Unravelling plant-microbe-microbe dynamics using *Arabidopsis thaliana* and synthetic communities of *Pseudomonas*

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität

Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Or Shalev Skriptchak

aus Haifa, Israel

Tübingen

2021

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

Tag der mündlichen Qualifikation:	29.06.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Detlef Weigel
2. Berichterstatter:	Prof. Dr. Nadine Ziemert

Acknowledgments

This thesis is based on a work of five years. Five years is a long period of time to encounter many good people, experiences, and to grow and develop (hopefully in the right direction). When I started, I was a young married man without children, hopping into an adventure in an anonymous town called Tuebingen. Five years later (a little less young) I am a father of two, and find this picturesque town as my second home (though my German is 'nicht so gut'). I am grateful I took this wonderful path.

To Detlef Weigel, thanks for giving me the chance to experience this adventure, and to be a part of this really great team. Not many are that fortunate. You were always decent and kind, though I sometimes did not make it easy for you staying in this mindset :). You made me wiser in many ways.

To Haim Ashkenazy, I have too many things to thank for. You taught me bioinformatics, critical thinking, writing and drinking Raki. I could not have asked for a more intelligent and inspiring colleague. But way beyond all of this, thanks for being a true dear friend and a wonderful person! (out of the many examples - you managed to keep a poker face after I broke my teeth. For a good 30 seconds)

To Derek Lundberg and Talia Karasov, you were my practical, everyday gurus. In particular, helped me to initiate everything. Which was not easy at all. I remember many days wanting to give up, considering shifting my profession to farming since I had no clue what I am doing here. You two were my lifebuoy. Thanks for contributing to my path.

To Nadine Ziemert and Ruth Ley, thanks for taking your time to guide me as members of my Thesis Advisory Committee. You helped me to find the weak spots in my research before they emerged into an irreversible problem. It greatly advanced my project, and prevented me

unnecessary guilt and suffering (which are the bread and butter of the workaholic). Also, much appreciation for both of you and to Eric Kemen for taking your time to review this thesis and to take a part in the defence.

Much appreciation to all of WeigelWorld! There are too many to name, but all of you are wonderful! You made this environment a community, and a place where I can be myself and be happy. At the end, this is life, and not the work we do or the papers we publish. I learned much about friendship and got to know fun facts from around the world (...yes, Chinese do eat everything). I am proud that I was a part of this team. Special thanks to Miriam Lucke for translating the summary from English to German, while requesting only a modest favour in return.

To my great great friends which are like family to me, both in Israel and Germany: Oded, Kruvi, Omer, Shir (a.k.a 'money time'), my brothers in arms Tomer, Golan, Shay and Vidi (a.k.a 'Subterranean Masquerade'), Polina and Fabi, the Verena and the David, madam Bridgit and sir Marco, and for many others which cannot fit this page… Grateful to have you in my life.

To my family who lives in Israel. My mother, my father, my sister Keren, nephew and nice Itay and Noya, my brother Elad. Tania and Alexey, Daria and Refael. I missed you all very much during our stay in Germany. Thanks for so much support!

Lastly, to my loving wife Taisa - you are the source. Nothing would mean anything without you. Thank you for sharing this long path with me. To my beautiful kids Aria and David - you are my heart. In every step I take, you are in my mind. I hope that by the time you can read this, you already sleep well at night. Perhaps I wish this to myself. Yassu - you are a stinky but a lovely dog. And a real Tuebinger.

Table of contents

Summary 8

Zusammenfassung 10

Introduction 13

General introduction and motivation 13

Individual plant-microbe interactions 14

Plant defensive system to control microbial colonization 14

The underlying coevolution between plants and microbes 16

The effect of individual microbes on plant health 17

Plants microbial ecology: shifting from individuals to communities 18

Technological aspects in microbial characterization 19

The microbiota of the phyllosphere: a random mixture or a well defined team? 21

Mechanisms of microbial antagonism 22

Microbe-microbe interactions and the plant microbiota 26

The reciprocal effects between the plant and the microbiota 29

Protective microbial interactions in plants 29

Protective microbe-host-microbe communication: interactions of microbes via the plant immune system 31

Aims and objectives 31

Chapter one 34

Protective host-dependent antagonism among Pseudomonas in the Arabidopsis phyllosphere 34

Chapter two 36

Commensal Pseudomonas protect Arabidopsis thaliana from a coexisting pathogen via multiple taxonomy-dependent mechanisms 36

Discussion 38

Disentangling microbe-microbe and host-microbe from the plant-microbe-microbe relationships 38

Microbial dark matter: commensal microbes in plants 40

The plant microbiome and its function in the age of whole genome sequencing 41

Conclusions and outlook 44

References 49

Thesis appendix I-II

Summary

Organisms affect and shape each other, both during their own lifespans and in evolutionary terms. The relationship between a host and its colonizing microbes can have major immediate and long-term effects on host health. Whether the colonizing microbes have a good impact or a bad one depends in part on how those microbes interact with each other. Overproliferation of pathogenic microbes is associated with negative impacts on the host health. This can be countered by protective microbes, which may suppress pathogenic ones. Taken together with the host immune system, the complicated host-microbe-microbe dynamics form a balance of fitness between the host and its microbes. Of specific interest are the dynamics between plants and microbes. Phytopathogens harm global agricultural production, yet are often somehow held in check in wild settings. In this thesis, I am studying the plant-microbe-microbe interface using *Arabidopsis thaliana* and its associated bacterial genus *Pseudomonas*, leveraging a collection of 1,524 *Pseudomonas* strains which were isolated from plants of the same geographic region.

In the first chapter, I focus on synergistic effects, studying the interactions between multiple coexisting pathogenic and commensal *Pseudomonas* strains with a panel of *Arabidopsis thaliana* accessions. By employing synthetic communities of genome-barcoded strains, I monitored the abundance of individual isolates in the context of communities, including exclusively commensal, exclusively pathogenic, and mixed commensal and pathogenic communities. I revealed that the inclusion of commensal members led to inhibition of pathogens, preventing the harmful effect on plant biomass. I associated these protective interactions with both microbe-microbe interactions and with the host transcriptomic signature. I found that the extent of plant protection varied with host genotype, further highlighting the role of the host in mediating protective interactions. Finally, I unravelled similar

genotype-specific effects on the microbial side, presenting how an individual *Pseudomonas* pathogenic isolate caused this differential protection effect.

In the second chapter, I investigate (i) the prevalence of protection against pathogens by commensal *Pseudomonas*, (ii) the taxonomic specificity of such protection and (iii) the bacterial elements in commensal *Pseudomonas* that lead to protection. To address these questions, I made systematic co-infections, pairing each of ninety-nine locally-isolated commensal isolates with a local *Pseudomonas* pathogen. The majority of these wild commensal *Pseudomonas* protected the plant to some extent. In particular, one taxonomic group was enriched for protective isolates. However, the ability to protect the plant varied between closely related strains, even within this protective group. I leveraged this variation to conduct a genome-wide association study (GWAS), pinpointing gene orthologs in presence-absence variation that are associated with the protective ability. Instead of a universal set, I found taxon-specific gene sets. According to gene annotation, these sets indicated different mechanisms of protection, including iron-uptake, antibiosis, and motility. Using gene deletion, I validated the role of a subset of genes, confirming a link between plant protection with three iron-uptake genes and one biofilm-related gene.

Collectively, this work advances our knowledge about how genetic diversity in both the microbe and the host affects the outcome of the interaction, disentangling different aspects of this complex system. Among the main conclusions of this work is that commensal bacteria are an important factor in maintaining plant health, acting via multiple competitive microbe-microbe mechanisms and via induction of the host immune response. Ultimately, application of such commensal bacteria to control pathogens may sustainably improve global agriculture.

Zusammenfassung

Organismen befinden sich in ständiger Wechselwirkung und beeinflussen sich gegenseitig, im Zuge ihres eigenen Lebenszyklusses und im Zusammenhang von evolutionären Aspekten. Die Beziehung zwischen Wirt und seinem darauf lebenden Mikroben kann spontane und langfristige Wirkungen auf die Gesundheits des Wirtes haben. Ob die kolonisierten Mikroben gute oder schlechte Einwirkungen auf den Wirt haben, hängt mit den Wechselwirkungen der Mikroben zusammen. Überstimulation von pathogenen Mikroben wirken sich negativ auf die Gesundheit des Wirtes aus. Wohingegen schützende Mikroben eine schützende Wirkung auf den Wirt ausüben können, indem sie das Wachstum von Pathogenen hemmen. Zusammen mit dem Immunsystem des Wirtes und der komplexen Dynamik von Mikrob-Mikrob Interaktionen entsteht ein Fitness Gleichgewicht zwischen dem Wirt und seinen Mikroben. Der Fokus hier liegt auf der Dynamik zwischen Pflanzen und seinen Mikroben. Phytopathogene schädigen die globale Produktion von Nutzpflanzen, dennoch ist es unklar, wie natürliche Umgebungen diese abwehren. Diese Thesis, befasst sich mit den Interaktionen zwischen Wirt-Mikrob-Mikrob an *Arabidopsis thaliana* und seinen assoziierenden *Pseudomonas* Genus, genauer eine Kollektion von 1.524 *Pseudomonas* Spezien die auf derselben Pflanze sowie von derselben geografischen Region isoliert wurden.

Das erste Kapitel umfasst die synergistischen Interaktionen zwischen unterschiedlichen zusammenlebenden pathogenen und kommensalen *Pseudomonas* Spezien, die auf unterschiedlichen *Arabidopsis thaliana* Linien vorkommen. Die synthetischen Gesellschaften wurden basierend auf ihr Genom mit unterschiedlichen Barcoden markiert, dadurch konnte ich die Fülle der individuellen Isolaten im Bezug auf die Gesellschaften analysieren. Aufgebaut waren die Gesellschaften in nur kommensal, nur Pathogenen und gemischen kommensal und pathogenen Gesellschaften. Die Einführung von kommensalen Mikroben hemmen die pathogenen Mikroben und können dadurch die pflanzliche Biomasse schützen. Diese Pflanzen

schützenden Wirkungen wurden wurde weiter untersucht mit zwischen bakteriellen Interaktionen und dem Wirts Transkriptom. Daraus schloss ich, dass der Pflanzenschutz sich unterscheidet zwischen den unterschiedlichen Pflanzen Genotypen. Dies bedeutet, dass der Wirt unterschiedliche schützende Interaktionen einhergeht. Zu guter letzt, beobachtete ich, die aehnliche Genotyp spezifischen Effekte auf der mikrobiellen Seite: wie ein einzelnes pathogenes *Pseudomonas* Isolat verschiedene schützende Effekte hervorrufen kann.

Im zweiten Kapitel, wurden die schützenden kommensalen *Pseudomonas* Isolate genauer untersucht auf (I) ihre Häufigkeit, (II) die spezifischen Taxa, und (III) die bakteriellen Elemente. Dafür wurden verschiedene systematische Co-infektionen von jedem 99 lokal isolierten kommensalen und pathogenen *Pseudomonas* Isolate miteinander zusammengestellt. Die Mehrheit von diesen wilden kommensalen Pseudomonas Isolaten schützen die Pflanzen zu einem gewissen Masse. Besonders auffallend war eine spezifische taxonomische Gruppe, die jedoch Variationen aufzeigte von ihrer schützenden Wirkung mit sehr nah verwandten Spezien, selbst wenn sie in derselben Gruppe auftauchten. Eine Genomweite-assoziation Studie (GWAS) zeigte, dass schützende Wirkung zurückzuführen ist zu einem taxon-spezifischen Gen Sets, die innerhalb einer orthologen Gruppe liegen. Diese Sets sind ein Indiz für die unterschiedlichen Mechanismen die zum Pflanzenschutz führen. Laut den Gen Beschreibungen, sind die Gensets relevant für unterschiedliche Schutzfunktionen, wie die Aufnahme von Eisen, Antibiose, und Motilität. Verschiedene Gendeletionen bestätigen die Konnektivität zwischen Pflanzenschutz und drei Genen die relevant sind für die Aufnahme von Eisen und eins dass zur Bildung von Biofilm beiträgt.

Zusammenfassend, diese Arbeit erweitert unser Wissen wie genetische Diversität einen Effekt hat auf mikrobielle und pflanzliche Ebene, im Sinne von unterschiedlichen Interaktionen und somit wurde einen weiteren Aspekt in diesem komplexen System entschlüsselt. Eine der Hauptaussage dieser Arbeit ist, dass kommensale Bakterien einen entscheidenden Faktor zum Pflanzenschutz beitragen, durch die wett kämpferischen Interaktionen zwischen den Mikroben

und durch die Aktivierung des Immunsystems vom Wirt. Die Anwendung von kommensalen Mikroben um Pathogene zu kontrollieren kann die Nachhaltigkeit der globalen Landwirtschaft stärken.

Introduction

General introduction and motivation

As far as we know, unicellular organisms were the first forms of life on planet earth, dating to about 3.5 billion years ago (Schopf 2006; Pearce et al. 2018). A current estimation dates the emergence of land plants to about 3 billion years later, thus around 500 million years ago (Morris et al. 2018), positioning plants as biological hosts that offer a niche for microbial colonization.

The associations between plants and microbes have reciprocal effects on both the organisms, ranging from beneficial (Finkel et al. 2017) to harmful (Mansfield et al. 2012; Dean et al. 2012). On an ecological scale, these relationships have an importance in shaping the global ecosystem (Langley and Hungate 2014; Vorholt 2012). In particular, they have had a substantial impact on agricultural systems, from the dawn of the neolithic revolution (Stukenbrock and McDonald 2008). Phytopathogens cause various diseases in crops, leading to an annual reduction of more than 10% in worldwide food supply (Strange and Scott 2005), and up to 30% in food-deficit areas (Savary et al. 2019). Besides these implications, dynamics between plants and colonizing microbes project general rules governing other host-microbe systems, such as humans and their colonizers, manifesting our health. Furthermore, microbes also interact directly with each other within the plant, indicating patterns that resemble host-free environments such as the ocean and soil.

Ultimately, understanding microbial colonization of plants facilitates our ability to improve food production in a sustainable fashion. Furthermore, it can advance our knowledge about the host-microbiome concept, promoting health care and global ecosystem preservation.

In this thesis, I am using the plant *Arabidopsis thaliana* and its colonizing bacterial genus *Pseudomonas* as a model for plant-microbe-microbe relationships, focusing on protective

microbial interactions in the phyllosphere (the aerial parts of the plant). I am detailing how genetic diversity affects these relationships, using different genotypes of plants and bacteria sampled from the same geographic region. I leverage this natural genetic diversity to understand the underlying mechanisms leading to pathogen mitigation and a healthy plant, addressing both the host- and microbial-related mechanisms. In a broader context, I am highlighting how the plant-microbe and microbe-microbe dimensions come into play in the multifarious relationship of plant-microbe-microbe, thus moving back and forth between the individual relationships and their ensemble.

Below I introduce current relevant knowledge in the field, placing my work in the appropriate scientific background.

Individual plant-microbe interactions

Wild plants are colonized by a vast diversity of microbes (Bai et al. 2015), manifesting multiple direct interactions between microbes and their hosting plant. In this section I will present the importance of individual plant-microbe interactions, detailing how they shaped the genetic architecture of both plants and microbes. Side-by-side, I will cover relevant literature to understand the extent to which these individual interactions reflect the complex relationships between plants and consortia of microbes; is the whole greater than the sum of its parts?

Plant defensive system to control microbial colonization

Plants evolved several lines of defense to protect themselves from microbial invasion. The first line of defense is physical. The leaf comprises a waxy cuticle that serves as an external barrier, preventing the entering of microbes (Serrano et al. 2014; Reina-Pinto and Yephremov 2009). Microbes can breach the cuticle layer and penetrate into the leaf via natural pores, such as stomata (Melotto, Underwood, and He 2008) and hydathodes (Cerutti et al. 2017), as well as via wounds. Some microbes can actively decompose the cuticle (Garrido et al. 2012). After

breaching the cuticle, microbes will face the cell wall - another physical barrier preventing microbes from exploiting plant resources (Lionetti and Métraux 2014; Miedes et al. 2014), acting also as a reservoir of antimicrobial compounds (Schulze-Lefert 2004). Important to note that both the cuticle and cell wall are not just passive physical barriers, but rather dynamic systems that sense the environment and react (Underwood 2012). This topic is interesting, but beyond the scope of this introduction.

The second line of plant defense comprises an active two-tiered immune system, which specializes in the recognition of microbes, and the activation of appropriate antibacterial pathways. The first tier includes pattern-recognition receptors (PRRs), which are cell surface proteins that recognize general microbial- or pathogen-associated molecular patterns shared across many taxa (MAMPs or PAMPs). These molecules are normally slowly evolving, for instance flagellin (Jones and Dangl 2006). Recognition of PAMPs will result in a specific immune response, referred to as PAMP-triggered immunity (PTI). The second tier of immunity is tailor-made for the recognition of pathogens or lower taxonomic ranks, via molecules that are rapidly evolving. It is actioned within the plant cell by a diverse family of nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins (McHale et al. 2006; Dodds and Rathjen 2010). These NBS-LRRs are also referred to as resistance (R) proteins. R proteins are triggered by bacterial effectors - proteins that are injected into the plant cell and in most cases suppress the initial PTI response, consequently conferring higher survivability of the invading microbe (Toruño, Stergiopoulos, and Coaker 2016). Both PTI and effector-triggered immunity (ETI) lead to the activation of similar antimicrobial mechanisms, such as a burst of reactive oxygen species (Nürnberger and Scheel 2001), though in different response amplitudes (Thomma, Nürnberger, and Joosten 2011; Naveed et al. 2020). The ETI response is considered as stronger. This is manifested by the stereotypical hypersensitive response (HR) - a programmed cell death encapsulating biotrophic microbes, hence preventing their spread within the leaf (Morel and Dangl 1997; Balint-Kurti 2019).

The distinction between ETI and PTI is becoming more and more blurred (Thomma, Nürnberger, and Joosten 2011; Lu and Tsuda 2021), and their cross-talk is still under research. What is clear is that the two mechanisms act together to confer higher survivability of plants in the wild (M.Yuan et al. 2021) .

The underlying coevolution between plants and microbes

During evolution, the antagonistic relationships between plants and microbes yielded co-evolutionary dynamics in which microbes evolved to invade, and plants to protect. At the molecular level, these dynamics are demonstrated by the gene-for-gene model, which describes how pairs of matching genes in a plant and a microbe determine the outcome of infection (Van der Biezen and Jones 1998). Thus, the plant response to a given microbe is often determined by the interaction of two molecules: a receptor (host-derived) and a ligand (microbe-derived). This concept was introduced in 1956 by H.H. Flor (Flor 1956), who examined the inheritance of disease in both flax (*Linum usitatissimum*) and its fungal pathogen *Melampsora lini*. Flor was the first to propose that resistance in plants is the result of matching pairs of individual genes, one in the host and one in the pathogen. In the course of time, numerous experimental evidences have supported Flor's theory (Leonelli et al. 2011; Petit-Houdenot and Fudal 2017; Kobayashi, Tamaki, and Keen 1989; van Dijk et al. 1999), conceptualizing it as a paradigm in plant pathology. Following Flor's concept, the 'zig-zag model' entangles evolution with the gene-for-gene model, presenting the genetic co-arms race over time between two molecules in two different organisms, leading to their diversification (Jones and Dangl 2006). In agreement with the zig-zag model, R genes in the plant species *A. thaliana* are highly polymorphic (Van de Weyer et al. 2019), as is the effector content in its pathogen *Pseudomonas syringae* (Dillon et al. 2019), which is a compatible *A. thaliana* colonizer. This high polymorphism resembles what could be a signature of a tight co-arms race between individual host R genes and their compatible bacterial effectors, as presented in the zig-zag model.

Nonetheless, the zig-zag model considers a scenario with one host and one microbe. In reality, multiple microbes colonize diverse hosts that coexist in the same population. In other words, plants from the very same population vary in their immune genes repertoire (Karasov, Shirsekar, et al. 2020). This diversity leads to balancing selection rather than a selective sweep, and reduces the importance of individual plants (Karasov, Shirsekar, et al. 2020; Karasov et al. 2014; Koenig et al. 2019). Therefore, plant resistance should be considered at the population level, dynamics which are not incorporated in the zig-zag model. Furthermore, although the zig-zag model is considered as a good expository model, it explicitly excludes major determinants of successful microbial colonization, even in the context of individual plants. Among these are the aforementioned simultaneous colonization of multiple microbes and the existence of symbiosis types other than parasitism (Pritchard and Birch 2014).

Consequently, the one dimensional interactions between matching genes in plants and microbes should have a limited predictability of colonization rate in most realistic scenarios. Notwithstanding, it is clear that the plant immune system has an important role in microbial colonization, especially regarding the colonization of phytopathogens.

The effect of individual microbes on plant health

The symbiotic interactions between plants and their colonizers range from mutualism, through commensalism, to parasitism (Vorholt 2012; Finkel et al. 2017). Often, high ratios of microbial to host cells are associated with pathogenicity, as was found in controlled infections (Velásquez et al. 2017; Innerebner, Knief, and Vorholt 2011; Fabro et al. 2011), and concluded about wild pathosystems (Karasov, Neumann, et al. 2020). This portrays a zero-sum game scenario in which microbial proliferation comes at the account of the plant fitness, probably by depleting its resources and reprogramming its physiological development. Nonetheless, there are microbes that increase the ability of the plant to uptake nutrients from the environment, as exemplified by nitrogen fixing bacteria (Jacoby et al. 2017; Hestrin et al. 2019), as well as microbes which

prevent abiotic stresses such as heavy metal detoxifiers (Mishra, Singh, and Arora 2017), presenting a win-win scenario.

Investigating how individual bacteria affect plant health is an important front to understand the function of wild microbiomes. In particular, this approach can provide a mechanistic understanding of microbial effects on plants. However, it is low in throughput, and mostly focuses on a small subset of stereotypical microbes. This stands in contrast to the fact that wild plants are colonized by consortia of microbes, provoking questions about the importance of such binary interactions. An alternative approach is to investigate host-microbe interaction in the context of multiple microbes. The added complexity is that microbes also affect each other. This is exemplified by pathogen-suppressing strains, as found among the genera *Sphingomonas* (Innerebner, Knief, and Vorholt 2011) and *Pseudomonas* (De Vleesschauwer et al. 2008; Ganeshan and Manoj Kumar 2005) .

Plants microbial ecology: shifting from individuals to communities

Both prokaryotes and eukaryotes colonize the plant below and above ground (Bai et al. 2015; Noble et al. 2020; Singh et al. 2019; Lundberg et al. 2012). Information about the identity and function of these microbes is important to understand principles governing microbial colonization of plants. One fundamental question regards the existence of a core microbiota, thus querying about the existence of a specific subset of microbes that repeatedly reside in plants. Such consistency would manifest particular functions that enable these microbes to colonize plants. One aspect of such functions is actually related to microbe-microbe interactions rather than plant-microbe, as microbes interact with each other, affecting the proliferation of their co-colonizers.

In this section, I will introduce current knowledge about the ecology of microbes within plants, starting with a chronological literature review about methods to identify and classify microbes. This will be used as the background for subsequent introductory topics.

Technological aspects in microbial characterization

The identification and classification of microbes are the very basics of microbiology, as even in the simplest study it is required to know which microbes are presented in a given sample. In 1667, the Dutch scientist Antonie van Leeuwenhoek published the first documented observation of microbes through a microscope (Lane 2015). Leeuwenhoek described the forms he noted as "animalcules", which translates from latin as "tiny animals". A crucial improvement in the visualization of microbes came in 1884 by Hans Christian Gram, in the form of bacterial staining (Gram and Friedlaender 1884). While Gram created this technique to better visualize bacteria in lung tissues, his method was subsequently used for bacterial classification, as only one out of two bacterial types can be stained, based on the cell wall type (Coico 2005). Besides microscopy, the cultivation of microbes greatly advanced microbiologists in the identification and classification of microorganisms. Around the same time Gram published his staining method, Angelina Hesse introduced the agar media (Hesse 1992), replacing older cultivating techniques such as food assortment- and gelatin-based media. The usage of agar media, combined with different nutrient sources, allowed microbiologists to cultivate a diverse set of microorganisms, while researching their macroscopic colonies properties. Overall, microscopy and cultivation methods were the available tools for microbial classification in the early days of microbiology, and are still in common use.

However, today we classify microbes primarily relying on their DNA sequence, either marker-gene- or whole-genome-based. The most common marker genes in use are the conserved 16S rDNA for bacterial profiling and the internal transcribed spacers (ITS) 1 and 2 or rDNA for eukaryotic profiling. Profiling is done by PCR-amplification of the gene of interest, followed by sequencing of the PCR product. The resulting sequences can be binned by their similarity to compose 'Operational Taxonomic Units' (OTUs), and an exact nomenclature can be then assigned to these units by mapping to existing databases (these act as a dictionary

between sequences and taxonomic assignments) (Liu et al. 2020). Microbial profiling based on marker genes has a few advantages, resulting from the low complexity of the data and its long-time usage. To name the main advantages: (i) the lack of need in sequence assembly, (ii) well maintained databases, and (iii) the low price, projected from the low coverage needed, thus allowing the affordable profiling of many samples simultaneously (Caporaso et al. 2011). On the other hand, marker-gene-based profiling, using short-read technologies, provides the confidence to taxonomically classify microbes only up to the genus level (Johnson et al. 2019; Earl et al. 2018), which stands in a big contrast to current understandings about differences in genetics even within a species (Barnes, Carter, and Lewis 2020; Tett et al. 2017; Yan et al. 2020). Considering the breadth of evolution in microbial populations - both from the angle of single nucleotide polymorphisms (SNPs) and presence-absence variation of genes (PAV) - marker-gene-based profiling allows only a very limited view of evolutionary understanding in microbes. Furthermore, since even highly related strains can differ in their activities (Barnes, Carter, and Lewis 2020), marker genes are not very suitable for predicting the function of profiled microbes.

An alternative approach is profiling by whole-genome-sequencing (WGS), which in principle is suitable for the analysis of both individual culturable bacteria and for consortia (although single-genome- and metagenomic-data are very different, as detailed below). In contrast to marker-gene profiling, WGS profiling allows the taxonomic classification of microbes up to the strain level, thus microbes within species can be distinguished (Ranjan et al. 2016). Additionally, WGS provides the full genetic architecture of microbes, enabling various gene-based analyses to estimate functions, as well as in-depth evolutionary dynamics in microbes (Levy et al. 2017; Karasov et al. 2018). WGS is generating orders of magnitude more data than marker-gene sequencing, thus is much pricier. Focusing on the plant phyllosphere, WGS is suitable for culturable bacteria, producing a comprehensive analysis for a subset of relevant microbes, as was done in *Pseudomonas* (Karasov et al. 2018). However, as

demonstrated in wild *A. thaliana*, WGS is not suitable for the analysis of microbes in plant leaf samples (Regalado et al. 2020). The main bottleneck is the high amount of host-related reads, diluting microbial reads to a negligible fraction. This limitation is specific to endophytic microbes - colonizers of the inner part of the leaf - which thus far cannot be appropriately separated from the host tissue. However, WGS of microbial communities from abiotic environments can yield enough data to produce metagenome assembled genomes (MAGs) (Tully, Graham, and Heidelberg 2018), therefore enabling the reconstruction of individual genomes from mixed communities.

In summary, DNA sequences of marker genes can be utilized to gain a general overview of the microbial profile in a given sample, affordably, and with a high throughput. Analysis of the exact microbial phylogeny, evolutionary dynamics and microbial functional annotation requires the utilization of WGS, which is more suitable to culturable microbes in the plant phyllosphere.

The microbiota of the phyllosphere: a random mixture or a well defined team?

Among the pioneering studies that employed culture-independent methods to appreciate microbial communities, Yang et al. (Yang et al. 2001) were the first to profile the microbiota of the phyllosphere. Using the technology available at the time, they utilized denaturing gradient gel electrophoresis (DGGE) with 16S rDNA primers, finding that the phyllosphere community was more complex than previously thought based on culture-based methods. They reported an overall number of only 17 unique bacterial sequences and three unique fungal. With the emergence of high throughput sequencing, many more reports confirmed the existence of hundreds of unique 16S rDNA sequences within individual leaf samples (Redford et al. 2010; Lambais et al. 2006). The accumulated knowledge led to fundamental questions regarding the consistency of the plant microbiota, with regards to geography, host genotype *et cetera*. The first comprehensive research addressing this concept confirmed that a subset of bacteria is consistently colonizing the plant roots, despite different soil sources, with a weak but significant

effect of the host genotype on the bacterial profile (Lundberg et al. 2012; Bulgarelli et al. 2012). Subsequent studies confirmed these findings, presenting a stereotypical microbial profile of various plant species, while noting differences that are associated with the field, season and host genotype (Bulgarelli et al. 2015; Brown et al. 2020; Regalado et al. 2020; Thiergart et al. 2020) .

The existence of a core microbiome in plant species indicates that only compatible microbes can colonize the plant. The effect of the host genotype on the microbiome profile highlights that compatibility is a function of the host genetic arsenal. Nonetheless, the effect of the intraspecific host genotype on the microbiota was repeatedly found as weak, both in the phyllosphere (Wagner et al. 2016; Brown et al. 2020) and the rhizosphere (Lundberg et al. 2012; Durán et al. 2018; Brown et al. 2020; Thiergart et al. 2020). These findings supposedly are in conflict with the aforementioned implications of the zig-zag model (Jones and Dangl 2006), which translates into a high polymorphism of R genes within plant species (Van de Weyer et al. 2019). I.e, since different host genotypes hold a diverse set of R genes, they should present substantial differences in the abundance of various microbes. However, the interactions between effectors secreted by phytopathogens and their matching R gene compose only one variable out of many in predicting the overall microbiome. In fact, the microbial profile of *A. thaliana* is a polygenic trait, comprising various genes that exceed the scope of R genes (Horton et al. 2014).

Regardless of the exact genes in play, plant genetics seems to play only a modest role in determining the overall microbiome profile. An additional dimension holds promise in helping to explain the composition of microbial assemblage in the plant - the antagonism and mutualism occurring between co-residing microbes.

Mechanisms of microbial antagonism

From the perspective of evolution, organisms that reproduce in the highest numbers win. Reproduction depends on the available resources, and resources in most environments are

practically finite. Thus, competition is common between organisms that reside together, specifically within microbial communities (Hibbing et al. 2010; Foster and Bell 2012) .

Microbes have developed multiple mechanisms to proliferate in the presence of other competitors. One of these mechanisms comprises a set of tools to inhibit or kill other microbes. The type VI secretion system (T6SS) in bacteria is a contact-dependent transmembrane contractile machine, responsible for the injection of effectors into other organisms living in proximity (Hood et al. 2010). While different T6SSs have various functions such as biofilm formation (Gallique et al. 2017) and metal uptake (Wang et al. 2015; Chen et al. 2016), the primary function of T6SSs is to toxify nearby bacterial opponents (Cianfanelli, Monlezun, and Coulthurst 2016; Durand et al. 2014). The importance of this function is manifested by its main role in shaping the overall microbiota, as was demonstrated in the human gut (Verster et al. 2017; Sana et al. 2016; Chatzidaki-Livanis, Geva-Zatorsky, and Comstock 2016). Recently, the contact-dependent type IV secretion system (T4SS) - that has a main role in transferring DNA to other bacterial cells (Grohmann et al. 2018) - was also found to be involved in the injection of toxic effectors into other bacteria. This was demonstrated, for example, in the plant pathogen *Xanthomonas citri* which antagonized other bacteria via T4SS, leading to the hypothesis that T4SS role in microbial competition is underestimated (Souza et al. 2015). Other contact-dependent mechanisms include the type VII secretion systems (T7SSs), the contact-dependent growth inhibition (CDI) systems, nanotubes and the outer membrane exchange (OME), which are all found to be involved in toxification of other microbes (Granato, Meiller-Legrand, and Foster 2019) .

Another way to fight rival microbes is by contact-independent toxin secretion systems. These systems comprise diffusible antimicrobial molecules that can inhibit or kill microbes in proximity by interfering with their basic biological activities (Kohanski, Dwyer, and Collins 2010). In most cases, the secreting bacteria will also hold the according resistance, to avoid harming themselves (i.e. similar clones), though other clone types may also gain the appropriate

resistance to confer protection (Harms, Maisonneuve, and Gerdes 2016). The diversity of antibiotics production and resistance is a result of a probable coevolutionary arms race between competing microbes (Granato, Meiller-Legrand, and Foster 2019). Antibacterial molecules can also be produced by fungi, as was demonstrated by multiple fungi colonizing the leaf of *Indigofera suffruticosa* (Dos Santos et al. 2015), noting that the competition in the phyllosphere involves interkingdom interactions.

Conquering micro-niches that are rich in nutrients can greatly increase survivability and growth, thus facilitating competition. Using chemoreceptors and mobility factors, microbes can perform chemotaxis: sense the environment, and migrate towards a more beneficial or less harmful environment (Wadhams and Armitage 2004). Adhesion (Schluter et al. 2015), the secretion of extracellular slimy polymers that stop competitors and help the producers to spread (Kim et al. 2014), and cell morphology (Smith et al. 2017) are all mobility tools that determine the movement of microbes towards rich environments, on the account of their competitors. Another communal property that increases the survivability of bacteria is biofilm formation. Biofilm is the aggregation of bacterial communities, embedded in a matrix of extracellular polymeric substances (EPS). Bacteria in biofilms have different properties than free-living cells, such as enhanced resources uptake and increased resistance to antimicrobial substances (Flemming et al. 2016). The decision to migrate or to aggregate is based on cell-to-cell communication, and specifically on the cell density of cooperative bacteria - a function called quorum sensing (Miller and Bassler 2001). In order to outcompete others, microbes can also interfere with the quorum sensing of rivals (Xavier and Bassler 2005), preventing them from making the right communal decisions.

The efficient uptake of available nutrients is another factor that determines success in a competitive environment. Microbes differ in their ability to acquire and decompose different nutrients from the environment. The set of various nutrition sources microbes can exploit constrain their success rate in different environments (Wawrik et al. 2005; Seth and Taga 2014).

Intermediate resources that can only be decomposed by a few microbial types relax the competitive burden, while nutrients that are compatible with all microbes pose a higher competition, especially when these are scarce (Seth and Taga 2014). A hallmark for the effect of scarce nutrients is the battle over iron. Iron is an essential micronutrient, having a major role in vital basic cellular activities such as respiration, DNA synthesis and resistance to reactive oxygen species (Andrews, Robinson, and Rodríguez-Quiñones 2003). While it is generally abundant on Earth, iron bioavailability is limited (Emerson, Roden, and Twining 2012), thus bacteria must compete over iron in most environments. Bacterial iron uptake is done by a few mechanisms, the main one is the secretion of iron scavenging molecules called siderophores and their active take up by matching receptors (Guerinot 1994). There is a vast diversity of siderophores and, in accordance, various siderophores receptors. Microbes differ in their siderophore set and receptors, a polymorphism which creates a differential success rate of iron sequestration (Kramer, Özkaya, and Kümmerli 2020). Some microbes cheat by stealing siderophores produced by others, reflecting the competition over iron (Kramer, Özkaya, and Kümmerli 2020).

Overall, microbes have an arsenal of functions to maintain their survival in a competitive environment. A few of these functions were presented in this introduction, focusing on antagonistic interactions. It is important to note that although functions related to cooperative relationships were ignored, they play a role in the proliferation of microbes in competitive conditions.

Competitive mechanisms dictate the nature of individual microbe-microbe interactions, which in turn affect the microbial composition. Next, I will present how individual microbial interactions come into play in consortia, which is the realistic scenario. I will highlight current knowledge in the field, both in host-free environments and in the plant phyllosphere.

Microbe-microbe interactions and the plant microbiota

Linking microbial composition of consortia to individual microbial interactions is a highly challenging task. Generally, there are two approaches: the bottom-up, in which the individual components are being studied to predict the ensemble, and the top-down, in which the individual components are extrapolated from the ensemble. Each approach has its pros and cons, and both are complementary for the understanding of microbial interactions in consortia.

The very basis of the bottom-up approach is the assumption that species in consortia interact in a pairwise manner (Faust and Raes 2012). In its simplest form, it excludes the non-additive higher-order effects - the unpredicted alteration of a pairwise interaction by a third species - which can complicate the final outcome (Billick and Case 1994; Wootton 1994). Dromann and Roxburgh (Dormann and Roxburgh 2005) were among the first to experimentally validate how pairwise interactions fit with community assemblage, albeit not in microbes. They used Lotka–Volterra (LV) equations (modelling classic predator-prey dynamics) to predict biomass and coexistence of plants in three- and seven-species plant communities (appreciating plant ecology rather than microbial ecology in plants). Overall, pairwise interactions failed to predict the outcome of consortia in their study, highlighting that the interaction of multiple organisms is not a straightforward result of the pairwise interactions. In more recent work, an alternative qualitative model was employed to accurately predict 90% of three-species microbial communities from pairs (Friedman, Higgins, and Gore 2017). Nonetheless, the pairwise model failed to predict microbial communities of seven- and eight-species. When outcomes of the trio competitions were incorporated, seven- and eight-species communities could be accurately predicted, demonstrating how non-additive effects by a third species can affect complex microbial communities (Friedman, Higgins, and Gore 2017).

Important to note that current pairwise-based equations produce an abstractive model that ignores the mechanisms of interaction. Thus, on one hand such equations require minimum input to predict the consortia (i.e., abundance change in pairwise interactions), but on the other

hand they cannot predict complex interaction which results from the mechanistic context. Inclusion of mechanistic information can reduce non-additive effects, and improve predictions of microbial assembly based on pairwise interactions (Momeni, Xie, and Shou 2017). The explanation is that non-additive effects often result from the high abstraction yielded by minimal models (e.g., which account only for abundance change between two organisms). Such models cannot explain complicated relationships between multiple microbes, as found in the wild.

The top-down approach is the symmetric complement for the bottom-up, aiming to predict pairwise microbial interactions using communal data. In most cases, multiple samples across time or space are analyzed for their microbial profile, and changes in the profile are indicating microbial interactions. By concluding all pairwise co-occurrence associations, a microbial network can be constructed. The network topology can reveal the importance of each microbe in the communal context (Faust and Raes 2012). This is demonstrated by microbial keystone taxa, which are microbes with high connectedness in the network, that significantly affect the microbiome structure and function, irrespective of their abundance in different samples or sampling time (Banerjee, Schlaeppi, and van der Heijden 2018). Such hub microbes were found both in the rhizosphere (Jiang et al. 2017; Yan et al. 2017) and the phyllosphere (Agler et al. 2016) of plants. Two hub microbes that were found in the phyllosphere of *A. thaliana* - *Albugo* and *Dioszegia* - are eukaryotes that heavily affect bacteria, representative of interkingdom interactions (Agler et al. 2016). Interkingdom relationships within the *A. thaliana* microbiota were demonstrated also in the rhizosphere, noting how bacteria alter the composition of diverse filamentous eukaryotes (Durán et al. 2018).

While the top-down approach presents a feasible way to research microbial interactions in complex communities, it is primarily based on correlations. Spurious correlations pose a major risk to the reliability of such results. The risk is even greater when considering purley compositional data such as relative abundance, often used to profile microbes in plants and other hosts. Relative abundance is constrained to a constant amount of microbes in a given

sample, and this mathematical property alone will often lead to false correlations (Gloor et al. 2017). To exemplify the issue, consider a community of two microbes: A and B. The abundance of microbe A in both samples #1 and #2 is 100 cells. The abundance of microbe B in sample #1 is 100 cells, while in sample #2 is 300 cells. In this example, the abundance of microbe B does not affect the abundance of microbe A. Nonetheless, a conversion to relative abundance will lead to a false negative association: microbe A having an abundance of 50% in sample #1 and 25% in sample #2, while microbe B comprise 50% of sample #1 and 75% of sample #2. The increase of microbe B could be associated with the decrease of microbe A, only due to a mathematical constraint. There are mathematical tools to reduce the risk of such false correlations (Friedman and Alm 2012; Morton et al. 2019). Another simple way to circumvent the problem is by normalizing the amount of bacterial DNA to host DNA. Thus, the host is used as a reference point to estimate the load of microbes. Indeed, when normalizing microbial abundance to the plant, either by metagenomics (Regalado et al. 2020) or by PCR that combines microbe and host markers (Lundberg et al. 2020), microbial networks in the phyllosphere were substantially altered. Nonetheless, the paradigmatic risk of concluding causality from associations (Hill 1965) remains. Thus, experimental validation of microbial associations is still required. Among the common experimental approaches to validate the effect of microbes in consortia is their exclusion from synthetic communities (SynComs). The same approach can also be used naively, without prior information about microbial networks, to estimate the effect of individual microbes in consortia (Vorholt et al. 2017). A systemic drop out of individual bacteria from SynComs revealed keystone species in the phyllosphere of *A. thaliana* (Carlström et al. 2019), demonstrating an alternative way to construct microbial networks which are experimentally validated. Despite this advantage, such an approach is cumbersome and low in throughput. Hence evidence may be inconclusive for wild microbiota, which are much more complex (Banerjee, Schlaeppi, and van der Heijden 2018). A current compromise is

the combination of co-occurrence networks with experimental validation of a few focal microbes (Durán et al. 2018; Agler et al. 2016) .

In conclusion, microbe-microbe interactions have a major role in shaping microbial communities, and specifically the phyllosphere microbiota. Combining microbe-microbe interactions with the individual plant-microbe interactions will yield the final plant-microbe-microbe relationships. In these relationships, all components affect each other, creating a complex network of interactions which eventually determine the fitness of both the host and its co-residing microbes.

The reciprocal effects between the plant and the microbiota

The plant plays a role as an encapsulating environment, providing the potential resources for microbes that colonize this niche. Microbes can overpopulate the plant, deplete its resources, and therefore lead to negative health impacts, as is common after the blooming of phytopathogens. Nonetheless, such events are quite rare in wild ecosystems (Karasov, Shirsekar, et al. 2020). In fact, wild *Arabidopsis thaliana* capacitate pathogenic microbes without compromising their health, due to the activity of other protective microbes (Durán et al. 2018). In other words, non-pathogenic microbes can affect the relationships of plants and pathogens. The general concept is not new; Billick and Case (Billick and Case 1994) previously pointed out that various kinds of relationships between two organisms, ranging from cooperative to competitive, are constrained by the biotic environment.

