the PhD has taken the time to listen to my talks and give very valuable feedback. I am also very grateful for the support she gave me during the last year and for creating a nice environment in the Department.

Teresa Mayer. Living and working in Germany was definitely easier with her help and friendship. I am very thankful that we started this process together and that we supported one another when needed.

Abdullah and Aminat for their hard work, and especially for their patience when I did not know how to proceed. Working with them was also fun and seeing how they developed their research projects and finalized their thesis was very motivating.

I am also thankful to all my other lab mates, for their valuable input and for keeping a nice environment in the office. I am especially grateful to Kerstin Unger and Jisna Jose for their friendship and support. To Jisna, thanks for the help with molecular work.

Dr. Nico Ueberschaar from the MS Platform for his valuable input and for finding the time to run my samples and answer my multiple questions.

Prof. Miriam Agler-Rosenbaum for her support, especially during the last years. The short time I was part of her group allowed me to see the great leader and person she is. Thanks also to DeDe Man and Ivan Schlembach from here group for their help with the experiments.

Christine Vogler (ILRS) and Katja Prafke (JSMC) for their continuous help and advice throughout these years. They were always available to answer my questions about German bureaucracy.

Katrin Krause, Christin Reichmann, Peggy Brand-Schön, Petra Mitscherlich, Robert Blumenstein and all the people from NG25 for helping whenever needed.

Sira Groscurth who encouraged me to apply to the position and then helped me throughout the application process, with moving into Germany, with multiple tips and finally by translating my abstract to German.

Dina Thäle-Hoffmann from the Psychosoziale Beratung and Anne-Christin Warskulat from Curiositas for providing me with great tools to cope with the uncertainty and challenges of the PhD.

To Martha and Kilian Krügel, for their friendship and support when most needed, for reminding me not to work too hard and for introducing me to bouldering and climbing. The wall constantly reminds me to breathe and to be present. It has taught me resilience, humbleness, and the awesome feeling of taking risks!

Quisiera agradecer también a la Universidad de Costa Rica, en especial a Karol Cordero de la OAICE, por el apoyo brindado en el último año para mi formación académica en Alemania. También agradecer a la Dra. Lidieth Uribe del CIA por motivarme a realizar los estudios de doctorado.

A “las guapas de Jena” y a los amigos en Halle, que hicieron de estos años una experiencia inolvidable. Gracias por el apoyo, los consejos y los momentos tan bonitos que hemos pasado juntos. Gracias en especial a Kathy y Dolly con quién pude compartir miedos e ilusiones del proceso.

A Jo porque aunque ya no está aquí, la tengo presente constantemente y sé que estaría orgullosa.

A mis papás y hermanos porque, aunque hemos estado lejos estos años, siempre los he sentido cerca. El apoyo y los consejos que me han dado han sido vitales para mantenerme a flote y para disfrutar la oportunidad de estar aquí.

Paúl, por motivarme a empezar este proceso y por creer en mí. Gracias por estar dispuesto a reinventarnos constantemente, por ser mi compañero de vida y aventuras y sobre todo por darme un espacio lleno de amor. La vida con vos es más bonita.

Gracias a Gael, por enseñarme lo que es importante, a vivir más despacio y más conscientemente.

Statement of authorship

I confirm I am familiar with the valid doctoral examination regulations and that this dissertation was done independently by me. Nothing else than the mentioned references and sources were used. I have not submitted this thesis as a test paper for any governmental or other form of scientific examination. Moreover, I have not tried to submit this thesis or a substantially similar version as a dissertation to another university. I did not receive any assistance from specialized consultants to conduct this work or write the document. No third party received either direct or indirect financial benefits from this work.

Ich bestätige, dass mir die gültige Promotionsordnung bekannt ist und dass diese Dissertation von mir selbstständig angefertigt wurde. Es wurden ausschließlich die genannten Referenzen und Quellen verwendet. Ich habe diese Arbeit nicht als Prüfungsarbeit zu einer amtlichen oder sonstigen wissenschaftlichen Prüfung eingereicht. Außerdem habe ich nicht versucht, diese Arbeit oder eine im Wesentlichen ähnliche Version als Dissertation an einer anderen Universität einzureichen. Ich habe keine Unterstützung von spezialisierten Beratern erhalten, um diese Arbeit durchzuführen oder das Dokument zu schreiben. Kein Dritter hat aus dieser Arbeit direkten oder indirekten finanziellen Nutzen gezogen.

Jena, 23 of January 2023

Mariana Murillo Roos

Curriculum vitae

Mariana Murillo Roos

Born: 03.08.1989, San José, Costa Rica
 06110 Halle, Germany
 mariana.murillo.roos@uni-jena.de

EDUCATION**PhD in Microbiology** **2018 to 2023**

Friedrich-Schiller-Universität

- Thesis: Chemical landscape in leaves – shaping microbial interactions and community structure. Supervisors: Prof. Dr. Erika Kothe and PhD Matthew Agler.

MSc in Biology **2014 - 2018**

University of Costa Rica

- Thesis: Biotechnological potential of *Umbelopsis isabellina* to produce lipids of industrial importance from pineapple plant residues. Supervisors: PhD Wei Liao (Michigan State University) and PhD Lidieth Uribe (University of Costa Rica).