In this section I will elaborate on the link between plant health and consortia of microbes, focusing specifically on protective microbes in the phyllosphere.

Protective microbial interactions in plants

There is increasing amount of evidence linking members from the plant microbiome to the mitigation of phytopathogens, both in the rhizosphere (Durán et al. 2018; Berendsen, Pieterse,

and Bakker 2012; Mendes et al. 2011) and the phyllosphere (Agler et al. 2016; Qin et al. 2019). The prevention of pathogen proliferation by other microbes is eventually resulting in plant protection. Such microbial protective interactions occur between species of the same taxonomic kingdom, for example among bacteria (Carlström et al. 2019), and between species of different taxonomic kingdoms (Durán et al. 2018; Mendes et al. 2011; Barda et al. 2015). Some protective microbes confer protection to a wide range of pathogens, while others are more pathogen-specific (Lugtenberg et al. 2013). This reflects the differential mechanisms and importance of microbes in determining plant health.

The majority of studies investigate the impact of individual protective microbes (Innerebner, Knief, and Vorholt 2011; Haas and Défago 2005; Barda et al. 2015; Gu et al. 2020). Nonetheless, there are a few examples demonstrating how a consortium of microbes acts as one protective unit in plants (Berendsen et al. 2018; Wei et al. 2015), exemplifying protective interactions that are consortium-dependent. Important to note is that as a general rule, the more complex microbial communities are, the more stable they are (A. Mougi and Kondoh 2012). Consistent with this ecological principle, diverse consortia better resist pathogen perturbations in plants (Ives, Klug, and Gross 2000; Morella et al. 2020). Hence, while individual protective microbes may reflect competitive mechanisms which are specific to mitigate certain pathogenic taxa, protection by consortia better reflect a general trend linking high diversity with community stability, and improved plant health (Berg et al. 2017).

Protective interactions in the plant can result from direct microbial warfare. For example, competition over iron drives the control of the pathogen *Ralstonia solanacearum* by microbiome members of tomato plants (Gu et al. 2020). However, protective interactions can also be host-mediated, as will be expanded on in the next chapter.

Protective microbe-host-microbe communication: interactions of microbes via the plant immune system

To understand how plant health is maintained by consortia of microbes, it is important to disentangle health effects derived from direct microbe-microbe interactions, which are local, and the indirect microbe-host-microbe interactions, which can be either local or systemic. Microbes interact in microenvironments, meaning that they can directly affect cells in their very proximity (Esser et al. 2015), be it the host or other microbes. In contrast, plants are multicellular organisms that generate both local and systemic signals.

Not only pathogens affect the plant immune system; non-pathogenic bacteria can also manipulate the host defensive response (Hacquard et al. 2017), either by activating or suppressing it (Vogel et al. 2016; Garrido-Oter et al. 2018; Lebeis et al. 2015). The interplay between commensals and the plant defense response can affect other microbes, including pathogens. For example, commensal microbes can prime the plant (a phenomenon also referred to as induced systemic resistance [ISR]), conferring resistance to multiple pathogens in distal tissues (Pieterse et al. 2014). The priming of plants by commensals seems to be a controlled process with an evolutionary history (Selosse, Bessis, and Pozo 2014). This principle is demonstrated by the recruitment of priming-microbes via root exudation, following encounters with the pathogens *Pseudomonas syringae* (J. Yuan et al. 2018) and *Hyaloperonospora arabidopsidis* (Berendsen et al. 2018).

These lines of evidence point to the importance of host-facilitated interactions, specifically highlighting how commensals mitigate pathogens via the host defense response, thus maintaining its health.

Aims and objectives

The general aim of this work is to link microbial interactions with plant health. I focus on both direct microbe-microbe and host-mediated interactions. I investigated these topics using the

model plant *Arabidopsis thaliana* and a collection of *Pseudomonas* sampled from south west Germany (Karasov et al. 2018).

In my first project, I employed synthetic communities of *Pseudomonas* to study how consortia of commensal and pathogenic bacteria affect each other - both within and among the groups - and how they affect the plant response and eventually its health. I genome-barcoded a selected set of *Pseudomonas* isolates, allowing the tracking of single isolates in a mixture. Thus, I characterize microbe-microbe interactions and appreciate their effect on the plant health at the intraspecific level. Another layer I addressed is the effect of the host genotype on bacterial interactions, and finally on plant health. I combine information about the plant weight, load of individual bacteria, host transcriptomic signature and pairwise bacterial suppression *in vitro* to portray a complete image of the plant-microbe-microbe interface. In particular, I focus on how the host facilitates protective microbial interactions and on how natural genetic variation affects plant-microbe-microbe interactions.

In my second project, I unravel the protective ability of the genus *Pseudomonas* by systematically testing pairwise interactions between one pathogenic lineage and a set of ~100 commensal *Pseudomonas*. This set spans the whole taxonomy of a three years collection in south-west Germany (Karasov et al. 2018). I leverage whole genome sequencing (WGS) data to perform a genome wide association study (GWAS), deciphering the protective bacterial gene orthologs that lead to protection in different *Pseudomonas* taxonomic subgroups.

One general topic to which I draw special attention in these projects relates to the synergistic effects of microbial consortia. Thus, I move back and forth between simple and complex bacterial communities, aiming to infer general conclusions about the ensemble effect.

Among the questions I address in this work: (i) How do multiple interacting commensal and pathogenic *Pseudomonas* affect plant health? (ii) To what extent does the plant affect bacterial interactions? (iii) How common is it for commensal *Pseudomonas* to protect *A.*

thaliana in the wild? And (iv) is there a global set of protective genes among commensal *Pseudomonas* or are these genes taxon-specific?

Studying these topics can improve our understanding about the role of the microbiota in maintaining plant health, as well as our conclusions about dynamics in the multifaceted host-microbe-microbe relationships.

Chapter one

Protective host-dependent antagonism among *Pseudomonas* in the *Arabidopsis* phyllosphere

Content of this chapter is published as:

Shalev, O., Karasov, T.L., Lundberg, D.S., Ashkenazy, H., Weigel, D. (2021). BioRxiv https://doi.org/10.1101/2021.04.08.438928*

*See *thesis appendix I* for an updated version, as submitted to *Nature Ecology & Evolution*.

Abstract

The plant microbiome is a rich biotic environment, comprising numerous taxa. The community structure of these colonizers is constrained by multiple factors, including host-microbe and microbe-microbe interactions, as well as the interplay between the two. While much can be learned from pairwise relationships between individual hosts and microbes, or individual microbes with themselves, the ensemble of interrelations between the host and microbial consortia may lead to different outcomes that are not easily predicted from the individual interactions. Their study can thus provide new insights into the complex relationship between plants and microbes. Of particular importance is how strain-specific such plant-microbe-microbe interactions are, and how they eventually affect plant health. Here, we test strain-level interactions in the phyllosphere between groups of co-existing commensal and pathogenic *Pseudomonas* among each other and with *A. thaliana*, by employing synthetic communities of genome-barcoded isolates. We found that commensal *Pseudomonas* prompted a host response leading to a selective inhibition of a specific pathogenic lineage, resulting in plant protection. The extent of plant protection, however, was dependent on plant genotype,

indicating that these effects are host-mediated. There were similar genotype-specific effects on the microbe side, as we could pinpoint an individual *Pseudomonas* isolate as the predominant cause for this differential interaction. Collectively, our work highlights how within-species genetic differences on both the host and microbe side can have profound effects on host-microbe-microbe dynamics. The paradigm that we have established provides a platform for the study of host-dependent microbe-microbe competition and cooperation in the *A. thaliana-Pseudomonas* system.

Contribution

O.S.S conceived and designed the research. D.S.L conceived bacterial barcoding and developed the method with O.S.S.. O.S.S performed the experiments and analyzed the results, O.S.S, T.L.K, D.S.L, H.A and D.W. discussed and interpreted the results. O.S.S wrote the first draft and the manuscript was written by O.S.S. and D.W, with input from all authors.

Author	**Author position**	**Scientific ideas**	**Data generation**	**Analysis & interpretation**	**Paper writing**
Or Shalev Skriptchak	1st	75%	98%	70%	70%
Talia L. Karasov	2nd	8%	1%	10%	4%
Derek S. Lundberg	3rd	7%	1%	5%	1%
Haim Ashkenazy	4th	3%	-	8%	5%
Detlef Weigel	5th	7%	-	7%	20%
Status in the publication process		In peer-to-peer review in *Nature Ecology & Evolution*			

Chapter two

Commensal Pseudomonas protect Arabidopsis thaliana from a coexisting pathogen via multiple taxonomy-dependent mechanisms

See *thesis appendix II*

Abstract

Plants are protected from pathogens by both their immune arsenal and colonizing microbes. In a recent *Pseudomonas* survey, a pathogenic lineage was reported to dominate wild *Arabidopsis thaliana* populations, while it was isolated from non-symptomatic plants. We previously reported that coexisting commensal *Pseudomonas* protect *A. thaliana* from this pathogenic lineage. However, it remained unclear (i) how common is protection among wild commensal *Pseudomonas* strains (ii) how taxon-specific is the protective ability within the *Pseudomonas* genus and (iii) what are the underlying bacterial genes and mechanisms leading to protection. Here, we address these questions by systemically co-infecting *A. thaliana* plants with an individual *Pseudomonas* pathogen and each of a hundred coexisting *Pseudomonas* commensals. We found that plant protection is a common function among non-pathogenic *Pseudomonas* taxa, and enriched in one lineage. Yet, we found a substantial variation in the protective ability among isolates of the same lineage, implying gene presence-absence polymorphism. We leveraged this variation to perform a genome-wide association study, discovering gene ortholog associated with plant protection. We found different sets of taxon-specific orthologs, highlighting various mechanisms that mitigate the pathogenic effect. By employing gene deletion, we validated the protective role of an ortholog subset, unique to one *Pseudomonas* lineage. We unraveled that iron acquisition and biofilm formation were the main mechanisms of plant protection in this lineage. Collectively, our work illustrates the importance of commensal bacteria in maintaining plant health, while presenting the diversity of underlying mechanisms.

Contribution

O.S. conceived the research. O.S. performed the experiments and analyzed the results. M.N. generated the knockout mutants. I.B. fine-tuned the image analysis algorithm for this study. O.S., H.A. and D.W. discussed and interpreted the results of this study. O.S. wrote the first draft and the manuscript was written by O.S. and D.W, with input from all authors.

Author	**Author position**	**Scientific ideas**	**Data generation**	**Analysis & interpretation**	**Paper writing**
Or Shalev Skriptchak	1st	85%	78%	80%	75%
Haim Ashkenazy	2nd	10%	10%	15%	10%
Manuela Neumann	3rd	-	10%	-	-
Ilja Bezrukov	4th	-	2%	-	-
Detlef Weigel	5th	5%	-	5%	15%
Status in the publication process		Unpublished, in preparation			

Discussion

Disentangling microbe-microbe and host-microbe from the plant-microbe-microbe relationships

The multiple pairwise interactions between microbes among each other and with the host create a highly connected network of relationships. It is crucial to differentiate direct and host-facilitated interactions to understand how the microbiota assembles in plants. While it is clear that microbial interactions occur *in planta*, the role of the host as a mediator is poorly understood, despite numerous lines of evidence of host-mediated interactions (Barda et al. 2015; De Vleesschauwer et al. 2008; Kamle et al. 2020). In other words, it is unclear how plant-specific the interactions *in planta* are.

Beneficial microbes prompt a protective response by the plant, leading to inhibition of pathogens (Barda et al. 2015; De Vleesschauwer et al. 2008). Furthermore, direct inhibitions between plant-associated microbes had been reported as well, inferred from *in vitro* experiments (Helfrich et al. 2018). Nonetheless, how these are orchestrated *in planta* has not been thoroughly studied. Rare exceptions are the studies of Duran et al. (Durán et al. 2018) and Gu et al. (Gu et al. 2020), which though not focusing directly on host-facilitated interactions demonstrates the relevance of *in vitro* direct microbial interactions *in planta*.

In this thesis, I present inconsistencies between microbial interactions *in vitro* and *in planta*, and specifically a differential ability of commensals to inhibit pathogens in these two environments (Chapter I; *Thesis appendix I*). Such differences manifest themselves in the importance of the plant as a mediator of microbial interactions, and in this case plant-protective interactions.

The host transcriptome signature provides strong evidence for how the plant may mediate these protective microbial interactions; only commensal *Pseudomonas* prompted a host immune response (Chapter I; *Thesis appendix I*). The lack of pathogen recognition denotes

a possible suppression of the plant defense system. Alternatively, it notes the inability of the tested plant genotype to recognize these pathogenic strains as different plant genotypes differentially recognize microbes (Haddadi, Larkan, and Borhan 2019). When pathogens were mixed with the commensal consortia, the plant immune response was activated (Chapter I; *Thesis appendix I*). This response can explain the suppression of some pathogenic *Pseudomonas* which could not be found *in vitro*, when tested against the commensal members. The set of induced defensive genes was exclusively commensal-dependent - i.e., insensitive to the presence of pathogens - and this may explain why the abundance of commensal isolates has not substantially changed when mixed with pathogens: commensals were exposed to a similar host defensive status, regardless of the presence of pathogens. Nonetheless, the mixture of commensals and pathogens led to the upregulation of many genes that were not expressed in each of the groups alone (Chapter I; *Thesis appendix I*). The synergistic effect derived from mixing pathogens with commensals demonstrates the importance of studying plant transcriptomics in the context of consortia, which is the realistic scenario. As far as I know, in addition to the work presented here, only a single study investigated the response of the plant to a consortia of microbes, using *Medicago truncatula* and two of its symbionts: rhizobia and mycorrhizal fungi (Afkhami and Stinchcombe 2016). In their study, infection with a consortia led to synergistic effects as well. The plant response to multiple microbes may be related to microbial assembly and plant health, and thus should be further investigated.

In the light of these findings, it seems important to complement studies focusing on microbial interactions *in planta* (Agler et al. 2016; Carlström et al. 2019) with host-free assays, and vice versa - to complement interactions found *in vitro* (Helfrich et al. 2018) by plant infections. Such a complementary approach will enable a full understanding of the plant-microbe and microbe-microbe vectors in the complex plant-microbe-microbe interactions, a need that was recently pointed out by others as well (Chaudhry et al. 2020).

Microbial dark matter: commensal microbes in plants

Many studies concentrated on the interactions between phytopathogens with plants, probably due to their obvious role in food production. Yet, most of the plant's microbiota seems to be non-pathogenic, commonly referred to as commensals (Vorholt 2012). The term 'commensal' is commonly interchangeable with 'non-pathogen' in host-microbe studies (Khan, Petersen, and Shekhar 2019; Durán et al. 2018; Maignien et al. 2014), and such usage deviates from the strict ecological definition of commensalism - a symbiosis where one organism gains something (i.e., the microbe) while the other gains nothing (i.e., the host) (Akihiko Mougi 2016). In fact, non-pathogens can also be mutualists, e.g., providing the plant benefits such as higher nutrient uptake or protection from pathogens. Therefore, the common usage of 'commensalism' in the discipline of plant microbial ecology combines 'commensalism' and 'mutualism' in a single phrase. The terminological difference is probably a result of our inability to characterize the functions of each of the numerous microbes in the plant, making it easier to just assume a microbe does nothing until it will be proven otherwise. This is especially true when considering that some functions may be only observed in specific biotic backgrounds, as in the case of protective microbes.

Here, I demonstrate that protection from a pathogen by non-pathogens is a common feature, at least in the genus *Pseudomonas* (Chapter II; *Thesis appendix II*). More often than not, commensal strains mitigated the effect of the tested pathogen. I demonstrated this phenomenon using a reductionistic system, testing pairwise interactions between a pathogenic *Pseudomonas* to putative commensal strains that span the whole taxonomy of a deep survey recently published (Karasov et al. 2018). In chapter I (*Thesis appendix I*) I found similar protective abilities of non-pathogenic *Pseudomonas* using consortia of commensal isolates, tested against consortia of pathogenic *Pseudomonas*, in soil-grown plants (unsterile conditions). The ability of commensal *Pseudomonas* to protect the plant in these different contexts and

experimental settings demonstrates a robust phenomenon. These findings highlight the unrealized potential of wild non-pathogens to maintain plant health.

Gu et al. (Gu et al. 2020) reported a strong ability of rhizosphere microbial members to protect the plant. They tested these members for inhibition of the pathogen *Ralstonia solanacearum* both *in vitro* and by infecting tomato plants. The importance of commensals in maintaining plant health was also demonstrated by inferring keystone species from microbial co-occurrence in wild *Arabidopsis thaliana* (Agler et al. 2016; Durán et al. 2018), implying that wild plants capacitate pathogens, and pathogen control is commensal-dependent. Not surprisingly, the importance of commensals in disease suppression is an emerging topic also in the study of humans (Khan, Petersen, and Shekhar 2019) .

The accumulating evidence notes that many bacteria referred to as commensals are actually mutualistic microbes. Can we predict such functions using sequencing data?

The plant microbiome and its function in the age of whole genome sequencing

As data about microbial genomes are exponentially growing, and annotations of genomes are improving, it is becoming clear that the rough classification by individual marker genes - as 16S and ITS - can predict microbial functions only to a very limited degree. This is especially true in bacteria, having a very flexible genome that is heavily shaped by horizontal gene transfer (HGT) (Garcia-Vallvé, Romeu, and Palau 2000). Such genomic variations often reflect functional variation, as even highly similar bacterial strains can differ in biochemical and other activities (Freschi et al. 2019; Heintz-Buschart and Wilmes 2018) .

Here, I document a substantial difference in the plant-protective ability among highly related commensal *Pseudomonas* (Chapter II; *Thesis appendix II*). The isolates used in this study were assigned to multiple taxonomic groups using a threshold of 99% similarity in their V3-V4 region of the 16S rDNA (Karasov et al. 2018), thus beyond the normal cutoff of 97% that defines species (Chan et al. 2012). Nonetheless, I found functional variation not just among the

taxonomic groups, but also within (Chapter II; *Thesis appendix II*). This means that isolates with almost identical 16S sequences have a differential plant-protective ability. Conclusively, even distinction of bacterial strains (thus, within species) is not enough to predict their effect on plants.

According to this, the mainstream description of bacteria by their 16S rDNA similarity is indeed too general to infer function. However, the microbiota of various plant species is conserved at the class level, with many consistent members at the genus level (Trivedi et al. 2020). Such consistency hints to qualities that are shared among isolates of the same taxonomic group, for example - the ability to colonize the plant. In agreement, my findings reveal how plant protection is a function of phylogeny, as one commensal lineage (99% similarity of V3-V4 16S rDNA) was enriched with protective isolates in comparison to other lineages (Chapter II; *Thesis appendix II*). It is thus puzzling to pinpoint what is shared and what differs between bacteria at a given taxonomic rank. This conflict challenges us in pinpointing the right taxonomic resolution to infer microbial functions of interest, e.g. effect on plant health or microbial assembly. To what taxonomic extent should we generalize microbial functions? Full genomes of bacteria provide the appropriate data to answer this question.

As I reported in the genome-wide association study I conducted, the exact bacterial gene set which leads to protection is taxon-specific (Chapter II; *Thesis appendix II*). This implies multiple protective mechanisms within the genus *Pseudomonas*, as each lineage protects the plant using a different set of genes. For example, iron competition was a prominent protective mechanism in one taxonomic group, and the causal set of iron-related genes was unique only to this lineage (Chapter II; *Thesis appendix II*). Therefore, microbial taxa do correlate with some specific abilities. However, there are differences also within lineages with similar 16S rDNA, hence between highly related bacteria. The iron-related gene set was present in only a subset of isolates within the relevant lineage. posing a challenge in generalizing isolates of this group as protective due to their high iron-competitiveness.

Similarities in genome content and overall sequence identity (considering both core and entire genome) are often conflicting (Bapteste et al. 2005; Gogarten, Doolittle, and Lawrence 2002; Puigbò, Wolf, and Koonin 2009), and combining the two is a matter of ongoing research (Avni and Snir 2020). Taken together with the findings presented here (Chapter II; *Thesis appendix II*), it seems risky to generalize functions by similarities in 16S or even by core genomes, as such generalizations may be too broad. The implications may be troubling, raising the need for sample-specific analysis, both in the case of individual bacteria and of consortia.

Although pricey, analysis of plant WGS-metagenomes provides a reliable base to predict functions of plant-associated microbes. Therefore, inferring functions from metagenomic data of plant microbiomes may build up our understanding about governing mechanisms in microbial assembly, as well as the eventual effect on plant health. A specific challenge in analysing plant phyllosphere samples is the inflation of host reads, which lead to negligible amounts of microbial reads (Regalado et al. 2020). If this challenge will be solved, the phyllosphere microbiome could be better assessed to understand its core-functions. This can be done by either processing metagenomic assembled genomes (MAGs), which are more suitable for long-reads sequencing data (Tsai et al. 2016), or by using unassembled data of short-reads which are enough for most annotations (Tamames, Cobo-Simón, and Puente-Sánchez 2019).

Linking annotated data from WGS with phylogeny can improve our understanding about the general role of dominant plant colonizers. A few pioneers have already taken such an approach (Ofek-Lalzar et al. 2014; Chapelle et al. 2016; Bulgarelli et al. 2015), providing insights about the genetic mechanisms which govern the stability of root-associated microbial consortia, and specifically of microbial members which antagonise pathogen invasion. However, to complement the link between phylogeny and functions, experimental validation should be coupled with such annotative data. Furthermore, the reciprocal host response to different microbial members should be investigated, as it affects the maintenance of plant health (Chapter I; *Thesis appendix I*).

The approach I took in chapter II (*Thesis appendix II*) can be expanded to tackle interactions between multiple commensals and multiple pathogens, and monitor their effect on plant health. Coupling data about bacterial genomes with the host transcriptome-response to these infections can enable the production of a plant-microbe network at the gene level. This network can be improved using multi-omics data (i.e. transcriptomics, proteomics and metabolomics), creating multiple networks that manifest the relationships between plant-microbe genes, proteins *et cetera*.

Eventually, such a high-depth overview can provide the appropriate level of information to reveal core microbial functions which maintain plant health in the wild, putting emphasis also on the interactive role of the host. As I repeatedly highlighted in this thesis, non-additive interactions among consortia of microbes provide an important and unpredictable dimension in the maintenance of plant health. The experimental approach I pursued can be expanded to investigate the associations of trios rather than pairs, advancing our understanding about how additive and non-additive interactions play a role in controlling pathogen invasion. This may provide an explanation for why individual protective strains often fail to present protective abilities in the field (Finkel et al. 2017).

Conclusions and outlook

Agriculture, the food source of our world, is the practice of creating anthropogenic ecosystems for the sake of humankind. The stability of agricultural ecosystems is considered to be low in comparison to wild ecosystems. Specifically, the ability of pathogens to dominate plant populations in agricultural settings - manifested by the relatively high disease rate in crops - is a signature of our inability to mimic the stability found in wild ecosystems (Karasov, Shirsekar, et al. 2020). Thus, in order to improve our agriculture, we should first learn how wild plants develop in balance with their microbes (Vannier, Agler, and Hacquard 2019).

My doctoral work revolved around different aspects of the plant-microbe-microbe relationships, using *A. thaliana* and *Pseudomonas* as a model, and concentrating on pathogen mitigation. One of the major topics I addressed is the facilitation of protective microbial interactions by the plant. By presenting differences in microbial interactions *in vitro* and *in planta*, I demonstrated that the plant has a substantial effect on microbial interactions which are related to plant protection (Chapter I; *Thesis appendix I*). Furthermore, dynamics in the plant transcriptome revealed association between exposure to commensal *Pseudomonas* and the activation of the host immune system. Coupling these results, it seems that the mediation of protective interactions by the plant is an active process, resulting from the expression of defensive genes. According to this interpretation, the plant has a direct role in pathogen inhibition - perhaps a systemic process - due to infection by commensal bacteria. This scenario is supported by other studies (De Vleesschauwer et al. 2008; Barda et al. 2015). On the other hand, I found in the same work that commensal members inhibit most pathogens *in vitro*, thus regardless of the host response (Chapter I; *Thesis appendix I*). It thus remained unclear what the exact mechanism of protection by commensal *Pseudomonas* was.

To address this question I conducted my second project, which takes an alternative route - aiming to reveal the protective bacterial elements in *Pseudomonas* (Chapter II; *Thesis appendix II*). I tested for multiple pairwise interactions between commensal *Pseudomonas* members and a representative pathogen. Following a genome wide association study (GWAS), sets of bacterial-related gene orthologs were associated with plant protection. The subset which I experimentally validated presented relatedness to iron-uptake and biofilm formation. These results indicate direct pathogen suppression *in planta*, a local event which requires proximity (Esser et al. 2015).

Taken together with the first project, this shows that both plant-mediated and direct interactions have a role in protective microbial interactions, prompted by commensal bacteria. Hence, protection against pathogens is an event which combines both local (direct interactions)

and systemic (plant response) tools, driven by the relationships of the plant with commensal bacteria. It is tempting to hypothesize that such relationships are a the outcome of long-term coevolutionary processes, leading to a win-win situation: the plant is protected and can capacitate pathogens (which are not taking over), while commensals are not being harshly restricted by the plant, as they are 'the least worst'. In other words, these so-called commensal microbes are actually beneficial symbionts, maintaining the plant health in the presence of microbial pathogens.

To advance agricultural applications of protective microbes, it is crucial to disentangle host-mediated protection from direct microbe-microbe interactions. Specifically, it is important to understand which mechanism dominates and when. When the plant is actively inhibiting pathogens after exposure to commensal strains, the main questions are (i) which microbes prompt such an immune response, (ii) what are the exact plant responses, and (iii) which pathogens are inhibited by it. I assume that different commensals can activate similar defensive mechanisms, considering the somewhat limited defensive arsenal of an individual plant. Thus, different microbial molecules will yield a few general downstream responses, and in principal these microbes could be replaced with the eliciting molecules for applicative purposes. On the other hand, if the plant is majorly a passive mediator of protective interactions, then the exact defense mechanism is microbial-related. Considering the broad range of microbes capacitating the plant - and the corresponding large pan genome - one can safely assume a tremendous amount of defense mechanisms that exist among microbes. Therefore, direct protective interactions will raise the need for tailor-made, rationally designed microbiomes for agricultural purposes, instead of focusing on elicitation of costley plant defenses.

In order to appreciate how active, host-mediated interactions affect direct microbe-microbe interactions, I suggest 'freezing' the responsive host status by lysing leaf tissues. As stated, a living leaf tissue offers a very dynamic and responsive environment to colonizing microbes. Leaf lysates offer a passive environment that is somewhat comparable to a

living tissue, reflecting the metabolic status of the tissue when it was last alive. Pathogenic microbes can be tested for their growth in lysates of leaves that were pre-challenged with microbes of interest. Therefore, it can be tested how the host response to a given microbe alters pathogen growth. Doing so systemically, using a range of exposure periods (e.g., immediately after infection, 24 hours after infection, *et cetera*), can quantitatively examine how the plant response to various microbes affects pathogen invasion. Moreover, competition between microbes can be conducted on such lysates, revealing microbe-microbe interactions in the static plant environment. Coupling metabolome analysis to such experiments will reveal if the plant indeed produces pathogen-inhibiting products due to commensal exposure, while pinpointing such products.

Extending this research direction, pairwise interactions can be decomposed into their plant-microbe and microbe-microbe components, so that microbial networks will be improved to mirror plant-microbe-microbe relationships. Previous microbial networks inferred from only from co-occurrences of microbes in the plant (Durán et al. 2018; Agler et al. 2016), meaning that it is not possible to approximate plant-mediation for a given pairwise interaction. Being able to do so will improve our understanding about how plant-specific protective interactions are.

Another aspect of microbe-host-microbe interactions is the underlying genetics. Similarly to the gene-for-gene model (Van der Biezen and Jones 1998), matching pairs of host and microbial elements can be associated to unravel mechanisms behind host-mediated interactions, and specifically of protective ones. Examples for questions that can be pursued in this context are: (i) Which microbial elements prompt a systemic resistance, and what are the responsive plant receptors? (ii) How taxon-specific are these elements and receptors? (iii) What is different in plant-microbe relationships between commensals and pathogens? Answering these paradigmatic questions, together with others, will enable us to have a deeper understanding of what makes protective microbes. Finally, genetic characterization of different microbes can be translated into function prediction of microbiomes, following recent trends in big data and

machine learning (Cammarota et al. 2020). Thus, protective candidates can potentially be identified *in silico* using metagenomic data.

Current practices in agriculture - our main food source for thousands of years - are now endangering our planet. In part, this is due to plant diseases which cause heavy annual yield losses (up to 30% in food-deficit areas; (Savary et al. 2019)), leading to extensive usage of pesticides (Aktar, Sengupta, and Chowdhury 2009) and the need of more land area to provide the global quota (Bommarco, Kleijn, and Potts 2013). Many of the challenges in crop health are related to the differences between natural and anthropogenic ecosystems (McCann 2020; Karasov, Shirsekar, et al. 2020). The robustness of wild plants is related to the interactions with the microbiota, and mimicking these dynamics can eventually improve global agriculture in a sustainable way. This approach will less likely lead to troubling ecological side effects, since it is relying on long-term evolutionary processes that shaped ecosystems to be as balanced as possible. In the age of climate change and other anthropogenic ecological hazards, human society has a crucial role in maintaining global health, reducing the ecological burden. Thus, we must follow sustainable agricultural developments.

References

Afkhami, Michelle E., and John R. Stinchcombe. 2016. "Multiple Mutualist Effects on Genomewide Expression in the Tripartite Association between Medicago Truncatula, Nitrogen-Fixing Bacteria and Mycorrhizal Fungi." *Molecular Ecology* 25 (19): 4946–62.

Agler, Matthew T., Jonas Ruhe, Samuel Kroll, Constanze Morhenn, Sang-Tae Kim, Detlef Weigel, and Eric M. Kemen. 2016. "Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation." *PLoS Biology* 14 (1): e1002352.

Aktar, Md Wasim, Dwaipayan Sengupta, and Ashim Chowdhury. 2009. "Impact of Pesticides Use in Agriculture: Their Benefits and Hazards." *Interdisciplinary Toxicology* 2 (1): 1–12.

Andrews, Simon C., Andrea K. Robinson, and Francisco Rodríguez-Quiñones. 2003. "Bacterial Iron Homeostasis." *FEMS Microbiology Reviews* 27 (2-3): 215–37.

Avni, Eliran, and Sagi Snir. 2020. "A New Phylogenomic Approach For Quantifying Horizontal Gene Transfer Trends in Prokaryotes." *Scientific Reports* 10 (1): 12425.

Bai, Yang, Daniel B. Müller, Girish Srinivas, Ruben Garrido-Oter, Eva Potthoff, Matthias Rott, Nina Dombrowski, et al. 2015. "Functional Overlap of the Arabidopsis Leaf and Root Microbiota." *Nature* 528 (7582): 364–69.

Balint-Kurti, Peter. 2019. "The Plant Hypersensitive Response: Concepts, Control and Consequences." *Molecular Plant Pathology* 20 (8): 1163–78.

Banerjee, Samiran, Klaus Schlaeppi, and Marcel G.A. van der Heijden. 2018. "Keystone Taxa as Drivers of Microbiome Structure and Functioning." *Nature Reviews. Microbiology* 16 (9): 567–76.

Bapteste, E., E. Susko, J. Leigh, D. MacLeod, R. L. Charlebois, and W. F. Doolittle. 2005. "Do Orthologous Gene Phylogenies Really Support Tree-Thinking?" *BMC Evolutionary Biology* 5 (May): 33.

Barda, Omer, Or Shalev, Shanee Alster, Kobi Buxdorf, Aviva Gafni, and Maggie Levy. 2015. "Pseudozyma Aphidis Induces Salicylic-Acid-Independent Resistance to Clavibacter Michiganensis in Tomato Plants." *Plant Disease* 99 (5): 621–26.

Barnes, Elle M., Erin L. Carter, and J. D. Lewis. 2020. "Predicting Microbiome Function Across Space Is Confounded by Strain-Level Differences and Functional Redundancy Across Taxa." *Frontiers in Microbiology* 11 (February): 101.

Berendsen, Roeland L., Corné M. J. Pieterse, and Peter A. H. M. Bakker. 2012. "The Rhizosphere Microbiome and Plant Health." *Trends in Plant Science* 17 (8): 478–86.

Berendsen, Roeland L., Gilles Vismans, Ke Yu, Yang Song, Ronnie de Jonge, Wilco P. Burgman, Mette Burmølle, Jakob Herschend, Peter A. H. M. Bakker, and Corné M. J. Pieterse. 2018. "Disease-Induced Assemblage of a Plant-Beneficial Bacterial Consortium." *The ISME Journal* 12 (6): 1496–1507.

Berg, Gabriele, Martina Köberl, Daria Rybakova, Henry Müller, Rita Grosch, and Kornelia Smalla. 2017. "Plant Microbial Diversity Is Suggested as the Key to Future Biocontrol and Health Trends." *FEMS Microbiology Ecology* 93 (5). https://doi.org/10.1093/femsec/fix050.

Billick, Ian, and Ted J. Case. 1994. "Higher Order Interactions in Ecological Communities: What Are They and How Can They Be Detected?" *Ecology* 75 (6): 1529–43.

Bommarco, Riccardo, David Kleijn, and Simon G. Potts. 2013. "Ecological Intensification: Harnessing Ecosystem Services for Food Security." *Trends in Ecology & Evolution* 28 (4): 230–38.

Brown, Shawn P., Michael A. Grillo, Justin C. Podowski, and Katy D. Heath. 2020. "Soil Origin and Plant Genotype Structure Distinct Microbiome Compartments in the Model Legume Medicago Truncatula." *Microbiome* 8 (1): 139.

Bulgarelli, Davide, Ruben Garrido-Oter, Philipp C. Münch, Aaron Weiman, Johannes Dröge, Yao Pan, Alice C. McHardy, and Paul Schulze-Lefert. 2015. "Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley." *Cell Host & Microbe* 17 (3): 392–403.

Bulgarelli, Davide, Matthias Rott, Klaus Schlaeppi, Emiel Ver Loren van Themaat, Nahal Ahmadinejad, Federica Assenza, Philipp Rauf, et al. 2012. "Revealing Structure and Assembly Cues for Arabidopsis Root-Inhabiting Bacterial Microbiota." *Nature* 488 (7409): 91–95.

Cammarota, Giovanni, Gianluca Ianiro, Anna Ahern, Carmine Carbone, Andriy Temko, Marcus J. Claesson, Antonio Gasbarrini, and Giampaolo Tortora. 2020. "Gut Microbiome, Big Data and Machine Learning to Promote Precision Medicine for Cancer." *Nature Reviews. Gastroenterology & Hepatology* 17 (10): 635–48.

Caporaso, J. Gregory, Christian L. Lauber, William A. Walters, Donna Berg-Lyons, Catherine A. Lozupone, Peter J. Turnbaugh, Noah Fierer, and Rob Knight. 2011. "Global Patterns of 16S rRNA Diversity at a Depth of Millions of Sequences per Sample." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 1 (March): 4516–22.

Carlström, Charlotte I., Christopher M. Field, Miriam Bortfeld-Miller, Barbara Müller, Shinichi Sunagawa, and Julia A. Vorholt. 2019. "Synthetic Microbiota Reveal Priority Effects and Keystone Strains in the Arabidopsis Phyllosphere." *Nature Ecology & Evolution* 3 (10): 1445–54.

Cerutti, Aude, Alain Jauneau, Marie-Christine Auriac, Emmanuelle Lauber, Yves Martinez, Serge Chiarenza, Nathalie Leonhardt, Richard Berthomé, and Laurent D. Noël. 2017. "Immunity at Cauliflower Hydathodes Controls Systemic Infection by Xanthomonas Campestris Pv Campestris." *Plant Physiology* 174 (2): 700–716.

Chan, Jacqueline Z-M, Mihail R. Halachev, Nicholas J. Loman, Chrystala Constantinidou, and Mark J. Pallen. 2012. "Defining Bacterial Species in the Genomic Era: Insights from the Genus Acinetobacter." *BMC Microbiology* 12 (December): 302.

Chapelle, Emilie, Rodrigo Mendes, Peter A. H. M. Bakker, and Jos M. Raaijmakers. 2016. "Fungal Invasion of the Rhizosphere Microbiome." *The ISME Journal* 10 (1): 265–68.

Chatzidaki-Livanis, Maria, Naama Geva-Zatorsky, and Laurie E. Comstock. 2016. "Bacteroides Fragilis Type VI Secretion Systems Use Novel Effector and Immunity Proteins to Antagonize Human Gut Bacteroidales Species." *Proceedings of the National Academy of Sciences of the United States of America* 113 (13): 3627–32.

Chaudhry, Vasvi, Paul Runge, Priyamedha Sengupta, Gunther Doehlemann, Jane E. Parker, and Eric Kemen. 2020. "Topic: Shaping the Leaf Microbiota: Plant-Microbe-Microbe Interactions." *Journal of Experimental Botany*, September. https://doi.org/ 10.1093/jxb/eraa417.

Chen, Wen-Jen, Tzu-Yen Kuo, Feng-Chia Hsieh, Pi-Yu Chen, Chang-Sheng Wang, Yu-Ling Shih, Ying-Mi Lai, Je-Ruei Liu, Yu-Liang Yang, and Ming-Che Shih. 2016. "Involvement of Type VI Secretion System in Secretion of Iron Chelator Pyoverdine in Pseudomonas Taiwanensis." *Scientific Reports* 6 (September): 32950.

Cianfanelli, Francesca R., Laura Monlezun, and Sarah J. Coulthurst. 2016. "Aim, Load, Fire: The Type VI Secretion System, a Bacterial Nanoweapon." *Trends in Microbiology* 24 (1): 51–62.

Coico, Richard. 2005. "Gram Staining." *Current Protocols in Microbiology* Appendix 3 (October): Appendix 3C.

Dean, Ralph, Jan A. L. Van Kan, Zacharias A. Pretorius, Kim E. Hammond-Kosack, Antonio Di Pietro, Pietro D. Spanu, Jason J. Rudd, et al. 2012. "The Top 10 Fungal Pathogens in Molecular Plant Pathology." *Molecular Plant Pathology* 13 (4): 414–30.

De Vleesschauwer, David, Mohammad Djavaheri, Peter A. H. M. Bakker, and Monica 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 (4): 1996–2012.

Dijk, K. van, D. E. Fouts, A. H. Rehm, A. R. Hill, A. Collmer, and J. R. Alfano. 1999. "The Avr (effector) Proteins HrmA (HopPsyA) and AvrPto Are Secreted in Culture from Pseudomonas Syringae Pathovars via the Hrp (type III) Protein Secretion System in a Temperature- and pH-Sensitive Manner." *Journal of Bacteriology* 181 (16): 4790–97.

Dillon, Marcus M., Renan N. D. Almeida, Bradley Laflamme, Alexandre Martel, Bevan S. Weir, Darrell Desveaux, and David S. Guttman. 2019. "Molecular Evolution of Pseudomonas Syringae Type III Secreted Effector Proteins." *Frontiers in Plant Science* 10 (April): 418.

Dodds, Peter N., and John P. Rathjen. 2010. "Plant Immunity: Towards an Integrated View of Plant-Pathogen Interactions." *Nature Reviews. Genetics* 11 (8): 539–48.

Dormann, Carsten F., and Stephen H. Roxburgh. 2005. "Experimental Evidence Rejects Pairwise Modelling Approach to Coexistence in Plant Communities." *Proceedings. Biological Sciences / The Royal Society* 272 (1569): 1279–85.

Dos Santos, Irailton Prazeres, Luís Cláudio Nascimento da Silva, Márcia Vanusa da Silva, Janete Magali de Araújo, Marilene da Silva Cavalcanti, and Vera Lucia de Menezes Lima. 2015. "Antibacterial Activity of Endophytic Fungi from Leaves of Indigofera Suffruticosa Miller (Fabaceae)." *Frontiers in Microbiology* 6 (May): 350.

Durand, Eric, Christian Cambillau, Eric Cascales, and Laure Journet. 2014. "VgrG, Tae, Tle, and beyond: The Versatile Arsenal of Type VI Secretion Effectors." *Trends in Microbiology* 22 (9): 498–507.

Durán, Paloma, Thorsten Thiergart, Ruben Garrido-Oter, Matthew Agler, Eric Kemen, Paul Schulze-Lefert, and Stéphane Hacquard. 2018. "Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival." *Cell* 175 (4): 973–83.e14.

Earl, Joshua P., Nithin D. Adappa, Jaroslaw Krol, Archana S. Bhat, Sergey Balashov, Rachel L. Ehrlich, James N. Palmer, et al. 2018. "Species-Level Bacterial Community Profiling of the Healthy Sinonasal Microbiome Using Pacific Biosciences Sequencing of Full-Length 16S rRNA Genes." *Microbiome* 6 (1): 190.

Emerson, David, Eric Roden, and Benjamin S. Twining. 2012. "The Microbial Ferrous Wheel: Iron Cycling in Terrestrial, Freshwater, and Marine Environments." *Frontiers in Microbiology* 3 (October): 383.

Esser, Daniel S., Johan H. J. Leveau, Katrin M. Meyer, and Kerstin Wiegand. 2015. "Spatial Scales of Interactions among Bacteria and between Bacteria and the Leaf Surface." *FEMS Microbiology Ecology* 91 (3). https://doi.org/10.1093/femsec/fiu034.

Fabro, Georgina, Jens Steinbrenner, Mary Coates, Naveed Ishaque, Laura Baxter, David J. Studholme,

Evelyn Körner, et al. 2011. "Multiple Candidate Effectors from the Oomycete Pathogen Hyaloperonospora Arabidopsidis Suppress Host Plant Immunity." *PLoS Pathogens* 7 (11): e1002348.

Faust, Karoline, and Jeroen Raes. 2012. "Microbial Interactions: From Networks to Models." *Nature Reviews. Microbiology* 10 (8): 538–50.

Finkel, Omri M., Gabriel Castrillo, Sur Herrera Paredes, Isai Salas González, and Jeffery L. Dangl. 2017. "Understanding and Exploiting Plant Beneficial Microbes." *Current Opinion in Plant Biology* 38 (August): 155–63.

Flemming, Hans-Curt, Jost Wingender, Ulrich Szewzyk, Peter Steinberg, Scott A. Rice, and Staffan Kjelleberg. 2016. "Biofilms: An Emergent Form of Bacterial Life." *Nature Reviews. Microbiology* 14 (9): 563–75.

Flor, H. H. 1956. "The Complementary Genic Systems in Flax and Flax Rust**Joint Contribution from the Field Crops Research Branch, Agricultural Research Service, United States Department of Agriculture and the North Dakota Agricultural Experiment Station." In *Advances in Genetics*, edited by M. Demerec, 8:29–54. Academic Press.

Foster, Kevin R., and Thomas Bell. 2012. "Competition, Not Cooperation, Dominates Interactions among Culturable Microbial Species." *Current Biology: CB* 22 (19): 1845–50.

Freschi, Luca, Antony T. Vincent, Julie Jeukens, Jean-Guillaume Emond-Rheault, Irena Kukavica-Ibrulj, Marie-Josée Dupont, Steve J. Charette, Brian Boyle, and Roger C. Levesque. 2019. "The Pseudomonas Aeruginosa Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity." *Genome Biology and Evolution* 11 (1): 109–20.

Friedman, Jonathan, and Eric J. Alm. 2012. "Inferring Correlation Networks from Genomic Survey Data." *PLoS Computational Biology* 8 (9): e1002687.

Friedman, Jonathan, Logan M. Higgins, and Jeff Gore. 2017. "Community Structure Follows Simple Assembly Rules in Microbial Microcosms." *Nature Ecology & Evolution* 1 (5): 109.

Gallique, Mathias, Victorien Decoin, Corinne Barbey, Thibaut Rosay, Marc G. J. Feuilloley, Nicole Orange, and Annabelle Merieau. 2017. "Contribution of the Pseudomonas Fluorescens MFE01 Type VI Secretion System to Biofilm Formation." *PloS One* 12 (1): e0170770.

Ganeshan, Girija, and A. Manoj Kumar. 2005. "Pseudomonas Fluorescens, a Potential Bacterial Antagonist to Control Plant Diseases." *Journal of Plant Interactions* 1 (3): 123–34.

Garcia-Vallvé, S., A. Romeu, and J. Palau. 2000. "Horizontal Gene Transfer in Bacterial and Archaeal Complete Genomes." *Genome Research* 10 (11): 1719–25.

Garrido-Oter, Ruben, Ryohei Thomas Nakano, Nina Dombrowski, Ka-Wai Ma, AgBiome Team, Alice C. McHardy, and Paul Schulze-Lefert. 2018. "Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia." *Cell Host & Microbe* 24 (1): 155–67.e5.

Garrido, Sharon Marie, Noriyuki Kitamoto, Akira Watanabe, Takahiro Shintani, and Katsuya Gomi. 2012. "Functional Analysis of FarA Transcription Factor in the Regulation of the Genes Encoding Lipolytic Enzymes and Hydrophobic Surface Binding Protein for the Degradation of Biodegradable Plastics in Aspergillus Oryzae." *Journal of Bioscience and Bioengineering* 113 (5): 549–55.

Gloor, Gregory B., Jean M. Macklaim, Vera Pawlowsky-Glahn, and Juan J. Egozcue. 2017. "Microbiome Datasets Are Compositional: And This Is Not Optional." *Frontiers in Microbiology* 8 (November): 2224.

Gogarten, J. Peter, W. Ford Doolittle, and Jeffrey G. Lawrence. 2002. "Prokaryotic Evolution in Light of

Gene Transfer." *Molecular Biology and Evolution* 19 (12): 2226–38.
Gram, Hans Christian Joachim, and Carl Friedlaender. 1884. *Ueber Die Isolirte Färbung Der Schizomyceten : In Schnitt-Und Trockenpräparaten* . Berlin: Theodor Fischer's medicinischer Buchhandlung.
Granato, Elisa T., Thomas A. Meiller-Legrand, and Kevin R. Foster. 2019. "The Evolution and Ecology of Bacterial Warfare." *Current Biology: CB* 29 (11): R521–37.
Grohmann, Elisabeth, Peter J. Christie, Gabriel Waksman, and Steffen Backert. 2018. "Type IV Secretion in Gram-Negative and Gram-Positive Bacteria." *Molecular Microbiology* 107 (4): 455–71.
Guerinot, M. L. 1994. "Microbial Iron Transport." *Annual Review of Microbiology* 48: 743–72.
Gu, Shaohua, Zhong Wei, Zhengying Shao, Ville-Petri Friman, Kehao Cao, Tianjie Yang, Jos Kramer, et al. 2020. "Competition for Iron Drives Phytopathogen Control by Natural Rhizosphere Microbiomes." *Nature Microbiology* 5 (8): 1002–10.
Haas, Dieter, and Geneviève Défago. 2005. "Biological Control of Soil-Borne Pathogens by Fluorescent Pseudomonads." *Nature Reviews. Microbiology* 3 (4): 307–19.
Hacquard, Stéphane, Stijn Spaepen, Ruben Garrido-Oter, and Paul Schulze-Lefert. 2017. "Interplay Between Innate Immunity and the Plant Microbiota." *Annual Review of Phytopathology* 55 (August): 565–89.
Haddadi, Parham, Nicholas J. Larkan, and M. Hossein Borhan. 2019. "Dissecting R Gene and Host Genetic Background Effect on the Brassica Napus Defense Response to Leptosphaeria Maculans." *Scientific Reports* 9 (1): 6947.
Harms, Alexander, Etienne Maisonneuve, and Kenn Gerdes. 2016. "Mechanisms of Bacterial Persistence during Stress and Antibiotic Exposure." *Science* 354 (6318). https://doi.org/10.1126/science.aaf4268.
Heintz-Buschart, Anna, and Paul Wilmes. 2018. "Human Gut Microbiome: Function Matters." *Trends in Microbiology* 26 (7): 563–74.
Helfrich, Eric J. N., Christine M. Vogel, Reiko Ueoka, Martin Schäfer, Florian Ryffel, Daniel B. Müller, Silke Probst, Markus Kreuzer, Jörn Piel, and Julia A. Vorholt. 2018. "Bipartite Interactions, Antibiotic Production and Biosynthetic Potential of the Arabidopsis Leaf Microbiome." *Nature Microbiology* 3 (8): 909–19.
Hesse, Wolfgang. 1992. "Walther and Angelina Hesse-Early Contributors to Bacteriology." *American Society for Microbiology*. https://jornades.uab.cat/workshopmrama/sites/jornades.uab.cat.workshopmrama/files/Hesse.pdf.
Hestrin, Rachel, Edith C. Hammer, Carsten W. Mueller, and Johannes Lehmann. 2019. "Synergies between Mycorrhizal Fungi and Soil Microbial Communities Increase Plant Nitrogen Acquisition." *Communications Biology* 2 (June): 233.
Hibbing, Michael E., Clay Fuqua, Matthew R. Parsek, and S. Brook Peterson. 2010. "Bacterial Competition: Surviving and Thriving in the Microbial Jungle." *Nature Reviews. Microbiology* 8 (1): 15–25.
Hill, A. B. 1965. "THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION?" *Proceedings of the Royal Society of Medicine* 58 (May): 295–300.
Hood, Rachel D., Pragya Singh, Fosheng Hsu, Tüzün Güvener, Mike A. Carl, Rex R. S. Trinidad, Julie M. Silverman, et al. 2010. "A Type VI Secretion System of Pseudomonas Aeruginosa Targets a Toxin to Bacteria." *Cell Host & Microbe* 7 (1): 25–37.

Horton, Matthew W., Natacha Bodenhausen, Kathleen Beilsmith, Dazhe Meng, Brian D. Muegge, Sathish Subramanian, M. Madlen Vetter, et al. 2014. "Genome-Wide Association Study of Arabidopsis Thaliana Leaf Microbial Community." *Nature Communications* 5 (November): 5320.

Innerebner, Gerd, Claudia Knief, and Julia A. Vorholt. 2011. "Protection of Arabidopsis Thaliana against Leaf-Pathogenic Pseudomonas Syringae by Sphingomonas Strains in a Controlled Model System." *Applied and Environmental Microbiology* 77 (10): 3202–10.

Ives, A. R., J. L. Klug, and K. Gross. 2000. "Stability and Species Richness in Complex Communities." *Ecology Letters* 3 (5): 399–411.

Jacoby, Richard, Manuela Peukert, Antonella Succurro, Anna Koprivova, and Stanislav Kopriva. 2017. "The Role of Soil Microorganisms in Plant Mineral Nutrition-Current Knowledge and Future Directions." *Frontiers in Plant Science* 8 (September): 1617.

Jiang, Yuji, Shuzhen Li, Rongpeng Li, Jia Zhang, Yunhao Liu, Lianfei Lv, Hong Zhu, Wenlong Wu, and Weilin Li. 2017. "Plant Cultivars Imprint the Rhizosphere Bacterial Community Composition and Association Networks." *Soil Biology & Biochemistry* 109 (June): 145–55.

Johnson, Jethro S., Daniel J. Spakowicz, Bo-Young Hong, Lauren M. Petersen, Patrick Demkowicz, Lei Chen, Shana R. Leopold, et al. 2019. "Evaluation of 16S rRNA Gene Sequencing for Species and Strain-Level Microbiome Analysis." *Nature Communications* 10 (1): 5029.

Jones, Jonathan D. G., and Jeffery L. Dangl. 2006. "The Plant Immune System." *Nature* 444 (7117): 323–29.

Kamle, Madhu, Rituraj Borah, Himashree Bora, Amit K. Jaiswal, Ravi Kant Singh, and Pradeep Kumar. 2020. "Systemic Acquired Resistance (SAR) and Induced Systemic Resistance (ISR): Role and Mechanism of Action Against Phytopathogens." In *Fungal Biotechnology and Bioengineering*, edited by Abd El-Latif Hesham, Ram Sanmukh Upadhyay, Gauri Dutt Sharma, Chakravarthula Manoharachary, and Vijai Kumar Gupta, 457–70. Cham: Springer International Publishing.

Karasov, Talia L., Juliana Almario, Claudia Friedemann, Wei Ding, Michael Giolai, Darren Heavens, Sonja Kersten, et al. 2018. "Arabidopsis Thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales." *Cell Host & Microbe* 24 (1): 168–79.e4.

Karasov, Talia L., Joel M. Kniskern, Liping Gao, Brody J. DeYoung, Jing Ding, Ullrich Dubiella, Ruben O. Lastra, et al. 2014. "The Long-Term Maintenance of a Resistance Polymorphism through Diffuse Interactions." *Nature* 512 (7515): 436–40.

Karasov, Talia L., Manuela Neumann, Alejandra Duque-Jaramillo, Sonja Kersten, Ilja Bezrukov, Birgit Schröppel, Efthymia Symeonidi, et al. 2020. "The Relationship between Microbial Population Size and Disease in the Arabidopsis Thaliana Phyllosphere." *bioRxiv*. https://doi.org/10.1101/828814.

Karasov, Talia L., Gautam Shirsekar, Rebecca Schwab, and Detlef Weigel. 2020. "What Natural Variation Can Teach Us about Resistance Durability." *Current Opinion in Plant Biology* 56 (August): 89–98.

Khan, Rabia, Fernanda Cristina Petersen, and Sudhanshu Shekhar. 2019. "Commensal Bacteria: An Emerging Player in Defense Against Respiratory Pathogens." *Frontiers in Immunology* 10 (May): 1203.

Kim, Wook, Fernando Racimo, Jonas Schluter, Stuart B. Levy, and Kevin R. Foster. 2014. "Importance of Positioning for Microbial Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 111 (16): E1639–47.

Kobayashi, D. Y., S. J. Tamaki, and N. T. Keen. 1989. "Cloned Avirulence Genes from the Tomato

Pathogen Pseudomonas Syringae Pv. Tomato Confer Cultivar Specificity on Soybean." *Proceedings of the National Academy of Sciences of the United States of America* 86 (1): 157–61.

Koenig, Daniel, Jörg Hagmann, Rachel Li, Felix Bemm, Tanja Slotte, Barbara Neuffer, Stephen I. Wright, and Detlef Weigel. 2019. "Long-Term Balancing Selection Drives Evolution of Immunity Genes in Capsella." *eLife* 8 (February). https://doi.org/ 10.7554/eLife.43606.

Kohanski, Michael A., Daniel J. Dwyer, and James J. Collins. 2010. "How Antibiotics Kill Bacteria: From Targets to Networks." *Nature Reviews. Microbiology* 8 (6): 423–35.

Kramer, Jos, Özhan Özkaya, and Rolf Kümmerli. 2020. "Bacterial Siderophores in Community and Host Interactions." *Nature Reviews. Microbiology* 18 (3): 152–63.

Lambais, M. R., D. E. Crowley, J. C. Cury, R. C. Büll, and R. R. Rodrigues. 2006. "Bacterial Diversity in Tree Canopies of the Atlantic Forest." *Science* 312 (5782): 1917.

Lane, Nick. 2015. "The Unseen World: Reflections on Leeuwenhoek (1677) 'Concerning Little Animals.'" *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370 (1666): 20140344.

Langley, J. Adam, and Bruce A. Hungate. 2014. "Plant Community Feedbacks and Long-Term Ecosystem Responses to Multi-Factored Global Change." *AoB Plants* 6 (July). https://doi.org/10.1093/aobpla/plu035.

Lebeis, Sarah L., Sur Herrera Paredes, Derek S. Lundberg, Natalie Breakfield, Jase Gehring, Meredith McDonald, Stephanie Malfatti, et al. 2015. "PLANT MICROBIOME. Salicylic Acid Modulates Colonization of the Root Microbiome by Specific Bacterial Taxa." *Science* 349 (6250): 860–64.

Leonelli, Lauriebeth, Jeffery Pelton, Allyn Schoeffler, Douglas Dahlbeck, James Berger, David E. Wemmer, and Brian Staskawicz. 2011. "Structural Elucidation and Functional Characterization of the Hyaloperonospora Arabidopsidis Effector Protein ATR13." *PLoS Pathogens* 7 (12): e1002428.

Levy, Asaf, Isai Salas Gonzalez, Maximilian Mittelviefhaus, Scott Clingenpeel, Sur Herrera Paredes, Jiamin Miao, Kunru Wang, et al. 2017. "Genomic Features of Bacterial Adaptation to Plants." *Nature Genetics* 50 (1): 138–50.

Lionetti, Vincenzo, and Jean-Pierre Métraux. 2014. "Plant Cell Wall in Pathogenesis, Parasitism and Symbiosis." *Frontiers in Plant Science* 5 (November): 612.

Liu, Yong-Xin, Yuan Qin, Tong Chen, Meiping Lu, Xubo Qian, Xiaoxuan Guo, and Yang Bai. 2020. "A Practical Guide to Amplicon and Metagenomic Analysis of Microbiome Data." *Protein & Cell*, May. https://doi.org/10.1007/s13238-020-00724-8.

Lugtenberg, Ben J. J., Natalia Malfanova, Faina Kamilova, and Gabriele Berg. 2013. "Microbial Control of Plant Root Diseases." In *Molecular Microbial Ecology of the Rhizosphere*, 575–86. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Lundberg, Derek S., Pratchaya Pramoj Na Ayutthaya, Annett Strauß, Gautam Shirsekar, Wen-Sui Lo, Thomas Lahaye, and Detlef Weigel. 2020. "Host-Associated Microbe PCR (hamPCR): Accessing New Biology through Convenient Measurement of Both Microbial Load and Community Composition." *Cold Spring Harbor Laboratory*. https://doi.org/ 10.1101/2020.05.19.103937.

Lundberg, Derek S., Sarah L. Lebeis, Sur Herrera Paredes, Scott Yourstone, Jase Gehring, Stephanie Malfatti, Julien Tremblay, et al. 2012. "Defining the Core Arabidopsis Thaliana Root Microbiome." *Nature* 488 (7409): 86–90.

Lu, You, and Kenichi Tsuda. 2021. "Intimate Association of PRR- and NLR-Mediated Signaling in Plant

Immunity." *Molecular Plant-Microbe Interactions: MPMI* 34 (1): 3–14.
Maignien, Loïs, Emelia A. DeForce, Meghan E. Chafee, A. Murat Eren, and Sheri L. Simmons. 2014. "Ecological Succession and Stochastic Variation in the Assembly of Arabidopsis Thaliana Phyllosphere Communities." *mBio* 5 (1): e00682–13.
Mansfield, John, Stephane Genin, Shimpei Magori, Vitaly Citovsky, Malinee Sriariyanum, Pamela Ronald, Max Dow, et al. 2012. "Top 10 Plant Pathogenic Bacteria in Molecular Plant Pathology." *Molecular Plant Pathology* 13 (6): 614–29.
McCann, Honour C. 2020. "Skirmish or War: The Emergence of Agricultural Plant Pathogens." *Current Opinion in Plant Biology* 56 (August): 147–52.
McHale, Leah, Xiaoping Tan, Patrice Koehl, and Richard W. Michelmore. 2006. "Plant NBS-LRR Proteins: Adaptable Guards." *Genome Biology* 7 (4): 212.
Melotto, Maeli, William Underwood, and Sheng Yang He. 2008. "Role of Stomata in Plant Innate Immunity and Foliar Bacterial Diseases." *Annual Review of Phytopathology* 46: 101–22.
Mendes, Rodrigo, Marco Kruijt, Irene de Bruijn, Ester Dekkers, Menno van der Voort, Johannes H. M. Schneider, Yvette M. Piceno, et al. 2011. "Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria." *Science* 332 (6033): 1097–1100.
Miedes, Eva, Ruben Vanholme, Wout Boerjan, and Antonio Molina. 2014. "The Role of the Secondary Cell Wall in Plant Resistance to Pathogens." *Frontiers in Plant Science* 5 (August): 358.
Miller, M. B., and B. L. Bassler. 2001. "Quorum Sensing in Bacteria." *Annual Review of Microbiology* 55: 165–99.
Mishra, Jitendra, Rachna Singh, and Naveen K. Arora. 2017. "Alleviation of Heavy Metal Stress in Plants and Remediation of Soil by Rhizosphere Microorganisms." *Frontiers in Microbiology* 8 (September): 1706.
Momeni, Babak, Li Xie, and Wenying Shou. 2017. "Lotka-Volterra Pairwise Modeling Fails to Capture Diverse Pairwise Microbial Interactions." *eLife* 6 (March). https://doi.org/10.7554/eLife.25051.
Morel, J. B., and J. L. Dangl. 1997. "The Hypersensitive Response and the Induction of Cell Death in Plants." *Cell Death and Differentiation* 4 (8): 671–83.
Morella, Norma M., Francis Cheng-Hsuan Weng, Pierre M. Joubert, C. Jessica E. Metcalf, Steven Lindow, and Britt 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 (2): 1148–59.
Morris, Jennifer L., Mark N. Puttick, James W. Clark, Dianne Edwards, Paul Kenrick, Silvia Pressel, Charles H. Wellman, Ziheng Yang, Harald Schneider, and Philip C. J. Donoghue. 2018. "The Timescale of Early Land Plant Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 115 (10): E2274–83.
Morton, James T., Clarisse Marotz, Alex Washburne, Justin Silverman, Livia S. Zaramela, Anna Edlund, Karsten Zengler, and Rob Knight. 2019. "Establishing Microbial Composition Measurement Standards with Reference Frames." *Nature Communications* 10 (1): 2719.
Mougi, Akihiko. 2016. "The Roles of Amensalistic and Commensalistic Interactions in Large Ecological Network Stability." *Scientific Reports* 6 (July): 29929.
Mougi, A., and M. Kondoh. 2012. "Diversity of Interaction Types and Ecological Community Stability." *Science* 337 (6092): 349–51.

Naveed, Zunaira Afzal, Xiangying Wei, Jianjun Chen, Hira Mubeen, and Gul Shad Ali. 2020. "The PTI to ETI Continuum in Phytophthora-Plant Interactions." *Frontiers in Plant Science* 11 (December): 593905.

Noble, Anya S., Stevie Noe, Michael J. Clearwater, and Charles K. Lee. 2020. "A Core Phyllosphere Microbiome Exists across Distant Populations of a Tree Species Indigenous to New Zealand." *PloS One* 15 (8): e0237079.

Nürnberger, T., and D. Scheel. 2001. "Signal Transmission in the Plant Immune Response." *Trends in Plant Science* 6 (8): 372–79.

Ofek-Lalzar, Maya, Noa Sela, Milana Goldman-Voronov, Stefan J. Green, Yitzhak Hadar, and Dror Minz. 2014. "Niche and Host-Associated Functional Signatures of the Root Surface Microbiome." *Nature Communications* 5 (September): 4950.

Pearce, Ben K. D., Andrew S. Tupper, Ralph E. Pudritz, and Paul G. Higgs. 2018. "Constraining the Time Interval for the Origin of Life on Earth." *Astrobiology* 18 (3): 343–64.

Petit-Houdenot, Yohann, and Isabelle Fudal. 2017. "Complex Interactions between Fungal Avirulence Genes and Their Corresponding Plant Resistance Genes and Consequences for Disease Resistance Management." *Frontiers in Plant Science* 8 (June): 1072.

Pieterse, Corné M. J., Christos Zamioudis, Roeland L. Berendsen, David M. Weller, Saskia C. M. Van Wees, and Peter A. H. M. Bakker. 2014. "Induced Systemic Resistance by Beneficial Microbes." *Annual Review of Phytopathology* 52 (June): 347–75.

Pritchard, Leighton, and Paul R. J. Birch. 2014. "The Zigzag Model of Plant-Microbe Interactions: Is It Time to Move On?" *Molecular Plant Pathology* 15 (9): 865–70.

Puigbò, Pere, Yuri I. Wolf, and Eugene V. Koonin. 2009. "Search for a 'Tree of Life' in the Thicket of the Phylogenetic Forest." *Journal of Biology* 8 (6): 59.

Qin, Chong, Jiemeng Tao, Tianbo Liu, Yongjun Liu, Nengwen Xiao, Tianming Li, Yabing Gu, Huaqun Yin, and Delong Meng. 2019. "Responses of Phyllosphere Microbiota and Plant Health to Application of Two Different Biocontrol Agents." *AMB Express* 9 (1): 42.

Ranjan, Ravi, Asha Rani, Ahmed Metwally, Halvor S. McGee, and David L. Perkins. 2016. "Analysis of the Microbiome: Advantages of Whole Genome Shotgun versus 16S Amplicon Sequencing." *Biochemical and Biophysical Research Communications* 469 (4): 967–77.

Redford, Amanda J., Robert M. Bowers, Rob Knight, Yan Linhart, and Noah Fierer. 2010. "The Ecology of the Phyllosphere: Geographic and Phylogenetic Variability in the Distribution of Bacteria on Tree Leaves." *Environmental Microbiology* 12 (11): 2885–93.

Regalado, Julian, Derek S. Lundberg, Oliver Deusch, Sonja Kersten, Talia Karasov, Karin Poersch, Gautam Shirsekar, and Detlef 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 (8): 2116–30.

Reina-Pinto, José J., and Alexander Yephremov. 2009. "Surface Lipids and Plant Defenses." *Plant Physiology and Biochemistry: PPB / Societe Francaise de Physiologie Vegetale* 47 (6): 540–49.

Sana, Thibault G., Nicolas Flaugnatti, Kyler A. Lugo, Lilian H. Lam, Amanda Jacobson, Virginie Baylot, Eric Durand, Laure Journet, Eric Cascales, and Denise M. Monack. 2016. "Salmonella Typhimurium Utilizes a T6SS-Mediated Antibacterial Weapon to Establish in the Host Gut." *Proceedings of the National Academy of Sciences of the United States of America* 113 (34): E5044–51.

Savary, Serge, Laetitia Willocquet, Sarah Jane Pethybridge, Paul Esker, Neil McRoberts, and Andy Nelson. 2019. "The Global Burden of Pathogens and Pests on Major Food Crops." *Nature Ecology & Evolution* 3 (3): 430–39.

Schluter, Jonas, Carey D. Nadell, Bonnie L. Bassler, and Kevin R. Foster. 2015. "Adhesion as a Weapon in Microbial Competition." *The ISME Journal* 9 (1): 139–49.

Schopf, J. William. 2006. "Fossil Evidence of Archaean Life." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 361 (1470): 869–85.

Schulze-Lefert, Paul. 2004. "Knocking on the Heaven's Wall: Pathogenesis of and Resistance to Biotrophic Fungi at the Cell Wall." *Current Opinion in Plant Biology* 7 (4): 377–83.

Selosse, Marc-André, Alain Bessis, and María J. Pozo. 2014. "Microbial Priming of Plant and Animal Immunity: Symbionts as Developmental Signals." *Trends in Microbiology* 22 (11): 607–13.

Serrano, Mario, Fania Coluccia, Martha Torres, Floriane L'Haridon, and Jean-Pierre Métraux. 2014. "The Cuticle and Plant Defense to Pathogens." *Frontiers in Plant Science* 5 (June): 274.

Seth, Erica C., and Michiko E. Taga. 2014. "Nutrient Cross-Feeding in the Microbial World." *Frontiers in Microbiology* 5 (July): 350.

Singh, Prashant, Sylvain Santoni, Audrey Weber, Patrice This, and Jean-Pierre Péros. 2019. "Understanding the Phyllosphere Microbiome Assemblage in Grape Species (Vitaceae) with Amplicon Sequence Data Structures." *Scientific Reports* 9 (1): 14294.

Smith, William P. J., Yohan Davit, James M. Osborne, Wook Kim, Kevin R. Foster, and Joe M. Pitt-Francis. 2017. "Cell Morphology Drives Spatial Patterning in Microbial Communities." *Proceedings of the National Academy of Sciences of the United States of America* 114 (3): E280–86.

Souza, Diorge P., Gabriel U. Oka, Cristina E. Alvarez-Martinez, Alexandre W. Bisson-Filho, German Dunger, Lise Hobeika, Nayara S. Cavalcante, et al. 2015. "Bacterial Killing via a Type IV Secretion System." *Nature Communications* 6 (March): 6453.

Strange, Richard N., and Peter R. Scott. 2005. "Plant Disease: A Threat to Global Food Security." *Annual Review of Phytopathology* 43: 83–116.

Stukenbrock, Eva H., and Bruce A. McDonald. 2008. "The Origins of Plant Pathogens in Agro-Ecosystems." *Annual Review of Phytopathology* 46: 75–100.

Tamames, Javier, Marta Cobo-Simón, and Fernando Puente-Sánchez. 2019. "Assessing the Performance of Different Approaches for Functional and Taxonomic Annotation of Metagenomes." *BMC Genomics* 20 (1): 960.

Tett, Adrian, Edoardo Pasolli, Stefania Farina, Duy Tin Truong, Francesco Asnicar, Moreno Zolfo, Francesco Beghini, et al. 2017. "Unexplored Diversity and Strain-Level Structure of the Skin Microbiome Associated with Psoriasis." *NPJ Biofilms and Microbiomes* 3 (June): 14.

Thiergart, Thorsten, Paloma Durán, Thomas Ellis, Nathan Vannier, Ruben Garrido-Oter, Eric Kemen, Fabrice Roux, et al. 2020. "Root Microbiota Assembly and Adaptive Differentiation among European Arabidopsis Populations." *Nature Ecology & Evolution* 4 (1): 122–31.

Thomma, Bart P. H. J., Thorsten Nürnberger, and Matthieu H. A. J. Joosten. 2011. "Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy." *The Plant Cell* 23 (1): 4–15.

Toruño, Tania Y., Ioannis Stergiopoulos, and Gitta Coaker. 2016. "Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners." *Annual Review of Phytopathology* 54 (August): 419–41.

Trivedi, Pankaj, Jan E. Leach, Susannah G. Tringe, Tongmin Sa, and Brajesh K. Singh. 2020. "Plant-Microbiome Interactions: From Community Assembly to Plant Health." *Nature Reviews. Microbiology* 18 (11): 607–21.

Tsai, Yu-Chih, Sean Conlan, Clayton Deming, NISC Comparative Sequencing Program, Julia A. Segre, Heidi H. Kong, Jonas Korlach, and Julia Oh. 2016. "Resolving the Complexity of Human Skin Metagenomes Using Single-Molecule Sequencing." *mBio* 7 (1): e01948–15.

Tully, Benjamin J., Elaina D. Graham, and John F. Heidelberg. 2018. "The Reconstruction of 2,631 Draft Metagenome-Assembled Genomes from the Global Oceans." *Scientific Data* 5 (January): 170203.

Underwood, William. 2012. "The Plant Cell Wall: A Dynamic Barrier against Pathogen Invasion." *Frontiers in Plant Science* 3 (May): 85.

Van der Biezen, E. A., and J. D. Jones. 1998. "Plant Disease-Resistance Proteins and the Gene-for-Gene Concept." *Trends in Biochemical Sciences* 23 (12): 454–56.

Van de Weyer, Anna-Lena, Freddy Monteiro, Oliver J. Furzer, Marc T. Nishimura, Volkan Cevik, Kamil Witek, Jonathan D. G. Jones, Jeffery L. Dangl, Detlef Weigel, and Felix Bemm. 2019. "A Species-Wide Inventory of NLR Genes and Alleles in Arabidopsis Thaliana." *Cell* 178 (5): 1260–72.e14.

Vannier, Nathan, Matthew Agler, and Stéphane Hacquard. 2019. "Microbiota-Mediated Disease Resistance in Plants." *PLoS Pathogens* 15 (6): e1007740.

Velásquez, André C., Matthew Oney, Bethany Huot, Shu Xu, and Sheng Yang He. 2017. "Diverse Mechanisms of Resistance to Pseudomonas Syringae in a Thousand Natural Accessions of Arabidopsis Thaliana." *The New Phytologist* 214 (4): 1673–87.

Verster, Adrian J., Benjamin D. Ross, Matthew C. Radey, Yiqiao Bao, Andrew L. Goodman, Joseph D. Mougous, and Elhanan Borenstein. 2017. "The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition." *Cell Host & Microbe* 22 (3): 411–19.e4.

Vogel, Christine, Natacha Bodenhausen, Wilhelm Gruissem, and Julia A. Vorholt. 2016. "The Arabidopsis Leaf Transcriptome Reveals Distinct but Also Overlapping Responses to Colonization by Phyllosphere Commensals and Pathogen Infection with Impact on Plant Health." *The New Phytologist* 212 (1): 192–207.

Vorholt, Julia A. 2012. "Microbial Life in the Phyllosphere." *Nature Reviews. Microbiology* 10 (12): 828–40.

Vorholt, Julia A., Christine Vogel, Charlotte I. Carlström, and Daniel B. Müller. 2017. "Establishing Causality: Opportunities of Synthetic Communities for Plant Microbiome Research." *Cell Host & Microbe* 22 (2): 142–55.

Wadhams, George H., and Judith P. Armitage. 2004. "Making Sense of It All: Bacterial Chemotaxis." *Nature Reviews. Molecular Cell Biology* 5 (12): 1024–37.

Wagner, Maggie R., Derek S. Lundberg, Tijana G. Del Rio, Susannah G. Tringe, Jeffery L. Dangl, and Thomas Mitchell-Olds. 2016. "Host Genotype and Age Shape the Leaf and Root Microbiomes of a Wild Perennial Plant." *Nature Communications* 7 (July): 12151.

Wang, Tietao, Meiru Si, Yunhong Song, Wenhan Zhu, Fen Gao, Yao Wang, Lei Zhang, et al. 2015. "Type VI Secretion System Transports Zn2+ to Combat Multiple Stresses and Host Immunity." *PLoS Pathogens* 11 (7): e1005020.

Wawrik, Boris, Lee Kerkhof, Jerome Kukor, and Gerben Zylstra. 2005. "Effect of Different Carbon

Sources on Community Composition of Bacterial Enrichments from Soil." *Applied and Environmental Microbiology* 71 (11): 6776–83.

Wei, Zhong, Tianjie Yang, Ville-Petri Friman, Yangchun Xu, Qirong Shen, and Alexandre Jousset. 2015. "Trophic Network Architecture of Root-Associated Bacterial Communities Determines Pathogen Invasion and Plant Health." *Nature Communications* 6 (September): 8413.

Wootton, J. Timothy. 1994. "THE NATURE AND CONSEQUENCES OF INDIRECT EFFECTS IN ECOLOGICAL COMMUNITIES." *Annual Review of Ecology and Systematics* 25 (1): 443–66.

Xavier, Karina B., and Bonnie L. Bassler. 2005. "Interference with AI-2-Mediated Bacterial Cell-Cell Communication." *Nature* 437 (7059): 750–53.

Yang, C. H., D. E. Crowley, J. Borneman, and N. T. Keen. 2001. "Microbial Phyllosphere Populations Are More Complex than Previously Realized." *Proceedings of the National Academy of Sciences of the United States of America* 98 (7): 3889–94.

Yan, Yan, Eiko E. Kuramae, Mattias de Hollander, Peter G. L. Klinkhamer, and Johannes A. van Veen. 2017. "Functional Traits Dominate the Diversity-Related Selection of Bacterial Communities in the Rhizosphere." *The ISME Journal* 11 (1): 56–66.

Yan, Yan, Long H. Nguyen, Eric A. Franzosa, and Curtis Huttenhower. 2020. "Strain-Level Epidemiology of Microbial Communities and the Human Microbiome." *Genome Medicine* 12 (1): 71.

Yuan, Jun, Jun Zhao, Tao Wen, Mengli Zhao, Rong Li, Pim Goossens, Qiwei Huang, et al. 2018. "Root Exudates Drive the Soil-Borne Legacy of Aboveground Pathogen Infection." *Microbiome* 6 (1): 156.

Yuan, Minhang, Bruno Pok Man Ngou, Pingtao Ding, and Xiu-Fang Xin. 2021. "PTI-ETI Crosstalk: An Integrative View of Plant Immunity." *Current Opinion in Plant Biology* 62 (March): 102030.

Protective host-dependent antagonism among *Pseudomonas* in the *Arabidopsis* phyllosphere

Or Shalev[1], Talia L. Karasov[1,2], Derek S. Lundberg[1], Haim Ashkenazy[1], Detlef Weigel[1,*]

[1]Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

[2]Present address: School of Biological Sciences, University of Utah, 84112 Salt Lake City, USA

[*]Corresponding author weigel@tue.mpg.de (D.W.)

Keywords: *Arabidopsis thaliana*, *Pseudomonas*, wild populations, microbial communities, plant health

Abstract

The plant microbiome is a rich biotic environment, comprising numerous taxa. The community structure of these colonizers is constrained by multiple factors, including host-microbe and microbe-microbe interactions, as well as the interplay between the two. While much can be learned from pairwise relationships between individual hosts and microbes, or individual microbes with themselves, the ensemble of interrelations between the host and microbial consortia may lead to different outcomes that are not easily predicted from the individual interactions. Their study can thus provide new insights into the complex relationship between plants and microbes. Of particular importance is how strain-specific such plant-microbe-microbe interactions are, and how they eventually affect plant health. Here, we test strain-level interactions in the phyllosphere between groups of co-existing commensal and pathogenic *Pseudomonas* among each other and with *A. thaliana*, by employing synthetic communities of genome-barcoded isolates. We found that commensal *Pseudomonas* prompted a host response leading to a selective inhibition of a specific pathogenic lineage, resulting in plant protection. The extent of plant protection, however, was dependent on plant genotype, indicating that these effects are host-mediated. There were similar genotype-specific effects on the microbe side, as we could pinpoint an individual *Pseudomonas* isolate as the predominant cause for this differential interaction. Collectively, our work highlights how within-species genetic differences on both the host and microbe side can have profound effects on host-microbe-microbe dynamics. The paradigm that we have established provides a platform for the study of host-

dependent microbe-microbe competition and cooperation in the *A. thaliana-Pseudomonas* system.

Introduction

Plants, like other complex organisms, host a diverse set of microbes. The assembly of these microbial communities is shaped both by host-microbe as well as microbe-microbe interactions. These interactions may be of any symbiotic type: mutualistic, commensalistic or parasitic, and are dictated by the balance of inhibition and facilitation of growth. As has been exemplified in many studies, interactions between organisms are not static, but rather a dynamic process that depends on the environment - both biotic [1,2] and abiotic [3,4] - as well as on evolutionary history [5,6].

Many aspects of the dynamic interactions between plants and microbes have been studied in considerable detail, not least because of their implications for agriculture and ecology. Colonization of the plant depends on the ability of microbes to grow on and in the host, but also on the antagonistic ability of the host to promote or restrict such microbial growth. In the case of pathogens, there is often a co-evolutionary arms race, in which plants evolve recognition and immune tools to restrict the growth of microbes, while microbes evolve evasion and an offensive arsenal to further populate the plant [7,8]. These co-evolutionary dynamics typically fuel the generation of genetic diversity within both host and microbe, and the dependence of microbial colonization and host health on intraspecific variation has been documented in numerous studies [5,9–11]. Nonetheless, the extent to which intraspecific host variation shapes overall microbial composition is minimal [3,4], with the most dramatic effects seen for specific taxa that are recognized by the immune system [12,13]. Instead, other environmental factors have a much larger influence on the overall composition [3,4], including other resident microbes [1,14,15]. Taken together, this suggests that successful colonizers reflect compatibility to grow in the presence of both the host and other microbes, and that this compatibility depends on their genetic makeup.

The colonizing microbes exert differential effects on host health - from harmful [16] to beneficial [17]. These effects are mainly related to microbial load, since overpopulation of the plant by microbes can negatively impact its health [9,18]. Nonetheless, the host has the genetic

arsenal to control the growth of some microbes, thus avoiding negative outcomes [7]. This raises questions about the ability of the host plant to differentially recognize and respond to a consortium of microbes with a range of functions, i.e. differentiating friend from foe in a complex assembly of microbial taxa. While there is a growing body of literature about host response to individual pathogens [19] and individual commensals [20], a more realistic scenario is the integrated host response to communities that include both diverse pathogens and diverse commensals.

In the same way, the numerous constraints resulting from multiple host-microbe and microbe-microbe interrelations create a complex system of relationships, making extrapolation of rules from simplistic systems likely difficult. For example, overpopulation of the plant by one microbe can result in negative health impacts, but these might be mitigated in the presence of other microbes [17,21]. While studies of microbe-microbe interactions *in planta* have paved the way for important findings about their impact on the overall community [14,15], the effect of the host on such microbial interactions has often not been considered, despite the host being able to affect these via direct host-microbe interactions [22]. Hence, the high degree of interconnectedness at the host-microbe-microbe interface calls for holistic research of this system, rather than tackling individual components, to unravel dynamics that result from the multiple constraints. Such an approach can be conducted using synthetic communities, which establish causality and not only associations between microbe-microbe and plant-microbe interactions [23].

In a previous study, Karasov and colleagues [11] surveyed *Pseudomonas* populations from leaves of wild *Arabidopsis thaliana* plants in south-west Germany. Among these, one lineage, which was highly pathogenic in axenic infections, often dominated endophytic microbial communities of *A. thaliana* leaves. Nonetheless, this lineage was isolated from plants without any visible disease symptoms, suggesting that other factors, including co-colonizing microbes, were mitigating the pathogenic phenotype. This includes other *Pseudomonas* lineages, which did not appear to have any significant impacts on host health when tested individually [11].

Here, we took advantage of our collection of wild *Pseudomonas* isolates to investigate intraspecific host-microbe-microbe dynamics by infecting *A. thaliana* plants with synthetic *Pseudomonas* communities. Specifically, we examined interactions between pathogenic and commensal *Pseudomonas* with the host leaves and with themselves, and the linkage of these to the host health. We found that the host facilitated protective commensal-pathogen interactions,

and revealed further complex interactions that could not be realized by studying host-microbe or microbe-microbe relationships individually.

Results

Genome barcoding of *Pseudomonas* isolates and experimental design

To test possible host-commensal-pathogen dynamics in local populations, we colonized six *A. thaliana* genotypes with synthetic bacterial communities composed of pathogenic and commensal *Pseudomonas* candidates. Pathogenicity classification was based on demonstrated pathogenic potential effects of different *Pseudomonas* lineages in the Karasov collection [11]. Only one lineage - which dominated local plant population - was associated with pathogenicity, both according to its negative impact on rosette weight and to visible disease symptoms [11]. This lineage was previously named "OTU5" (Operational Taxonomic Unit number 5) [11]. We henceforth call "ATUE5" (isolates sampled from Around TUEbingen, group 5) to all isolates that share a common 16S rDNA sequence in the V3-V4 region, previously associated with pathogenicity, and "non-ATUE5" to all other *Pseudomonas* from the Karasov collection [11]. We interchangeably use the terms pathogens and ATUE5, as well as commensals and non-ATUE5. We used host genotypes that originated from the same host populations from which the Pseudomonads were isolated - neary Tübingen, Germany (**Figure 1A**), aiming to reflect interactions between coexisting hosts and microbes.

Overall, seven pathogenic *Pseudomonas* and seven commensal isolates were chosen, prioritizing those with the highest estimated abundance in the field (**Figure 1B**). The abundance was estimated by the number of similar isolates (defined as nucleotide sequence divergence less than 0.0001 in their core genome) sampled in the original survey. Thus, the chosen isolates act as representatives for other similar isolates. In total, all 14 *Pseudomonas* isolates were classified as belonging to four OTUs, following 16S rDNA clustering at 99% sequence identity. Because of the high relatedness of several of the isolates, we could not rely upon a single endogenous genetic marker to distinguish between isolates. Instead, we genome-barcoded each of the isolates. We employed the mini-Tn7 system [24] to insert a single-copy of a 22 bp long unique sequence, flanked by universal priming sites, into the chromosome of each isolate (Illustration in **Figure S1A**). We validated the sequence of all barcodes in the corresponding isolates using Sanger sequencing (**Table S1**), and confirmed barcode integration by barcode-specific PCR

(**Figure S2A**). Furthermore, we confirmed that barcode-amplification yielded the expected products when PCR-amplified from DNA extracted from infected *A. thaliana* individuals (**Figure S2B**). While barcoding slightly impaired the growth rates of the isolates P3 and P4, the majority of barcoded bacteria exhibited similar growth dynamics as the non-barcoded parental strains when tested in Lysogeny Broth (LB) medium (**Figure S3**).