Professional Licentiate in Agronomy **2012 - 2014**

University of Costa Rica

- Thesis: Biogas production from organic wastes in lab-scale reactors operated under two temperatures. Supervisor: PhD Lidieth Uribe (University of Costa Rica)

BSc in Agronomy **2007 - 2012**

University of Costa Rica

WORK EXPERIENCE**Agronomist** **2014 - 2018**

Instituto Nacional de Innovación y Transferencia en Tecnología Agropecuaria, Ministerio de Agricultura y Ganadería, Costa Rica.

Undergraduate research intern

- Biosystems Department of Michigan State University **2016**
- Biosystems Department of Michigan State University **2012**
- Centro de Investigaciones en Biología Celular y Molecular, Universidad de Costa Rica **2012 – 2013**
- Laboratorio de Microbiología Agrícola, Universidad de Costa Rica **2011 – 2014**

PROFESSIONAL DEVELOPMENT**Courses/Training** **2018 - 2022**

- Jena School for Microbial Communication, Graduate Academy of Friedrich-Schiller-Universität and International Max Planck Research School
 - “Career Planning”, “Critical argumentation and reasoning”, “Dark Side of Science”, “Project Management”, “Scientific Writing”, “Functional Assays are Not ‘Just-So’ Stories”, “Untargeted metabolomics using high-resolution LC-MS”, “The basics of light and fluorescence microscopy”, “Bioinformatic Tools in Mass spectrometry”

PUBLICATIONS

- **Murillo-Roos M.**; Syed M Abdullah H.; Debbar M.; Ueberschaar N.; Agler M. Cross feeding niches among commensal leaf bacteria are shaped by the interaction of strain-level diversity and resource availability. *ISME J* 16, 2280–2289
- **Murillo-Roos, M.**; Uribe-Lorío, L.; Fuentes-Schweizer, P.; Vidaurre-Barahona, D.; Brenes-Guillén, L.; Jiménez, I.; Arguedas, T.; Liao, W.; Uribe, L. (2022). Biogas production and microbial communities of mesophilic and thermophilic anaerobic co-digestion of animal manures and food wastes in Costa Rica. *Energies*, 15(9), 3252.

- Mayer T, Mari A, Almario J, **Murillo-Roos M**, Syed M Abdullah H, 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(6), 1952-1965.
- Chen, A., Guan, Y. J., Bustamante, M., Uribe, L., Uribe-Lorío, L., **Roos, M. M.**, & Liu, Y. (2020). Production of renewable fuel and value-added bioproducts using pineapple leaves in Costa Rica. *Biomass and Bioenergy*, 141, 105675.
- Rojas-Sossa J.P., **Roos, M.M.**, Uribe, L; Lorío, L.U., Marsh, T., Larsen, N., Miranda, A., Solís, K., Kirk, D., Liao, W. 2017. Effects of coffee processing residues on anaerobic microorganisms and corresponding digestion performance. *Bioresource Technology*, 245, 714-723.
- Chen, R., **Roos, M. M.**, Zhong, Y., Marsh, T., Roman, M. B., Ascencio, W. H., ... & Chavarria, J. A. M. (2016). Responses of anaerobic microorganisms to different culture conditions and corresponding effects on biogas production and solid digestate quality. *Biomass and Bioenergy*, 85, 84-93.
- Rojas, C., Valverde, R., Rollins, A. W., & **Roos, M. M.** (2016). What can myxomycetes tell us about floricolous microbial systems?. *Nova Hedwigia*, 103(3-4), 3-4.

CONFERENCE AND SYMPOSIUM PARTICIPATION

- *The leaf nutrient landscape can drive bacterial diversity by differentially selecting on specific metabolic interactions.* International Society of Microbial Ecology (ISME), Laussane, Switzerland. 2022. (Poster presentation)
- *Metabolic niche partitioning during cross-feeding is linked to diversity in leaf bacterial communities.* Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM). Online. 2022. (Oral presentation)
- *Linking the leaf microbiome to leaf metabolite landscape.* Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM). Leipzig, Germany. 2021. (Poster presentation).
- *Chemical landscapes in leaves as a driver of microbe-microbe interactions and community structure.* Strasburger-Workshop: Diversity of chemical and molecular signaling networks in plant-microbe interaction. Bonn, Germany. 2020. (Oral presentation).

RECOGNITIONS

- **Siegfried Czapski Publication Prize**, Jena Alliance, Friedrich-Schiller-Universität **2022**
- **Poster Award**, Vereinigung für Allgemeine und Angewandte Mikrobiologie **2020**

SERVICE AND LEADERSHIP

- **Program *Agentes de Cambio*, Friedrich Ebert/Fesamerica** **2010**
 - Socio-political program promoting inclusive participation of civil society.
- **Volunteer at *Un Techo para mi país*.** **2007-2011**
 - Working with impoverished communities to pursue transitional housing

LANGUAGES

- Spanish: native language
- English: fluent writing and speaking (TOEFL 110 out of 120)
- German: B1