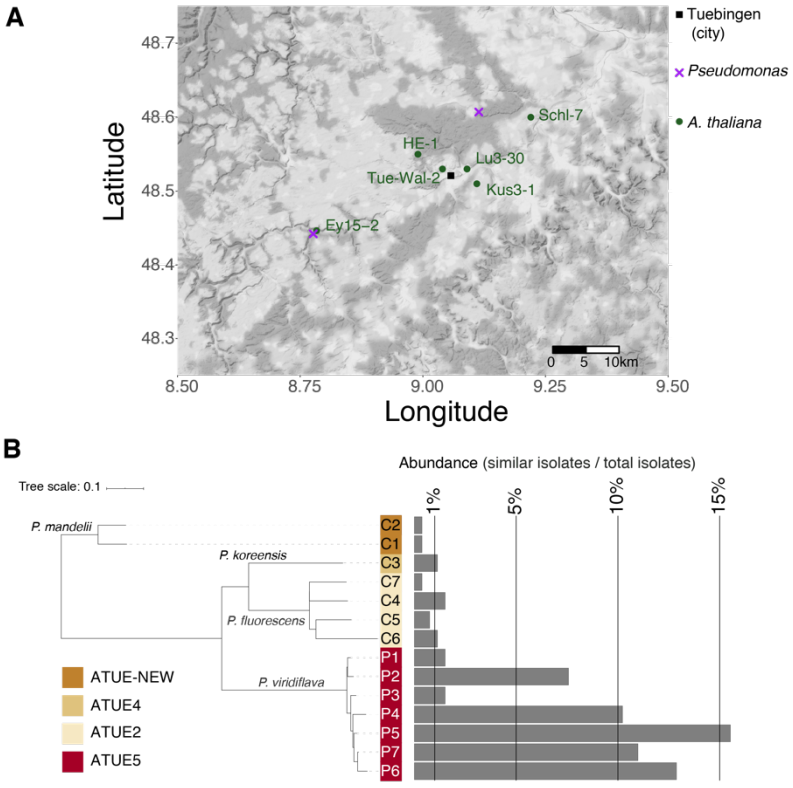


Figure 1. Study system. A. Location of original *A. thaliana* and *Pseudomonas* sampling sites around Tübingen (Germany). **B.** Taxonomic representation of the 14 *Pseudomonas* isolates used, and their respective abundance in the 1,524 strains of the Karasov collection [11]. Isolates were binned according to similarity (divergence < 0.0001 in core genome). Taxonomic assignment is indicated for each ATUE group (corresponding to a specific OTU in [11]). 'P' - Pathogen candidate. 'C' - Commensal candidate.

Next, we constructed three synthetic communities using the barcoded isolates: An exclusively pathogenic synthetic community, comprising the seven ATUE5 isolates (hereafter 'PathoCom'), an exclusively commensal synthetic community, comprising the seven non-ATUE5 isolates (hereafter 'CommenCom'), and a joint synthetic community comprising all 14 isolates - both pathogens and commensals (hereafter 'MixedCom'). Isolates were mixed in an equimolar fashion, and their absolute starting concentration was identical in each synthetic community.

Thus, the inoculum of the MixedCom with 14 isolates had twice the total number of bacterial cells as either the PathoCom or CommenCom inoculum.

The community experiments were conducted in plants grown on soil in the presence of other microbes. Our decision to perform experiments on non-sterile soil stemmed from initial observations that the infection outcomes of plants grown on soil were more consistent with the outcomes observed in the field than infections of axenically grown plants. Specifically, our initial isolation of the focal bacterial strains was done from plants in the field that were alive and not obviously diseased [11]. In the lab, axenic infections with these strains showed rapid and dramatic phenotypic effects on the plants, often killing the plants as early as three days-post-infection (**Figure S4**). In contrast, soil-grown plants displayed only mild disease symptoms and decreased size 12 dpi (**Figure S4**), phenotypes more consistent with those observed in the field.

To more closely mimic natural infections, which likely occur through the air, we chose to infect plants by spraying with bacterial suspension, rather than direct leaf infiltration, as is common for testing of leaf pathogenic bacteria in *A. thaliana*. Twenty one days after sowing, we spray-infected the leaves of soil-grown *A. thaliana* plants raised in growth chambers with the three synthetic communities, and with bacteria-free buffer (hereafter 'Control'). Twelve days after infection (dpi), we sampled the fresh rosettes, weighed them and extracted DNA (see Methods). Subsequently, we coupled barcode-specific PCR and qPCR. We included an amplicon from an *A. thaliana*-specific genomic sequence in the qPCR assay, which allowed us to approximate the absolute abundance per isolate, i.e., the ratio of isolate genome copies to plant genome copies (**Figure S1B**).

Host-genotype effects on composition of synthetic communities

The six *A. thaliana* genotypes used in this study were originally sampled from the same geographic region (**Figure 1A**) - a maximum of 40 km apart. They were all from the area from which the *Pseudomonas* strains used were isolated [11], and they were also all from the same host genetic group [25]. In accordance, we expected that host genotype would have little, if any effect on the composition of our synthetic communities of local *Pseudomonas* isolates. However, while not large, there was a significant effect of host genotype, explaining 5 to 12% of compositional variation in the different communities, as determined by permutational multivariate

analysis of variance (PERMANOVA) with Bray–Curtis distances (**Table 1**). For comparison, the batch effect (between the different experiments) explained up to 26% of compositional variation.

Analysis of similarities (ANOSIM) within each experiment indicated similar trends as PERMANOVA - with the genotype having a significant effect on isolate composition in each synthetic community (**Table S2A**).

Table 1. Permutational multivariate analysis of variance (PERMANOVA) based on Bray-Curtis distances, for compositions of the 14 barcoded bacteria in treated hosts. The analysis was constrained by the host genotype and the experiment batch ('exp') to estimate their effect on the explained variance. n=170 for PathoCom, n=151 for CommenCom, and n=182 for MixedCom.

Treatment	Df	Sum Sq	Pseudo-F	R2	Pr(>f)	Variation source
PathoCom	5	4.83	4.67	0.1199	**0.0005**	Genotype
	1	1.68	8.11	0.0417	**0.0005**	Exp
	5	1.08	1.04	0.0268	0.3973	Genotype:Exp
	158	32.68	NA	0.8116	NA	Residuals
	169	40.27	NA	1.0000	NA	Total
CommenCom	5	2.36	3.88	0.0839	**0.0005**	Genotype
	1	7.32	60.19	0.2604	**0.0005**	Exp
	5	1.53	2.52	0.0545	**0.0030**	Genotype:Exp
	139	16.89	NA	0.6012	NA	Residuals
	150	28.1	NA	1.0000	NA	Total
MixedCom	5	2.17	1.99	0.0456	**0.0020**	Genotype
	1	6.29	28.89	0.1324	**0.0005**	Exp
	5	2.03	1.86	0.0427	**0.0020**	Genotype:Exp
	170	36.99	NA	0.7793	NA	Residuals
	181	47.47	NA	1.0000	NA	Total

In bold, statistically significant relationships ($P \leq 0.05$).

We then examined bacterial composition clustering according to host genotype, by applying multilevel pairwise comparison using adonis (pairwise adonis, based on Bray-Curtis distances). Some pairs of genotypes differed in their effects on all three communities (**Table S2B**), an observation that was supported by nonmetric multidimensional scaling (NMDS) ordination of bacterial composition in each treatment (**Figure S5A**). The cumulative load of all isolates was associated with the loading on the NMDS1 axis (Pearson's $r > 0.99$ and p-value $< 2.2e^{-16}$, for all three communities), suggesting that a part of the compositional differences between host genotypes was due to absolute rather than relative abundance. In agreement, we observed differences in total bacterial load between the host genotypes, and the nature of the differences was treatment-dependent (**Figure S5B**).

How do these different community compositions affect plant growth?

Host-genotype dependent pathogenicity, growth promotion or protection

PathoCom infection caused plants to grow less than control plants, during the 12 days of the experiment (**Figure 2**; **Figure S6**). In two out of the six host genotypes - Lu3-30 and TueWal-2 - weight decrease was milder, indicating a certain level of resistance to the PathoCom members (mean difference to control: Lu3-30 -29.1 mg [-59.3, -1.4], TueWal-2 -30.0 mg [-46.4, -13.4], Kus3-1 -77.2 mg [96.4, 54.2], Schl-7 -93.1 mg [123.5, 67.7], Ey15-2 -92.5 mg [116.4, 66.0] and HE-1 -53.9 mg [82.6, 27.0], with 95% confidence intervals in brackets). To validate that the effect of the PathoCom on plant weight was due to bacterial activity, and not merely a host response to the inoculum (e.g., PAMP-triggered immunity), we infected plants with heat-killed PathoCom. We found a minor weight decrease in three out of the six genotypes, but the overall contribution to weight reduction was small (**Figure S7**; heat-killed PathoCom accounts for 14% of the variation explained by the living PathoCom in the model shown).

In contrast to PathoCom, infections with CommenCom led to a slight increase in fresh weight, suggesting plant growth promotion activity or alternatively protection from resident environmental pathogens (**Figure S6A**). This effect was independent on the host genotype (**Figure S6B**).

Importantly, the negative growth effects of the PathoCom were greatly reduced in the MixedCom experiment. Plants infected with MixedCom grew to a similar extent as the control, with the exception of the genotype Ey15-2, which continued to suffer a substantial weight reduction when infected by the mixed community (**Figure 2**; mean difference to Control = -48.5 mg, [-74.8, -22.6] at 95% confidence interval). Nonetheless, this reduction was less than that caused on Ey15-2 by PathoCom. Hence, co-colonization of pathogenic *Pseudomonas* with commensals led to enhanced growth, while the magnitude was host-genotype dependent.

These results support the role of ATUE5 strains as pathogenic, and provide additional evidence for protection against ATUE5 by commensal *Pseudomonas* strains that coexist with ATUE5 in nature. Next, we wanted to learn whether and how changes in bacterial abundance or shifts in *Pseudomonas* community composition led to differential impacts on growth of the infected plants.

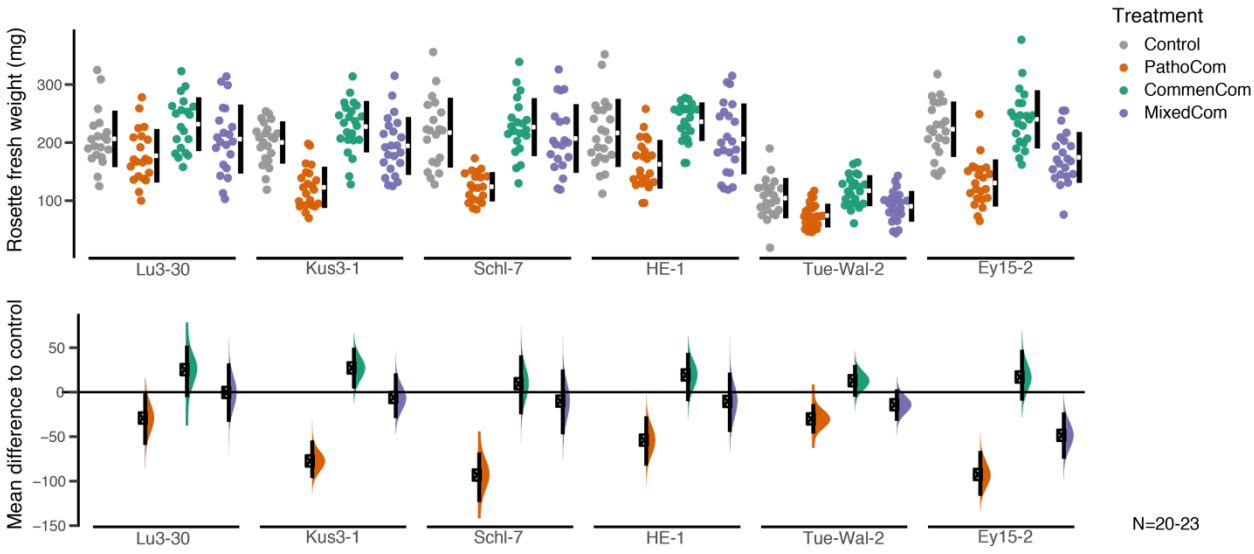


Figure 2. Commensal *Pseudomonas* protect the plant in a host-dependent manner. Each of the six *A. thaliana* genotypes used in this study was treated with Control, PathoCom, CommenCom and MixedCom. Fresh rosette weight was measured 12 dpi. The top panel presents the raw data, the breaks in the black vertical lines denote the mean value of each group, and the vertical lines themselves indicate standard deviation. The lower panel presents the mean difference to control, inferred from bootstrap sampling [26][27], indicating the distribution of effect sizes that are compatible with the data. 95% confidence intervals are indicated by the black vertical bars. Shown here are the results of one experiment. A second experiment gave similar results. n=20-23.

Differences in bacterial load and impact per a given load of pathogenic and commensal *Pseudomonas*

We hypothesized that the total cumulative load of all barcoded isolates (i.e., regardless of the identity of the colonizing isolates) should be a significant explanatory variable for weight differences among treatments. We based this expectation on the association previously found between abundance in the field and pathogenicity for similar *Pseudomonas* isolates [11].

Contrary to our hypothesis, we found that while the differences in plant weight between treatments were considerable, the bacterial loads of MixedCom and PathoCom were not significantly different from one another (**Figure 3A**). This result implies that plant weight is also a function of bacterial composition, and not load *per se*. In agreement with this inference, the load-weight relationships were found to be treatment-dependent, indicating that weight can be better predicted by load within a treatment than among treatments (difference in expected log-scaled predictive density = -52.9 and in standard error = 9.4 when comparing the model weight ~

treatment * $\log_{10}$(isolates load) + genotype + experiment + error to the same model without the treatment factor, using leave-one-out cross-validation; see Methods).

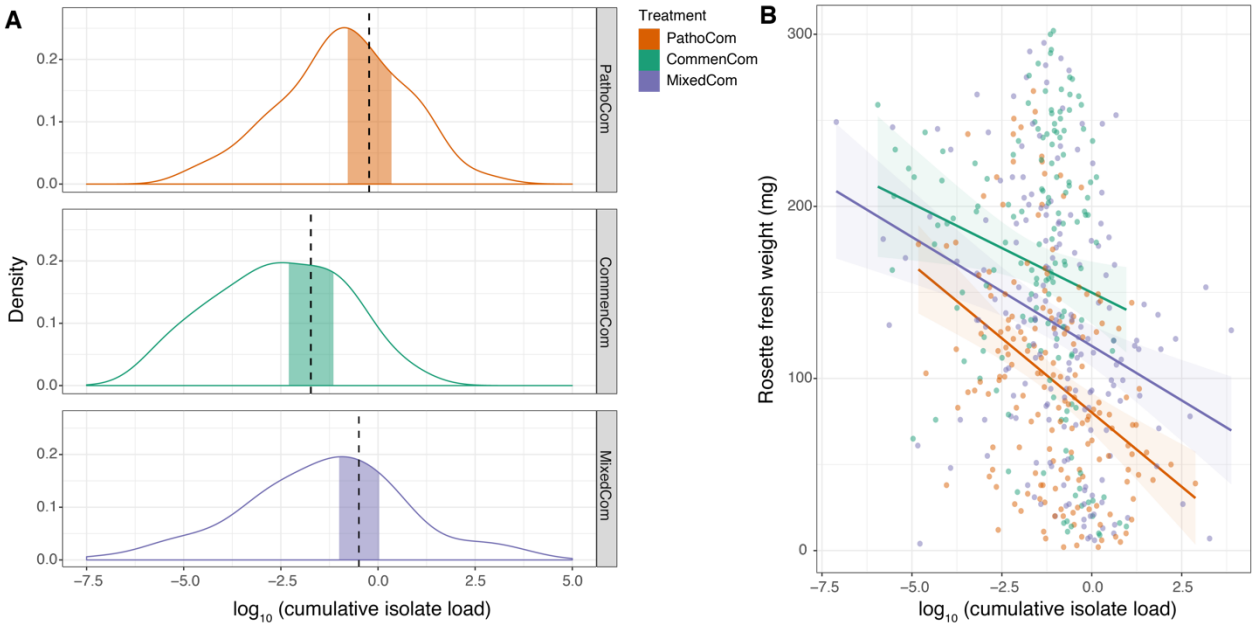


Figure 3. Plants are more tolerant to commensals than pathogens. A. Density plot of $\log_{10}$(bacterial load) for the three synthetic communities. Vertical dashed lines indicate means, and the shaded areas 95% credible intervals of the fitted parameter, following the model $\log_{10}$(bacterial load) ~ treatment + genotype + experiment + error. **B.** Correlation of $\log_{10}$(bacterial load) with rosette fresh weight. Shaded areas indicate 95% confidence intervals of the correlation curve; Bacterial load was defined as the cumulative abundance of all barcoded isolates that constituted a synthetic community. n=170 for PathoCom, n=151 for CommenCom, and n=182 for MixedCom.

Notably, we noticed that the regression slope of PathoCom was more negative than the regression slope of CommenCom, suggesting that ATUE5 isolates had a stronger negative impact on weight per bacterial cell than non-ATUE5 isolates (**Figure 3B**; **Figure S8A;** CommenCom mean effect difference to PathoCom: 12.0 mg [4.4,19.5], at 95% credible interval of the parameter $\log_{10}$(isolates load) * treatment). From the reciprocal angle, that of the host, it can be seen that plants were less tolerant to ATUE5 isolates than non-ATUE5 isolates. MixedCom presented a regression slope between the two exclusive synthetic communities, implying that the impact on plant growth resulted from both groups - ATUE5 and non-ATUE5 (MixedCom mean effect difference to PathoCom: 4.8 mg [-1.6,11.8], at 95% credible interval of the parameter $\log_{10}$(isolates load) * treatment). Lastly, we observed differential regression slopes

between the host genotypes, particularly among Pathocom- and CommenCom-infected hosts, revealing differential tolerance levels to the same *Pseudomonas* isolates (**Figure S8B-C**).

We have described two general differences between pathogenic and commensal *Pseudomonas*: (i) on average pathogens have a greater impact per a given load on plant growth, and (ii) they can reach higher titers in *A. thaliana* leaves. Together, this points to dual effects of pathogens on plant health. In order to explain how commensal non-ATUE5 isolates were able to mitigate the harmful impact of pathogenic ATUE5 in MixedCom, we next addressed the bacterial compositionality in MixedCom-infected hosts.

Protection by commensal members and host-mediated pathogen suppression

Given that (i) MixedCom-infected plants grew better than PathoCom-infected plants (**Figure 1A**; **Figure S6A**), (ii) there was no considerable difference in total load between PathoCom- and MixedCom-infected plants (**Figure 3A**), and (iii) pathogens were found to cause more damage per cell (**Figure 3B**; **Figure S8A**), we expected commensal members to dominate MixedCom.

Consistent with our expectations, the composition of MixedCom was more similar to CommenCom than PathoCom (**Figure 4A**). We then analyzed the change in bacterial abundance due to the mixture of pathogens and commensals at the isolate level. We compared the absolute abundance of each isolate among the treatments: Pathogenic isolates were compared between PathoCom and MixedCom, and commensals between CommenCom and MixedCom. In general, the abundance of pathogens was significantly lower in MixedCom, while the abundance of commensals was either similar or slightly higher in MixedCom (**Figure 4B**). Thus, the mixture of pathogens and commensals led to pathogen suppression, while commensal load was largely unchanged in MixedCom compared to CommenCom. Thus, non-ATUE5 isolates appear to be more competitive in the MixedCom context than ATUE5 isolates. The abundance change of each isolate in the presence of additional community members was similar among the host genotypes, implying that commensal-pathogen interactions were majorly a general trait, possibly independent of the host (**Figure S9, Table S3**).

We therefore tested for direct, host-independent interactions among isolates with an *in vitro* growth inhibition assay (Methods). Each of the 14 isolates was examined for growth inhibition against all other isolates, covering all possible combinations of binary interactions. In total, three strains out of the 14 had inhibitory activity; all were non-ATUE5 (**Figure 4C**). Specifically, C4 and

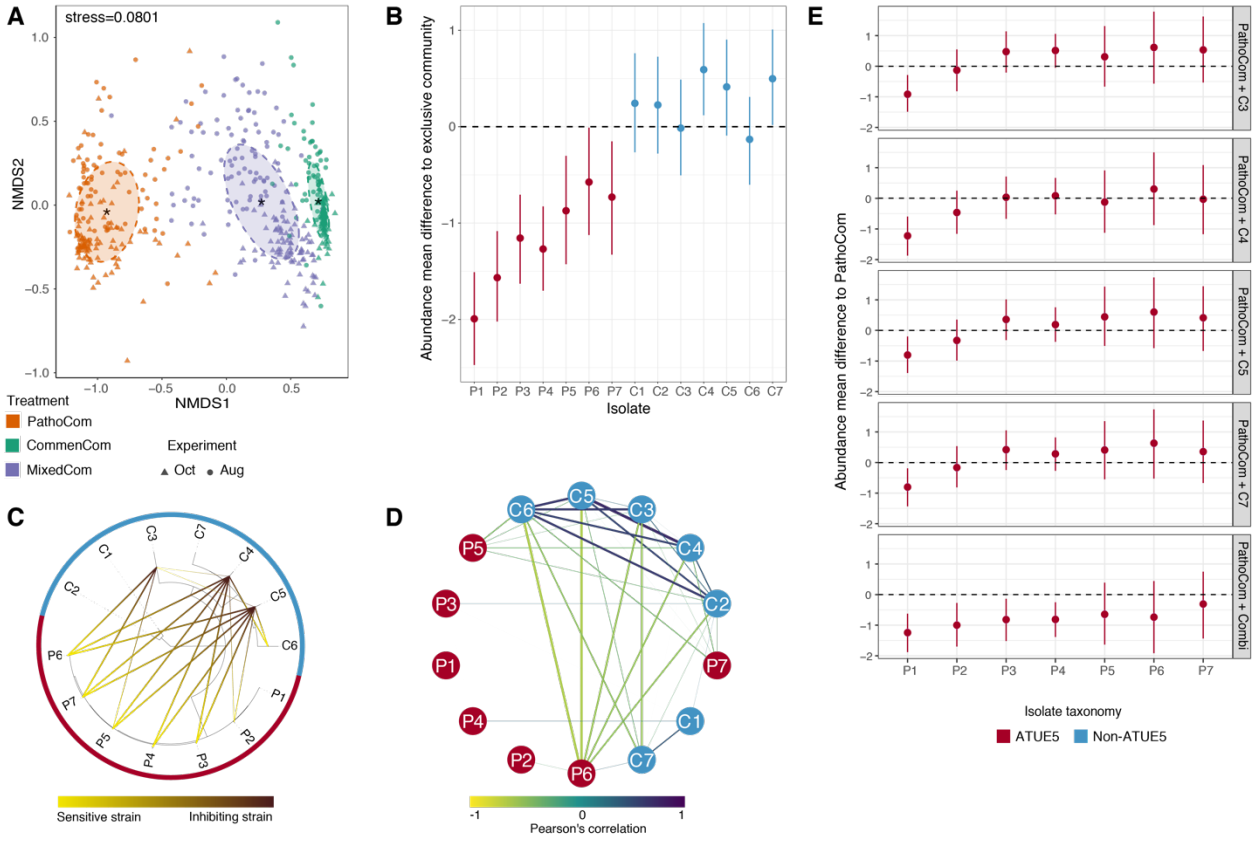


Figure 4. Differential inhibition patterns of pathogens by commensals *in vitro* and *in planta* A. Nonmetric multidimensional scaling (NMDS) based on Bray-Curtis distances between samples infected with the three synthetic communities, across two experiments. The abundance of all 14 barcoded isolates was measured in all communities, including PathoCom and CommenCom, which contained only 7 of the 14 isolates, to account for potential cross contamination and to avoid technical bias. Oct=October, Aug=August. n=170 for PathoCom, n=151 for CommenCom, and n=182 for MixedCom. **B.** Abundance change of the 14 barcoded isolates in MixedCom when compared to their exclusive community, in infected plants (i.e. PathoCom for ATUE5, and CommenCom for non-ATUE5). Abundance mean difference was estimated with the model $\log_{10}$(isolate load) ~ treatment * experiment + error, for each individual strain. Thus, the treatment coefficient was estimated per isolate. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter. **C.** Taxonomic representation of the 14 barcoded isolates tested *in vitro* for directional interactions. Ring colors indicate the bacterial isolate classification, ATUE5 or non-ATUE5. Directional inhibitory interactions are indicated from yellow to black. The experiments were repeated three times, with two technical replicates. Only inhibitions observed in at least in two independent experiments and in both technical replicates were considered. **D.** Correlation network of relative abundances of all 14 barcoded isolates in MixedCom-infected plants. Strengths of negative and positive correlations are indicated from yellow to purple. Boldness of lines is also indicating the strength of correlation, and only correlations > |±0.2| are shown. Node colors indicate the bacterial isolate

classification, ATUE5 or non-ATUE5. **E.** *in planta* Abundance change of the seven ATUE5 isolates in non-ATUE5 inclusive treatments, in comparison to PathoCom. Abundance mean difference was estimated with the model $\log_{10}$(isolate load) ~ treatment * experiment + error, for each individual strain. Thus, the treatment coefficient was estimated per isolate. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter. 'Combi' - combination of the isolates C3,C4,C5 and C7. n=23.

C5 showed the same inhibition pattern: Both inhibited all pathogenic isolates but P1, and both inhibited the same two commensals, C6 and, only weakly, C3. C3 inhibited a total of three ATUE5 isolates: P5, P6 and P7. In summary, the *in vitro* assay provides evidence that among the tested *Pseudomonas*, direct inhibition was a trait unique to commensals, and susceptible bacteria were primarily pathogens. This supports the notion that ATUE5 and non-ATUE5 isolates have divergent competition mechanisms, or at least differ in the strength of the same mechanism.

The *in vitro* results recapitulated the general trend of pathogen inhibition found among treatments *in planta*. Nevertheless, we observed major discrepancies between the two assays. First, P1 was not inhibited by any isolate in the host-free assay (**Figure 4C**), though it was the most inhibited member *in planta*, among the communities (**Figure 4B**). Second, no commensal isolate was inhibited *in plana*, among communities (**Figure 4B**), while two commensals - C3 and C6 - were inhibited *in vitro* (**Figure 4C**). Both could suggest an effect of the host on microbe-microbe interactions. To explore such effects, we analysed all pairwise microbe-microbe abundance correlations within MixedCom-infected hosts. When we used absolute abundances, all pairwise correlations were positive, also in CommenCom and PathoCom (**Figure S10A**), consistent with there being a positive correlation between absolute abundance of individual isolates and total abundance of the entire community (**Figure S11**), i.e., no isolate was less abundant in highly colonized plants than in sparsely colonized plants. It indicates that there does not seem to be active killing of competitors *in planta* in the CommenCom, which is probably not surprising. With relative abundances, however, a clear pattern emerged, with a cluster of commensals that were positively correlated, possibly reflecting mutual growth promotion, and several commensal strains being negatively correlated with both P6 and C7, possibly reflecting unidirectional growth inhibition (**Figure 4D**). We did not observe the same correlations within CommenCom among commensals and within PathoCom among pathogens as we did for either subgroup in MixedCom, reflecting higher-order interactions (**Figure S10B**).

The *in planta* patterns, measured in complex communities, did not fully recapitulate what we had observed *in vitro*, with pairwise interactions. We therefore investigated individual

commensal isolates for their ability to suppress pathogens *in planta*, and also tested the entourage effect. We focused on the three commensals C3, C4 and C5 ,which had directly inhibited pathogens *in vitro*, and C7, which had not shown any inhibition activity *in vitro*, as control. We infected plants with mixtures of PathoCom and each of the four individual commensals, as well as PathoCom mixed with all four commensals. Since pathogen inhibition seemed to be independent of the host genotype, we arbitrarily chose HE-1. Regardless of the commensal isolate, only P1 was significantly suppressed in all commensal-including treatments (**Figure 4E**), with P2,P3 and P4 being substantially inhibited only by the mixture of all four commensals. Together with the lack of meaningful inhibitory difference between individual commensals, this indicates that pathogen inhibition was either a function of commensal dose, or a result of interaction among commensals.

An important finding was that four commensal strains had much more similar inhibitory activity *in planta* than *in vitro*, and that the combined action was greater than the individual effects. Together, this suggested that the host contributes to the observed interactions between commensal and pathogenic *Pseudomonas*. To begin to investigate this possibility, we studied potential host immune responses with RNA sequencing.

Defensive response elicited by non-ATUE5 inferred from host transcriptome changes

For the RNA-seq experiment, we treated plants of the genotype Lu3-30 with the three synthetic communities, and also used a bacteria-free control treatment. We sampled the treated plants at three and four days after infection (dpi), thus increasing the ability to pinpoint differentially expressed genes (DEGs) between treatments that are not highly time-specific. Exploratory analysis indicated that the two time points behaved similarly, and they were combined for further in-depth analysis.

We first looked at DEGs in a comparison between infected plants and control; with PathoCom, there were only 14 DEGs, with CommenCom 1,112 DEGs, and MixedCom 1,949 DEGs, suggesting that the CommenCom isolates, which are also present in the MixedCom, elicited a host stronger response than the PathoCom members. Furthermore, the high number of DEGs in MixedCom - higher than both PathoCom and CommenCom together - suggest a synergistic response derived from inclusion of both PathoCom and CommenCom members. Alternatively, this could also be a consequence of the higher initial inoculum in the 14-member

MixedCom than either the 7-member PathoCom or 7-member CommenCom, or a combination of the two effects (**Figure 5A-B; Figure S12**).

The genes induced by the MixedCom fell into two classes: Group 5 (**Figure 5A-B**) was also induced, albeit more weakly, by the CommenCom, but not induced by the PathoCom. This

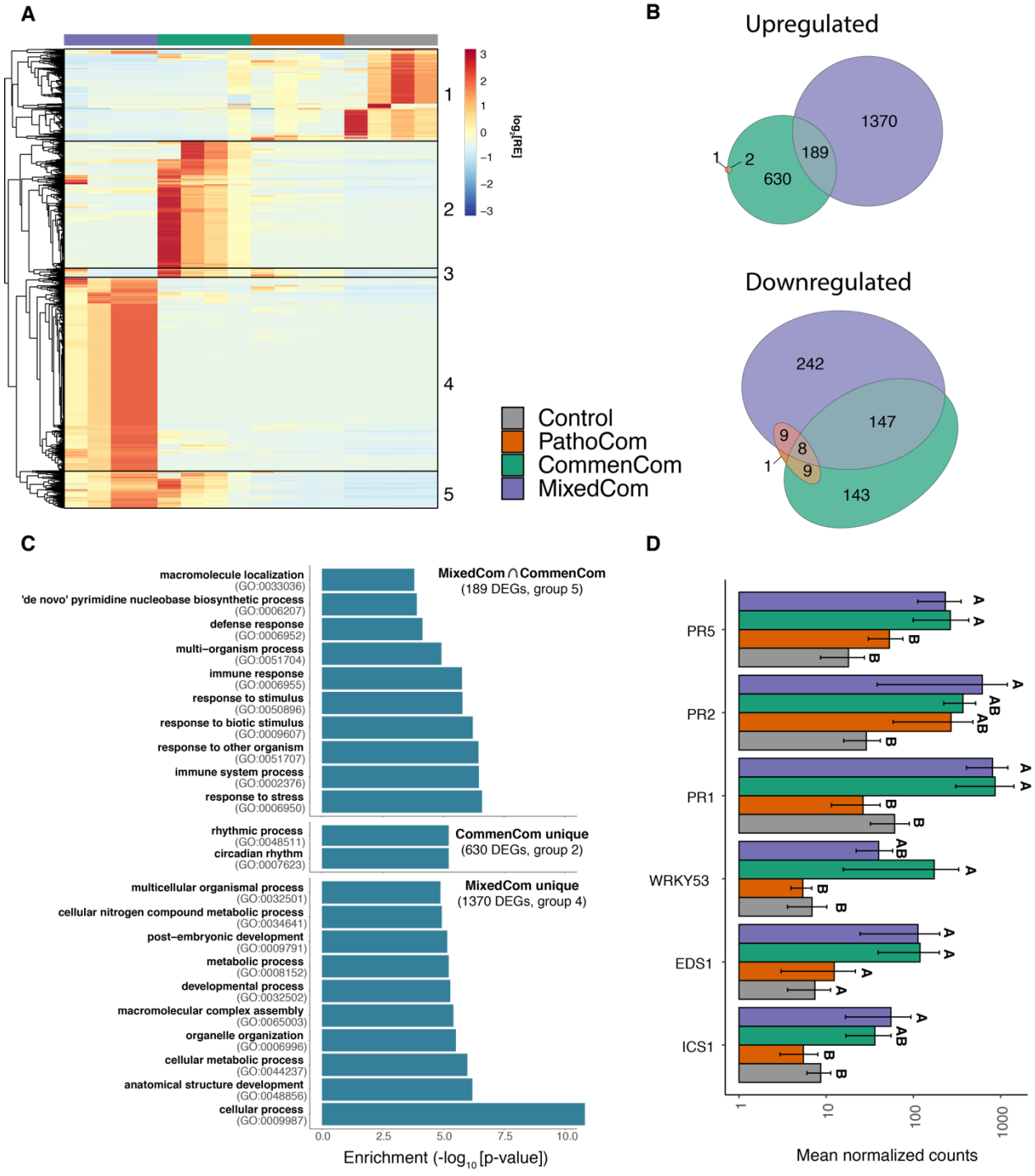


Figure 5. Only commensal members elicit a host-defensive response. A. Relative expression (RE) pattern of 2,727 differentially expressed genes (DEGs) found in at least one of the comparisons of CommenCom, PathoCom and MixedCom with control. DEGs were hierarchically clustered. **B.** Euler diagram of DEGs in PathoCom-, CommenCom- and MixedCom-treated plants, compared with control (log2[FC]>|±1|; FDR<0.05; two-tailed Student's *t*-test followed by Benjamini-Hochberg correction). **C.** Overrepresented GO terms in upregulated DEG subsets: CommenCom and MixedCom intersection (189 DEGs), CommenCom unique (630 DEGs) and MixedCom unique (1,370 DEGs). Only the top ten non-redundant GO terms are presented; for the full lists of overrepresented GO terms and expression data, see Table S4 and Supplementary Data 1. **D.** Expression values of six defense marker-genes. Mean ± SEM. Groups sharing the same letter are not significantly different (Tukey-adjusted, P>0.05); n=4.

group was overrepresented for non-redundant gene ontology (GO) categories linked to defense (**Figure 5C**) and most likely explains the protective effects of commensals in the MixedCom. Specifically, among the top ten enriched GO categories in the shared MixedCom and CommenCom set, eight relate to immune response or response to another organism ('defense response', 'multi–organism process', 'immune response', 'response to stimulus', 'response to biotic stimulus', 'response to other organism', 'immune system process', 'response to stress')(**Figure 5C**).

Group 4 was only induced in MixedCom, either indicating synergism between commensals and pathogens, or being a consequence of the higher initial inoculum. This group included a small number of redundant GO categories indicative of defense, such 'salicylic acid mediated signaling pathway', 'multi-organism process', 'response to other organism' and 'response to biotic stimulus' (**Table S4**). Moreover, the MixedCom response cannot simply be explained by synergistic effects or commensals suppressing pathogen effects, since there was a prominent class, Group 2, which included genes that were induced in the CommenCom, but to a much lesser extent in the PathoCom or MixedCom. From their annotation, it was unclear how they can be linked to infection (**Figure 5C**).

About 500 genes (Group 1) that were downregulated by all bacterial communities are unlikely to contain candidates for commensal protection (**Figure 5A**).

Cumulatively, these results imply that the CommenCom members elicited a defensive response in the host regardless of PathoCom members, while the mixture of both led to additional responses. To better understand if selective suppression of ATUE5 in MixedCom infections may have resulted from the recognition of both non-ATUE5 and ATUE5 (reflected by

a unique MixedCom set of DEGs) or solely non-ATUE5 (a set of DEGs shared by MixedCom and CommenCom), we examined the expression of key genes related to the salicylic acid (SA) pathway and downstream immune responses. Activation of the SA pathway was previously related to increased fitness of *A. thaliana* in the presence of wild bacterial pathogens, a phenomenon which was attributed to an increased systemic acquired resistance (SAR) [28]. We observed a general trend of higher expression in MixedCom- and CommenCom-infected hosts for several such genes (**Figure 5D**). Examples are *PR1* and *PR5*, marker genes for SAR and resistance execution. Therefore, according to the marker genes we tested, non-ATUE5 elicited a defensive response in the host, regardless of ATUE5 presence.

We conclude that the expression profile of non-ATUE5 infected Lu3-30 plants suggests an increased defensive status, supporting our hypothesis regarding host-mediated ATUE5 suppression. We note, however, that ATUE5 suppression was not associated with full plant protection (thus control-like weight levels) in all plant genotypes. One, Ey15-2, was only partially protected by MixedCom (**Figure 2**), despite levels of pathogen inhibition being not very different from other host genotypes (**Figure S9**).

Lack of protection in the genotype Ey15-2 explained by a single pathogenic isolate

The fact that Ey15-2 was only partially protected by MixedCom (**Figure 2**), manifest the importance of the host genotype in plant-microbe-microbe interactions, and reflecting dynamics between microbes and plants in wild populations. We wanted to reveal the cause for this differential interaction.

Our first aim was to rank compositional variables in MixedCom according to their impact on plant weight, regardless of host genotype. Next, we asked whether any of the top-ranked variables could explain the lack of protection in Ey15-2. With Random Forest analysis, we estimated the weight-predictive power of all individual isolates in MixedCom, as well as three cumulative variables: Total bacterial abundance, total ATUE5 abundance, and total non-ATUE5 abundance. We found that the best weight-predictive variable was the abundance of pathogenic isolate P6, followed by total bacterial load and total ATUE5 load - which were probably confounded by the abundance of P6 (**Figure 6A**). In agreement, P6 was the dominant ATUE5 in MixedCom (**Figure 6B**, **Figure S13A**). We thus hypothesized that the residual pathogenicity in MixedCom-infected Ey15-2 was caused by P6.

Although P6 grew best in Ey15-2, the difference to most other genotypes was not significant (**Figure S13B**). However, P6 was particularly dominant in Ey15-2 (**Figure 6B**).

Given that pathogen load in Ey15-2 was driven to a substantial extent by P6, we assumed that this isolate had a stronger impact on the weight of this genotype than in others. We experimentally validated that removal of P6 restored protection, when Ey15-2 was infected with MixedCom (**Figure 6C**). To confirm that restored protection was due to the interaction of commensals with the five other pathogenic isolates (P1-P5), rather than simply removal of P6, we also treated Ey15-2 with PathoCom only, but lacking P6. The removal of P6 did not diminish the negative weight impact of PathoCom (P1-P5) (**Figure S14**), implying that it was indeed the interaction between commensals with five out of six pathogenic isolates that mitigated the harmful effect of pathogens in Ey15-2 plants.

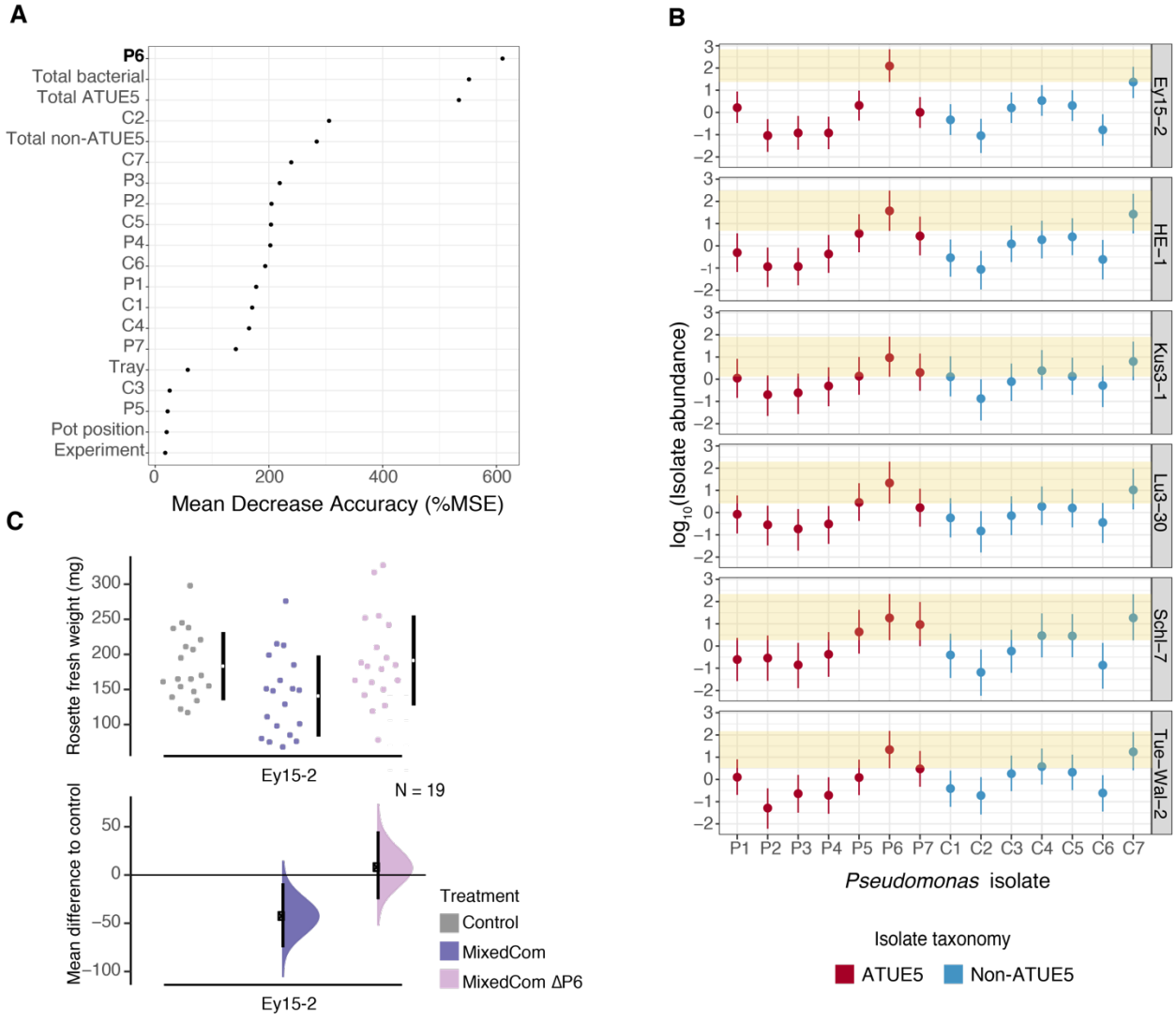


Figure 6. The effect of the isolate P6 on weight in MixedCom-infected hosts, and particularly on the host Ey15-2. A. Relative importance (mean decrease accuracy) of 20 examined variables in weight prediction of MixedCom-infected hosts, as determined by Random Forest analysis. The best predictor was abundance of isolate P6. 'Total Bacterial', 'Total ATUE5' and 'Total non-ATU5' are the cumulative abundances of the 14 isolates, 7 ATUE5 isolates, and 7 non-ATUE5 isolates, respectively. **B.** Abundance of P6 compared with the other 13 barcoded isolates in MixedCom-infected hosts, across the six *A. thaliana* genotypes used in this study. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter, following the model $\log_{10}$(isolate load) ~ isolate * experiment + error. Each genotype was analyzed individually, thus the model was utilized for each genotype separately. Shaded area denotes the 95% credible intervals for the isolate P6. **C.** Fresh rosette weight of Ey15-2 plants treated with Control, MixedCom and MixedCom without P6 (MixedCom ΔP6). Fresh rosette weight was measured 12 dpi. The top panel presents the raw data, with the breaks in the vertical black lines denoting the mean value of each group, and the vertical lines themselves indicating standard deviation. The lower panel presents the mean difference to control, plotted as bootstrap sampling [26,27], indicating the distribution of effect size that is compatible with the data. 95% confidence intervals are indicated by the black vertical bars. n=19.

Collectively, these results reveal the outcome of direct host-microbe interactions in the context of multiple microbes. Furthermore, they illustrate how plant genotype affects colonization by microbes, and how this may lead to plant health outcomes.

Discussion

In this work, we aimed to understand how complex interactions between closely related *Pseudomonas* strains affect plant health, considering host-microbe, microbe-microbe and host-microbe-microbe relationships. Not surprisingly, we found that genetics mattered at all levels: membership of *Pseudomonas* strain in commensal or pathogenic clade, genetic variation within each *Pseudomonas* clade, and genetic diversity among *A. thaliana* host strains. Commensal *Pseudomonas* can protect *A. thaliana* from the effects of pathogenic *Pseudomonas* by reducing their proliferation within the plant. However, although this was a general phenomenon, one *A. thaliana* genotype was only partially protected, and this was due to this genotype being particularly susceptible to a specific *Pseudomonas* pathogen. Together, this demonstrates how the host environment can affect microbe-microbe interactions.

The importance of protective interactions for plant health has been demonstrated in both agricultural and wild contexts [1,21,29]. Our results reveal the extreme specificity of these

interactions, with closely related pathogenic isolates interacting differently with protective strains We found that upon co-infection with a mixture of pathogens and commensals, pathogens were preferentially suppressed. Perhaps our most important finding was that different plant responses induced by commensals, pathogens and mixed communities. Specifically, commensals, but not pathogens induced transcriptome signatures of defense, and these changes further enhanced in the presence of pathogens. In addition, there were sets of genes that were no longer induced when plants were infected by the mixed community rather than only commensals, as well as sets of genes specifically induced only by the mixed community. This suggests not only that microbe-microbe interactions alter the plant response, but also that these altered plant responses are causal for the differential proliferation of commensals and pathogens in plants affected with mixed communities. Synergistic host responses were previously demonstrated in *Medicago truncatula*, following infections with two symbionts - rhizobia and mycorrhizal fungi [30]. These findings support the hypothesis that the complex interplay between the plant immune system and the microbiota goes beyond the individual plant-pathogen interactions, eventually leading to microbial homeostasis [31]. The exact mechanism behind the synergistic effect we describe must still be investigated, though known cases of host-dependent protective interactions provide plausible explanations. For example, early exposure to harmless rhizosphere microbes can prime the plant to suppress at a later time point a broad range of pathogens even in distal tissues, a phenomenon known as induced systemic resistance (ISR) [28].

Another strength of our study is that we used naturally co-occurring biological material, namely strains of *A. thaliana* host and *Pseudomonas* bacteria that had been isolated from a single geographic area. Our results help to explain why the *Pseudomonas* pathogens used here, which are lethal in mono-associations, seem to cause only limited disease in the field [11], namely their effects being modified by other microbes, including other *Pseudomonas* strains.

A limitation of the current study was that we examined only a few commensal isolates, and tested them mostly in complex mixtures. A next logical step will be to test the protective effects of individual commensal *Pseudomonas* strains from the local Tübingen [11] collection, to explore (i) how common protection by commensal *Pseudomonas* is, (ii) how much it depends on the genotype of the pathogen, and (iii) what the genes are that support protection.

We used pathogenic isolates that share over 99% of their 16S rDNA signature, and are highly similar in their core genome [11]. Nonetheless, we found functional differences, relating to both host-microbe and microbe-microbe interactions, exemplified by an individual pathogenic *Pseudomonas* isolate that both dominated the mixed synthetic communities, and that caused a

lack of protection in one host genotype. In agreement, Karasov and colleagues [11] had already found that members of this clade of *Pseudomonas* differ substantially in their ability to cause disease in mono-associations.

Friedman and colleagues [32] accurately predicted microbial community structures in the form of trios based on information about pairwise interactions. How easily, however, higher-order communities can be predicted from pairwise interactions, remains to be seen, although recent statistical advances are promising [33,34]. The genome-barcoding method we developed allows strain-level tracking, and thus can be implemented to understand multistrain community assembly. However, in its current format, it is limited to low-throughput studies, mainly due to the cumbersome cloning and transformation serial process. An alternative is presented by high-throughput experiments that combine whole-genome sequencing with statistical reconstitution of known haplotypes [35,36], and which could be employed to study the dynamics of more complex communities. A growing body of literature is revealing effects that can only be found by the ensemble of relationships. For example, in inflammatory bowel disease [37] disease has been linked to changes in microbial community structure rather than to an individual microbe. Another example is provided by plant beneficial consortia, in which only microbial mixtures, but not any single strain triggered pathogen suppression [38,39].

Further advancements in understanding of the plant-microbe-microbe complex in the light of plant health can improve our agriculture practices, allowing the development of more sustainable plant protection methods [40–42].

Methods

Plant material

The plant genotypes HE-1, Lu3-30, Kus3-1, Schl-7, Ey15-2 and Tue-Wal2 were used in this study, all originally collected from around Tuebingen, Germany. More details, including stock numbers, can be found in Table S5. Seeds were sterilized by overnight incubation at –80°C, followed by ethanol washes (shake seeds for 5-15min in solution containing 75% EtOH and 0.5% Triton-X-100, and then wash seeds with 95% EtOH and let them dry in a laminar flow hood). Seeds were stratified in the dark at 4°C for 6-8 days prior to planting on potting soil (CLT Topferde; www.einheitserde.de). Plants were grown in 60-pots trays (Herkuplast Kubern, Germany), in which compatible mesh-net pot baskets were inserted, to allow for subsequent relocation of the pots. All plants were grown in short days (8 h of light) at 23°C. Light was applied

554 using Cool White Deluxe fluorescent bulbs, at 125 to 175 µmol m-2 s-1. Relative humidity was
555 set to 65%.
556
557 **Barcoding *Pseudomonas* isolates**
558 Excluding the *E. coli* strains that were used for cloning, all 14 bacterial isolates used in this study
559 were classified as *Pseudomonas* and collected from two locations around Tuebingen (Germany)
560 by Karasov and colleagues [11]. Full list, including metadata can be found in Table S1. The
561 procedure of genome-barcoding of the 14 bacterial isolates included random barcodes
562 preparation, cloning the barcodes into pUC18R6KT-mini-Tn7T-Km plasmid and co-
563 transformation of bacteria with the recombinant pUC18R6KT-mini-Tn7T-Km plasmid and pTNS2
564 helper plasmid (both plasmids from [24]). Preparation of barcodes and the flanking priming sites
565 was done by double stranding two overlapping single strand oligos: One that contains restriction
566 sites, universal priming site, 16 random nucleotides and an overlapping region (Bar1), and
567 another oligo that contains the reverse complement overlapping region, the second universal
568 priming site and restriction sites (Bar2), as illustrated in **Figure S15**; Detailed oligo list in Table
569 S6. The two overlapping single strand oligos were mixed in an equi-molar fashion (5ng each, 2µL
570 in total), together with 0.2 µL Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich,
571 MA, USA), 1x Q5 5x reaction buffer and 225 µM dNTP in a total reaction volume of 20 µL. The
572 mixture went through a double stranding reaction using a thermocycler (Bio-Rad Laboratories,
573 Hercules, CA, USA), with the following conditions: 95°C for 40 s, 55°C for 60 s and elongation at
574 72°C for 3 min. The resulting product was cloned into pUC18R6KT-mini-Tn7T-Km plasmid, using
575 the restriction enzymes XhoI and SacI and ligation with T4 DNA-Ligase (Thermo Fisher Scientific,
576 USA). Standard restriction and ligation were conducted, as instructed by the manufacturer
577 protocol. Pir1 competent *E. coli* (Thermo Fisher Scientific, USA) cells were transformed with the
578 ligation product, and subsequently plated on selective Lysogeny broth (LB) agar (1.75%) with 50
579 ng/mL Kanamycin and 100 ng/mL. Bacterial colonies were validated as successful transformants
580 by PCR with the primers p1 and p2 that are specific for the foreign DNA (detailed oligo list in
581 Table S6). Positive colonies were grown in LB overnight and then used for subsequent plasmid
582 isolation (GeneJET Plasmid Miniprep Kit; Thermo Fisher Scientific, USA). About 150
583 pUC18R6KT-mini-Tn7T-Km recombinant plasmids were stored at -4°C, each is expected to
584 contain a unique barcode. Sanger sequencing was conducted on a subset of the plasmid library
585 using the primer p1, to validate their barcodes sequence. 14 validated barcodes-inclusive

plasmids were randomly selected for the barcoding of the 14 isolates, and these were used for co-transformation together with the plasmid pTNS2 to genome-barcode the selected 14 *Pseudomonas* isolates, as described in [24]. Briefly, *Pseudomonas* strains were grown overnight in LB, pelleted and washed with 300mM sucrose solution to create electrocompetent cells, and were finally electroporated with the recombinant pUC18R6KT-mini-Tn7T-Km (barcodes inclusive) and pTNS2 in a ratio of 1:1. Transformed *Pseudomonas* isolates were grown on selective LB-agar media with 30 mg/mL Kanamycin, and colonies were validated by PCR with the primers p1 and p2 (detailed oligo list in Table S6; Gel electrophoresis results in Figure S2A). Positive colonies were grown in LB with 30 mg/mL Kanamycin overnight, and one portion was stored at -80°C in 25% glycerol, while the other portion was used for DNA extraction (Puregene DNA extraction kit; Invitrogen, USA), followed by Sanger sequencing to validate the barcodes sequences (sequences detailed in Table S1).

Barcoded and WT isolates growth comparison assay

To compare the growth of the 14 barcoded bacteria with their respective WT, both barcoded and WT isolates were grown overnight in Lysogeny broth (LB) and 10 mg/mL Nitrofurantoin (antibiotic in which all isolated *Pseudomonas* can grow), diluted 1:10 in the following morning and grown for 3 additional hours until they entered log phase. Subsequently, bacteria were pelleted at 3500 g and resuspended in LB to a concentration of OD_{600} = 0.0025, in a 96-wells format plate with a transparent, flat bottom (Greiner Bio One, Austria). Finally, the plate was incubated in a plate reader at 28°C while shaking, for 10 hours (Robot Tecan Infinite M200; Tecan Life Sciences, Switzerland). OD_{600} was measured in one hour intervals.

Synthetic communities infections and plant sampling

All synthetic communities were prepared as followed: The relevant barcoded isolates were grown overnight in Lysogeny broth (LB) and 30 mg/mL Kanamycin, diluted 1:10 in the following morning and grown for 3 additional hours until they entered log phase, pelleted at 3500 g, resuspended in 10 mM $MgSO_4$ and pelleted again at 3500 g to wash residual LB, and resuspended again in 10 mM $MgSO_4$ to a concentration of OD_{600} = 0.2, creating a stock solution per isolate for subsequent mixtures. Next, the relevant barcoded isolates were mixed to a final solution with a concentration of OD_{600} = 0.0143 per isolate. Thus, the total concentration per synthetic community was OD_{600} = (0.0143 * isolates number), e.g. PathoCom and CommenCom which

618 comprised 7 isolates had a total concentration of OD_{600} = ~0.1, and MixedCom which comprised
619 14 isolates had a total concentration of OD_{600} = ~0.2. The prepared volume for any synthetic
620 community was calculated by the function: Final volume = number of plants to infect * 2.5ml.
621 Control treatment was sterile 10 mM $MgSO_4$ solution. Heat-killed PathoCom was made by
622 incubating a portion of the living PathoCom in 100°C for two hours. In Murashige and Skoog
623 infections, PathoCom was diluted 1:10, thus infections were done using O.D. 0.01. All solutions
624 with synthetic communities were stored at 4°C overnight, and infections were conducted in the
625 morning of the following day.
626 Infections in the Murashige and Skoog (MS) sterile system were done as described by
627 Karasov and colleagues [11]. In brief, 12-14 days old plants were infected by drip-inoculating
628 200 µl of the corresponding treatment onto the whole rosette.
629 The leaves of soil-grown plants were spray-infected, 21 days post sowing. Spraying was
630 done with an airbrush (BADGER 250-1; Badger Air-Brush Co., USA), and each plant was sprayed
631 on both the abaxial and adaxial side for about 1.5 s each. Plants of the same treatment group
632 were placed together in 60-pots trays (Herkuplast Kubern, Germany), in which compatible mesh-
633 net pot baskets were pre-inserted to allow for subsequent relocation of the pots. After the
634 treatment, the transportable pots were reshuffled in new 60-pots trays to form a full randomized
635 block design, thus each tray contained plants from all treatments, in equal amounts. The
636 randomized trays were covered with a transparent lid to increase humidity (Bigger Greenhouse-
637 60x40cm; Growshop Greenbud, Germany). Four days post infection, two built-in openings in the
638 lids were opened to allow for better air flow and to limit humidity. Eight days post infection, lids
639 were removed. Twelve days post infection, the rosettes of all treated plants were detached using
640 sterilized scalpel and tweezers, weighted, washed from epiphytes (sterile distilled water, 70%
641 EtOH with 0.1% Triton X-100 and then again with sterile distilled water), dried using sterilized
642 paper towels and sampled in 2ml screw cap tubes prefilled with Garnet sharp particles 1mm
643 (Roth, Germany). Tubes with the sampled plants were flash freezed in liquid N_2, and stored in -
644 80°C.
645
646 **DNA extraction, barcodes PCR and qPCR**
647 Frozen sampled plants were used for DNA extraction suitable for metagenomics, using a
648 protocol that was previously described by karasov and colleagues [11]. Briefly, the samples were
649 subjected to bead-beating in the presence of 1.5% sodium dodecyl sulfate (SDS) and 1mm

garnet rocks, followed by SDS cleanup with 1/3 volume 5M potassium acetate, and then SPRI beads.

The resulting DNA was used for a two step PCR. The first PCR step amplified the genome-integrated barcodes and added short overhangs, using the primer p3 and the primers p4-p9. The latter are different versions of one primer with frameshifting nucleotides, allowing for better Illumina clustering, and thus sequencing quality, following the method described by [43] (2013; Detailed oligo list in Table S6). Each primer frameshift version was used for a different PCR plate (i.e. 96 samples). The second step primed the overhangs to Illumina adapters for subsequent sequencing, using standard Illumina TruSeq primer sequences. Unique tagging of PCR samples was accomplished by using 96 indexing primers, combined with the six combinations of frameshift primers in the first PCR (as detailed in [43]), allowing demultiplexing of up to 576 samples in one Illumina lane. The first PCR was done in 25 µL reactions containing 0.125 µL TaqI DNA polymerase (Thermo Fisher Scientific, USA), 1x Taq1 10x reaction buffer, 0.08 µM each of forward and reverse primer, 225 µM dNTP and 1.5 µL of the template DNA. The first PCR was run for 94°C for 5 min followed by 10 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 1 min, and a final 72°C for 5 min. 5 µL of the first PCR product was used in the second PCR with tagged primers including Illumina adapters, in 25 µL containing 0.25 µL Q5 high-fidelity DNA polymerase (New England Biolabs, USA), 1x Q5 5x reaction buffer, 0.08 µM forward and 0.16 µM of reverse (tagging) primer and 200 µM dNTP. The final PCR products were cleaned twice using SPRI beads in a 1:1 bead to sample ratio, and eluted in 15 µL. Samples were combined into one library in an equimolar fashion. Final libraries were cleaned twice using SPRI beads in a 0.6:1 bead to sample ratio to clean the primers from the product, and were finally eluted in half of their original volume. Samples were sequenced by a MiSeq instrument (Illumina), using a 50 bp single-end kit.

In order to estimate the ratio of barcoded *Pseudomonas* to plant chromosomes, two qPCR reactions were conducted - one which is specific to the barcodes, and the other which is plant-specific, targeting the gene GIGANTEA which is normally found in one copy. For barcodes-specific qPCR, the primers p10 and p11 were used, and for plant-specific qPCR the primers p12 and p13 (Table S6). qPCR reactions were done in 10 µL reactions containing x1 Maxima SYBR green qPCR master mix x2, 0.08 µM each of forward and reverse primer and 1 µL of template DNA. All qPCR reactions were run for 94°C for 2 min followed by 94°C for 15 s and 60°C for 1

min in a BioRad CFX384 Real Time System (Biorad, USA) qPCR machine. Reactions were done in triplicates.

In vitro directional suppression assay

All 14 barcoded isolates *in vitro* pairwise interactions were tested following the method described in Helfrich et al. [44], while adjusting the conditions to better fit *Pseudomonas*. Briefly, the 14 barcoded isolates were grown in LB with 30 mg/mL Kanamycin overnight, diluted 1:10 the following morning and regrown. One portion was taken from each isolate after 3 hours (when entering the log phase), diluted to a final concentration of OD_{600} = 0.001 in 15ml LB with 1% agar and immediately poured into a square plate to form a uniform layer containing the test strain. Another portion pelleted at 3500 g, washed from residual LB in 10 mM $MgSO_4$, pelleted again at 3500 g in 10 mM $MgSO_4$ with half of the original volume. Roughly 1 µL of each strain was printed onto the solidified agar layer containing the putative sensitive strain. Inhibitory interactions were estimated after 1-2 days incubation at 28°C by documenting observable halos. The strength of inhibitions was assessed by the halo size as previously described [44]).

RNA-sequencing

Plants from the genotype Lu3-30 were infected with Control, PathoCom, CommenCom and MixedCom as described below. Sampling was conducted three and four days post infection, two replicates per treatment in each time point, thus four samples per treatment in total. Plants were sampled using sterilized scalpel and tweezers and were immediately placed in 2ml screw cap tubes prefilled with Garnet sharp particles 1mm (Roth, Germany), flash freezed in liquid N_2 and stored in -80°C. RNA extraction was conducted on the frozen samples as previously described [45]. Briefly, a guanidine hydrochloride buffer was added to grounded frozen and rosettes, followed by phase separation and sediments removal. Combined with 96% EtOH, the solution was loaded onto a plasmid DNA extraction column (QIAprep Spin Miniprep Kit; Qiagen), and went through several washes before elution of the RNA. mRNA enrichment and sequencing libraries were prepared as previously described [46]. Briefly, mRNA enrichment was done using NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, USA), followed by heat fragmentation. Next, First strand synthesis (SuperScript II reverse transcriptase; Thermo Fisher Scientific, USA), and second strand synthesis (DNA polymerase I; New England Biolabs, USA) were conducted, and subsequently end repair (T4 DNA polymerase, Klenow DNA

polymerase and T4 Polynucleotide Kinase ;New England Biolabs, USA) and A-tailing (Klenow Fragment; New England Biolabs, USA). Nextera-compatible universal adapters [47] were ligated to the product (T4 DNA ligase; New England Biolabs, USA), and i5 and i7 PCR amplification was done (Q5 polymerase; New England Biolabs, USA). Size selection and DNA purification were made using SPRI beads. Samples were sequenced by a HiSeq3000 instrument (Illumina), using a 150 bp paired-end kit.

Sampling locations map, phylogenetics and isolates abundance in the field

Information about sampling locations of the six *A. thaliana* used in this study was retrieved from the 1001 genome project [25], and *Pseudomonas* sampling locations were retrieved from Karasov and colleagues [11]. The map was plotted using the "ggmap" function of the ggmap R package [48].

Phylogenetic analysis of the 14 selected *Pseudomonas* isolates was done using their core genomes, as they were previously published [11]. Maximum-likelihood phylogenies were constructed with RAxML (v.0.6.0) using GTR+Gamma model [49], and visualization was done by iTOL [50]. The abundance in the field of the selected isolates was estimated by binning similar isolates using a threshold of divergence less than 0.0001 in the core genome. The mean number of substitutions per site taken from the estimated branch length for the core-genome based phylogeny calculated by RAxML. Lastly, the number of binned isolates was divided by the total number of isolates surveyed by Karasov and colleagues [11].

Growth analysis of WT and barcoded isolates

Growth of both WT and barcoded isolates was analyzed using the function "SummarizeGrowthByPlate" from the Growthcurver R package[51]. The change of barcoded isolates in comparison to their corresponding WT in growth rate, carrying capacity and area under the curve, was calculated by the model: Growth quantity ~ strain type (i.e. WT/barcoded).

Plant weight analysis

All rosette fresh weight analyses and visualizations were done using the function "dabest" of the dabestr R package [26,27].

Combining barcode PCR and qPCR results to estimate bacterial load per isolate

All reads from barcode-PCR sequencing were mapped against a custom barcodes database (Table S1) using the algorithm BWA-MEM[52] (version 0.7.17-r1188), and a count matrix of all 14 isolates for every plant sample was created. Samples with less than a total of 200 hits were discarded or resequenced (mean=15709.8). Counts were transformed to proportions by dividing the counts of each isolate in the total hits per sample, resulting in relative abundance matrix.

qPCR results were analyzed using the software Bio-Rad CFX Manager (v3.1) with default parameters. Quantification cycle (Cq) values smaller than 32 were discarded, and barcoded bacterial load was determined by the equation $bacteri\ \llcorner l\ load\ = \frac{2.057^{-(barcode\ Cq)}}{2.027^{-(GIGANTEA\ Cq)}}$. The exponent bases (2.057 and 2.027) were adjusted according to primer efficiency - as determined by a calibration curve derived from a series of dilutions. The relative abundance matrix was factorized by bacterial load (relative abundance multiplied by bacterial load, per isolate) to manifest the ratio of bacterial to plant chromosomes per barcoded isolate.

Regression analysis

All posterior distributions of focal factors were estimated using the function “stan_glm” in the R package rstanarm ([53] or “lmBF” in the R package BayesFactor [54]). In both functions, default priors were used. In “stan_glm” default iteration number was used, and in “lmBF” 10,000 iterations were used. In all figures, the median, as well as 2.5% and 97.5% (95% credible intervals) of the posterior distribution were presented for each factor of interest. The exact model for every analysis is presented in the figure legend, as well as the selected references for comparison.

To compare the effect of individual predictors in a model, the full model was compared to a different model, lacking the predictor of interest (e.g. genotype). The comparison was conducted by a leave-one-out cross validation, using the function “loo_compare” in the R package Loo [55]. This Bayesian-based model comparison provides an estimate for the importance of a predictor in explaining the data. Leave-one-out cross validation improves the estimate in comparison to the common Akaike information criterion (AIC) and deviance information criterion (DIC) [55].

Variance partitioning of microbial community composition

NMDS analyses were conducted using the function "metaMDS" in the R package vegan [56], adjusting dissimilarity index to Bray-Curtis (method = "bray"), number of dimensions to 3 (k=3) and maximal iterations to 200 (trymax=200). Permutational multivariate analysis of variance (PERMANOVA) was conducted using the function "adonis", and analysis of similarities (ANOSIM) was conducted using the function "anosim" in the R package vegan [56]. Both were adjusted to Bray-Curtis dissimilarity index (method = "bray") and 2000 permutations (permutations = 2000). Multilevel pairwise comparison using adonis was conducted using the function "pairwise.adonis2" in the R package pairwiseAdonis [57].

Isolate-isolate interactions network

All pairwise isolate-isolate Pearson correlations were calculated using the function "rcorr" in the R package Hmisc [58], and visualization was done with Cytoscape 3.7.0 [59].

RNA-sequencing analysis

Reads from RNA sequencing were mapped against the *A. thaliana* reference TAIR10 using STAR (v.2.6.0; [60] with default parameters. Transcript counts matrix was done using featureCounts [61], while restricting counts to exons only (-t exon). Differential gene expression (DEG) analysis was conducted using DESeq2 (v.1.22.2; [62]), using the model 'gene_expression ~ Treatment + Time_point'. Genes with average counts of less than five were excluded from the analysis. Zero counts were converted to one to allow for the log conversion in unexpressed genes. Genes with log2FoldChange>|±1| and FDR<0.05 (two-tailed Student's *t*-test followed by Benjamini-Hochberg correction) were defined as DEGs. Euler diagrams were created using the function "euler" in the R package eulerr [63]. Statistically overrepresented GO terms were identified using the BiNGO plugin (v3.0.3) for Cytoscape [64]. Summarization and the removal of redundant overrepresented GO terms was done with the web server REVIGO [65] to extract the main trends found in the long full output by BiNGO (full list in Table S4).

Statistical analysis

All statistical analyses were performed using the R environment version 3.5.1, unless mentioned otherwise. Sample sizes were not predetermined using statistical methods.

Acknowledgments

We thank Alba Gonzalez, Stefan Petschak, Wangsheng Zhu, Wei Yuan, Sonja Kersten, Sergio Latorre, Julian Regalado, Anjar Wibowo, Bridgit Waithaka and Thanvi Srikant for helping with plant sampling. Special thanks to Christine Vogel for her detailed guidance regarding the *in vitro* assay, and for Christoph Ratzke for his input about the manuscript. This work was supported by fellowships from DAAD (O.S.), Funding was provided by HFSP long-term fellowships (LT000348/2016-L, T.L.K.; LT000565/2015-L, D.S.L.), an EMBO long-term fellowship (LRTF 1483-2015, T.L.K.) and Alexander von Humboldt Foundation (H.A.), and by DFG SPP DECRyPT and the Max Planck Society.

Data and materials availability

RNA sequencing data have been deposited with the European Nucleotide Archive (ENA) under study accession number PRJEB41069. The genome-barcoded *Pseudomonas* isolates used in this study can be obtained upon request.

Competing Interests

The authors declare no competing interests.

Author Contributions

O.S. conceived and designed the research. D.S.L. conceived bacterial barcoding and together with O.S. developed the method. O.S. performed the experiments and analyzed the results, O.S., T.L.K., D.S.L., H.A. and D.W. discussed and interpreted the results. O.S. wrote the first draft and the manuscript was written by O.S. and D.W., with input from all authors.

References

1. Durán P, Thiergart T, Garrido-Oter R, Agler M, Kemen E, Schulze-Lefert P, et al. Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. Cell. 2018;175: 973–983.e14.

2. Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE. Cross-species comparisons of

host genetic associations with the microbiome. Science. 2016;352: 532–535.

3. Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL, et al. Diversity and heritability of the maize rhizosphere microbiome under field conditions. Proc Natl Acad Sci U S A. 2013;110: 6548–6553.

4. Thiergart T, Durán P, Ellis T, Vannier N, Garrido-Oter R, Kemen E, et al. Root microbiota assembly and adaptive differentiation among European Arabidopsis populations. Nat Ecol Evol. 2020;4: 122–131.

5. Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, et al. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 Pseudomonas syringae isolates. PLoS Pathog. 2011;7: e1002132.

6. Fiegna F, Moreno-Letelier A, Bell T, Barraclough TG. Evolution of species interactions determines microbial community productivity in new environments. ISME J. 2015;9: 1235–1245.

7. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444: 323–329.

8. Gururani MA, Venkatesh J, Upadhyaya CP, Nookaraju A, Pandey SK, Park SW. Plant disease resistance genes: Current status and future directions. Physiol Mol Plant Pathol. 2012/4;78: 51–65.

9. Velásquez AC, Oney M, Huot B, Xu S, He SY. Diverse mechanisms of resistance to Pseudomonas syringae in a thousand natural accessions of Arabidopsis thaliana. New Phytol. 2017;214: 1673–1687.

10. Lakkis S, Trotel-Aziz P, Rabenoelina F, Schwarzenberg A, Nguema-Ona E, Clément C, et al. Strengthening Grapevine Resistance by Pseudomonas fluorescens PTA-CT2 Relies on Distinct Defense Pathways in Susceptible and Partially Resistant Genotypes to Downy Mildew and Gray Mold Diseases. Front Plant Sci. 2019;10: 1112.

11. Karasov TL, Almario J, Friedemann C, Ding W, Giolai M, Heavens D, et al. Arabidopsis thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales. Cell Host Microbe. 2018;24: 168–179.e4.

12. Haney CH, Samuel BS, Bush J, Ausubel FM. Associations with rhizosphere bacteria can confer an adaptive advantage to plants. Nat Plants. 2015;1. doi:10.1038/nplants.2015.51

13. Walters WA, Jin Z, Youngblut N, Wallace JG, Sutter J, Zhang W, et al. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. Proc Natl Acad Sci U S A. 2018;115: 7368–7373.

14. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, et al. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. PLoS Biol. 2016;14: e1002352.

15. Carlström CI, Field CM, Bortfeld-Miller M, Müller B, Sunagawa S, Vorholt JA. Synthetic microbiota reveal priority effects and keystone strains in the Arabidopsis phyllosphere. Nat Ecol Evol. 2019;3: 1445–1454.

872 16. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant
873 pathogenic bacteria in molecular plant pathology. Mol Plant Pathol. 2012;13: 614–629.

874 17. Innerebner G, Knief C, Vorholt JA. Protection of Arabidopsis thaliana against leaf-pathogenic
875 Pseudomonas syringae by Sphingomonas strains in a controlled model system. Appl
876 Environ Microbiol. 2011;77: 3202–3210.

877 18. Xin X-F, Nomura K, Aung K, Velásquez AC, Yao J, Boutrot F, et al. Bacteria establish an
878 aqueous living space in plants crucial for virulence. Nature. 2016;539: 524–529.

879 19. Nobori T, Wang Y, Wu J, Stolze SC, Tsuda Y, Finkemeier I, et al. Multidimensional gene
880 regulatory landscape of a bacterial pathogen in plants. Nat Plants. 2020;6: 883–896.

881 20. Vogel C, Bodenhausen N, Gruissem W, Vorholt JA. The Arabidopsis leaf transcriptome
882 reveals distinct but also overlapping responses to colonization by phyllosphere commensals
883 and pathogen infection with impact on plant health. New Phytol. 2016;212: 192–207.

884 21. Mendes R, Kruijt M, de Bruijn I, Dekkers E, van der Voort M, Schneider JHM, et al.
885 Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science.
886 2011;332: 1097–1100.

887 22. Mendes LW, Raaijmakers JM, de Hollander M, Mendes R, Tsai SM. Influence of resistance
888 breeding in common bean on rhizosphere microbiome composition and function. ISME J.
889 2018;12: 212–224.

890 23. Vorholt JA, Vogel C, Carlström CI, Müller DB. Establishing Causality: Opportunities of
891 Synthetic Communities for Plant Microbiome Research. Cell Host Microbe. 2017;22: 142–
892 155.

893 24. Choi K-H, Schweizer HP. mini-Tn7 insertion in bacteria with single attTn7 sites: example
894 Pseudomonas aeruginosa. Nat Protoc. 2006;1: 153–161.

895 25. 1001 Genomes Consortium. Electronic address: magnus.nordborg@gmi.oeaw.ac.at, 1001
896 Genomes Consortium. 1,135 Genomes Reveal the Global Pattern of Polymorphism in
897 Arabidopsis thaliana. Cell. 2016;166: 481–491.

898 26. Calin-Jageman RJ, Cumming G. The New Statistics for Better Science: Ask How Much,
899 How Uncertain, and What Else is Known. Am Stat. 2019;73: 271–280.

900 27. Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A. Moving beyond P values: data analysis
901 with estimation graphics. Nat Methods. 2019;16: 565–566.

902 28. Traw MB, Kniskern JM, Bergelson J. SAR increases fitness of Arabidopsis thaliana in the
903 presence of natural bacterial pathogens. Evolution. 2007;61: 2444–2449.

904 29. Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health.
905 Trends Plant Sci. 2012;17: 478–486.

906 30. Afkhami ME, Stinchcombe JR. Multiple mutualist effects on genomewide expression in the
907 tripartite association between Medicago truncatula, nitrogen-fixing bacteria and mycorrhizal
908 fungi. Mol Ecol. 2016;25: 4946–4962.

31. Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay Between Innate Immunity and the Plant Microbiota. Annu Rev Phytopathol. 2017;55: 565–589.

32. Friedman J, Higgins LM, Gore J. Community structure follows simple assembly rules in microbial microcosms. Nat Ecol Evol. 2017;1: 109.

33. Maynard DS, Miller ZR, Allesina S. Predicting coexistence in experimental ecological communities. Nat Ecol Evol. 2020;4: 91–100.

34. Momeni B, Xie L, Shou W. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. Elife. 2017;6. doi:10.7554/eLife.25051

35. Kessner D, Turner TL, Novembre J. Maximum likelihood estimation of frequencies of known haplotypes from pooled sequence data. Mol Biol Evol. 2013;30: 1145–1158.

36. Burghardt LT, Epstein B, Guhlin J, Nelson MS, Taylor MR, Young ND, et al. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. Proc Natl Acad Sci U S A. 2018;115: 2425–2430.

37. Matsuoka K, Kanai T. The gut microbiota and inflammatory bowel disease. Semin Immunopathol. 2015;37: 47–55.

38. Berendsen RL, Vismans G, Yu K, Song Y, de Jonge R, Burgman WP, et al. Disease-induced assemblage of a plant-beneficial bacterial consortium. ISME J. 2018;12: 1496–1507.

39. Wei Z, Yang T, Friman V-P, Xu Y, Shen Q, Jousset A. Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. Nat Commun. 2015;6: 8413.

40. Chaudhry V, Runge P, Sengupta P, Doehlemann G, Parker JE, Kemen E. Topic: Shaping the leaf microbiota: plant-microbe-microbe interactions. J Exp Bot. 2020. doi:10.1093/jxb/eraa417

41. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant-microbiome interactions: from community assembly to plant health. Nat Rev Microbiol. 2020;18: 607–621.

42. Finkel OM, Castrillo G, Herrera Paredes S, Salas González I, Dangl JL. Understanding and exploiting plant beneficial microbes. Curr Opin Plant Biol. 2017;38: 155–163.

43. Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL. Practical innovations for high-throughput amplicon sequencing. Nat Methods. 2013;10: 999–1002.

44. Helfrich EJN, Vogel CM, Ueoka R, Schäfer M, Ryffel F, Müller DB, et al. Bipartite interactions, antibiotic production and biosynthetic potential of the Arabidopsis leaf microbiome. Nat Microbiol. 2018;3: 909–919.

45. Yaffe H, Buxdorf K, Shapira I, Ein-Gedi S, Moyal-Ben Zvi M, Fridman E, et al. LogSpin: a simple, economical and fast method for RNA isolation from infected or healthy plants and other eukaryotic tissues. BMC Res Notes. 2012;5: 45.

46. Cambiagno DA, Giudicatti AJ, Arce AL, Gagliardi D, Li L, Yuan W, et al. HASTY modulates

945 miRNA biogenesis by linking pri-miRNA transcription and processing. Mol Plant. 2020.
946 doi:10.1016/j.molp.2020.12.019

947 47. Rowan BA, Seymour DK, Chae E, Lundberg DS, Weigel D. Methods for Genotyping-by-
948 Sequencing. In: White SJ, Cantsilieris S, editors. Genotyping: Methods and Protocols. New
949 York, NY: Springer New York; 2017. pp. 221–242.

950 48. Kahle D, Wickham H. ggmap: Spatial Visualization with ggplot2. R J. 2013;5: 144–161.

951 49. Stamatakis A, Ludwig T, Meier H. RAxML-III: a fast program for maximum likelihood-based
952 inference of large phylogenetic trees. Bioinformatics. 2005;21: 456–463.

953 50. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation
954 of phylogenetic and other trees. Nucleic Acids Res. 2016;44: W242–5.

955 51. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics
956 from microbial growth curves. BMC Bioinformatics. 2016;17: 172.

957 52. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
958 Bioinformatics. 2009;25: 1754–1760.

959 53. Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via
960 Stan. 2020. Available: https://mc-stan.org/rstanarm

961 54. Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. Package
962 "BayesFactor." 2018). Available: https://github.com/richarddmorey/BayesFactor/issues

963 55. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out
964 cross-validation and WAIC. Stat Comput. 2017;27: 1413–1432.

965 56. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan:
966 Community Ecology Package. 2019. Available: https://CRAN.R-project.org/package=vegan

967 57. Martinez Arbizu P. pairwiseAdonis: Pairwise multilevel comparison using adonis. 2020.

968 58. Harrell FE Jr. Package Hmisc. 2020.

969 59. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software
970 environment for integrated models of biomolecular interaction networks. Genome Res.
971 2003;13: 2498–2504.

972 60. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast
973 universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21.

974 61. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning
975 sequence reads to genomic features. Bioinformatics. 2014;30: 923–930.

976 62. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
977 seq data with DESeq2. Genome Biol. 2014;15: 550.

978 63. Micallef L, Rodgers P. eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses.

PLoS One. 2014;9: e101717.

64. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 2005;21: 3448–3449.

65. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6: e21800.

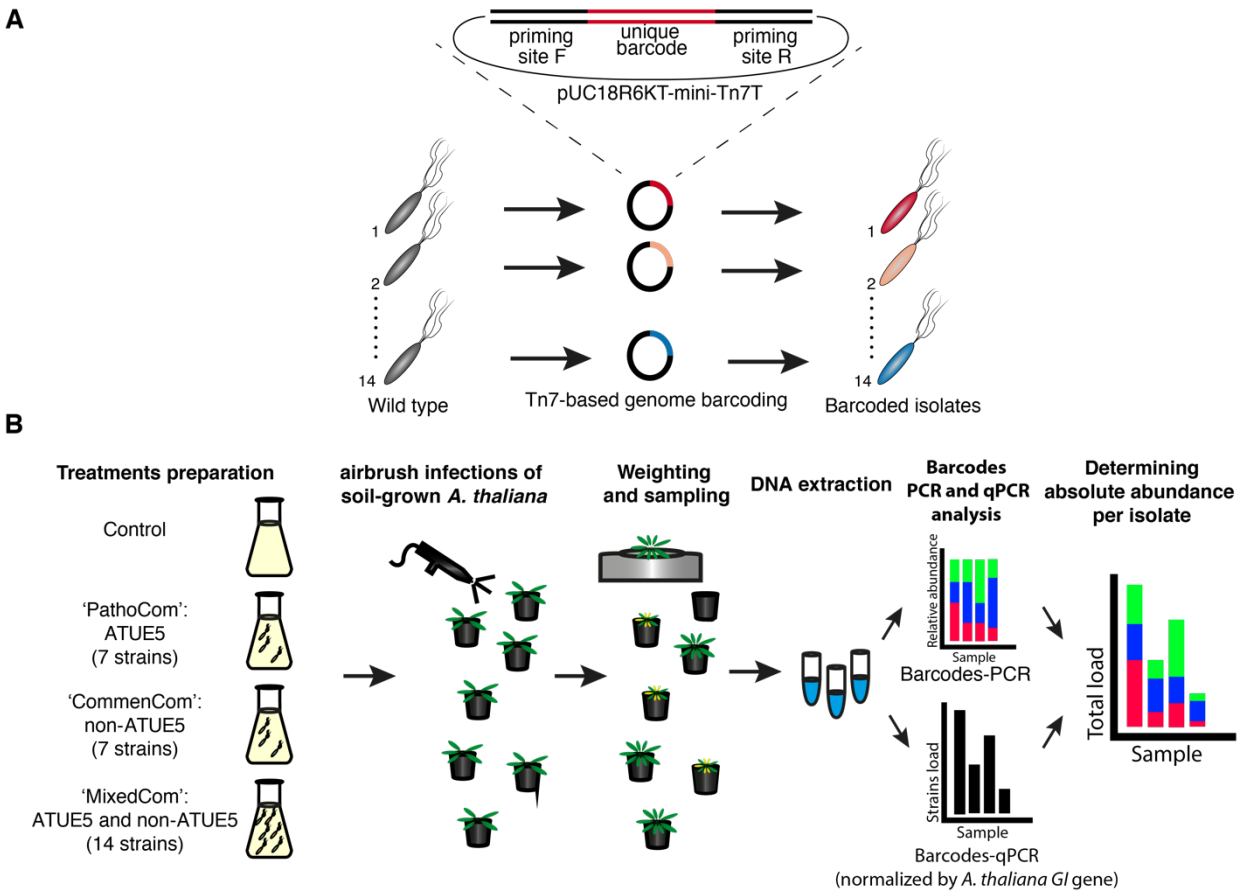


Figure S1. Illustration of (A) bacterial barcoding and (B) experimental design.

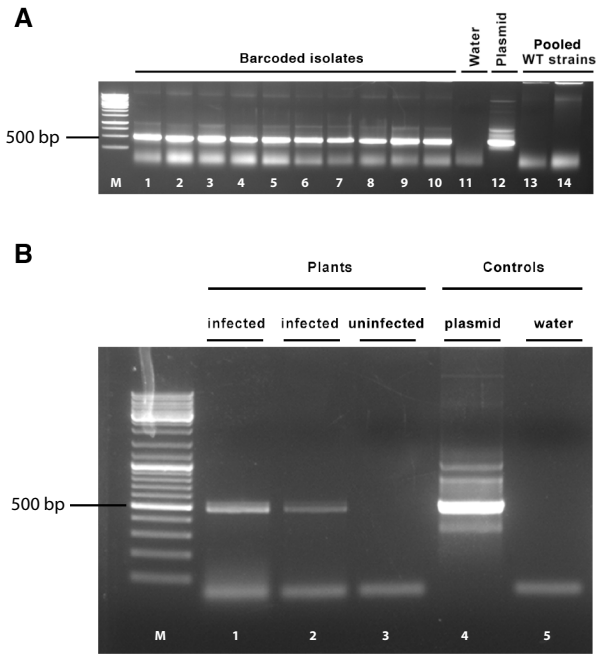


Figure S2. Validation of barcode integration and barcode-PCR specificity by agarose gel electrophoresis of PCR amplified products. A. Validation of barcode integration to chosen isolates. Lanes 1–10 used DNA from examined barcoded isolates, lane 11 is water (negative control), lane 12 is the pUC18R6KT-mini-Tn7T plasmid into which a barcode was cloned (positive control), and lanes 13-14 are replicates of the 14 pooled parental (wild-type, WT) isolates. **B.** Validation of barcode-PCR specificity. Lanes 1-2 used DNA from plants infected with the 14 barcoded bacteria, lane 3 from an uninfected plant, lane 4 pUC18R6KT-mini-Tn7T plasmid (positive control), and lane 5 is water (negative control). Both infected and uninfected plants were grown in non-sterile conditions; barcode-specific primer sets yielded expected product sizes of 522 bp. Lane M, DNA size marker. 500 bp marker indicated.

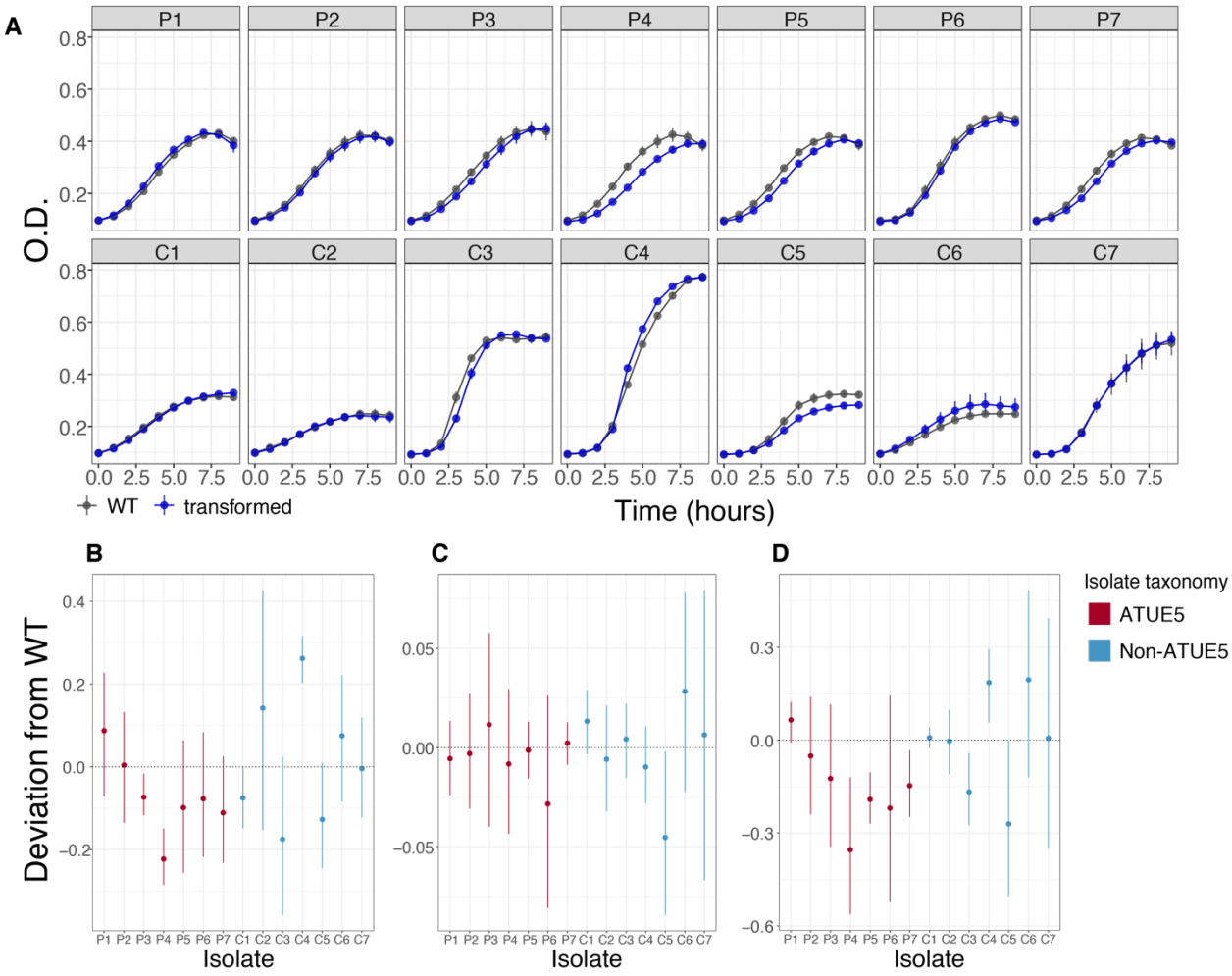


Figure S3. Comparison of growth characteristics between non-barcoded wild-type (WT) isolates and their barcoded derivatives. A. Growth curves of the 14 WT parents and their barcoded derivatives in Lysogeny broth (LB) over 10 hours, with OD_{600} recorded hourly. Mean ± SD, n=3. The change of barcoded isolates in comparison to their corresponding parents in growth rate (**B**), carrying capacity (**C**), and area under the curve (**D**) is shown. All three growth parameters were derived from the original growth curves. Dotted line signifies the non-barcoded parental baseline for a given quantity. Mean ± 95% cdl, n=3.

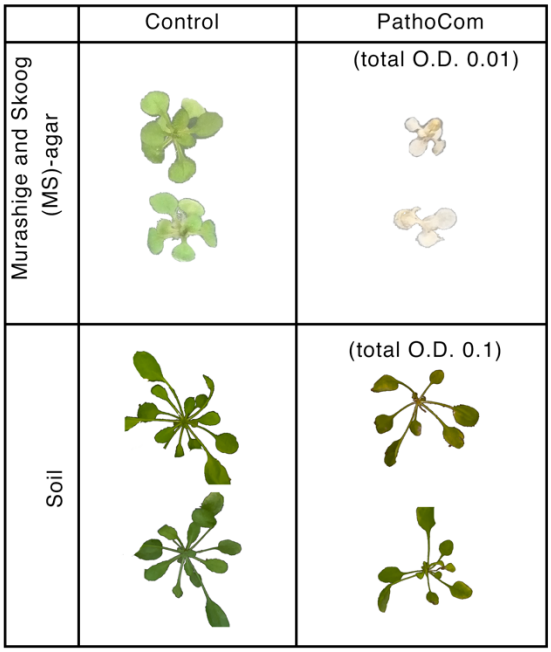


Figure S4. Illustrative photos of control- and PathoCom-treated plants, grown in either MS-agar (sterile) or soil (unsterile). In both systems, the genotype Ey15-2 was used. For the MS-agar system, photos were taken 3-dpi, for the soil system 14-dpi. Sizes of plants are comparable within each system, but not between. Because images in the soil system were taken and parsed by pot automatically by a high-throughput imaging pipeline, some plant images were cropped.

Table S2. A. Analysis of similarities (ANOSIM) based on Bray-Curtis distances for compositions of the 14 barcoded bacteria in treated hosts. The analysis was constrained by the host genotype in each experiment batch (exp) to estimate its effect on the explained variance. **B.** Multilevel pairwise comparison of barcoded bacteria compositions for the different *A. thaliana* genotypes, using adonis based on Bray-Curtis distances. Data derived from one representative experiment (October).

A

Treatment	R2	Pr(>f)	Experiment
PathoCom	0.0630	**0.0175**	August
	0.1792	**0.0005**	October
CommenCom	0.0622	0.0615	August
	0.1761	**0.0005**	October
MixedCom	0.0538	**0.0265**	August
	0.0951	**0.0005**	October

In bold, statistically significant relationships ($P \leq 0.05$).

B

Treatment	R2	Pr(>f)	Genotype1	Genotype2
PathoCom	0.1491	0.005	Ey15-2	HE-1
	0.2561	0.001	Ey15-2	Lu3-30
	0.1236	0.01	Ey15-2	Schl-7
	0.0804	0.044	HE-1	Kus3-1
	0.1678	0.001	HE-1	Lu3-30
	0.1313	0.001	Kus3-1	Lu3-30
	0.0757	0.049	Kus3-1	Schl-7
	0.0865	0.002	Lu3-30	Schl-7
	0.1986	0.001	Lu3-30	Tue-Wal-2
CommenCom	0.0813	0.035	Ey15-2	Lu3-30
	0.1285	0.028	Ey15-2	Tue-Wal-2
	0.1330	0.004	HE-1	Lu3-30
	0.2450	0.001	HE-1	Tue-Wal-2
	0.2197	0.001	Kus3-1	Lu3-30
	0.1305	0.005	Kus3-1	Schl-7
	0.1709	0.007	Kus3-1	Tue-Wal-2
	0.4401	0.001	Lu3-30	Tue-Wal-2
	0.3788	0.001	Schl-7	Tue-Wal-2
Mixedcom	0.0825	0.012	Ey15-2	Lu3-30
	0.0818	0.01	Ey15-2	Tue-Wal-2
	0.0673	0.027	HE-1	Tue-Wal-2
	0.1660	0.001	Kus3-1	Lu3-30
	0.2283	0.001	Lu3-30	Tue-Wal-2

Only statistically significant comparisons are presented ($P \leq 0.05$).

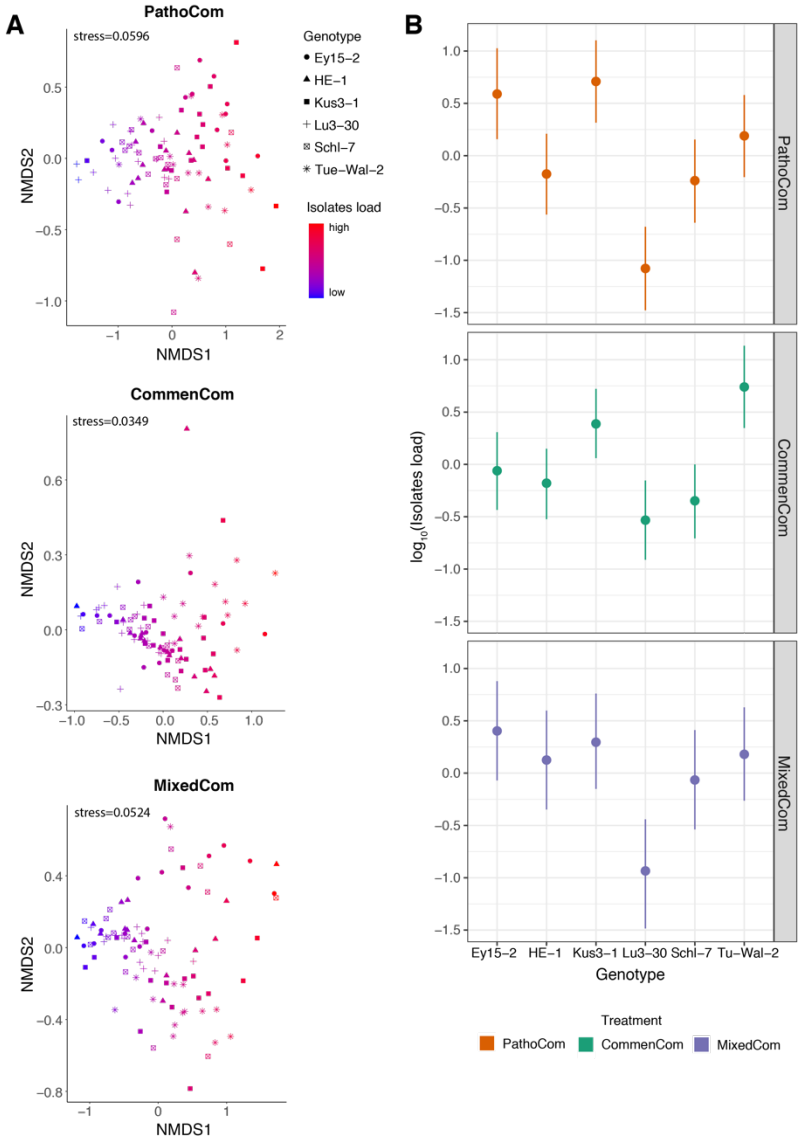


Figure S5. Comparison of composition and load of the 14 barcoded isolates on different *A. thaliana* genotypes. A. Nonmetric multidimensional scaling (NMDS) based on Bray-Curtis distances between six *A. thaliana* genotypes, in one representative experiment (October). Each synthetic community was analyzed separately. The abundance of all 14 barcoded isolates was considered, also among PathoCom and CommenCom to account for cross contaminations and technical distortions. Shapes denote the different genotypes, and bacterial load is indicated from blue to red. **B.** Isolate load of the six *A. thaliana* genotypes, among the three synthetic communities. Isolate load was defined as the cumulative abundance of all barcoded isolates that composed a synthetic community. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter, following the model $\log_{10}$(isolates load) ~ genotype + experiment + error.

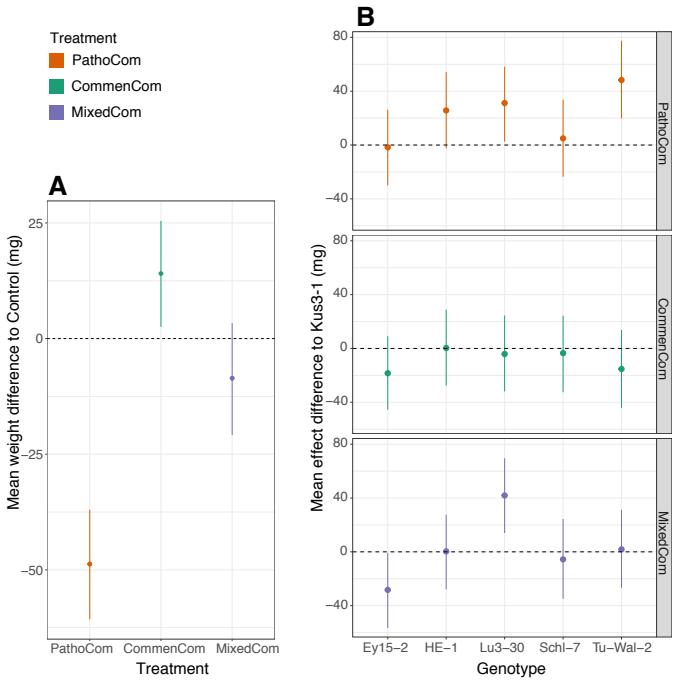


Figure S6. Effects of treatment and treatment-by-genotype on fresh rosette weight. Both effects were assessed using the model weight ~ treatment * genotype + experiment + error. **A.** Mean weight difference of plants infected with each of the three synthetic communities relative to control - i.e., the treatment coefficients. **B.** Mean treatment effect differences between the six *A. thaliana* genotypes used in this study - i.e., the treatment * genotype coefficients. Kus3-1 was randomly selected as a reference; dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter.

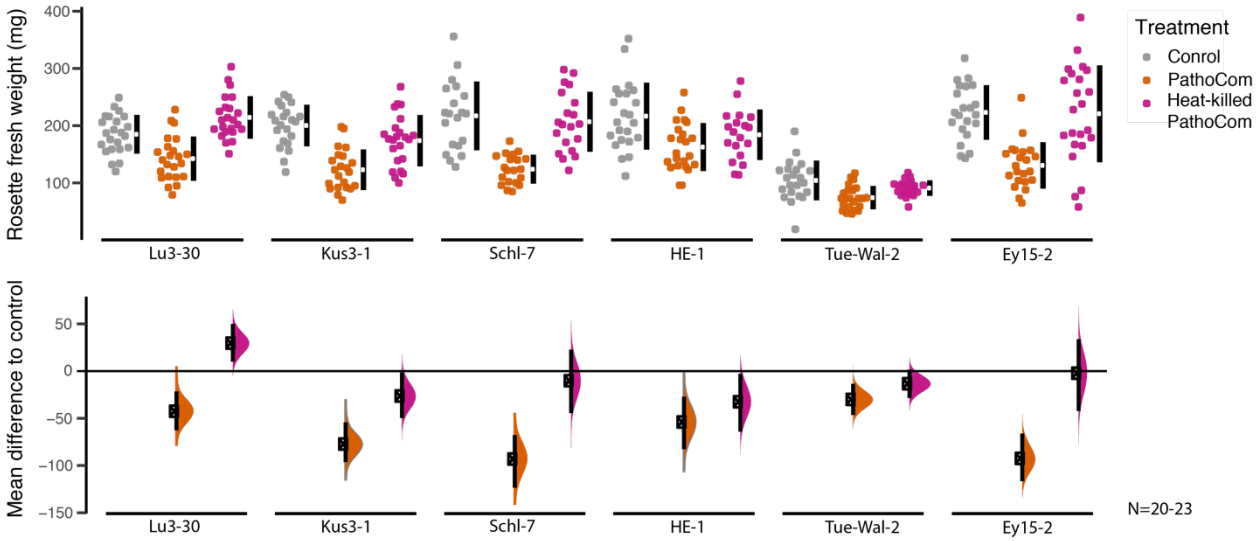


Figure S7. Fresh rosette weight of plants treated with Control, PathoCom or heat-killed PathoCom. Each of the six *A. thaliana* genotypes used in this study was treated with control, PathoCom and heat-killed PathoCom inoculum, and fresh rosette weight was measured 12 dpi. The top panel presents the raw data, the breaks in the vertical black lines denote the mean value of each group, and the vertical lines themselves indicate standard deviation. The lower panel presents the mean differences to control, plotted as bootstrap sampling [26,27], indicating the distribution of effect sizes that are compatible with the data. 95% confidence intervals are indicated by the black vertical bars.

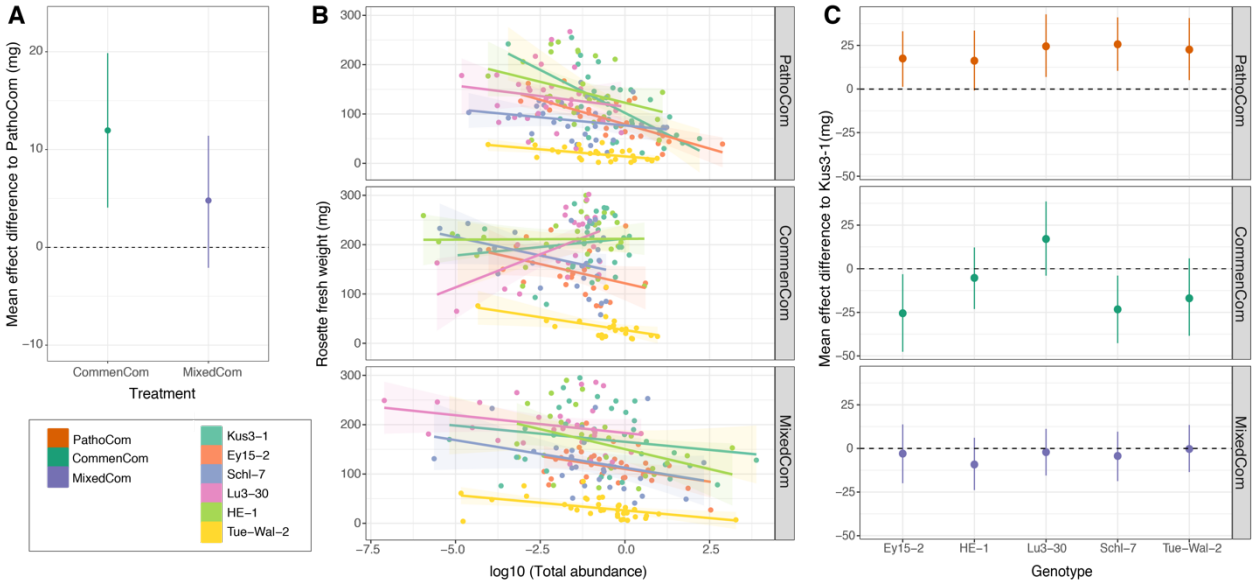


Figure S8. Effect of total load on weight, per treatment and genotype. A. Mean slope difference of the three synthetic communities. The slope difference indicates the effect of the treatment on the correlation between weight and isolate load - i.e. treatment * $\log_{10}$(cumulative isolate load) - following the model weight ~ treatment * $\log_{10}$(cumulative isolate load) + genotype + experiment + error. PathoCom was selected as a reference. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter. Related to Fig 3B. **B.** Correlation of $\log_{10}$(cumulative isolate load) with rosette fresh weight, for each of the genotypes within each of the three synthetic communities. Shaded areas indicate 95% confidence intervals of the correlation. Color codes in the bottom left box, on the right. **C.** Mean slope difference of the six *A. thaliana* genotypes used in this study. The slope difference indicates the effect of the genotype on the correlation between weight and isolate load - i.e. genotype * $\log_{10}$(cumulative isolate load) - following the model weight ~ genotype * $\log_{10}$(cumulative isolate load) + experiment + error. Each treatment was analyzed individually, thus the model was utilized for each treatment separately. Kus3-1 was randomly selected as a reference. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter. Related to panel B. n=170 for PathoCom, n=151 for CommenCom, and n=182 for MixedCom. n=77-94 for the six *A. thaliana* genotypes.

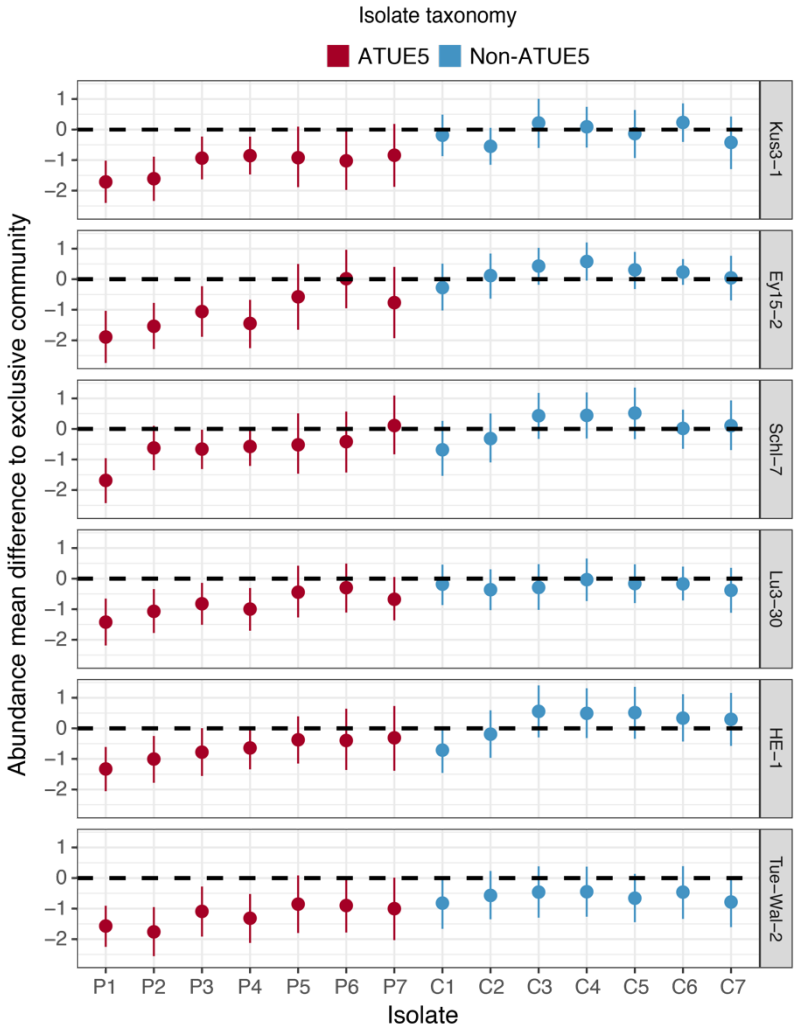


Figure S9. Effect of host genotype on abundance changes of the 14 barcoded isolates in MixedCom, when compared to their exclusive community (i.e., PathoCom for ATUE5 and CommenCom for non-ATUE5). Abundance effect mean differences were estimated with the model $\log_{10}$(isolate load) ~ genotype * treatment * experiment + error for each individual strain. Thus, the genotype * treatment coefficient was estimated per each barcoded isolate. Dots indicate medians, and vertical lines 95% credible intervals of the fitted parameter.

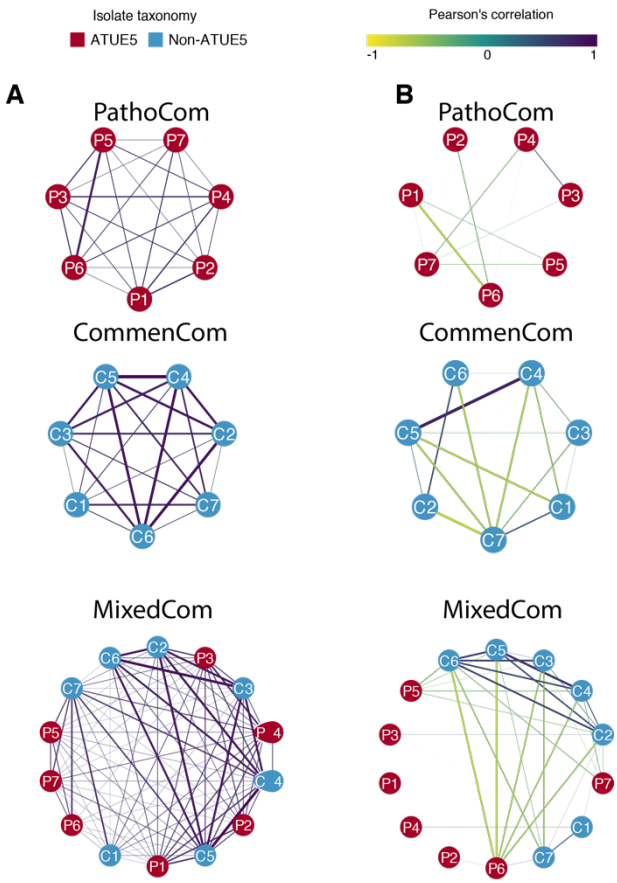


Figure S10. Correlation networks of barcoded bacteria. **A.** Correlation networks of absolute abundance in PathoCom, CommenCom and MixedCom. **B.** Correlation networks of relative abundance in PathoCom and CommenCom. Strengths of negative and positive correlations are indicated from yellow to purple. Boldness of lines is related to the strength of correlation, and only correlations > |±0.2| are shown. Node colors indicate the isolate classification: ATUE5 or non-ATUE5.

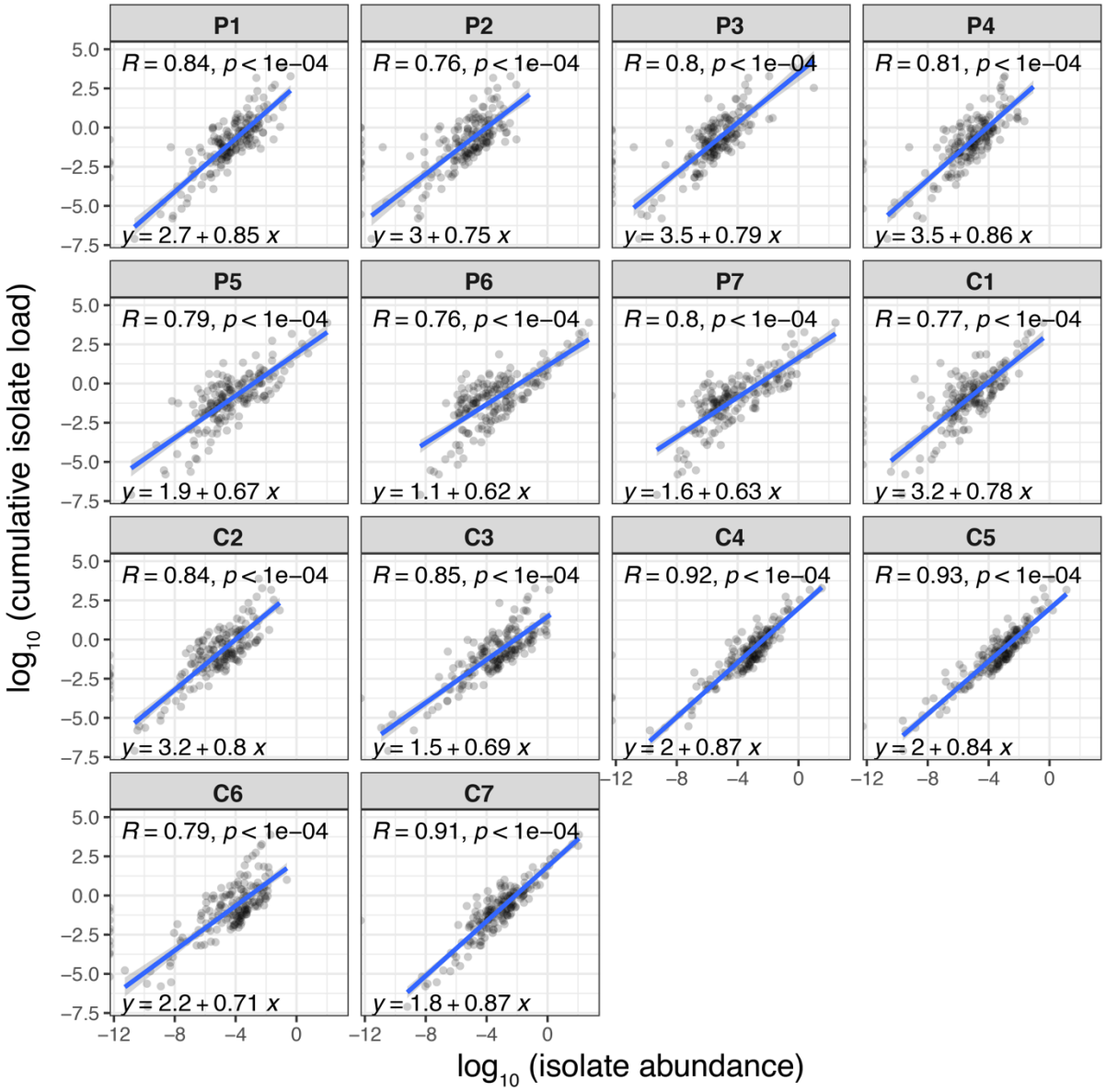


Figure S11. Correlations between the absolute abundance of each isolate and the cumulative bacterial abundance in MixedCom. Each panel represents an individual isolate. Pearson correlation (R) and p-value (p) are stated at the top, and the matching linear equation at the bottom of each panel. Shaded areas indicate 95% confidence intervals of the correlation curve. N=

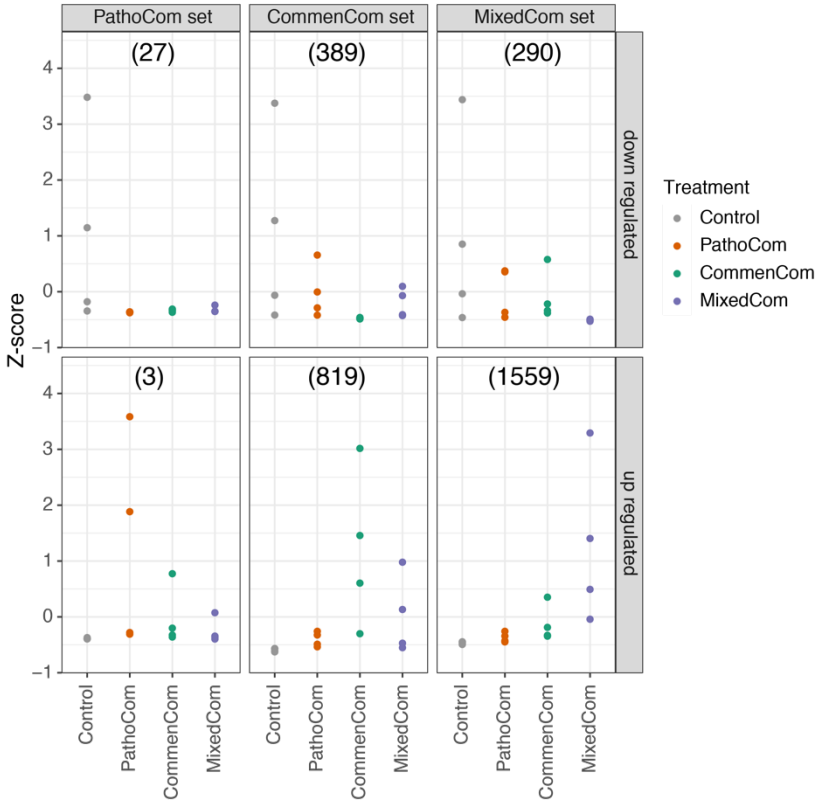


Figure S12. Comparison of PathoCom, CommonCom and MixedCom DEGs across treatments. The average z-score is presented for each sample. Downregulated and upregulated DEGs were analyzed separately. In brackets - the number of DEGs in each category. n=4.

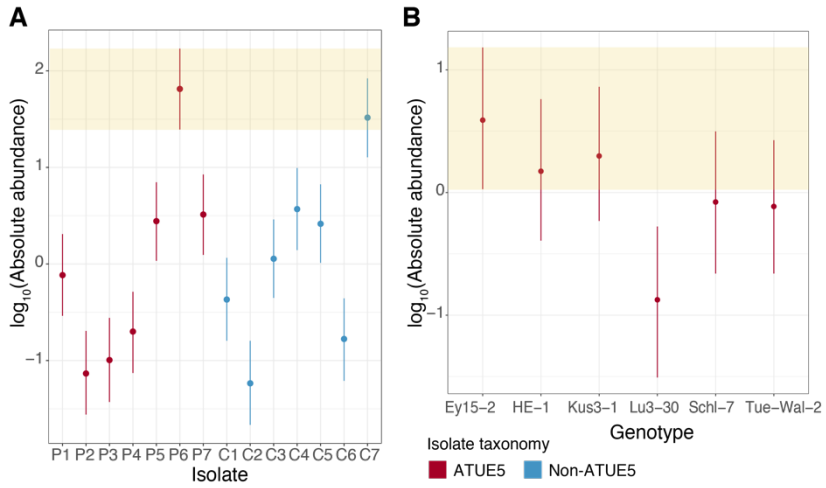


Figure S13. The abundance of P6 in MixedCom-infected hosts. A. Abundance of P6 compared with the other 13 barcoded bacteria in MixedCom-infected hosts, for all host genotypes. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter, following the model $\log_{10}$(isolate load) ~ isolate * experiment + error. Shaded area denotes the 95% credible intervals of the isolate P6. **B.** The abundance of P6 in MixedCom-infected hosts, compared between the six *A. thaliana* genotypes used. Dots indicate the medians, and vertical lines 95% credible intervals of the fitted parameter, following the model $\log_{10}$(isolate load) ~ genotype * experiment + error. Shaded area denotes the 95% credible intervals of the host genotype Ey15-2.

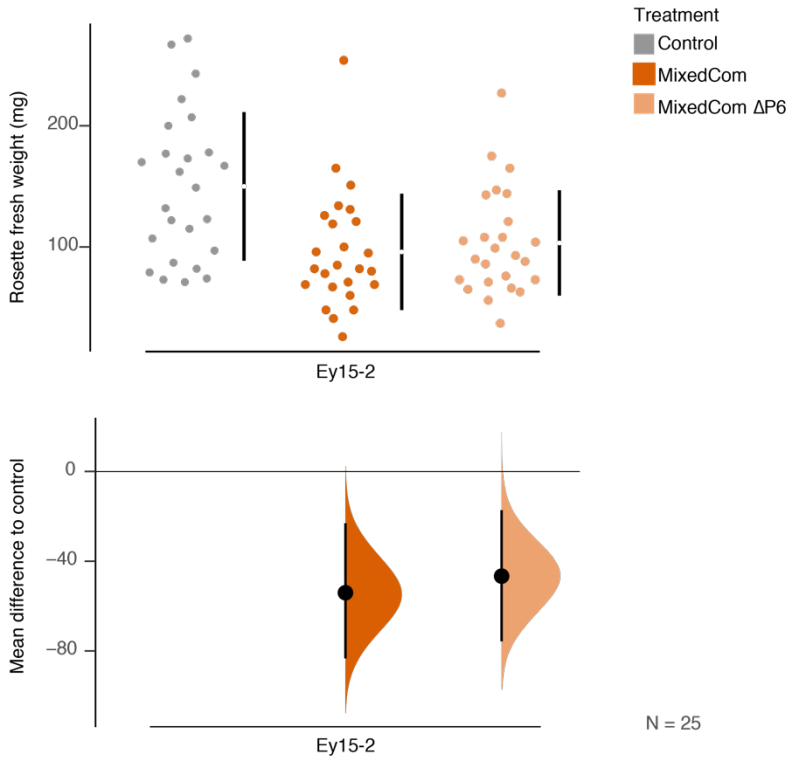


Figure S14. Fresh rosette weight of Ey15-2 plants treated with Control, PathoCom or PathoCom without P6 (PathoCom ΔP6). Fresh rosette weight was measured 12 dpi. The top panel presents the raw data, the breaks in the vertical black lines denote the mean value of each group, and the vertical lines themselves indicate standard deviation. The lower panel presents the mean differences to control, plotted as bootstrap sampling [26,27], indicating the distribution of effect sizes that are compatible with the data. 95% confidence intervals are indicated by the black vertical bars. n=25.

1279

1280 **Figure S15. Illustration of barcodes design.** Two single-stranded oligos were synthesized: 'Bar1' and

1281 'Bar2'. N symbolizes random nucleotides.

Commensal *Pseudomonas* protect *Arabidopsis thaliana* from a coexisting pathogen via multiple taxonomy-dependent mechanisms

Or Shalev[1], Haim Ashkenazy[1], Manuela Neumann[1], Ilja Bezrukov[1], Detlef Weigel[1,*]

[1]Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

[*]Corresponding author weigel@tue.mpg.de (D.W.)

Keywords: *Arabidopsis thaliana*, *Pseudomonas*, wild populations, microbial communities, plant health

Abstract

Plants are protected from pathogens not only by their immune arsenal, but often also by colonizing commensal microbes. In a recent survey of plant-colonizing *Pseudomonas* strains, a cryptically pathogenic lineage was reported to dominate wild *Arabidopsis thaliana* populations. While the wild plant hosts seemed to be asymptomatic, infection trials in the laboratory demonstrated that members of this lineage can have a negative impact on plant growth. In the wild, this pathogenic lineage coexists with commensal *Pseudomonas* strains that can mitigate the fitness penalties imposed by the pathogens in the laboratory. Here we address (i) how common protection is among wild commensal *Pseudomonas* strains, (ii) how taxon-specific such protective ability within the *Pseudomonas* genus, and (iii) what the underlying bacterial genes and mechanisms for protection are. To do so, we systemically co-infected *A. thaliana* plants with an individual *Pseudomonas* pathogen and each of ninety-nine *Pseudomonas* commensals. We found plant protection to be a common function among non-pathogenic *Pseudomonas* taxa. While enriched in one specific lineage, there is also substantial variation in the protective ability among isolates of this lineage. These functional differences do not align with core-genome phylogenies, suggesting repeated gene inactivation or loss as causal. Using genome-wide association, we discovered that different bacterial genes are linked to plant protection in each lineage. We validated a protective role of several lineage-specific genes by gene inactivation, highlighting iron acquisition and biofilm formation as prominent mechanisms of plant protection

in this *Pseudomonas* lineage. Collectively, our work illustrates the importance of functional redundancy in plant protective traits across an important group of commensal bacteria.

Introduction

The health of a plant depends to a large extent on its resident microbiota. The effect of individual microbes on plant health has been extensively investigated since the dawn of phytopathology[1], mainly focusing on pathogens, considering their profound effect on global agriculture and food supply [2]. The ability of phytopathogens to overpopulate plants, leading to disease, is reflected in the reciprocal ability of the plant to recognize and control these pathogens [3]. Nonetheless, in recent years there has been an increasing realization that the plant relies not only on its immune system, but also on other resident microbes [4–7].

For example, in natural settings, the health of the ephemeral plant *Arabidopsis thaliana* is associated with the presence of several bacterial members, preventing the onset of filamentous microbes driven diseases [5]. Similar patterns have also been found in controlled settings, with suppression of bacterial pathogens by other bacteria [8,9]. These protective agents can have several modes of action, including (i) activation of systemic defences that spread throughout the plant [10], (ii) outcompeting pathogens over nutrients [11,12], and (iii) direct antibiosis [13,14]. These mechanisms are non-exclusive and can operate simultaneously.

These studies have greatly advanced our understanding of protective interactions in plants, even if the investigated pathogens and protective microbes do not always coexist in the wild. An exception is a recent study [15], in which individual members of native tomato microbiomes were tested for suppression of the soil-borne pathogen *Ralstonia solanacearum*. These authors focused on siderophore production and competition for iron as a known mechanism for microbe-microbe competition, revealing a link to pathogen inhibition in the plant rhizosphere. This study [15] also highlighted the insights that could be gained from investigating pathogen protection in a phylogenetic framework.

Recently, Karasov et al. [16] conducted a large scale survey of *Pseudomonas* in south west Germany, in which 1,524 isolates were sampled from wild *A. thaliana* plants. One pathogenic lineage dominated this collection, although other commensal strains were found as well. By employing synthetic communities of commensal and pathogenic *Pseudomonas* from this collection, we previously revealed that commensal strains protected *A. thaliana* from pathogenic *Pseudomonas* strain that co-occur in the wild [17]. In a related study, commensal

Pseudomonas was found to often outcompete pathogenic *Pseudomonas syringae* isolates sampled from the same plant [18].

These studies point to the importance of understanding interactions among wild *Pseudomonas* strains in maintaining plant health, and how this might be enhanced by an understanding of the underlying genetic mechanisms. Here, we leveraged a local *Pseudomonas* collection [16] to examine the extent of protection conferred by commensal *Pseudomonas* against co-existing pathogenic *Pseudomonas*, and to discover some of the underlying mechanisms. Using a high throughput image-based assay, we monitored outcomes of systematic co-infections of a diverse set of commensal *Pseudomonas* with a focal *Pseudomonas* pathogen, covering the entire phylogeny of commensal *Pseudomonas* isolates in this collection. We found that protection by commensals was a common feature, although it was enriched in a specific taxon. We discovered bacterial genes for plant protection using genome-wide association and comparative genomics. Using knockout mutants, we experimentally validated the role of several candidate genes in plant protection, establishing a link to iron uptake and biofilm formation in the mitigation of phyllosphere pathogens.

Results

Systemic co-infections of commensal *Pseudomonas* with an individual pathogen

To examine the ability of commensal *Pseudomonas* strains to protect host plants from members of the pathogenic *Pseudomonas* lineage, we made use of a local isolate collection [16]. We henceforth refer to an operational taxonomic unit (OTU) as reported in that study as "ATUE" (isolates from Around TUEbingen), and following previous findings [16,17], we refer to the lineage ATUE5 as pathogenic, and to all non-ATUE5 lineages as commensals.

We grew plants on MS agar and monitored plant growth and health by extracting the number of green pixels from images over time (illustration in **Figure 1A**). Green pixel count and rosette fresh weight were strongly correlated (**Figure S1**; R^2=0.92, Pval < 2.2e-16), validating the use of green pixels as a proxy for the plant biomass.

We use a conservative definition of plant protection, considering protective strains as candidates leading to normal plant growth (comparable with uninfected plants) in the presence of a pathogen. To estimate how common the ability of non-ATUE5 strains to protect against the

impact of pathogenic *Pseudomonas* is, we infected plants with a panel of non-ATUE5 isolates in the presence of an common ATUE5 representative, strain 'p4.C9' (hereafter will be also referred to as 'pathogen'). The pathogen was chosen because it is dominant over other ATUE5 strains but at the same time highly susceptible to the presence of non-ATUE5 strains, a phenomenon that correlated with plant protection [17]. We excluded highly similar isolates (Jaccard distance ≥ 0.99 in gene content). From a total of 151 non-ATUE5 strains in our local collection [16], we initially chose a subset of 127 non-ATUE5 isolates, from which we were able to revive 99 isolates (**Figure 1B; Table S1**). We included three isolates from the pathogenic ATUE5 clade as control (**Table S1**), resulting in 102 strains that were tested in co-infections with the pathogen p4.C9.

One of the 102 strains - p5.F2 - was known to suppress ATUE5 strains inside plants [17], and was used as 'protective control'. To confirm that protection was due to bacterial activity and not merely a host response to the inoculum (e.g., PAMP-triggered immunity), we carried out infections with a heat-killed protective control. As expected, co-infection with the protective strain resulted in normal plant growth, while treatment with the pathogen by itself or co-inoculation with the heat-killed protective strain impaired growth (**Figure 1C**; **Figure S2**; mean growth [Δ7dpi-0dpi green pixels]: control 29,227 [15,721, 40,662], pathogen 3,328 [-10,350, 14,759], pathogen + heat-killed protective -5,222 [-16,410, 6,211] and protective 29,585 [16,323, 40,812], at 95% confidence interval).

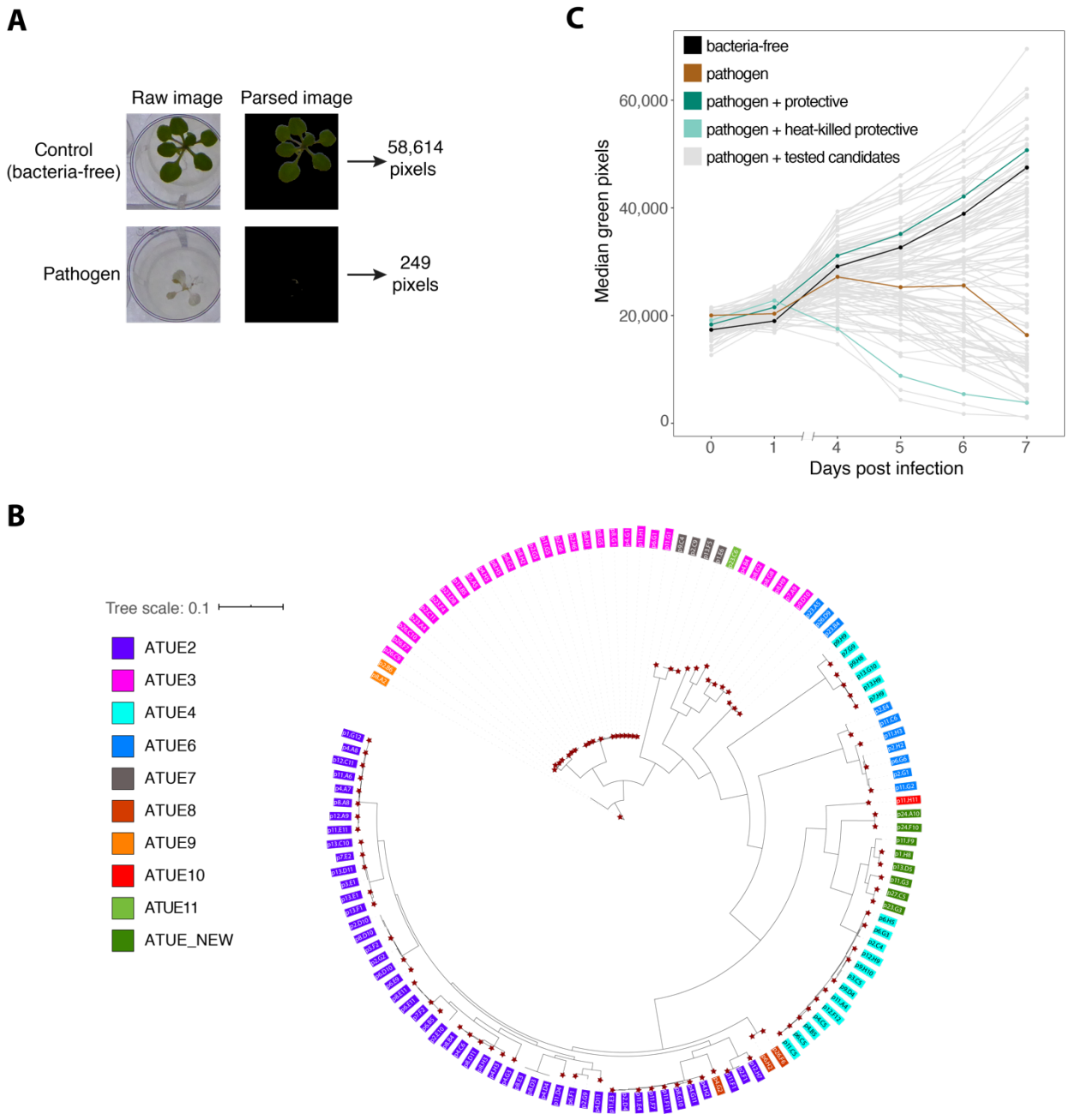


Fig 1. Panel of potentially protective *Pseudomonas*, and experimental design. A. Illustration of image processing to enumerate plant green pixels, approximating plant biomass. **B.** Phylogenetic tree of 127 representative non-ATUE5 strains (i.e., putative commensals) sampled from south-west Germany [16]. All other non-ATUE5 isolates in this collection are represented in this core collection by a strain with which they shared ≥99% of genes (Jaccard distance). Colors indicate the ATUE group (as previously determined[16]), and asterisks the 99 strains used here. **C.** Daily median of plant green pixels among the different treatments. Control treatments: Bacteria-free buffer, pathogen only, and co-infections of

pathogen with the protective strain and the heat-killed protective strain. Plant growth was measured daily by imaging. 8 replicates per treatment.

Plant protection is ubiquitous, but also lineage-specific

The co-inoculations of non-ATUE5 strains with the ATUE5 pathogen yielded a span of results, ranging from reduced to enhanced growth in comparison to uninfected plants, with protection being common (**Figure 2A)**.

Protection was unevenly distributed among the three highly sampled ATUE groups - ATUE2, ATUE3 and ATUE4 - with ATUE2 being the most protective mean growth (Δ[7 dpi]-[0 dpi] green pixels, 95% CI): ATUE2 22,231 [19,522, 24,908], ATUE3 14,120 [10,780, 17,389], ATUE4 10,868 [6,844, 14,937], and bacteria-free 27,069 [19,882, 34,264]) (**Figure 2B**).

Within each ATUE group, there was considerable variation in protective ability, even among strains with highly similar genomes (**Figure 2C**). In some cases, sister strains had opposite activities,: one providing protection and the other having no effect (e.g., ATUE2 strains p11.F1 and p12.H7; mean growth of p11.F1 25,554 pixels [8,303, 42,567] and p12.H7 -28,914 pixels [-45,789, -11,837]).

Even within ATUE5, some strains had protective ability. We had chosen the three ATUE5 strains because in a previous set of experiments they were less competitive than our focal ATUE5 pathogen p4.C9 [17]. Surprisingly, two of them also mitigated the pathogen effect, causing normal plant growth. This further exemplifies the differential functions within ATUE groups (**Figure 2C**).

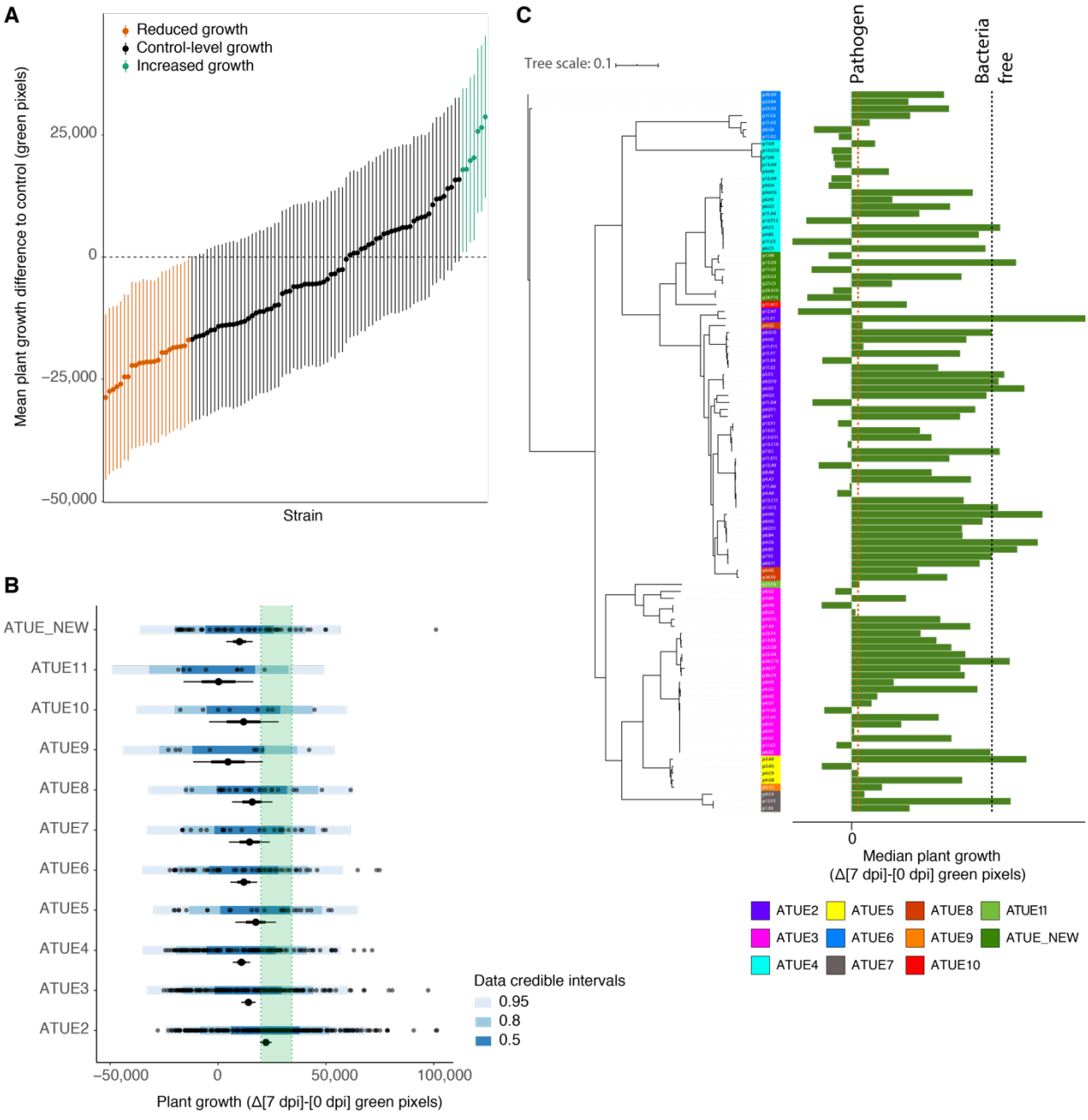


Fig 2. Protection by different *Pseudomonas* strains is common, and enriched in the ATUE2 lineage. **A.** Mean plant growth difference to control, after co-infection with different *Pseudomonas* strains and the focal p4.C9 pathogen. Growth was measured as the change in green pixels between the day of infection until day 7 days later. Vertical lines indicate 95% credible intervals, and dots indicate the median. The dashed horizontal line signifies the baseline, average growth after bacteria-free treatment (i.e., negative control). 8 replicates each. **B.** Plant growth after co-infections, binned by ATUE group. Growth was measured as the change in green pixels between the day of infection until day 7 days later. In each ATUE group, raw data for individual replicates are shown with dots; the overlain shades of blue indicate the data

credible interval, as presented on the bottom right. The mean growth for each ATUE group is also shown; dot indicates for the median, thin horizontal line for 95% and thick line for 67% credible interval. For the number of strains in each ATUE group, see **Table S1**. **C.** Median plant growth after co-infections, ordered by phylogeny. Colors indicate ATUE groups [16]. Medians of plant growth in mono-association with pathogen p4.C9 or without infection indicated by dashed vertical lines.

ATUE2 protective genes are lineage-specific

That functional variation could not be explained by phylogeny (considering both topology and branch length) suggested that variation in gene content, possibly due to horizontal gene transfer, was causal for protection. We therefore examined associations between the presence/absence (P/A) of gene orthology groups with plant protection.

There were 32,753 gene orthology groups that were not shared by all ATUE strains used for coinfections, but were found in more than one strain. We examined the correlation between plant growth and the presence of each of the individual orthology groups using treeWAS [19], a tool for genome-wide association studies (GWAS) in bacteria. Because treeWAS accounts for population structure, it will remove true positives that are highly correlated with population structure [20]. We therefore also used Spearman's rank correlation coefficient (SC) to search for genes that were correlated with protection. SC was also useful to differentiate a global from a taxon-specific signal. To this end, we calculated SCs separately among the highly sampled ATUE2, ATUE3 and ATUE4 groups. For the SC test, we only considered the difference in median green pixels between 0 and 7 dpi. For treeWAS, we used four different plant growth metrics (**Table S2;** Methods). The extent of agreement between the four phenotypic metrics was used as another indicator for robustness of each association (**Table S2**). We overlapped the results from all four analyses and removed genes with a negative SC (SC<0), leaving us with 14 strong candidates for plant protection (**Table S2**). The nine genes with the highest SC values (Rho 0.37-0.46) were unique to ATUE2 (**Figure 3A**).

The existence of a gene set unique to one taxon implies that plant protection by *Pseudomonas* is not driven by a global mechanism, but rather by taxon-specific. To validate this assumption, we ran the same GWAS analysis on each of the main ATUE groups separately, i.e., independently analyzed the subsets of strains related to ATUE2, ATUE3 and ATUE4. In total, 95 genes with positive SC were significantly associated with plant protection: 14 genes generated by the full set of strains, 46 genes by ATUE4, 35 genes by ATUE3 and no significant hits were

found in the ATUE2 dataset alone. Except for a single gene, we found no overlap among hits of the different ATUE datasets, providing further evidence for taxon-specificity (**Figure S3**).

Around 50% of all protection-associated genes (47/95) were annotated as 'Uncharacterized protein' or had no hit in either TrEMBL or Swiss-Prot databases (**Table S2**). Out of the well annotated genes, we observed a few noteworthy putative functions related to direct microbe-microbe interactions: three iron-uptake-related genes unique to ATUE2 (GC00000450_7: 'TonB_C domain-containing protein', GC00000032_87: 'Putative iron(III) dicitrate sensor protein FecR' and GC00000050_54: 'Probable RNA polymerase sigma factor FecI'), a gene related to resistance to antimicrobial peptides presented in ATUE3 (GC00000089_55: 'UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase' [*arnB*][21]), and an antitoxin (GC00007392_r1_1: 'Antitoxin FitA') alongside mobility-related genes (GC00002204_5: 'Twitching motility protein PilT', GC00000715_r1_r1_2: 'Pilus assembly protein PilW') found in ATUE4. These putative functions indicate the diversity of protective mechanisms among commensal *Pseudomonas*.

We observed that some significant genes had identical SC values in the examined ATUE group. Furthermore, these genes were all presented together in the same strains. Thus, we hypothesized that they may be in genetic linkage, i.e., forming genomic islands. We tested for chromosomal proximity between all the significant genes, and found clusters of proximal genes, confirming our hypothesis (**Table S2**; details in Methods). In these clusters, we found evidence for the horizontal gene transfer elements, for example phage elements combined with bacterial genes (proximity clusters 3 and 4 in **Table S2**).

Cumulatively, these results portray a scenario in which plant protection by commensal *Pseudomonas* is driven by multiple, clade-specific mechanisms which were horizontally transferred.

As described above, of the 14 protection-associated genes found among all strains, nine were unique to ATUE2 (**Figure 3A**). Albeit we found no significant hits when we analyzed ATUE2 subset separately, strains holding these nine genes were more protective not only when compared among all strains, but also within ATUE2 (Plant mean growth for strains with zero genes in non-ATUE2 was 10,791 [7,541, 14,078], with zero genes in ATUE2 was 12,780 [6,891, 18,821] and with at least one gene in ATUE2 was 29,299 [22,567, 35,983], at 95% confidence interval) (**Figure 3B-C**). Moreover, we found that the effect of these nine genes was additive within ATUE2, thus plant growth was associated with the number of genes presented in a given

strain (**Figure S4**; R^2=0.55, Pval=0.0004). These results highlight the importance of these nine genes in plant protection both among all strains, and within ATUE2, despite the lack of statistical signal in the ATUE2 subset.

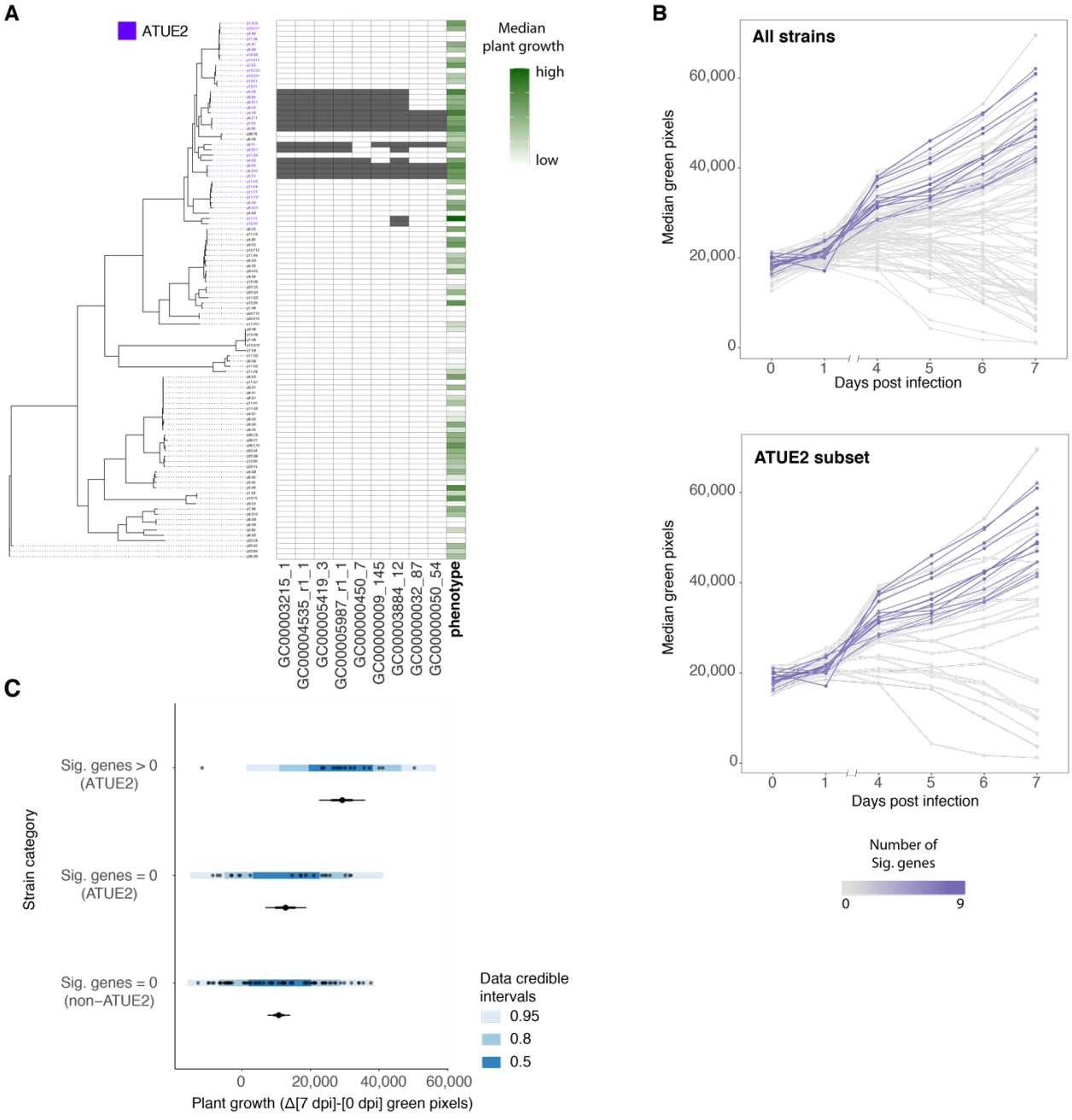


Fig 3. Genes with the highest association to protection are unique to ATUE2. A. Presence / absence variation of the nine genes with the highest association to protection. Strains are ordered by their phylogeny. Strains belonging to the ATUE2 group are indicated by magenta color. Gene presence is indicated by grey filling, and absence by white. Median plant growth was measured by the median change in green pixels between day 7 post infection to the day of infection, and is indicated by shades of green.

B. Daily median of plant green pixels among all strains (top panel) or the ATUE2 subset (bottom panel). Shades of magenta indicate the number of protective genes present in each strain, out of the nine gene candidates that are detailed in panel A. Plants were assessed daily by imaging. 8 replicates per treatment. **C.** Plant growth after co-infections with the pathogen p4.C9, binned by (i) the presence of at least one gene from the set of the nine protective genes (presented in panel A) in a given commensal strain and (ii) affiliation to the ATUE2 group. Plant growth was measured by the change in green pixels between day 7 post infection to the day of infection. In each group, raw data for individual replicates are shown with dots; the overlain shades of blue indicate the data credible interval, as presented on the bottom right. The mean growth for each group is also shown; dot indicates for the median, thin horizontal line for 95% and thick line for 67% credible interval. n = 8 per strain, while the number of strains differed among the categories as follows: zero genes in non-ATUE2 = 67, zero genes in ATUE2 = 20, and at least one gene in ATUE2 = 16.

Plant protection by ATUE2 is driven by iron acquisition and biofilm formation

We selected the nine ATUE2-unique candidates for gene deletion to validate their role in plant protection. The protective strain 'p5.F2' was used as a representative for the ATUE2 clade, therefore we deleted each of the nine gene candidates in this strain. Since two genes were found in chromosomal proximity (GC00000032_87 and GC00000050_54), we treated them as one functional unit and deleted the whole cluster. Thus, we deleted a total of eight loci in p5.F2, comprising nine genes.

Out of the eight knockout mutants, three lost their ability to mitigate the effect of the pathogen p4.C9 following co-infections (henceforth 'non-protective mutants'), while the wild type strain protected the plant as expected (**Figure 4A; Figure S5A**). These three non-protective mutants did not affect plant weight when tested individually, similarly to the wild type (**Figure S6A-B**), implying that the weight reduction after co-infections was due to the pathogen or to the interaction with the pathogen, rather than the knockout mutants themselves.

Out of these three non-protective mutants, two comprise the deletion of three genes that are annotated as iron-related (GC00000450_7: 'TonB_C domain-containing protein' [22,23], and a cluster of two genes - GC00000032_87: 'Putative iron(III) dicitrate sensor protein FecR' [24] and GC00000050_54: 'Probable RNA polymerase sigma factor FecI' [25]), while the third - ΔGC00003884_12 - is annotated as 'Uncharacterized protein' (**Table S2**). We noted an exceptional phenomenon in the tubes in which ΔGC00003884_12 was grown - formation of chunks, although bacteria were grown by shaking (Methods). Thus, we presumed that the

deleted gene in question is related to biofilm formation. To test this, we grew the WT and the mutated strain ΔGC00003884_12 overnight, and removed the tubes from shaking, placing them at room temperature for static incubation. After one hour, we noted sedimentation in the ΔGC00003884_12 tube (formation of clear liquid at the top and aggregates at the bottom) (**Figure S7**), unlike the WT strain which presented a homogenous opacity level. This phenotype implies that at least in the tested conditions, ΔGC00003884_12 has an augmented biofilm control in comparison to the WT strain, thus tends to form aggregates faster.

To validate the role of ΔGC00000032_87 // GC00000050_54 and ΔGC00000450_7 mutants in iron uptake, we performed an iron-deficiency growth assay *in-vitro*. The WT strain, the two iron knockout mutants and the pathogen were grown in LB with increasing levels of the iron chelator 2'2'-dipyridyl. We also tested the biofilm-related knockout mutant ΔGC00003884_12 in the same assay, but noted an exceptional growth curve shape for this strain, regardless of chelator levels (**Figure S8**), in agreement with the aggregates phenomenon we observed in the tubes (**Figure S7**). Consequently, we excluded the mutant ΔGC00003884_12 from the *in vitro* assay analysis. Both the knockout mutants ΔGC00000032_87 // GC00000050_54 and ΔGC00000450_7, and the pathogen, presented reduced growth in LB without chelator, when compared to the WT strain (**Figure 4B**; **Figure S8; Figure S9A;** mean difference to control: pathogen -4,424 AUC [-5,299, -3,551], ΔGC00000450_7 -3,526 AUC [-4,412, -2,634] and ΔGC00000032_87 // GC00000050_54 -2,469 AUC [-3,337, -1,587], with 95% confidence intervals in brackets). This confirms that both loci have a role in the growth of the WT strain p5.F2, regardless of iron availability. We then examined how the growth of focal mutants is affected by increasing iron chelator concentration, i.e. by decreasing iron levels. While the increasing chelator led to a similar reduction in growth between the WT and the knockout mutant ΔGC00000032_87 // GC00000050_54, both the pathogen and the mutant ΔGC00000450_7 presented a sharper growth reduction - hence were more sensitive to iron deficiency (**Figure 4B**; **Figure S8; Figure S9B;** mean slope difference to control: pathogen -5.6 AUC [-10.5, -0.7], ΔGC00000450_7 -10.1 AUC [-15.1, -5.3] and ΔGC00000032_87 // GC00000050_54 -0.4 AUC [-5.3, 4.3], with 95% confidence intervals in brackets). These findings provide further evidence for the involvement of iron acquisition in plant protection by ATUE2 strains, and specifically imply that ATUE2 strains protect the plant by outcompeting the pathogen over iron.

Taken together, these results imply that ATUE2 can protect *A. thaliana* from the coexisting pathogenic ATUE5 via mechanisms related to iron uptake, and seemingly also to

biofilm formation. Do these mechanisms act specifically against the local ATUE5 or a wider pathogen range? To test this, we co-infected the protective WT and the three non-protective mutants with *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*). *Pst* was not observed in the local *A. thaliana* populations from which ATUE2 was sampled [16], hence it is not assumed to coexist with ATUE2 in these populations. Furthermore, it is a different species, as ATUE5 is classified as *Pseudomonas viridiflava [16]*. The WT was protective against *Pst*, leading to control-level growth (**Figure 4C; Figure S5B**). This reflects that ATUE2 protective ability is not restricted to the pathogenic lineage ATUE5, but rather works against a wider range of pathogenic *Pseudomonas*. However, the exact protective mechanism against the two pathogens seems to differ, as only two of the three tested mutants exhibited a mild reduction in protection (ΔGC00000450_7 and ΔGC00003884_12) (**Figure 4C; Figure S5B**); therefore while not resulting in control-level growth, their co-infections led to a higher weight in comparison to *Pst* alone (mean difference to *Pst* alone: ΔGC00000450_7 28,466 pixels [11,343, 45,597], ΔGC00000032_87 // GC00000050_54 33,790 pixels [16,677, 50,862] and ΔGC00003884_12 31,687 pixels[14,814, 48,263], with 95% confidence intervals in brackets). This contrasts the co-infections with ATUE5 pathogen, in which co-infections with each of the three mutants resulted in pathogen-level weight (mean difference to p4.C9 alone: ΔGC00000450_7 -4,949 pixels [-23,916, 15,161], ΔGC00000032_87 // GC00000050_54 3,133 pixels [-15,959, 21,910] and ΔGC00003884_12 -14,601 pixels[-33,678, 4,581], with 95% confidence intervals in brackets). These differences imply that distinct genes in ATUE2 cause protection against ATUE5 and *Pst*.

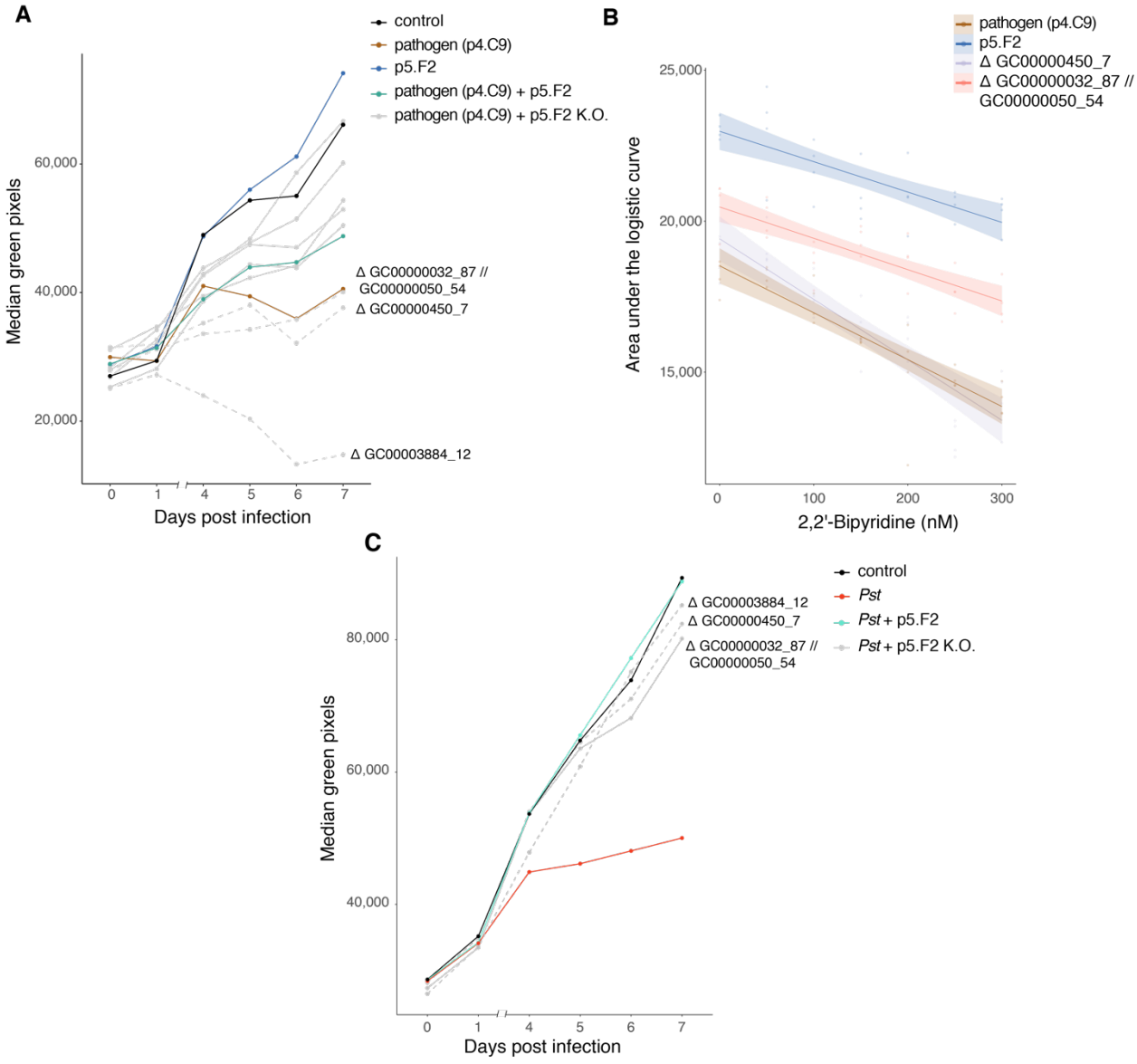


Fig 4. Three knockout mutants failed to confer protection, two associates with iron uptake. A. Daily median of plant green pixels after treatment with control, p4.C9 pathogen, the protective strain p5.F2, mixture of the pathogen and p5.F2, and mixture of the pathogen and each of eight p5.F2 knockout mutants ('p5.F2 K.O.'). Dashed lines note three treatments comprising p5.F2 knockout mutants that lost their ability to protect the plant ('Δ GC00000032_87 // GC00000050_54', 'Δ GC00000450_7' and 'Δ GC00003884_12'), as analyzed in Figure S5A. Plants were assessed daily by imaging. Experiments were performed two times with similar results. n=20. **B.** *In vitro* growth of the pathogen, p5.F2 and two knockout mutants ('Δ GC00000032_87 // GC00000050_54' and 'Δ GC00000450_7') as a function of 2,2'-bipyridine concentration. Each bacterial strain was grown in LB supplemented with seven different 2,2'-bipyridine concentrations (0, 50, 100, 150, 200, 250 and 300 nM), and OD_{600} was monitored for 10 hours (more details in Figure S5 and Methods). The logistic area under the growth curve was extracted as a proxy for bacterial

growth. Shaded area indicates 95% confidence intervals of the regression curve. n=4 for each bacterial strain grown in the corresponding 2,2'-bipyridine concentration. **C.** Daily median of plant green pixels after treatment with control, *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*), mixture of *Pst* and p5.F2, and mixture of *Pst* and each of the three tested p5.F2 knockout mutants ('p5.F2 K.O.') - the three mutants that lost their ability to protect the plant against the ATUE5 pathogen. Dashed lines note two treatments comprising p5.F2 knockout mutants that lost their ability to protect the plant ('Δ GC00000032_87 // GC00000050_54' and 'Δ GC00000450_7'), as analyzed in Figure S5B. Plants were assessed daily by imaging. Experiments were performed two times with similar results. n=20.

Discussion

In this work, we systematically tested the ability of non-pathogenic *Pseudomona*s to protect the plant from a pathogenic *Pseudomona*s lineage, using a set of isolates sampled from the same spatio-temporal space. We found that plant protection is a common function among the non-pathogenic *Pseudomonas* taxa, while unraveling differences between highly-related strains. Our genome wide association study (GWAS) implied taxon-specific gene sets which associate with plant protection. We selected the most promising gene set for knockout mutants validation, demonstrating that iron acquisition and biofilm formation play a major role in the mitigation of phyllosphere pathogens by protective bacteria.

Our findings highlight the link between bacterial phylogeny to plant protectiveness. We document a *Pseudomonas* lineage enriched in protective isolates, alongside high variability within each bacterial lineage. Similar patterns were described in a recent study focused on bacterial mitigation of *Ralstonia solanacearum*, a tomato pathogen [15]. Such similarities imply the generality of these trends in plant-commensal-pathogen ecosystems, in particular when considering that different hosts and microbes were focused.

Our work confirms the role of iron uptake in plant protection, pointing out how commensal *Pseudomonas* outcompete a pathogen over this nutrient, leading to plant protection. These evidences stand in line with current knowledge about the importance of iron competition in microbial communities [15,26–29] and about the link between iron availability to plant pathogenesis [30,31]. The weight of bacterial competition over iron is driven by its biological importance and environmental scarceness as it is an essential micronutrient [32] with limited bioavailability [33], fueling diversification of iron uptake instruments [34]. Therefore, it is not surprising that iron competition dictates microbial interactions within the plant, which in turn can affect host health.

Beyond the scope of iron acquisition, we experimentally validated a link between bacterial biofilm formation to plant protection, which as far as we know have yet to be described. It is clear that bacterial aggregation affects bacterial activities, altering competition with other bacteria [35]. Therefore, it is somewhat trivial that biofilm control can influence the competition between commensal and pathogenic *Pseudomonas* in the plant. Considering our findings and this simple explanation, we suggest further investigation of the role of biofilm control in shaping plant microbial communities and health.

In this study, we utilized comparative genomics to pinpoint protective genes among various *Pseudomonas* commensal lineages. Our results imply that the underlying protective mechanism is clade-dependent, according to homology-based annotations. For example, iron uptake was associated with only a single taxonomic lineage among the eleven tested lineages. Considering that we examined interactions between commensals of the same genus to a single individual pathogen, we are surprised to find such a diversity of taxon-specific gene-sets, rather than a global set of protective genes. These findings hint at the complexity of competitive interactions among the whole microbiota, and perhaps note a signature of a coevolutionary arms race between pathogens and commensals.

An additional complexity we note is the differential pathogen-specificity of the same protective genes. We found differences when tested the same knockout mutants against two different pathogens: the geographically coexisting *Pseudomonas* pathogen (corresponding to *Pseudomonas viridiflava*) and the exogenous *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*). Although the tested protective strain conferred protection against both pathogens, the knockout mutants that lost protection against the native pathogen were largely protective against *Pst*. This points out that protective strains may act via various mechanisms, tailor-made against a wide repertoire of pathogens.

Albeit our study involved bacteria sampled from the same local survey [16], it is too reductionist to conclude the exact role of these specific bacteria in wild settings, as only pairwise interactions in gnotobiotic conditions were tested. However, the underlying protective mechanisms we revealed in the tested strains should mirror their functionality in nature to a certain extent. We tested only one pathogen against the whole commensal *Pseudomonas* panel. Hence, it is beyond the scope of this study to understand the pathogen range of the protective genes found in the GWAS. Consequently, we suggest to upscale the experimental effort taken here, following the general approach we implemented. Examining interactions between multiple

pathogens with multiple commensals would better manifest the importance of protective genes, evaluating their pathogen range. Such upscaling may require new methodological improvements to reach a higher throughput.

Together with recent studies [5,36,37], our work emphasizes the importance of commensal microbes in maintaining plant health. Ultimately, the approach we took in this study can advance rational microbiome design for the sake of disease prevention in agriculture [38]. Understanding the stabilizing factors in wild plant-microbiota systems can enable their mimicry in agricultural systems - the food source of the world. In the heart of such emerging trends [38–40] stands a safe bet for the human kind - relying on evolution to pave the way for stable anthropogenic ecosystems.

Methods

Plant material

The plant genotype Ey15-2 was used in this study. It was originally collected from Eyach (Germany), near the sites in which *Pseudomonas* were isolated [16], and was found to represent a genetic background common to this region [41]. Seeds were sterilized by overnight incubation at –80°C, followed by ethanol washes (shake seeds for 5-15min in solution containing 75% EtOH and 0.5% Triton-X-100, and then wash seeds with 95% EtOH and let them dry in a laminar flow hood). Seeds were then stratified in the dark at 4°C for seven days, on ½ MS-agar media including vitamins and MES buffer (Duchefa, M0255). Plants were grown in 1.8 mL ½ MS-agar in 24-well plates. All plants were grown in long days (16 h of light) at 23°C, in controlled percivals (Plant Climatics, CLF CU-36L5).

Plant protection assay

All tested bacteria were taken from a previous *Pseudomonas* collection [16], except for *Pseudomonas syringae* pv. *tomato* DC3000 which was taken from our common-bacteria stock. All bacterial treatments were diluted to a concentration of OD_{600}=0.01 per strain. Thus, in the co-infections the total bacterial concentration was OD_{600}=0.02, in a solution comprising a mixture of two strains (the pathogen and the candidate strain), while in treatments of individual strains the total concentration was OD_{600}=0.01. Bacteria-free control was 10 mM $MgSO_4$.

Bacterial treatments were prepared as follows: The relevant isolates were grown overnight in Lysogeny broth (LB) and 10 mg/mL Nitrofurantoin (antibiotic in which all isolated

410 *Pseudomonas* can grow), diluted 1:10 in the following morning and grown for 3 additional hours
411 until they entered log phase. Subsequently, bacteria were pelleted at 3500 g, resuspended in 10
412 mM $MgSO_4$, pelleted again at 3500 g to wash residual LB, and resuspended again in 10 mM
413 $MgSO_4$ to a concentration of OD_{600}=0.02, creating a stock solution for each isolate for
414 subsequent mixtures (with either 10 mM $MgSO_4$ or a solution containing another strain, in both
415 cases resulting in OD_{600}=0.01 per strain).
416 All plants were treated with the relevant treatment ten days post-stratification. Infections
417 were done by drip-inoculation, pipetting 100 µl onto the whole rosette, and plates were
418 subsequently sealed using a Micropore tape (3M, Germany). The plants were photographed on
419 the day of infection (before the actual infection), one day post-infection (day 1), and consecutively
420 from day 4 to day 7 post-infection, using a tripod-mounted Canon PowerShot G12 digital camera
421 and built-in flash. Lids kept close to maintain sterility, while a backlight was used to avoid
422 reflections due to the flash light. Green pixels approximation was done using the same pipeline
423 as Karasov etl al. [42]. In brief, individual plants were segmented from the background using lab
424 color space thresholds, followed by morphological-based noise removal. Lastly, GrabCut
425 postprocessing was applied, resulting in a list of plant IDs and the corresponding green pixel
426 count. The script was written in Python 3.6, bash using OpenVN 3.1.0 and scikit-image 0.13.0.
427
428 **Bacterial gene knockout**
429 Deletion of native bacterial genes was done on the strain 'p5.F2' [16]. First, 300bp flanking
430 sequences of the relevant genes were extracted (p5.F2 full genome was previously published
431 [16]). Subsequently, restriction sites were added between the two 300bp flanking sequences
432 (XhoI / AvrII / PmeI / PmlI), and the resulting sequence was synthesized and cloned into
433 pDEST2T18ms vector (https://www.addgene.org/72647/) by Twist Bioscience (USA; Full
434 sequence list in Table S3).
435 Gentamicin resistance gene (GmR) was amplified using the template pUC18-mini-Tn7T-
436 Gm-lux (https://www.addgene.org/64963/), using the primers 5' cccgagctcatgcatgatcg 3' and 5'
437 ccggacgatcgaattgggg 3', and the reaction consisted of 25 µL containing 0.25 µL Q5 high-fidelity
438 DNA polymerase (New England Biolabs, USA), 1x Q5 5x reaction buffer, 0.08 µM forward and
439 0.16 µM of reverse (tagging) primer and 200 µM dNTP (PCR was run for 98°C for 30 s, followed
440 by 30 cycles of 98°C for 15 s, 58°C for 20 s, 72°C for 2 min, and a final 72°C for 2 min). The
441 amplified fragment was gel-purified (GeneJET Gel Extraction Kit; Thermo Scientific, USA), and

ligated between the 300bp flanking regions, utilizing the added PmeI restriction site and T4 DNA-Ligase (Thermo Fisher Scientific, USA), and following the standard restriction-ligation protocol as instructed by the manufacturer. Thus, the resulting fragment for each gene was 5' 300bp_upstream - GmR - 300bp_downstream 3', cloned in a pDEST2T18ms vector. DH5α competent *E.coli* were transformed with the ligation product, and subsequently plated on selective Lysogeny broth (LB) agar (1.75%) with gentamicin (10 ug/mL) and tetracycline (5 ng/mL). Bacterial colonies were picked and used as a template for a colony-PCR, using primers matching the foreign DNA (Full primer list in Table S4), to validate successful transformants. PCR reaction consisted of 25 µL containing 0.25 µL Q5 high-fidelity DNA polymerase (New England Biolabs, USA), 1x Q5 5x reaction buffer, 0.08 µM forward and 0.16 µM of reverse (tagging) primer and 200 µM dNTP (PCR was run for 98°C for 30 s, followed by 30 cycles of 98°C for 30 s, matching annealing temperature (Table S4) for 30 s, 72°C for 2 min, and a final 72°C for 2 min). The same conditions applied to all following colony PCRs.

Pseudomonas strain p5.F2 (i.e. 'recipient') was subjected to triparental mating with the positive DH5α colonies, carrying the pDEST2T18ms with the relevant insert for each gene (i.e. 'donor'), and a helper HB101 E. coli, carrying the plasmid pRK2013 that facilitates mobilization. Cultures of the donor, recipient and helper strains were grown overnight (shaking, 37°C for donor and helper, 28°C for the recipient) in LB with the relevant antibiotics (gentamicin 10 ug/mL and tetracycline 5 ng/mL for the donor, 100 ug/ml nitrofurantoin for the recipient and kanamycin 50 ug/ml for the helper). In the following morning, cultures were diluted 1:10, and were grown for additional 2-4 hours. The three strains were then mixed in a single tube, followed by immediate centrifugation (3000 g), removal of supernatant and the resuspension in 1 ml LB. The same washing procedure was used again, only that the cells were resuspended in 100 ul LB. The mixture of cells was inoculated onto LB plates to allow mating, and incubated for 72 hours in 28°C. The resulting blob was scraped off into 500 ul 10mM $MgSO_4$ and centrifuged for 2 minutes at 3500 g. Supernatant was removed, and the cells were resuspended in 1 ml of 10mM $MgSO_4$. A 100ul of the resulting mixture was plated onto LB-agar selective media including 100 ul/mL nitrofurantoin, gentamicin 10 ug/mL and tetracycline 5 ng/mL. The plates were incubated at 28°C for 2-3 days until visible colonies appeared. Colonies were conducted to PCR with matching primers to validate successful p5.F2 transformants, and to validate the absence of the native gene (should be replaced by GmR due to homologous recombination of the 300bp flanking arms; Full primer list in **Table S4**). Successful knockout were considered after the expected band size

474 appeared in gel electrophoresis (Table S4), and after the amplified fragment was gel-purified
475 (GeneJET Gel Extraction Kit; Thermo Scientific, USA) and Sanger-sequenced, presenting the
476 expected GmR sequence. Subsequently, successful colonies were plated in 5% sucrose and
477 gentamicin 10 ug/mL media, to remove the pDEST2T18ms plasmid, utilizing a sucrose
478 counterselection marker (SacB) found only in the plasmid. Finally, visible colonies were examined
479 for the absence of SacB, following PCR (same reaction and conditions as for GmR PCR, primer
480 list in **Table S4**) and gel electrophoresis, and re-examined for the absence of the native gene
481 and for the presence of GmR. Successful candidates were stocked in 25% glycerol and kept in
482 -80°C.
483
484
485 **Iron sensitivity assay**
486 Growth sensitivity to iron level was assayed as previously described [43]. Briefly, each strain was
487 grown in 200 ul, LB in the presence of increasing 2,2'-dipyridyl concentration, ranging between
488 0-300 nM. Bacteria were diluted to a starting concentration of OD_{600} = 0.05, in 96-wells format
489 plate (Greiner Bio One, Austria). Plates were incubated in a plate reader (Robot Tecan Infinite
490 M200; Tecan Life Sciences, Switzerland) at 28 °C while shaking for 10 hours, and OD_{600} was
491 measured in 30 minutes intervals.
492
493 **Phylogenetic analysis**
494 All phylogenies were done based on the core genome (ortholog presence > 70%) of the relevant
495 strains in the analysis, and were constructed with IQtree (v1.6.10; using the parameters: -mset
496 LG -st AA -nt 16 -ntmax 16) [44], which analyse phylogeny based on maximum likelihood.
497 Visualization was done by iTOL [45], except for the combined tree-heatmap plots, in which the
498 R function 'ggtree' (from the package 'ggtree') was used [46].
499
500 **Regression analysis**
501 Posterior distributions of the relevant predictors were approximated using the function
502 "stan_glm" in the R package rstanarm [47]. Default settings were used (including prior
503 distribution), unless more iterations were required (Rhat > 1.1), hence the number of iterations
504 was increased until a sufficient number reached (Rhat < 1.1). In all figures, the median and 95%
505 credible intervals (2.5% and 97.5%) of the posterior distribution were presented. The response

506 variable and predictors, as well as the reference for comparison, are explained in the figure
507 legend for each analysis.
508
509 **Bacterial genome-wide association with plant protection**
510 Gene orthology clusters were retrieved from Karsov etl al. [16], as well as the phyletic pattern of
511 gene presence/absence (P/A) for each strain used in this study. Gene orthologs were then
512 associated with plant growth using treeWAS [19], a dedicated tool for bacterial GWAS.
513 Four different sets of strains were used, in four separate analyses: All strains, ATUE2,
514 ATUE3 and ATUE4. In each set the resulting maximum likelihood phylogeny was used in
515 treeWAS. To gain higher sensitivity, each set of strains was analyzed using four different metrics,
516 manifesting plant growth. All significant hits per set of strains were conjugated, and the number
517 of tests in which an ortholog resulted as significant was counted, and is indicated in **Table S2**.
518 The four plant growth metrics that were used were: The median change in green pixels between
519 the last day of experiment (7 days post-infection) to the day of infection ('median growth'), a
520 bayesian-derived approximation of the 'median growth' ('cdl50', i.e. the posterior distribution
521 median of the 'median growth'), a binary categorization based on cdl50 (cdl =< 0 - i.e. zero
522 growth, and cdl > 0 - i.e. any growth) and the area under the curve while accounting for all
523 sampled time points (0-1, 4-7 days post-infection), using the value 'auc_l' after running the R
524 function 'growthcurver' [48]. The four metrics that were used are also detailed in the legend of
525 **Table S2**.
526 Heatmaps were produced and combined with phylogenetic trees using the R function
527 'gheatmap' (from the package 'ggtree') [46].
528 To test for proximity between significant genes (i.e. formation of genomic islands), we
529 examined the location of the corresponding genes in strains that presented them all. We
530 accounted for the gene ID provided by Prokka [49], that includes the order of genes within a
531 contig (e.g. gene_1 and gene_3 are separated on contig_X only by gene_2). Only genes that were
532 found on the same contig, and separated by no more than one gene were considered as
533 consecutive. Thus, clusters of proximal genes were formed in each strain. We examined for
534 disagreements between the strains, and could find no such disagreements. We further validated
535 the proximity of these genes by examining the number of bases differing each consecutive pair
536 of genes, finding individuals to hundreds.
537

Gene orthologs annotation

To assign function to each of the studied proteins we first used the MSA of the protein orthology group to build a representative HMM model using HMMER (version 3.2.1). Next, the resulting HMMs were used to search for homologues proteins using hmmsearch against the UniProt Knowledgebase (Swiss-Prot+TrEMBL); release 2020_05. The function of the best significant hit, which covered at least 80% of the query protein, was transferred to the query protein.

Analysis of growth in iron-sensitivity assay

Growth of all bacteria was analyzed using the R function 'SummarizeGrowthByPlate' from the Growthcurver R package[48]. The logarithmic area under the curve ('auc_l') was extracted and compared between the WT, knockout strains of interest and the pathogen.

Statistical analysis

All statistical analyses were performed using the R environment version 3.5.1, unless mentioned otherwise. Sample sizes were not predetermined using statistical methods.

Acknowledgments

This work was supported by fellowships from DAAD (O.S.) and Alexander von Humboldt Foundation (H.A.), and by DFG SPP DECRyPT and the Max Planck Society.

Data and materials availability

Knockout mutants made in this study can be obtained upon request.

Competing Interests

The authors declare no competing interests.

Author Contributions

O.S. conceived the research. O.S. performed the experiments and analyzed the results. M.N. generated the knockout mutants. I.B. fine-tuned the image analysis algorithm for this study. O.S.,

H.A. and D.W. discussed and interpreted the results of this study. O.S. wrote the first draft and the manuscript was written by O.S. and D.W, with input from all authors.

References

1. Galloway BT. Plant Pathology; A Review of the Development of the Science in the United States. Agric Hist. 1928;2: 49–60.
2. Strange RN, Scott PR. Plant disease: a threat to global food security. Annu Rev Phytopathol. 2005;43: 83–116.
3. Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. Science. 2013;341: 746–751.
4. Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay Between Innate Immunity and the Plant Microbiota. Annu Rev Phytopathol. 2017;55: 565–589.
5. Durán P, Thiergart T, Garrido-Oter R, Agler M, Kemen E, Schulze-Lefert P, et al. Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. Cell. 2018;175: 973–983.e14.
6. Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. Trends Plant Sci. 2012;17: 478–486.
7. Compant S, Samad A, Faist H, Sessitsch A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. J Advert Res. 2019;19: 29–37.
8. Haas D, Défago G. Biological control of soil-borne pathogens by fluorescent pseudomonads. Nat Rev Microbiol. 2005;3: 307–319.
9. Innerebner G, Knief C, Vorholt JA. Protection of Arabidopsis thaliana against leaf-pathogenic Pseudomonas syringae by Sphingomonas strains in a controlled model system. Appl Environ Microbiol. 2011;77: 3202–3210.
10. Kamle M, Borah R, Bora H, Jaiswal AK, Singh RK, Kumar P. Systemic Acquired Resistance (SAR) and Induced Systemic Resistance (ISR): Role and Mechanism of Action Against Phytopathogens. In: Hesham AE-L, Upadhyay RS, Sharma GD, Manoharachary C, Gupta VK, editors. Fungal Biotechnology and Bioengineering. Cham: Springer International Publishing; 2020. pp. 457–470.
11. Spadaro D, Droby S. Development of biocontrol products for postharvest diseases of fruit: The importance of elucidating the mechanisms of action of yeast antagonists. Trends Food Sci Technol. 2016;47: 39–49.
12. Raaijmakers JM, Sluis L van der, Bakker PAHM, Schippers B, Koster M, Weisbeek PJ. Utilization of heterologous siderophores and rhizosphere competence of fluorescent Pseudomonas spp. Can J Microbiol. 1995;41: 126–135.

599 13. Raaijmakers JM, Mazzola M. Diversity and natural functions of antibiotics produced by
600 beneficial and plant pathogenic bacteria. Annu Rev Phytopathol. 2012;50: 403–424.

601 14. Bernal P, Allsopp LP, Filloux A, Llamas MA. The Pseudomonas putida T6SS is a plant
602 warden against phytopathogens. ISME J. 2017;11: 972–987.

603 15. Gu S, Wei Z, Shao Z, Friman V-P, Cao K, Yang T, et al. Competition for iron drives
604 phytopathogen control by natural rhizosphere microbiomes. Nat Microbiol. 2020;5: 1002–
605 1010.

606 16. Karasov TL, Almario J, Friedemann C, Ding W, Giolai M, Heavens D, et al. Arabidopsis
607 thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary
608 Timescales. Cell Host Microbe. 2018;24: 168–179.e4.

609 17. Shalev O, Karasov TL, Lundberg DS, Ashkenazy H, Weigel D. Protective host-dependent
610 antagonism among Pseudomonas in the Arabidopsis phyllosphere. bioRxiv. 2021. p.
611 2021.04.08.438928. doi:10.1101/2021.04.08.438928

612 18. Humphrey PT, Satterlee TT, Whiteman NK. Competitive hierarchies, antibiosis, and the
613 distribution of bacterial life history traits in a microbiome*. bioRxiv. 2020. p.
614 2020.06.15.152272. doi:10.1101/2020.06.15.152272

615 19. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies
616 in microbes that accounts for population structure and recombination. PLoS Comput Biol.
617 2018;14: e1005958.

618 20. Didelot X, Lawson D, Darling A, Falush D. Inference of homologous recombination in
619 bacteria using whole-genome sequences. Genetics. 2010;186: 1435–1449.

620 21. Segev-Zarko L-A, Kapach G, Josten M, Klug YA, Sahl H-G, Shai Y. Deficient Lipid A
621 Remodeling by the arnB Gene Promotes Biofilm Formation in Antimicrobial Peptide
622 Susceptible Pseudomonas aeruginosa. Biochemistry. 2018;57: 2024–2034.

623 22. Shultis DD, Purdy MD, Banchs CN, Wiener MC. Outer membrane active transport:
624 structure of the BtuB:TonB complex. Science. 2006;312: 1396–1399.

625 23. Braun V, Hantke K. Recent insights into iron import by bacteria. Curr Opin Chem Biol.
626 2011;15: 328–334.

627 24. Van Hove B, Staudenmaier H, Braun V. Novel two-component transmembrane
628 transcription control: regulation of iron dicitrate transport in Escherichia coli K-12. J
629 Bacteriol. 1990;172: 6749–6758.

630 25. Mahren S, Braun V. The FecI extracytoplasmic-function sigma factor of Escherichia coli
631 interacts with the beta' subunit of RNA polymerase. J Bacteriol. 2003;185: 1796–1802.

632 26. Cordero OX, Ventouras L-A, DeLong EF, Polz MF. Public good dynamics drive evolution of
633 iron acquisition strategies in natural bacterioplankton populations. Proc Natl Acad Sci U S
634 A. 2012;109: 20059–20064.

635 27. Andersen SB, Marvig RL, Molin S, Krogh Johansen H, Griffin AS. Long-term social

636 dynamics drive loss of function in pathogenic bacteria. Proc Natl Acad Sci U S A.
637 2015;112: 10756–10761.

638 28. Barber MF, Elde NC. Buried Treasure: Evolutionary Perspectives on Microbial Iron Piracy.
639 Trends Genet. 2015;31: 627–636.

640 29. Kramer J, Özkaya Ö, Kümmerli R. Bacterial siderophores in community and host
641 interactions. Nat Rev Microbiol. 2020;18: 152–163.

642 30. Expert D, Franza T, Dellagi A. Iron in Plant–Pathogen Interactions. In: Expert D, O'Brian
643 MR, editors. Molecular Aspects of Iron Metabolism in Pathogenic and Symbiotic Plant-
644 Microbe Associations. Dordrecht: Springer Netherlands; 2012. pp. 7–39.

645 31. Aznar A, Chen NWG, Thomine S, Dellagi A. Immunity to plant pathogens and iron
646 homeostasis. Plant Sci. 2015;240: 90–97.

647 32. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. FEMS
648 Microbiol Rev. 2003;27: 215–237.

649 33. Emerson D, Roden E, Twining BS. The microbial ferrous wheel: iron cycling in terrestrial,
650 freshwater, and marine environments. Front Microbiol. 2012;3: 383.

651 34. Niehus R, Picot A, Oliveira NM, Mitri S, Foster KR. The evolution of siderophore production
652 as a competitive trait. Evolution. 2017;71: 1443–1455.

653 35. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an
654 emergent form of bacterial life. Nat Rev Microbiol. 2016;14: 563–575.

655 36. Santhanam R, Luu VT, Weinhold A, Goldberg J, Oh Y, Baldwin IT. Native root-associated
656 bacteria rescue a plant from a sudden-wilt disease that emerged during continuous
657 cropping. Proc Natl Acad Sci U S A. 2015;112: E5013–20.

658 37. Kwak M-J, Kong HG, Choi K, Kwon S-K, Song JY, Lee J, et al. Rhizosphere microbiome
659 structure alters to enable wilt resistance in tomato. Nat Biotechnol. 2018.
660 doi:10.1038/nbt.4232

661 38. Cordovez V, Dini-Andreote F, Carrión VJ, Raaijmakers JM. Ecology and Evolution of Plant
662 Microbiomes. Annu Rev Microbiol. 2019;73: 69–88.

663 39. Herrera Paredes S, Gao T, Law TF, Finkel OM, Mucyn T, Teixeira PJPL, et al. Design of
664 synthetic bacterial communities for predictable plant phenotypes. PLoS Biol. 2018;16:
665 e2003962.

666 40. Parnell JJ, Berka R, Young HA, Sturino JM, Kang Y, Barnhart DM, et al. From the Lab to
667 the Farm: An Industrial Perspective of Plant Beneficial Microorganisms. Front Plant Sci.
668 2016;7: 1110.

669 41. Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, et al. Local-scale
670 patterns of genetic variability, outcrossing, and spatial structure in natural stands of
671 Arabidopsis thaliana. PLoS Genet. 2010;6: e1000890.

42. Karasov TL, Neumann M, Duque-Jaramillo A, Kersten S, Bezrukov I, Schröppel B, et al. The relationship between microbial biomass and disease in the Arabidopsis thaliana phyllosphere. Cold Spring Harbor Laboratory. 2019. p. 828814. doi:10.1101/828814

43. Watson RJ, Millichap P, Joyce SA, Reynolds S, Clarke DJ. The role of iron uptake in pathogenicity and symbiosis in Photorhabdus luminescens TT01. BMC Microbiol. 2010;10: 177.

44. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32: 268–274.

45. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44: W242–5.

46. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. McInerny G, editor. Methods Ecol Evol. 2017;8: 28–36.

47. Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via Stan. 2020. Available: https://mc-stan.org/rstanarm

48. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics. 2016;17: 172.

49. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30: 2068–2069.

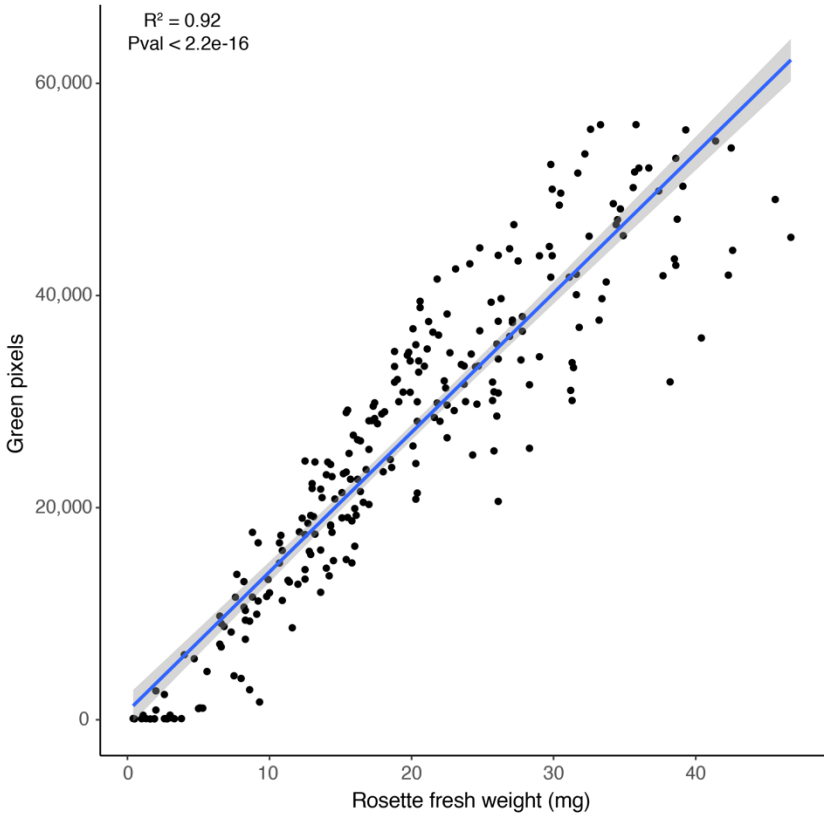


Fig S1. Correlation between green pixels count and rosette fresh weight. Plants were randomly sampled 7 days post-infection (end of experiment) for weight measurement. Green pixels count retrieved from the same day of sampling. Blue line indicates the regression line, and the shaded area indicates 95% confidence interval. n=251.

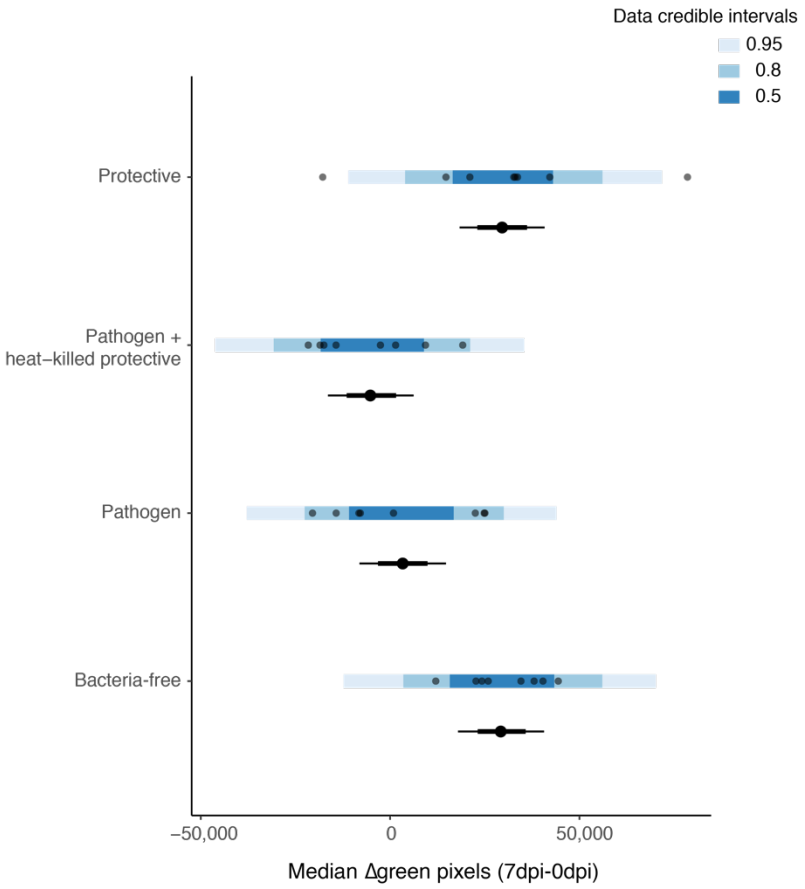


Fig S2. Plant growth in control treatments. Plant growth in the control treatments: Bacteria-free, pathogen and co-infections of the pathogen with the protective strain and heat-killed protective strain. Growth was measured by the median change in green pixels between day 7 post-infection to the day of infection. In each treatment, raw data is described by the dots at the top panel, and the shades of blue indicate data credible intervals. The bottom panel of each treatment manifests the mean growth: Dot indicates the median of the credible interval, while the vertical line indicates 95% credible interval. n=8.

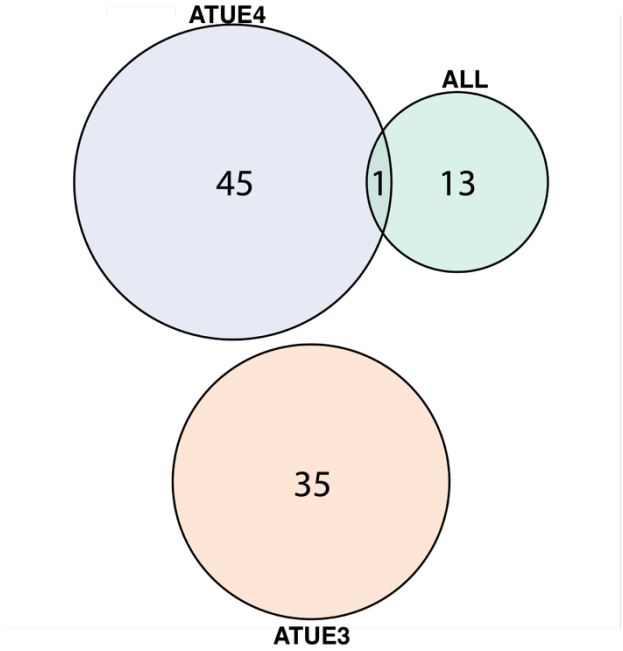
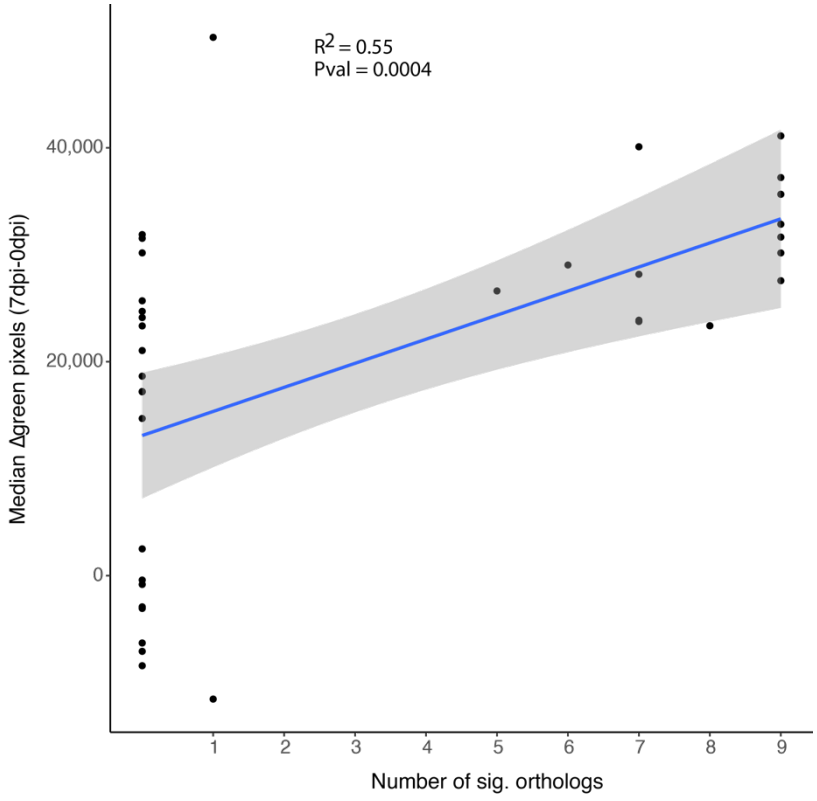


Fig S3. Venn diagram of significant gene genes in three strain sets. Significant genes, as determined by treeWAS analysis[19] after including all strains, and the subsets of ATUE2, ATUE3 and ATUE4. Four analyses were conducted, using four different metrics to manifest plant growth (**Table S2**; Methods). Only significant hits with positive Spearman correlation (>0) were considered. No significant hit with positive correlation was found after analyzing the ATUE2 subset.



745
746
747 **Fig S4. Correlation between median plant growth and the number of significant genes presented in**
748 **a given strain, in the ATUE2 subset.** Plant growth was measured by the median change in green pixels
749 between day 7 post infection to the day of infection. The curated gene gene set was considered, as
750 presented in figure 3A. Blue line indicates the regression line, and the shaded area indicates 95%
751 confidence interval. n=8 per each of 36 strains.
752
753

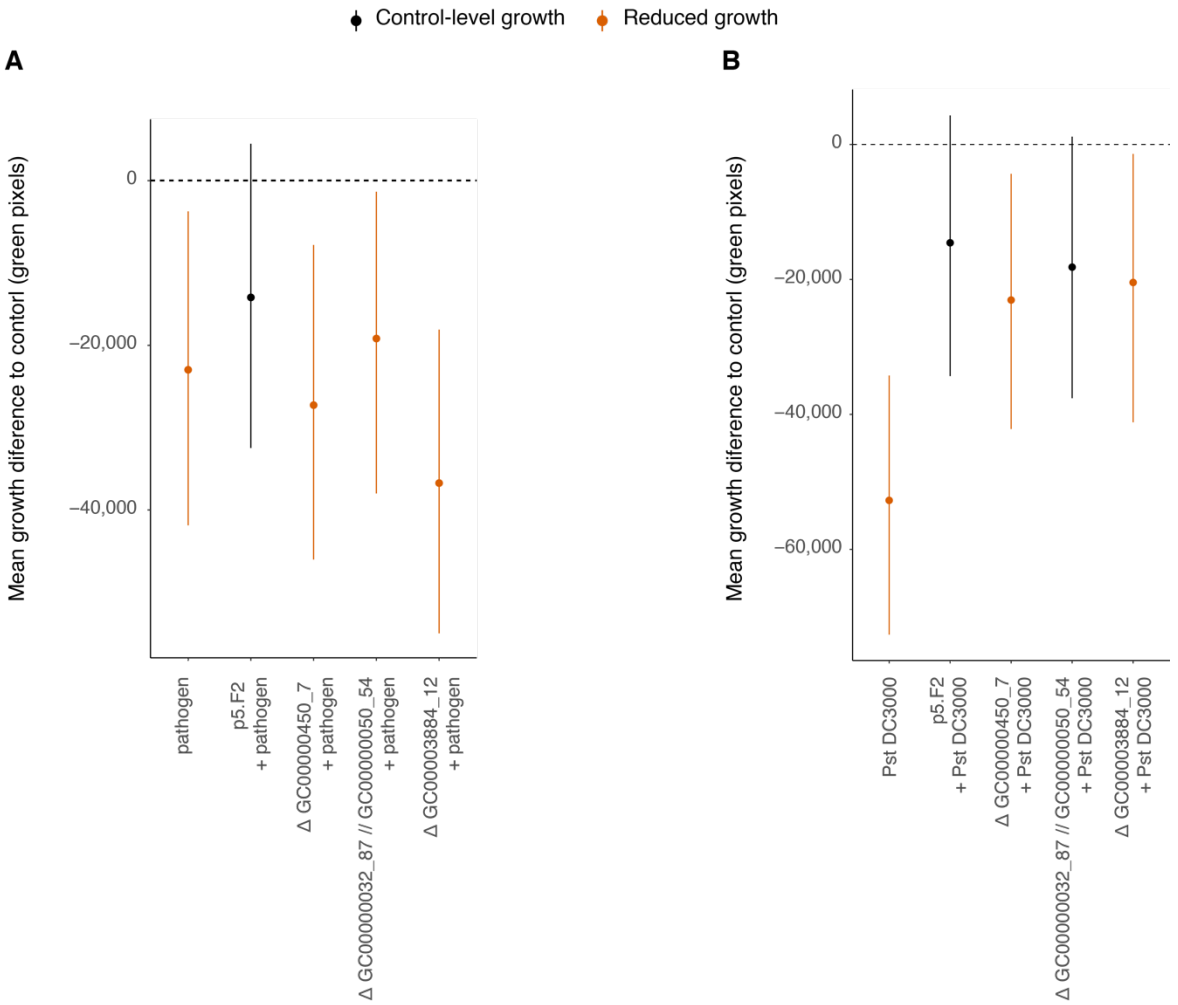


Fig S5. Co-infections of the wild type p5.F2, and each of three tested knockout mutants with (A) ATUE5 pathogen and (B) *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*). Presented is the mean plant growth difference to control (bacteria-free treatment), signified by the dashed horizontal line. Growth was measured by the median change in green pixels between day 7 post infection to the day of infection. Vertical lines indicate 95% credible intervals of the mean, while dots indicate the median. Colors indicate the growth difference of co-infections to control, as determined by the overlap between the baseline and 95% credible intervals. Experiment performed two times with similar results. n=20. Related to Figure 4.

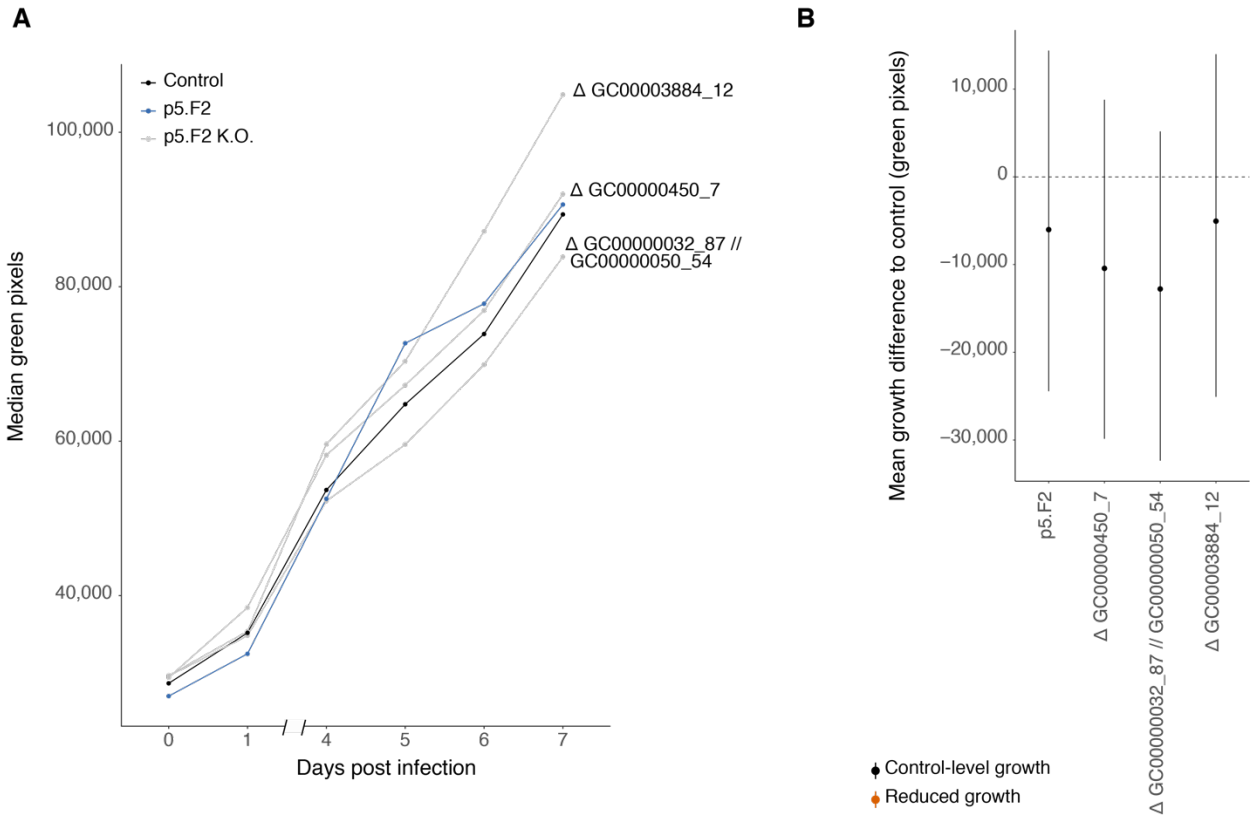


Fig S6. Individual infections with the wild type and three non-protective knockout mutants. A. Daily median of plant green pixels after treatment with control, p5.F2 and the three non-protective knockout mutants ('p5.F2 K.O.'). Plants were assessed daily by imaging. n=20. **B.** Mean plant growth difference to control after treating plants with the protective strain p5.F2 and the three p5.F2 knockout that lost their protective ability ('Δ GC00000032_87 // GC00000050_54', 'Δ GC00000450_7' and 'Δ GC00003884_12'). Growth was measured by the median change in green pixels between day 7 post infection to the day of infection. Vertical lines indicate 95% credible intervals of the mean, while dots indicate the median. The dashed horizontal line signifies the baseline, which manifests plant growth after bacteria-free treatment (i.e. 'control'). Colors indicate the growth difference of co-infections to control, as determined by the overlap between the baseline and 95% credible intervals. n=20. Experiments were performed two times with similar results.

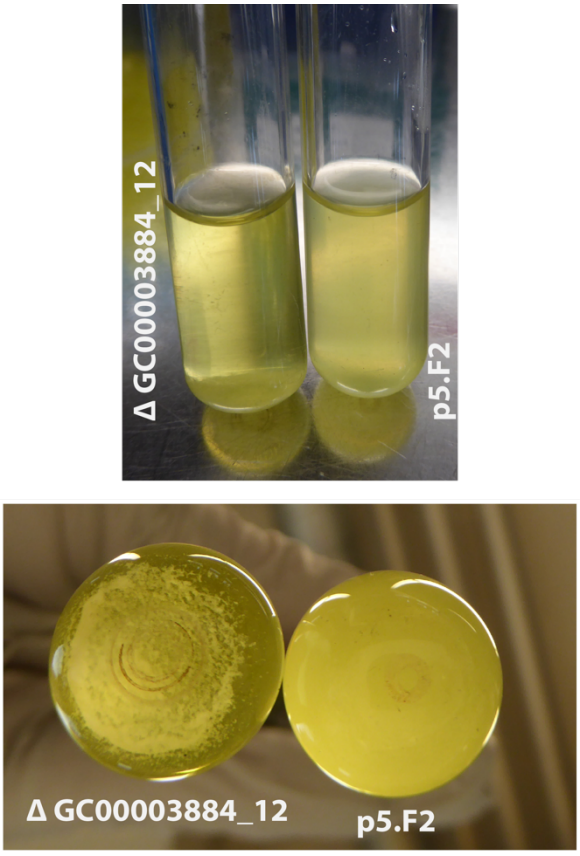


Fig S7. Representative images of LB-grown p5.F2 and its knockout mutant ΔGC00003884_12, one hour after static incubation. Experiment performed independently three times with similar results.

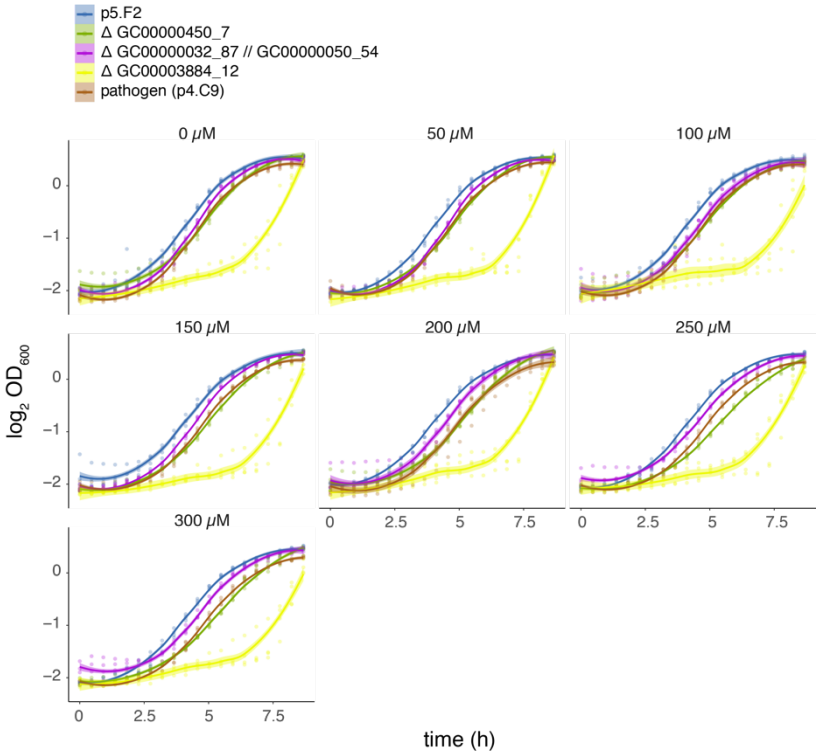


Fig S8. Growth curves of the wild type strain p5.F2, focal three knockout mutants and the pathogen in increasing 2,2'-bipyridine concentration. Bacteria were grown in LB medium, supplemented with seven different 2,2'-bipyridine concentrations (0, 50, 100, 150, 200, 250 and 300 nM). Growth was monitored for 10 hours in 30 minutes intervals. Shaded area indicates 95% confidence intervals of the regression curve. n=4 per strain.

Fig S9. Regression analysis of bacterial growth as a function of 2,2'-bipyridine concentration. The growth of the two p5.F2 knockout mutants ('Δ GC00000032_87 // GC00000050_54' and 'Δ GC00000450_7') and the pathogen were compared to the wild type protective strain p5.F2. The logistic area under the growth curve was extracted as a proxy for bacterial growth. The linear model {growth ~ strain*chelator} was analyzed, while 'chelator' states 2,2'-bipyridine concentration. **A.** Growth difference without supplementation of 2,2'-bipyridine (thus, the difference to p5.F2 of the variable 'strain' while 'chelator' is zero, i.e. the intersection with the y axis). **B.** Difference in the effect of 2,2'-bipyridine on growth, by strain (thus, the difference to p5.F2 of the interaction between the variables 'strain' and 'chelator', i.e. the slope of each strain). Vertical lines indicate 95% credible intervals of the mean, while dots indicate the median. Dashed horizontal lines signify the baseline, which manifests the wild type strain p5.F2. Colors indicate the difference to p5.F2, as determined by the overlap between the baseline and 95% credible intervals. n=4 PER STRAIN. related to **figure 4B**.